International Journal of Combined Research & Development (IJCRD)
eISSN:2321-225X;pISSN:2321-2241 Volume: 4; Issue: 5; May -2015

Forecasting Using Time Series Models

Dr. J Katyayani[1], M Jahnavi[2] Pothugunta Krishna Prasad [3]
[1]Professor, Department of MBA, SPMVV, Tirupati, India
[2]Assistant Professor, Koshys Institute of Management Studies,Bangalore
[3]Assistant Professor, Koshys Institute of Management Studies, Bangalore

Abstract: The importance of statistical forecasting is in its ability to discover hidden knowledge from databases. In this paper, forecasting is implemented in the BSE indices, three time series algorithms are used, namely-simple moving average forecasting, weighted average forecasting and exponential smoothing method. Data collected for three years i.e., from Jan 2010 to Dec 2013. The result showed that weighted moving is efficient in predicting BSE Sensex compared to simple moving average and exponential smoothing as far as mean absolute deviation, mean sum error, residual mean sum error and mean absolute percentile error are concerned.

1.1 Introduction:

Forecasting can be broadly considered as a method or a technique for estimating many future aspects of a business or other operations. It is one of the rapidly growing fields in business world which is mainly based on strong statistical theory and application. Over the past few years, there has been lot of research work carried out in predicting the future and making better decisions. This process of research has been led to several developments in forecasting methods. Many of these methodologies have been developed based on sophisticated statistical techniques available in the theory.

In management and administrative fields, the need for planning is vital because the lead time for decision making ranges for several years to few days or hours to few seconds. Forecasting is an important aid in effective and efficient planning. Forecasting is a common statistical task where it evaluates decisions about scheduling of several activities like production, transportation, assignments etc., provides an insight to long-term strategic planning. Forecasting is an important aid in effective and efficient planning. Forecasting is a common statistical task where it evaluates decisions about scheduling of several activities like production, transportation, assignment etc., and provides an insight to long term strategic planning. Forecasting is all about doing some predictions about the future as accurately as possible, given all the information available including historical data and knowledge of any future events that might impact the forecast. This becomes an integral part of the decision-making activities of all the fields and domains as it can play an important role in many areas like production planning, sales inventory and demand analysis.

In general, forecasting methods can be divided into two broad categories viz., qualitative and quantitative. Qualitative forecasting techniques generally employ the judgment of experts in the appropriate field to generate predictions. Quantitative forecasting methods are used when historical data on variables of interest are available. These methods are based on an analysis of historical data over a particular period of time for the specific variable of interest and possibly other related time series. Depending on the requirements and other constraints, one can choose between qualitative and quantitative methods of forecasting. The accuracy of the predicted values are mainly because of this vital selection only.

Before 1960, little empirical research was done on forecasting methods. Since then, the literature has grown rapidly, especially in the area of quantitative forecasting. There were several theories proposed in various forecasting models and models and many research studies are published based on that.

1.2 Basics of forecasting methodology:

Forecasting methodology has its basis from planning related activities in the business establishment. For better plan implementations, we require better future prediction which will be acquired through proper forecasting techniques. Each and every forecasting methodology involves the following components.

1.2.1. Objectives of forecasting:

This refers to the list of aims of the forecasting study and also the need for having the forecasting analysis.

1.2.2. List of available variables:

It is the list of variables which are associated with the current study and having some direct or indirect impact on the forecasting variables.

1.2.3. Assumptions about the variables:

Assumptions regarding the distribution, availability, minimum and maximum limits for all the variables under study must be specified.

1.2.4. Data availability in any form:

The format in which data available and scope of retrieving them to the required format is needed.

1.3 Basic steps in forecasting process:

Basic steps in forecasting are list of consecutive steps one should follow while carrying any research study. The steps list is exhaustive but it may vary from research study to domain. There are five basic steps in any forecasting problem which is outlined as follows:

1.2.1 Objective Identification:

The problem identifying right objective for the given forecasting problem is the first step in any forecasting process. It involves deep understanding of how the forecasts will be used, who is the end user of the forecasts, and how the forecasting function fits within the organization. It is always advisable to spend considerable amount of time in finalizing these objective before starting with the actual task.

A forecaster has a great deal of work to do to properly define the forecasting aims before giving the forecasting results. The research study becomes very effective only when there is perfect mapping of objectives defined and the results obtained.

1.2.2 Collecting research data:

Usually we can find two types of data related to any research study and they are (1) numerical data on facts of the research study and (2) the comments from expertise in the concerned field or area. Both kind of data needs to gathered while doing the research since combination of both kinds will give more efficient output for the research study.

In the case of numerical data, generally we use the historical data on the relevant variables for the specific period of time and which needs to complete in all respects in the case of time series data, there should not be any irregular intervals in the time variable and each time series should be complete over the specific period.

1.2.3 Pilot analysis:

The preliminary data analysis done on the actual data to study the basic characteristics of data is called pilot analysis. It involves graphing of some data variables to know their distribution, to compute some descriptive statistics associated with data variables. The purpose of doing this analysis at this stage is to get the familiarity with the data regarding its behavior, existing patterns and relationships among the data variables etc.

In many research studies, the pilot analysis helps to identify the outliers of the data variables by plotting simple charts on the data in which analyst can take decision. This stage is to avoid or modify them to get good models. Also, this analysis helps to arrive some basic decisions on the data models because we come to know about the types of relationships among the variables as well as data behavior.

1.2.4 Choosing the best model:

This step includes the selection and fitting of several quantitative forecasting models with different parameters. Based on the recommendations of pilot study, user can restrict himself from large variety of models and can fit handful of suitable models for the data.

Each fitted models is a showcase of relationships among the variables and based on certain set of explicit and implicit assumptions. A model usually involves one or more parameters which must be forecasted or predicted using the known historical data. Selection of history data and variables for consideration purely lies on the expertise of the researcher and based on the current model used for forecasting. Along with fitted forecasts, this step also involves finding the various errors like MAPE, MSE, RMSE and MAD etc., which are getting used in the selection of best models among the fitted models.

1.2.5 Selection criteria for the model selection:

Criteria of model selection are the most critical part of any research study. In the literature there are several models selection criteria are available and one has to choose very intelligently the best method of model selection.

As general practice, usage of combination of two or more selection criteria is the best option to decide about the good model. As we know that a single measure always cannot give accurate overview of performance of the model, combined usage of two criteria will give good insights on the performance of the model. In the time series data, one or two measure of errors will give the good models among the class of fitted models to the data.

1.2.6 Evaluating the forecasting model:

Once the model selection has been finalized then the process of forecasting starts actually for the experimental data and forecasts has been generated. Now, one should evaluate the selected model based on forecasts and its practical applicability in the real time. Many a time a model is good theoretically and may not works properly for the real data in which again there is an evaluation process of selected model.

In the evaluation process, actual fitted model is tested for its pros and cons in the real time applications. If there is any requirement of refining of the model raised, then the model is changed accordingly and applied to the real data. It is the fact that none of the models can predict all possible real scenarios that will arise in the real time and hence this evaluation process will help to overcome that difficulty.

Objective of the study:

1. Calculating the deviation, mean error and percentile error using three different basic forecasting techniques
2. Determination of optimum smoothing constant value
3. Determination of best forecasting technique from the basic forecasting techniques.

1.3 Quantitative forecasting time series methods:

Time series models predict on the assumption that the future is a projection of the past. They look at what has happened over a period of time and use a series of past data to make a forecast for the future.

1.3.1 Time series forecasting methods:

A time series a time-ordered sequence of observations taken at regular intervals over a period of time (eg.,a hourly, daily, weekly, monthly, quarterly or annually). The data may be measurement of demand, earning, profits, outputs, productivity, consumer price index etc.

1.3.2 Decomposition of a Time series:

Analysis of time series data requires the analyst to identify the underlying behavior of the series. This can be done by plotting the data with time on the 'X' axis, and data on 'Y-axis and visually examining the plot. One or more patterns might appear. They are trends, seasonal variations, cycles and random or irregular variations.

Trend refers to gradual, long term, upward or downward movement in the data over time. Changes in income, population, age distribution or cultural views may account for such movements.

Seasonality refers to short term, fairly regular variations related to factors such as weather, holidays, vacation etc., seasonal variations can be daily, weekly or monthly.

Cycles are wavelike variations of more than one year duration or which occur every several years. They are usually tied with business cycle related to a variety of economic, political or agricultural conditions.

1.3.3 Techniques for averaging:

Historical data contain a certain amount of random variation, or noise which tends to observe systematic movement in the data. It is desirable to completely remove any randomness from the data and leave only real variation such as changes in demand.

Averaging techniques smooth fluctuations in a time series so that the forecast can be based

Random variations are residual variations which are blips in the data caused by chance and unusual situations which cannot be predicted.

1.4.3.1 Moving Average Method:

A moving average forecast uses a number of most recent historical actual data values to generate a forecast. The moving average for 'n' number of periods in the moving average is calculated as:

$$\text{Moving average} = \frac{\sum \text{demand in previous n periods}}{N}$$

N may be 3, 4, 5 or 6 periods for 3, 4, 5 or 6 period moving averages.

The **"simple moving average method"** is used to estimate the average of a demand time series and remove the effects of random fluctuation. It is most useful when demand has no pronounced trend or seasonal fluctuation.

In the **weighted moving average method** each historical demand in the moving average can have its own weight and the sum of the weight equals to one. For *example*, in a 3 period weighted moving average model, the most recent period might be assigned a weight of 0.50, the second most recent period might be assigned a weight of 0.30 and the third most recent period with a weight of 0.20. Then forecast,

$$F_{t+1} = \frac{0.5\, D_t + 0.3\, D_{t-1} + 0.2\, D_{t-2}}{\text{Sum of weights } (0.5+0.3+0.2)}$$

The average is obtained by multiplying the weight of each period by the value for that period and by adding the products together.

The advantage of weighted moving average methods is that it allows emphasizing recent demand over earlier demand.

1.4.3.2 Exponential smoothing method:

It is a sophisticated weighted moving average method that is still relatively easy to understand and use. It requires only three items of data: this periods forecast, the actual demand for this period and α which is referred to as smoothing constant and having a value between 0 and 1. The formula used is:

Next periods forecast = This period forecast + (α*(This periods actual demand-This periods forecast))

Selecting a smoothing constant is basically a matter of judgment or trial and error. The commonly used values for α range from 0.05 to 0.5.

1.5 Numerical measures of errors:

Numerical measures of errors play an important role in diagnostic study of time series analysis. The forecast accuracy can be identified and compared by using these numerical measures of errors only. The measures are most commonly dependent on the residual which is defined as the difference between actual and predicted value. The following are some of the important measures of the numerical measures of errors.

1.5.1 Mean squared error (MSE):

The mean squared error is useful to understand how close the predicted value is from its original value. It relay the concepts of bias, precision and accuracy in statistical forecasting. The average of the square of the difference between the desired response and the actual system output is known as Mean squared error. The smaller the error, closer the forecast to the actual data. It is obtained from

$$\text{MSE} = \frac{\sum_{t=1}^{N} e_t^2}{N}$$

1.5.2 Root Mean square error: (RMSE)

The RMSE measures variance of the error. In other words the difference between forecast and corresponding observed values are squared and then averaged over the sample. Finally, the square root of the average is taken. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when we need to identify the large residuals in the data.

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{N} e_t^2}{N}}$$

1.5.3 Mean Absolute Percentage Error (MAPE):

MAPE is considered as a measure of accuracy in the time series data and it is always expressed in percentage. The absolute difference between actual and predicted value is summed and averaged to minimize the effect of bulk data. Then it is converted to percentage in its final form. The expression for MAPE is given by

$$\text{MAPE} = \frac{\sum \frac{e_t}{y_t}}{N} * 100$$

Here I calculated the forecasted value and also four different types of errors for three years data i.e., BSE S&P Sensex from Jan 2010 to Dec 2013 by using simple moving average, weighted moving average and also by exponential smoothing constant. In simple moving average the accuracy level depends on number of months past data considered for calculation of forecasted value, here we computed prediction by using 3months, 5 months, 7 months and 10 months simple moving average.

2011					2012					2013				
	3 months	5 months	7 Months	10 Months		3 months	5 months	7 months	10 months		3 months	5 months	7 months	10 months
MAD	256.1092	300.8175	340.2417	389.5204	MAD	159.3225	195.5776	340.2417	263.9527	MAD	220.5033	207.23	287.8101	329.4969
MSE	26681.04	136719.1	171046.1	222981.3	MSE	39499.57	56567.8	73312.84	102336.7	MSE	75946.43	98528.29	120412	160346
RMSE	163.3433	369.7554	413.5772	472.209	RMSE	198.745	14.09769	270.7634	319.9011	RMSE	275.5838	313.8922	347.0044	400.4322
MAPE	1.461293	1.691793	1.88491	2.225309	MAPE	0.91079	1.118305	1.279606	1.507513	MAPE	1.124594	1.312458	1.466837	1.68073

From the above table it is clearly observed that deviation value is increasing with increase of number of past months in the calculation of simple moving average. So we can say that compared to all 3, 5, 7 and 10 months moving average, 3 months moving average is efficient.

Now consider weighted moving average table.

YEAR	MAD			MSE			RMSE			MAPE		
	3 Months	5 Months	7 Months	3 Months	5 Months	7 Months	3 Months	5 Months	7 Months	3 Months	5 Months	7 Months
2011	144.0173	157.8455	179.9442	77696.94	88720.69	107902.3	278.7417	297.8602	328.4848	1.317238	1.406032	1.545907
2012	143.1506	156.0429	177.5405	32718.66	37570.34	46905.93	180.883	193.8307	216.5778	0.818674	0.892398	1.014743
2013	197.696	212.0815	235.4157	62366.18	68964.13	81662.87	249.7322	262.6102	285.7672	1.008682	1.081733	1.200428

The deviation value is increasing with increase of number of months in weighted moving average. From the above we can say that 3 months weighted moving average is good compared to 5months and 7 months weighted moving average method because of less deviation in all of three years data.

Consider exponential smoothing constant method in predicting the prediction value.

Year	MSE							MAD							MAPE						
	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.2	0.25	0.3	0.35	0.4	0.45	0.5
2011	167374.3	138797.6	119337.4	105230.4	94553.5	86221.2	79571.71	335.0078	306.0494	283.9109	266.489	252.9848	242.5466	233.7529	1.903301	1.740497	1.615823	1.517705	1.44164	1.382038	1.331601
2012	74876.52	60484.62	51179.74	44762.23	40123.37	36651.22	33984.28	226.4613	202.9152	185.8918	172.7316	162.5636	154.2726	147.3771	1.293826	1.139606	1.062483	0.987295	0.929278	0.882018	0.842741
2013	117615.9	99128.12	86594.88	77663.9	71027.04	65919.35	61873.65	281.7214	258.7759	241.6844	228.4415	217.3185	207.7849	199.6419	1.43643	1.319463	1.232375	1.164956	1.108403	1.059993	1.0186

We calculated the deviation by using different values of smoothing constant ranges from 0.2 to 0.5, as the value of smoothing constant increases the deviation decreases up to 0.5 then after the deviation increases. So optimum smoothing constant is 0.5. Now we are comparing deviation values of 3 months simple moving average, 3 months weighted moving average and exponential smoothing constant at alpha=0.5, to determine out of this three which technique is effective in prediction.

Year	3 months simple moving average				exponential smoothing constant at alpha=0.5				3 months weighted moving average			
	MAD	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE
2011	256.1092	26681.04	163.3433	1.461293	233.7529	79571.71	282.0846	1.331601401	144.0173	77696.94	278.7417	1.317238
2012	159.3225	39499.57	198.745	0.91079	147.3771	33984.28	184.3483	0.842741023	143.1506	32718.66	180.883	0.818674
2013	220.5033	75946.43	275.5838	1.124594	199.6419	61873.65	248.7441	1.018600384	197.696	62366.18	249.7322	1.008682

It is clear that the deviation is less in weighted 3 months moving average compared to simple moving average and exponential smoothing constant.

Reference:

1. CORDEIRO, C. – NEVES, M. M.: Bootstrap and Exponential Smoothing Working together in Forecasting Time Series, Proceedings in Computational Statistics, Physica-Verlag, 891–899, 2008.
2. JARETT, J.: Business Forecasting Methods, Cambridge, MA: Basil Blackwell, 1991.
3. MONTGOMERY, D. C. – JOHNSON, L. A. – GARDINER, J. S.: Forecasting and Time Series Analysis, McGraw-Hill, Inc., 1990, ISBN 0-07-042858-1.
4. OSTERTAGOVA´, E.: Applied Statistics, [in Slovak], Elfa, Koˇsice, 2011, ISBN 978-80-8086-171-1.
5. LEWIS, C. D.: Industrial and Business Forecasting Methods, London, Butterworths, 1982.
6. SANJOY, K. P.: Determination of Exponential Smoothing Constant to Minimize Mean Square Error and Mean Absolute Deviation, Global journal of research in engineering, Vol. 11, 2011.