

## Autoregressive Modeling for Trend Fitting and Forecasting For the Gold Price Rate

P.Rajalakshmi. M. Sc; M. Phil; Dr. Lilly George .M. Sc; M. Phil, Ph .D

(Research scholar, Department of statistics, St. Joseph's college (Autonomous), Trichy, India)

(Asst Professor, Department of statistics, St. Joseph's college (Autonomous), Trichy, India)

**Abstracts:**

To select a suitable model for the gold rate data

To predicts the nature of gold rate next some years

**Data collected:** The data is secondary data collected the data shows trend in the Price History of Gold (10gms) over the given period of time. Given below is year wise Price of Gold in India in last 88 year. the data is collected year wise in 1925-2012.

**Keywords:**

**Statistical tools:** Trend line, Forecasting with auto regressive model, ARIMA model

**Software:** SPSS (Statistical packages for social science)

### I. Introduction

**Forecasting with autoregressive models:**

Autoregressive models use the possible autocorrelation between the time series values to predict values of the time series in future periods. Particularly, a first order autocorrelation refers to the magnitude of association between consecutive observations in the time series. The time series autoregressive model that expresses first order correlation is

$$Y_t = B_0 + B_1 Y_{t-1} + \epsilon_t$$

A p-order autocorrelation refers to the size of correlation between values p periods apart. The autoregressive model that expresses a p-order time series is

$$Y_t = B_0 + B_1 Y_{t-1} + B_2 Y_{t-2} + \dots + B_p Y_{t-p} + \epsilon_t$$

The autoregressive forecasting model takes advantage of all the information within p periods apart from period t in order to build a forecast for period t+1. The coefficients  $B_0, B_1, B_{t-1}, \dots, B_{t-p+1}$  are estimated using regression analysis. Note that the time series itself is used multiple times as “independent variables”. The following chart demonstrates the use of the time series to perform the regression analysis. Note how the data is organized.

$Y_t$	$Y_{t-1}$	$Y_{t-2}$	$Y_{t-3}$
a			
b	a		
c	b	a	
d	c	b	a
e	d	c	b
f	e	d	c

In this chart the regression can be run from the fourth row down; therefore the first three rows of information are lost. This model tries to predict the value of  $Y_t$  using  $Y_{t-1}, Y_{t-2},$  and  $Y_{t-3}$  as explanatory variables (for example,  $Y_5 = e$  in the fourth row is explained by the values  $Y_4 = d, Y_3 = c,$  and  $Y_2 = b$ ).

**Determining the number of lagged variables:**

The problem of losing some information as demonstrated above when running an autoregressive model might become serious if the data set is not large while the order ‘p’ used is. One does not want to include too many period in a model if they do not contribute significant information. We can use several methods to discover the periods relevant to the autoregressive model.

- Observing the autocorrelation matrix
- Testing the autocorrelation between  $Y_t$  and  $Y_{t-p}$

○ Observing the autoregressive model

We select periods that show high correlation to  $Y_t$ . If the correlation between  $Y_{t-p}$  and  $Y_t$  is low, the period t-p is eliminated from the data and the correlation matrix is recalculated, this time with respect to the first t-p-1 columns. We stop the elimination process when all the periods are sufficiently correlated with  $Y_t$ .

## **II. Review of Literature**

So far the nature, advantages and limitations of different time-series models generally adopted have been critically reviewed. Now comparison with reference to forecast accuracy of different models and their applications in the past are discussed below.

Christ (1951) in his work indicated that when structural changes are taking place in the economy, econometric models are not superior to timeseries approaches. 3.1 Application of Forecasting Models in the Past Gerra (1959) presented a series of behaviour relations and identities which were believed to stimulate the basic economic system for the egg industry. He indicated that in using the equations fitted (an econometric model) to forecast values of variables in the egg industry beyond the years for which equations were fitted, better estimates of the annual quantity variable (domestic egg consumption, egg production on farms, average number of layers on farms, and the number sold) were obtained from simultaneous equation approach, while better estimates for some variables like storage movement and price variables were obtained by least square method.

Suits (1962) while presenting an econometric model of the U.S. economy demonstrated its use as a forecasting instrument and explored its implications for policy analysis. He divided the presentation in to two parts. Part -I deals with the general nature o f econometric models using a highly simplified schematic example, illustrating how forecasts were made with a model, how a model could be modified to permit the introduction of additional information and judgment, and how short -run and long-run policy multipliers were derived from the inverse of the model. Pat-II presented 32 equation in economics.

Bluestone (1963) applied traditional methods of seasonal adjustment and data filtering to estimate broiler cycles. He employed the Bureau of census method for seasonal adjustment and found what he called "4 welldefined cycles". For broiler chick placements a cycle of approximately 27 months is needed and for broiler prices a cycle of approximately 30 months was suggested. Peaks in broiler prices by about 21 months at these cycles. Tobin and Arthur (1964) used a low-pass filter (simple Moving Average) of six months length for broiler chick prices and of twelve months length for hatchery supply flocks. The resulting filtered series revealed cycles of approximately 30 months for both series. A time difference of 12 to 18 months existed between the peaks of the two series for the 30 months cycle. Similar findings were also observed by Gross and Ray.

### **Comparison among Time-Series Models**

Lirby (1966) compared three different time-series methods viz. moving averages, exponential smoothing, and regression. He found that in terms of month-to-month forecasting, horizon was increased to six months. The regression models included was found to be the best method for longer term forecasts of one year or more. Sleckler found that econometric models were not entirely successful in improving the accuracy in forecasting. Schmitz and watts (1970) used parametric modeling to forecast wheat yields in the United States, Canada, Australia and Argentina. The essence of this approach was that the data were used for identifying the estimation of the random components in the form of moving average and autoregressive process. It did not identify and measure the structural relationship as was attempted when forecasting with econometric models. They used exponential smoothing to forecast yields in United States and Canada. They also compared the forecasting accuracy between parametric modeling and exponential smoothing.

Leuthold et.al (1970) In their study of forecasting daily hog price and quantities used Theil's inequality coefficient for comparing the predicative accuracy of the different forecasting approaches. For price forecast to hog market they compared econometric model, random walk model, and mean model and for supply forecasts they compared econometric model, random walk model, mean model and time-series models. They concluded that the data required for time series modeling was the concerned data on the variable to be forecasted, whereas for econometric models data are needed on both the regressor and regressand. Therefore the forecasts using econometric model are slightly better than those using a stochastic non casual frame-work. Further, the cost of making slightly greater error in using the latter will be less than the additional cost involved in setting up an econometric model and collecting the data.

Chambers et.al. (1971) in their study of "how to choose the right forecasting technique" discussed time -series analysis also. They discussed the different forecasting techniques viz, qualitative method, time -series analysis, and projection (moving average, exponential smoothing, Box-Jenkins and trend projections), and casual methods (regression model, econometric model, input-output model, leading indicator and life-cycle analysis). For each method they provided description, accuracy, identification or turning points, typical application and requirement of data. They tried to explain the potential of forecasting to the manager focusing special attention on sales forecasting for products of Corning glass works as these have matured through the product life cycle. They indicated that the need to-day is not for better forecasting methods but the better application of the techniques at hand. Similar findings were also observed by Gross and Rain.

Both Reid and Newbold (1971) concluded that the Box-Jenkins approach of ARIMA models gave more accurate results than exponential smoothing or step-wise regression methods. Soliman (1971) worked out several major relationships that explained the behaviour of the United States turkey industry in 1946-66. He constructed the model that consisted of four structural equations and he also used four different estimation techniques to derive the values of the structural parameters. He further observed that no one method proved superior to another with respect to their forecasting ability against the observed data in the post 1966 period.

Kulshreshtha (1971) developed a model for predicting the monthly egg production in Canada at least six months in advance. He indicated that the production process of eggs could be studied in two alternative ways: (i) as a single-step analysis, where the egg-production of a particular month was determined by economic and technical factors which affect the profitability of the enterprise and (ii) as a multi-step analysis in which the role of the above mentioned factors was not estimated at the final state of production but rather the whole production process was divided into finite, technically feasible stages, some of which might be the decision making ones. The role of economic and technical factors was then estimated by one or more of these stages of production of eggs. He adopted the multi-step method.

Cooper (1972) concluded from his study that econometric models are not in general superior to purely mechanical (time-series) method of forecasting. Similar findings were also observed by Gross and Krampf. Myers (1972) in his study on "Combining Statistical Techniques with Economic Theory for Commodity Forecasting" concluded that in the problems associated with determining future hog prices and marketing alternative, a combination of spectral analysis, autoregression and multiple

regression analysis and recursive system could be used. In forecasting, however, a model can only be evaluated by its ability to provide adequate answers to pricing problems and also only incidentally by internal statistical properties. Naylor et al (1972) made more extensive and detailed comparison of alternative methods and examined Box-Jenkins approach in contrast to Wharton econometric model for the year 1963 through 1967. They observed that the accuracy of ARMA models of Box-Jenkins methodology was considerably better than the accuracy of Wharton econometric model.

Nelson (1972) compared econometric (regression) and time-series (ARMA) methods for a longer time horizon. He concluded that the simple ARMA models are relatively more robust with respect to post sample predictions than the complex econometric models. If the mean square error is an appropriate measure of loss, an unweighted assessment clearly indicated that a decision maker will be better off relying simply on ARMA predictions in the post sample period i.e. in the forecasting phase.

Adam (1973) reported that the factors including the number of observations in the series, seasonality of the data, the number of periods in the time horizon are to be forecasted to the extent in which randomness in the series and others had a substantial impact on the accuracy and performance of individual forecasting models.

Kimball and Gutierrez (1973) adopted Mincer-Zarnwitz technique of the goal of forecasting the minimization of the mean square error (MSE) i.e. the squared difference between the actual and forecast values, which is a measure of dispersion around the line of perfect forecast. They indicated that a least square straight line must be fitted to a scatter diagram of actual realization and estimates. They gave the idea that on the overall forecast accuracy, i.e. the square root of the MSE has been computed and expressed as a percent of the actual mean value. They also said that  $R^2$  is not a reliable guide since it merely represents errors explained by a linear adjustment of the forecast series. Forman and Klein (1973) concluded that no single econometric model is overwhelmingly superior to others. They recognized that differences may exist in the forecasting performance for single items or over a limited time horizon but on the average these differences in accuracy do not consistently favour one model to another.

Groff (1973) observed from his study that Box-Jenkins methodology gave results that were approximately equal in accuracy or slightly worse than those achieved by using exponential smoothing.

Roy and Johnson (1973) developed seven simultaneous equation systems for wholesale shell-egg price and estimated it by using three-stage least squares method. The seven-equation quarterly models for wholesale shell-egg prices indicated that all estimated relations were consistent with economic or logical expectations. Price forecasts developed from the models appeared to be satisfactory, especially for the first three quarters of the calendar year. Methodologically price forecasts based on three-stage least squares equations were more accurate than those developed from the corresponding two-stage least squares equations. They suggested that one should re-estimate the model periodically after taking into account the changes in the structural parameters. This helps a lot to maintain accuracy of prediction.

Larson (1974) applied a well defined and thoroughly-examined (theoretically) model known as Larson's harmonic model to the egg industry. The harmonic model is the result of a small but important modification of the cobweb model. Hartman (1974) examined the often discussed instability of the price of eggs by the use of spectral analysis and confirmed the existence of an egg cycle. He found that while

explaining the cycle, the familiar cobweb model is inferior to an alternative model. The alternative model proposed by Nerlove had producers planning changes in their production levels in response to prices, with the changes being realized only after a lag. He gave economic justifications for such producer behaviour to continue even though it causes instability. He further discussed recent destabilizing factors which were not related to the cycle.

New Bold and Granger (1974) concluded that Box-Jenkins' approach of ARIMA models gave more accurate results than exponential smoothing or step-wise regression methods. Leuthold (1974) observed that cash cattle prices were found to be the more accurate indicators of subsequent cash cattle prices conditions than the future prices for distant contracts. He applied an explicit method of determining the ability of future prices to estimate subsequent cash prices with the model  $F_{Pt} = a + b F_{Pt-I}$  where  $F_{Pt}$  is the closing future price for a contract at delivery and  $F_{Pt-I}$  is the future price for the  $i$ th month before maturity. He calculated mean squared error (MSE) to gain more information about the efficiency of future prices in indicating the final cash price on weekly corn and cattle prices. The same conclusions were drawn by Geurts and Ibrahim.

Cromarty and Myers (1975) in their study argued that the work done by Connel and company on egg product prices provided a good practical example of working with partial equations to forecast prices in an appropriate time period. Research for egg products had been concentrated on forecasting price levels during the spring peak period.

Asbridge and Sappington (1976) fitted a regression model using currently known data to predict the price of broilers in 11<sup>th</sup> week following the prediction. This 11<sup>th</sup> week lead is to give producers an indication of the price that they can expect when they market the broilers from the time that the fertile eggs are put into incubator.

Lakshminarayan et al. (1977) developed the following form of Box-Jenkins model.  $Z_t = Z_{t-1} + 0.70 \Delta Z_{t-12} - 0.15 \Delta Z_{t-13}$ . to forecast the broiler chicken production for the year 1976. The mean absolute percentage of error was under 5 percent while the error in the total production for the year was 1.7 percent. The forecast followed the pattern of the actual data. The actual production was always within the 50 percent confidence limits of the forecast.

Sullivan et al. (1977) used the data regarding ambulance company calls responded to by the EMS company over 12 months. He compared the estimated and original for the models, double Moving Average, Single Moving Average, Single Exponential smoothing and Double Exponential Smoothing, Triple Exponential smoothing, and Winter's Method. In his case Double Exponential Smoothing can be given the minimum mean square error and Winter's as the second. When they performed the same analysis for quarterly data for plant and equipment expenditures over seven years the winter's method gave the best with the minimum Mean Square Error. The Triple exponential Smoothing model came next.

Haynes and Keyyon (1978) developed price forecasting models to predict broiler prices in one, two and three quarters in advance. The models were turn forecasting models in that all values of independent variables are estimated or known when forecasts were made. The ability of the model to forecast was evaluated out-side the data base to estimate the equations. The results indicated that all the three models predicted were better than no-change extrapolation. The models which predict broiler price two and three quarters in advance are better for future market.

Hernandez-Estrada (1978) estimated demand parameter co-efficient for future use in price forecasting or price projection models for graded eggs. He analysed the problem with a theoretical model which postulated the existence of a kinked demand curve for graded eggs and a regression was applied to estimate the effects of several variables on this. The result of the study using 1965-76 data indicated that price of graded eggs is inversely related to the quantity of graded eggs.

Kersen (1978) used statistics on hatching eggs put into incubation in GFR and Netherlands during the base period 1966-77 to measure production decisions. Turning points, prediction accuracy of trends and deviation of predictions from reality were used to evaluate export forecasts.

### **III. Autoregressive Modeling For Trend Fitting And Forecasting For The Gold Price Data**

In order to choose a proper model to understand the pattern of gold price, we choose ARIMA model of the rate of gold.

The result of the ARIMA model given below:

**Fifth-Order Auto Regressive**

YEAR	GOLD PRICE	RECODE	LAG1	LAG2	LAG3	LAG4	LAG5
1925	18.87	0					
1926	18.43	1	18.43				
1927	18.37	2	18.37	18.43			
1928	18.37	3	18.37	18.37	18.43		
1929	18.43	4	18.43	18.37	18.37	18.43	
1930	18.05	5	18.05	18.43	18.37	18.37	18.43
1931	18.18	6	18.18	18.05	18.43	18.37	18.37
1932	23.06	7	23.06	18.18	18.05	18.43	18.37
1933	24.05	8	24.05	23.06	18.18	18.05	18.43
1934	28.81	9	28.81	24.05	23.06	18.18	18.05
1935	30.81	10	30.81	28.81	24.05	23.06	18.18
1936	29.81	11	29.81	30.81	28.81	24.05	23.06
1937	30.18	12	30.18	29.81	30.81	28.81	24.05
1938	29.93	13	29.93	30.18	29.81	30.81	28.81
1939	31.74	14	31.74	29.93	30.18	29.81	30.81
1940	36.04	15	36.04	31.74	29.93	30.18	29.81
1941	37.43	16	37.43	36.04	31.74	29.93	30.18
1942	44.05	17	44.05	37.43	36.04	31.74	29.93
1943	51.05	18	51.05	44.05	37.43	36.04	31.74
1944	52.93	19	52.93	51.05	44.05	37.43	36.04
1945	62	20	62	52.93	51.05	44.05	37.43
1946	83.87	21	83.87	62	52.93	51.05	44.05
1947	88.62	22	88.62	83.87	62	52.93	51.05
1948	95.87	23	95.87	88.62	83.87	62	52.93
1949	94.17	24	94.17	95.87	88.62	83.87	62
1950	99.18	25	99.18	94.17	95.87	88.62	83.87
1951	98.05	26	98.05	99.18	94.17	95.87	88.62
1952	76.81	27	76.81	98.05	99.18	94.17	95.87
1953	73.06	28	73.06	76.81	98.05	99.18	94.17
1954	77.75	29	77.75	73.06	76.81	98.05	99.18
1955	79.18	30	79.18	77.75	73.06	76.81	98.05
1956	90.81	31	90.81	79.18	77.75	73.06	76.81
1957	90.62	32	90.62	90.81	79.18	77.75	73.06
1958	95.38	33	95.38	90.62	90.81	79.18	77.75
1959	102.56	3	102.56	95.38	90.62	90.81	79.18
160	111.87	35	111.87	102.56	95.38	90.62	90.81
1961	119.35	36	119.35	111.87	102.56	95.38	90.62
1962	119.75	37	119.75	119.35	111.87	102.56	95.38
1963	97	38	97	119.75	119.35	111.87	102.56
1964	63.25	39	63.25	97	119.75	119.35	111.87
1965	71.75	40	71.75	63.25	97	119.75	119.35
1966	83.75	41	83.75	71.75	63.25	97	119.75

1967	102.5	42	102.5	83.75	71.75	63.25	97
1968	162	43	162	102.5	83.75	71.75	63.25
1969	176	44	176	162	102.5	83.75	71.75
1970	184.5	45	184.5	176	162	102.5	83.75
1971	193	46	193	184.5	176	162	102.5
1972	202	47	202	193	184.5	176	162
1973	278.5	48	278.5	202	193	184.5	176
1974	506	49	506	278.5	202	193	184.5
1975	540	50	540	506	278.5	202	193
1976	432	51	432	540	506	278.5	202
1977	486	52	486	432	540	506	278.5
1978	685	53	685	486	432	540	506
1979	937	54	937	685	486	432	540
1980	1330	55	1330	937	685	486	432
1981	1800	56	1800	1330	937	685	486
1982	1645	57	1645	1800	1330	937	685
1983	1800	58	1800	1645	1800	1330	937
1984	1970	59	1970	1800	1645	1800	1330
1985	2130	60	2130	1970	1800	1645	1800
1986	2140	61	2140	2130	1970	1800	1645
1987	2570	62	2570	2140	2130	1970	1800
1988	3130	63	3130	2570	2140	2130	1970
1989	3140	64	3140	3130	2570	2140	2130
1990	3200	65	3200	3140	3130	2570	2140
1991	3466	66	3466	3200	3140	3130	2570
1992	4334	67	4334	3466	3200	3140	3130
1993	4140	68	4140	4334	3466	3200	3140
1994	4598	69	4598	4140	4334	3466	3200
1995	4680	70	4680	4598	4140	4334	3466
1996	5160	71	5160	4680	4598	4140	4334
1997	4725	72	4725	5160	4680	4598	4140
1998	4045	73	4045	4725	5160	4680	4598
1999	4234	74	4234	4045	4725	5160	4680
2000	4400	75	4400	4234	4045	4725	5160
2001	4300	76	4300	4400	4234	4045	4725
2002	4990	77	4990	4300	4400	4234	4045
2003	5600	78	5600	4990	4300	4400	4234
2004	5850	79	5850	5600	4990	4300	4400
2005	7000	80	7000	5850	5600	4990	4300
2006	8400	81	8400	7000	5850	5600	4990
2007	10800	82	10800	8400	7000	5850	5600
2008	12500	83	12500	10800	8400	7000	5850
2009	14500	84	14500	12500	10800	8400	7000
2010	18500	85	18500	14500	12500	10800	8400



2011	26500	86	26500	18500	14500	12500	10800
2012	31799	87	31799	26500	18500	14500	12500

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	31.061	1.885		16.477	.000
	LAG1	-.005	.003	-1.165	-1.885	.063
	LAG2	-.001	.005	-.238	-.251	.803
	LAG3	.007	.009	1.002	.827	.411
	LAG4	.002	.010	.272	.231	.818
	LAG5	.008	.006	.862	1.285	.203
a. Dependent Variable: RECODE						

The fitted ARIMA model is

$$\hat{Y}_i = 31.061 - 0.057Y_{i-1} - 0.01Y_{i-2} + 0.007Y_{i-3} + 0.002Y_{i-4} + 0.008Y_{i-5}$$

Here we take 5th order coefficient and test for significance.

The hypotheses used for this test are,

$$\text{Test statistic } t = \frac{a_5 - A_5}{S a_5}$$

**Interpretation:**

here lag5 is not significant, since the value 0.203 greater than 0.05 level of significance. So we consider up to lag4 and test for significance.

**Fourth-Order Auto Regressive:**

$$\hat{Y}_i = 30.696 - 0.006Y_{i-1} + 0.000Y_{i-2} + 0.005Y_{i-3} + 0.012Y_{i-4}$$

Here we take 4th order coefficient and test for significance.

The hypotheses used for this test are,

$$H_0 : A_4 = 0$$

$$H_1 : A_4 \neq 0$$

$$\text{Test statistic } t = \frac{a_4 - A_4}{S a_4}$$

**Interpretation:**

here lag4 is significant, since the value is 0.05 equal to 0.05 level of significance. So conclude that the fourth-order parameter of the autoregressive model is significant and can be accepted.

The model building approach has led to the selection fourth order autoregressive model as the most appropriate for the given data.

Using the estimate  $a_0 = 30.696$   $a_1 = -0.006$   $a_2 = 0.000$   $a_3 = 0.005$   $a_4 = 0.012$ .

**3.1.1 Fitting of straight line Trend  $Y = a + bx$ .**

Let us choose a straight line trend of the form  $Y = a + bx$  to the following data.

The result are given in the following table:

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-3145.170	843.527		-3.729	.000
	code_x	132.628	16.789	.648	7.900	.000

a. Dependent Variable: goldrate

Model Summary and Parameter Estimates

Dependent Variable: gold rate

Equation	Model Summary					Parameter Estimates			
	R Square	F	df1	df2	Sig.	Constant	b1	b2	b3
Linear	.421	62.408	1	86	.000	-3.145E3	132.628		
Quadratic	.696	97.240	2	85	.000	2.522E3	-275.942	4.738	

Cubic	.815	123.075	3	84	.000	-1.719E3	348.532	-13.457	.140
-------	------	---------	---	----	------	----------	---------	---------	------

The independent variable is code\_x.

From above table, we get the linear trend line as  $\hat{Y} = 3145.170 + 132.628x$ . Based on the above fitted model, the predicted values are obtained. Thus using

$$\chi^2 \text{ analysis for testing for goodness of data, } \chi^2 = \sum \left[ \frac{(Y - \hat{Y})^2}{\hat{Y}} \right] \sim \chi^2(n-1)df$$

$$\chi^2 = (226813.74 - 226813.556)^2 / 226813.74 = 0.00000149$$

**Inference:**

From the above chi-square test the calculated value is 0.000, which is less than table value  $\chi^2$  0.05 that we conclude that the linear curve as line is not a suitable model for the gold rate data.

**3.1.2 Fitting of second degree parabola of the form**

$$Y = a + bx + cx^2$$

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2522.141	891.747		2.828	.006
	code_x	-275.942	48.155	-1.349	-5.730	.000
	Quadratic	4.738	.540	2.065	8.772	.000

From above table, we get the second degree curve line as  $\hat{Y} = 2522.141 - 275.942x + 4.738x^2$ . based on the above fitting model, the predicted values are obtained. Thus using  $\chi^2$  analysis for testing for goodness of data

$$\chi^2 = \sum \left[ \frac{(Y - \hat{Y})^2}{\hat{Y}} \right] \sim \chi^2(n-1)df$$

$$= (226813.74 - 226757.55)^2 / 226757.55 = 0.0139$$

**Inference:**

From the above chi-square test the calculated value is 0.0139, which is less than table value  $\chi^2$  0.05 that we conclude that the second degree parabola curve as line is not a suitable model for the gold rate data.

**3.1.3 Fitting of third degree cubic of the form  $Y = a + bx + cx^2 + dx^3$  :**

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1718.638	907.953		-1.893	.062
	code_x	348.532	93.127	1.704	3.743	.000
	Quadratic	-13.457	2.516	-5.866	-5.349	.000
	cubic	.140	.019	5.050	7.338	.000

a. Dependent Variable: goldrate

From above table, we get the third degree curve as  $\hat{Y} = 1718.638 + 348.532x - 13.457x^2 + 0.14x^3$ . Based on the above fitting model, the predicted values are obtained. Thus using

$$\chi^2 \text{ analysis for testing for goodness of data } \chi^2 = \sum \left[ \frac{(Y - \hat{Y})^2}{\hat{Y}} \right] \sim \chi^2(n-1)df$$

$$= 14.36$$

**Inference:**

From the above chi-square test the calculated value is 14.36, which is less than table value  $\chi^2$  0.05 that we conclude that the cubic third degree curve as line is not a suitable model for the gold rate data



#### **IV. Findings And Conclusion**

Based on the statistical analysis of selecting a proper model for the gold rate data no model such as linear, quadratic, cubic etc, is suitable for the data. Here lag4 is significant, since the value is 0.05 equal to 0.05 level of significance. So conclude that the fourth-order parameter of the autoregressive model is significant and can be accepted. The model building approach has led to the selection fourth order autoregressive model as the most appropriate for the given data.

#### **Reference**

- [1]. Fundamentals of applied statistics ,S.C.GUPTA, V.K.KAPOOR
- [2]. Statistics for management Richard I.LEVIN, David S.RUBUN, The university of north Carolina chapel hill.
- [3]. Applied statistics, Parimal mukhopadhyay(2011)
- [4]. Basic Econometric, Damodar N. Gujarati(third edition)
- [5]. Harvey A.C.(1981), Time series models, new York:john wiley sons,
- [6]. 1998,Time series models for business and economic forecasting,Cambridge university.
- [7]. Statistical methods for practice and research, Ajai S. Gaur, Sanjaya S. Gaur 2009
- [8]. [www.google.co.in](http://www.google.co.in)
- [9]. <http://en.wikipedia.org>