www.ijramr.com

ISSN : 2350-0743

International Journal of Recent Advances in Multidisciplinary Research

Vol. 03, Issue 07, pp.1655-1658, July, 2016

REVIEW ARTICLE

VARIOUS STATISTICAL METHODS FOR TIME SERIES FORECASTING

***Amandeep Kaur**

Department of Mathematics, Guru Nanak Dev University, Amritsar-143005, India

ARTICLE INFO

Article History:

Received 14th April, 2016
Received in revised form
21st May, 2016
Accepted 18th June, 2016
Published online 31st July, 2016

Keywords:

Forecasting,
Planning, Future,
Growth and Statistical methods.

ABSTRACT

The need Forecasting is an important part of planning and future growth. Many of an organization's decisions are based on the prediction of future unknown events. This paper describes some statistical methods that can be used for time series forecasting.

INTRODUCTION

Forecasting means to calculate or predict some future events based on the study of past and present data. In business, industry and government, policymakers need the future behaviour of many events before they make decisions. Their decisions depend on the forecast and they expect these forecasts to be accurate. A forecast system is needed to make such predictions.

Following are some areas those require forecasting

- **Marketing managers:** They use sales forecasts to determine optimal sales force allocations, set sales goals, and plan promotions and advertising.
- **Production planners:** They need forecasts in order to: schedule production activities, order materials, establish inventory levels and plan shipments.
- **The bank**: Banks have to forecast too. Demands of various loans and deposits Money and credit conditions so that it can determine the cost of money it lends.
- **Public administrators:** They also must make forecasts for budgeting purposes.
- **Universities:** These forecast student enrolments, cost of operations, and, in many cases, the funds to be provided by tuition and by government appropriations.
- **The personnel department:** It requires a number of forecasts in planning for human resources. Trends that affect such variables as labor turnover, retirement age, absenteeism, and tardiness need to be forecast for planning and decision making.

Corresponding author: Amandeep Kaur,

Department of Mathematics, Guru Nanak Dev University, Amritsar-143005, India

Forecasting methods generally assume that the same underlying causal relationship that existed in the past will continue to prevail in the future. In other words, most of our techniques are based on historical data.

Forecasting methods can be broadly classified into two categories: quantitative and qualitative. Qualitative forecast methods are based on human judgement, opinions and non mathematical.

These methods are useful when historical data either are not available or are scarce. These methods are simple and easy to use but usually they give little or no information about the accuracy of the forecast. On the other side quantitative methods are based on mathematical or statistical models. These methods are used when the historical data are available. Once the underlying model has been chosen, the corresponding forecasts are determined automatically. These quantitative methods can be further classified as deterministic or statistical. In deterministic models the relationships between the variable of interest, Y, and predictor variables $X_1, X_2, \dots X_t$ is determined exactly;

$$Y = f(X_1, X_2, \dots X_t, \beta_1, \beta_2, \dots \beta_n)$$

But stochastic models involve some randomness or uncertainty i.e., of the form

$$Y = f(X_1, X_2, \dots X_t, \beta_1, \beta_2, \dots \beta_n) + noise$$

Where noise or error component is a realization from a certain probability distribution. That is why these models are also called statistical or probabilistic methods. In this paper, only quantitative forecast methods have been discussed.

A forecast system consists of some steps which are explained below

- Forecaster must first identify the problem for which the forecast is to be made. This is the most difficult aspect of the forecaster's task. A forecaster has a great deal of work to do to properly define the forecasting problem, before any answer can be provided.
- Next, it is necessary to collect historical data of the items of interest. The historical data is used to construct a model which can be used for forecasting.
- In the next step, collected data is analysed to inspect what do the data tell us? We plot the graph of the data for visual inspection. Then we compute some simple descriptive statistics e.g. mean, standard deviation, correlation etc. We also use decomposition analysis to check the relative strengths of trend, seasonality, cycles and to indentify unusual data points. Such analysis will help us to suggest a class of methods for forecasting.
- The next step involves choosing and fitting several forecasting methods. Each method is based on a set of assumptions and usually involves one or more parameters which must be fitted using the known historical data. The adequacy of the fitted model must be checked. If the model is unsatisfactory, it has to be respecified. This process must be repeated until a satisfactory model is found.
- Once a model has been selected and its parameters estimated appropriately, the model is to be used to make forecasts. The stability of the forecast model can be assessed by checking the forecasts against the exact observations. Forecast errors can be calculated, and possible changes in the model can be detected.

The rest of the paper is structured as follows; the next section describes various statistical models for predictions such as regression analysis, moving average methods and some advanced methods such as ARIMA methods and neural network. The last section concludes the papers.

Statistical methods for Prediction

Regression Analysis

Regression analysis is widely used for prediction. It is concerned with modeling the relationships among variables. It quantifies how a dependent variable is related to a set of independent (predicted) variables.

The regression model, in general, can be written as

$$y_t = f(x_t;\ \beta) + \epsilon_t \qquad (1)$$

Where $f(x_t;\ \beta)$ is a mathematical function of the n independent variables $x_t = (x_{t1}, x_{t2}, x_{t3}\ \ldots x_{tn})'$ and unknown parameters $\beta = (\beta_1, \beta_2, \ldots \beta_m)'$. The model (1) is probabilistic, since the error term $\in_t$ is a random variable. It is assumed that

- Its mean, $E(\in_t) = 0$ and its variance $V(\in_t) = \sigma^2$, are constant and independent from t.
- $Cov(\in_t, \in_{t-k}) = E(\in_t, \in_{t-k}) = 0\ for\ all\ k \neq 0$ that is, the errors $\in_t$ are uncorrelated. Also errors come from a normal distribution.

There are two types of regression models: (1) Linear models (2) Non linear models. Models those are linear in the parameters as well as independent variables, are linear models and can always be written as

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \quad + \beta_p x_{tp} + \epsilon_t$$

The set of the p independent variables $X_1, X_2, \ldots, X_p$ can be either original predictor variables or functions. Examples of linear models are:

(i) $y_t = \beta_0 + \epsilon_t$ (Constant mean model)
(ii) $y_t = \beta_0 + \beta_1 x_t + \epsilon_t$ (Simple linear regression model)
(iii) $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \epsilon_t$(Linear model with two independent variables)

Non linear models are those which are non linear in parameters and independent variables. Examples are:

(i) $y_t = \beta_0 \exp(\beta_1 x_t) + \epsilon_t$ (Exponential growth model)
(ii) $y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \epsilon_t$(Quadratic model)

The problem of estimating parameters β from the given historical data is to choose parameter estimates $\hat{\beta} = (\hat{\beta}_1, \ldots \hat{\beta}_m)'$ such that the fitted function $f(x_t;\ \hat{\beta})$ is close to the observations. The parameter estimates that minimize the sum of squared deviations

$$S(\boldsymbol{\beta}) = \sum_{t=1}^{n} [y_t \quad f(x_t; \boldsymbol{\beta})]^2$$

are called the least squares estimates and denoted

Averaging Methods

Averaging methods are suitable for the stationary time series data where the series is in equilibrium around a constant value (the underlying mean) with a constant variance over time. The various averaging methods are discussed below:

Mean

It uses the average of all the historical data as the forecast

$$F_{n+1} = \frac{1}{n} \sum_{t=1}^{n} y_t$$

This method is appropriate when there is no noticeable trend or seasonality. When the new data becomes available, the forecast for time n+2 is the new mean including the previously obtained data plus this new observation

$$F_{n+2} = \frac{1}{n+1} \sum_{t=1}^{n+1} y_t$$

Moving Averages (MA)

The moving average provides a simple method for smoothing the past values to estimate trend-cycle component. Taking an average of the points near observation provide a reasonable estimate of the trend-cycle at that observation.

The average eliminates some randomness in the data. The moving average for time period t is the mean of the "k" most recent observations. The smaller the number k, the more weight is given to recent periods. The greater the number k, the less weight is given to more recent periods. The constant number k is specified at the outset. A large k is desirable when there are wide, infrequent fluctuations in the series. A small k is most desirable when there are sudden shifts in the level of series. For quarterly data, a four-quarter moving average, MA (4), eliminates or averages out seasonal effects. For monthly data, a 12-month moving average, MA (12), eliminate or averages out seasonal effect. Equal weights are assigned to each observation used in the average. Each new data point is included in the average as it becomes available, and the oldest data point is discarded.

A moving average of order k, MA (k) is the value of k consecutive observations.

$$F_{t+1} = \frac{(y_t + y_{t-1} + \quad + y_{t-k+1})}{k}$$

here k is the number of terms in the moving average. The moving average model does not handle seasonality very well although it can do better than the total mean.

Exponential smoothing methods

This method provides an exponentially weighted moving average of all previously observed values. Here the aim is to estimate the current level and use it as a forecast of future value. The formula for exponential smoothing is

$$F_{t+1} = \alpha y_t + (1 \quad \alpha)F_t$$

Here α is a smoothing constant.

The forecast F_{t+1} is based on weighting the most recent observation y_t with a weight α and weighting the most recent forecast F_t with a weight of $1 \quad \alpha$. The value of smoothing constant α must be between 0 and 1. If stable predictions with smoothed random variation is desired then a small value of α is desired. If a rapid response to a real change in the pattern of observations is desired, a large value of α is appropriate. To estimate α, Forecasts are computed for α equal to .1, .2, .3, …, .9 and the sum of squared forecast error is computed for each. The value of α with the smallest RMSE is chosen for use in producing the future forecasts. Holt's two parameter exponential smoothing method is an extension of simple exponential smoothing. It adds a growth factor (or trend factor) to the smoothing equation as a way of adjusting for the trend. Three equations and two smoothing constants are used in the model. The exponentially smoothed series or current level estimate is

$$L_t = \alpha y_t + (1 \quad \alpha)(L_{t-1} + b_{t-1})$$

The trend estimate is

$$b_t = \beta(L_t \quad L_{t-1}) + (1 \quad \beta)b_{t-1}$$

Forecast m periods into the future is

$$F_{t+m} = L_t + mb_t$$

The weight α and β can be selected subjectively or by minimizing a measure of forecast error such as RMSE. Large weights result in more rapid changes in the component and small weights result in less rapid changes. Winter's exponential smoothing model is the second extension of the basic Exponential smoothing model. It is used for data that exhibit both trend and seasonality. It is a three parameter model that is an extension of Holt's method. An additional equation adjusts the model for the seasonal component.

Autoregressive Integrated Moving Average (ARIMA)

ARIMA model, also known as Box-Jenkins models, is widely used in time series forecasting because of its flexibility in representing different time series. The ARIMA model is usually denoted as ARIMA (p, q, d). Here p is the number of autoregressive orders that specify which previous values from the series are used to predict current values. The order of differencing, d is applied to the series before estimating model. The series with trends are nonstationarity and ARIMA modeling assumes stationarity. So, differencing is necessary when trends are present and is used to remove their effect. The number of moving average orders, q, specify how deviations from the series mean for previous values are used to predict current values. The ARIMA model assumes that the future values of a time series have functional relationship with past and current observations and white noise. So the underlying process that generates the time series has the following form

$$y_t = \delta + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \quad + \varphi_p y_{t-p} + \epsilon_t \quad \theta_1 \epsilon_{t-1} \quad \theta_2 \epsilon_{t-2} \dots \quad \theta_q \epsilon_{t-q}$$

i.e. the actual value y_t depends on its p previous values and q previous random error terms ϵ_t. In the ARIMA model $\varphi_i(i = 1,2,\dots p)$ and $\theta_j(j = 1,2,\dots,q)$ are called autoregressive and moving average operators respectively. Our main task of the ARIMA model building is to determine the appropriate model order (p, q, d). There are three steps to build a suitable ARIMA model.

***Model identification*:** First we identify the stationarity of the data by plotting the graph of the data. Then we estimate the order of p and q.

***Parameter estimation*:** Here we estimate the parameters by imposing some conditions.

***Diagnostic checking*:** After the estimation of the parameters the Ljung-Box Q statistic is used to check the overall adequacy of the fitted model

Neural network model

Neural networks are a class of flexible nonlinear models that can discover patterns adaptively from the data. They have been widely used as a promising method for the time series forecasting. They can be an important candidate for seasonal and trend time series forecasting. Being a flexible modeling tool, neural networks can model any type of relationship in the data with high accuracy. With neural networks, no specific assumptions need to be made about the model and the underlying relationship is determined solely through data mining. This data driven approach is one of the most important advantages of neural networks in solving many complex real world forecasting problems. Although neural networks are inherently nonlinear models, they are capable of modeling linear processes as well. As neural networks are universal function approximators, it is natural to use them to directly model seasonal and trend variations. Although many types of neural network models have been proposed, the most popular model for time series forecasting is the feedforward network model. Here the inputs nodes are the previous lagged observations while the output provides the forecast for the future value. Hidden nodes with appropriate nonlinear transfer functions are used to process the information received by the input nodes. The model can be written as

$$y_t = \alpha_0 + \sum_{j=1}^{n} \alpha_j f(\sum_{i=1}^{m} \beta_{ij} y_{t-1} + \beta_{0j}) + \epsilon_t$$

Where m is the number of input nodes, n is the number of hidden nodes, f is sigmoid transfer function. $\{\alpha_j, j = 0,1, \ldots n\}$ is a vectors of weights from the hidden to output nodes and $\{\beta_{ij}; i = 0,1, \ldots m; j = 1,2, \ldots n\}$ are weights from the input to hidden nodes. Here α_0 and β_{0j} are the weights of arcs leading from the bias terms which have values always equal to 1. In the time series forecasting context, neural networks can be perceived as equivalent to nonlinear autoregressive models. Lags of the time series, potentially together with lagged observations of explanatory variables, are used as inputs to the network. During training pairs of input vectors and targets are presented to the network. The network output is compared to the target and the resulting error is used to update the network weights. Neural network training is a complex nonlinear optimisation problem, and the network can often get trapped in local minima of the error surface. In order to avoid poor quality results, training should be initialised several times with different random starting weights and biases to explore the error surface more fully.

Conclusion

This short paper shows the various statistical techniques for time series forecasting. Some of the techniques are quite simple and rather inexpensive to develop and use. Others are extremely complex, require significant amounts of time to develop, and may be quite expensive. Some are best suited for short-term projections, others for intermediate or long-term forecasts. There are some techniques to select forecasting models such as (a) Is forecasting for short-run or long-run purposes? (b) How much data is available and is data stationary or non stationary? (c) How much accuracy is desired? (d) What is the cost associated with developing the forecasting model, compared with potential gains resulting from its use?

REFERENCES

Box, G. and Jenkins, G. 1970. Time series analysis: Forecasting and control, San Francisco, 1970.

Conejo, A., Plazas, A. M., Espinola, R. and Molina, A. 2005. Day-Ahead Electricity Price Forecasting using the wavelet transforms and ARIMA models, *IEEE Trans. Power Syst.*, Vol. 20, No. 2.

Gencay, R., Selcuk, F. and Whitcher, B. 2001. An Introduction to Wavelets and Other Filtering Methods in Finance and Economics. San Diego, CA: Academic Press.

Mcnames, J., Local averaging for chaotic time series prediction, Neurocomputing, 48, 279- 297, 2002

Peter Zhang, G. Min Qi. 2005. Neural network forecasting for seasonal and trend time series. European Journal of Operational Research 160, 501-514.

Skander, 2002. On the use of the wavelet decomposition for time series prediction, Neurocomputing 48, 267-277

Soltani, D. Boichu, P. Simard, S. Canu, 2000. The long- term memory prediction by multiscale decomposition, Signal processing 80, 2195-2205.

Weigand, A.S., Huberman, B. A., Rumelhart, D.E. 1990. Predicting the future : connectionist approach, *International Journal of Neural System* 1, 193-209.

****