Dr. Akram M. Chaudhry
19th Dec. , 2012

Forecasting

Concepts & Applications

In all spheres of life we, as individuals or as a group, always need to foresee future events of interest for better planning and optimizing our decisions. For example :

i) an investor would like to know future outcome (dividend) of a stock before purchasing its shares.
ii) a business organization would like to know possible sales of its new product in the near future for planning purpose.
iii) both, individuals and organizations would be interested to know shapes of new things to appear in future, such as, types of computers, cars and cameras.

Looking at these examples one can say that a forecast is a prediction of an event to occur or not to occur in future.

Before learning an art of foreseeing future let us see what forecasting is ?

19.1 **Guess and Common Sense**

Guess is an opinion formed on basis of a hypothesis (or supposition). It may or may not be true as it is based on uncertain knowledge. There is an old saying :

"To guess is cheap but to guess wrongly is expensive"

It means that if we attempt to foresee future with out careful (or calculated) thought then we may end up in disaster which, in turn, may cost us too much. It is better to be on safe side rather than exposing ourselves to risk.

Common sense is a practical good sense gained by experience through the passage of life.

There is a legend about a man who was sitting on a tree and cutting the branch on which he was sitting. A passerby, on noticing this non sense of the man advised him not to do so, as naturally, with the tree branch the person will also fall down and hurt himself. The man did not pay attention to the passerby's advice and continued cutting the branch of the tree until the branch was cut and he fell down . After falling down he ran to the passerby who did not go very far and asked him how he knew that he is going to fall down. He told the passerby that definitely he seems to be a foreteller and requested him to tell him about his doom day. On his insistence the passerby told the man that the day he will climb the tall tree and do the same nonsense that he did just now that day will be his last day of survival. Being afraid of his death the wood cutter stopped cutting the trees.

Prediction and Forecasting

The words: prediction and forecasting are synonymous to each other and, quite often, are used interchangeably. However, some sort of difference do exists between the concepts of these two words which obviously is clear from the following definitions of these words.

Prediction is foretelling (telling in advance) of occurrence or non occurrence of future events, based on prophecy, knowledge and experience.

Forecasting is an art of foreseeing future through the scientific perceptions of Statistics or otherwise.

19.2 Types of Forecasts

Three types of forecasts are usually made on a phenomenon of interest, called, short term , medium term and long term forecasts. These types are explained as follows.

i) **Short Term Forecasts**

The short term forecasts are made for immediate or near future time horizon, usually, up to two time steps (days, weeks, months, years).

ii) **Medium Term Forecasts**

The medium term forecasts are made for intermediate time horizons, such as, 3 to 5 steps ahead (weeks, months, years)

iii) **Long Term Forecasts**

The long term forecasts are made for distant future (in years), usually, 6 and more years

19.3 Approaches to Forecasting

There are three main approaches to forecasting.

i) Quantitative
ii) Qualitative
iii) Mixed (Quantitative-Qualitative)

These approaches are explained as follows.

19.3.1 Quantitative Approach to Forecasting

In this approach emphasize is laid on analysis of time series (data recorded in chronological order) employing statistical or non statistical models. For these models, first, the parameters of the models, are estimated by using an appropriate estimation technique, and then forecasts are generated through a well defined forecast function. Some of these, commonly used models are as stated as follows.

Statistical Models

- Average Models
- Exponential Smoothing Models
- ARMA (Auto-regressive Moving Average) Models
- Linear Regression Models
- Linear Dynamic System Models
- State Space Models

Non-Statistical Models

- Econometric Models
- Market Research Models

It is difficult to say whether statistical or non statistical models, in general, are better than the others as both types of models have their own merits and demerits.

In this text, however, the discussion is confined to the statistical models as these are related to the subject being discussed.

Before elaborating the quantitative approach to forecasting and its models let us go through the following preliminaries.

19.4 Time

Nothing is as transitory as time. It reigns supreme, passes and can't be held back. It determines the rise and fall of all things. Its mode, however, is relevant to subjects. It passes quickly, if some one is in a nice company and passes slowly if some one is confronted with awful events. In any case, it always has been a source of fascination. Since the dawn of civilization people have been devising various tools to measure it, express and describe it within their domains. But the question is at all what is it? Is it just a passage through which all of we transit or more than that?. Before replying to this question, let us look at some definitions of time.

Time is intangible, mysterious and valuable.

19.4.1 Time Series and its Components

A time series is a sequence of observations or instant records on an event of interest, recorded over equal passage of time, such as, hours, days, weeks, months and year in a chronological order. When analyzed, it reveals valuable information about phenomenon being studied and leads forecasters to foresee future in a meaningful manner. Before going in to detail, let us first see what a time series may be made of?

A time series may be considered as an ensemble of four or few of the components: Trend, Seasonal , Cyclical and Random Shocks. These components are explained as follows.

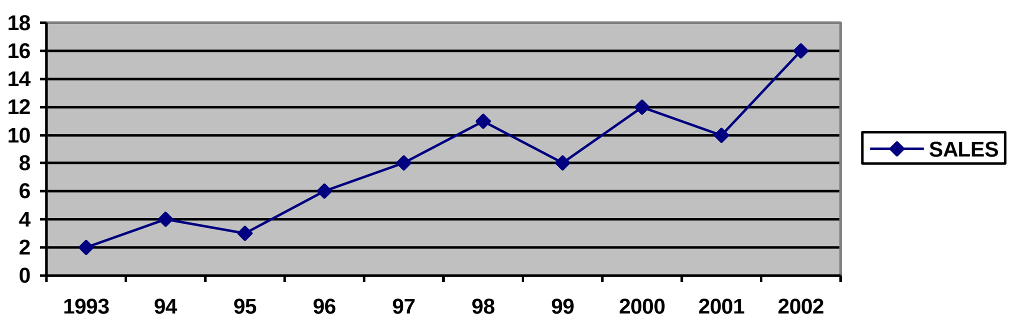
Trend

The Trend is a long term movement component of a time series that shows change in level of the data that underlies the growth or decline of a phenomenon being studied. This component is described by a straight line or a curve and is written by the symbol T_t.

The growth in the level of time series reflects positive (or upward) trend and the decline in the level of series indicates negative (or downward) trend. If the trend of data is neither positive nor negative then the time series is known as stationary.

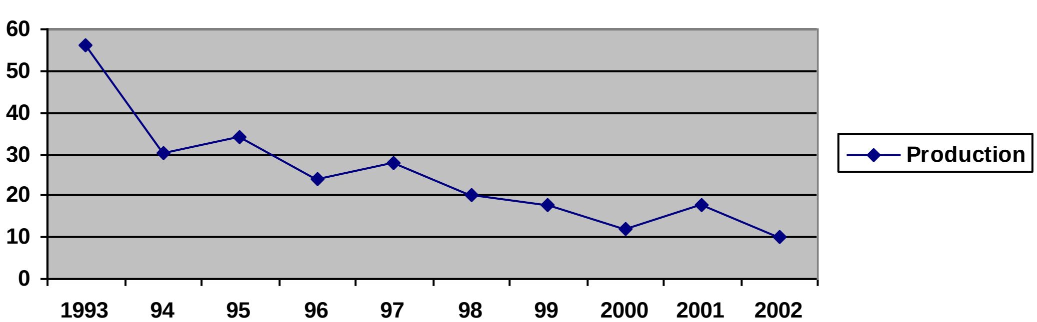
For example, the sale (in million $) of a company for the last 10 years :

Year:	1993	94	95	96	97	98	99	2000	01	02
Sale:	2	4	3	6	8	11	8	12	10	16



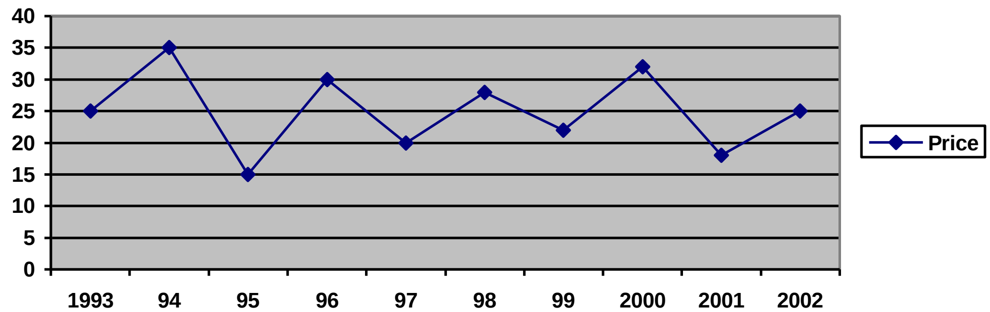
show the positive trend in sale of the company; whereas, the production of gas (in billion cubic feet) in a country over the last 10 year:

Year:	1993	94	95	96	97	98	99	2000	01	02
Production:	56	30	34	24	28	20	18	12	18	10



Indicate the negative trend in the production of gas. These are the examples of positive and negative trend. However, in some cases, there may not exist any trend at all. For example, the price (in BD) of certain commodity in Bahrain for the last 10 years:

Year:	1993	94	95	96	97	98	99	2000	01	02
Price:	25	35	15	30	20	28	22	32	18	25



show no trend (in general, neither increase nor decrease in price of the commodity) at all.

Seasonal Variations
As the name implies this component of time series is concerned with the study of changes in the data due to seasons within the duration of a year. The pattern of changes that is regular in nature is expected to repeat year after year during the same period of time. This component, denoted by the symbol S_t , is typically found in weekly, monthly or quarterly data.

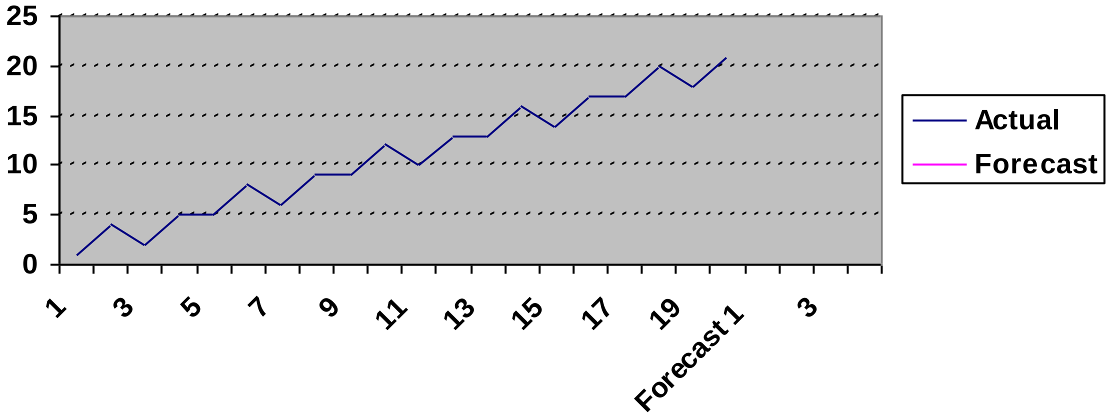
For example, an increase in sale of woolen clothes in winter season reflects an impact of winter on the sale. This impact is represented by the seasonal component of time series.

As a numerical example, consider the quarterly sale of certain commodity, over the last five years.

Quarters

Year	**Q_1**	**Q_2**	**Q_3**	**Q_4**
2000	1	4	2	5
2001	5	8	6	9
2002	9	12	10	13
2003	13	16	14	17
2004	17	20	18	21

The graph of these quarterly sales is



Cyclical Fluctuations
This component of time series, usually abbreviated by C_t, represents wavelike upward and downward fluctuations (or cycles) around the trend. These fluctuations which are less regular than the seasonal pattern appear after a long period of time. The arise due to changes in business activities, economic conditions and so on.

For example, a fashion cycle is believed to be repeated in 15 to 20 years, an economic cycle in 25 to 35 years and a business cycle in 20 to 30 years.

As an example, consider 40 years of data as shown below :

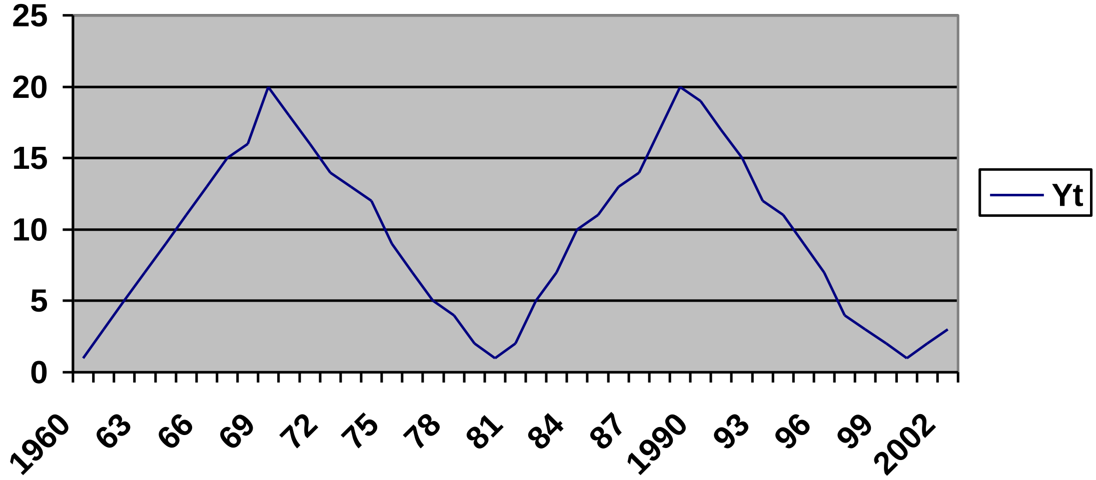
Year	**1960**	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**
Y_t	1	3	5	7	9	11	13	15	16	20

Year	**1972**	**1973**	**1974**	**1975**	**1976**	**1977**	**1978**	**1979**	**1980**	**1981**
Y_t	18	16	14	13	12	9	7	5	4	2

Year	**1982**	**1983**	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**	**1991**
Y_t	1	2	5	7	10	11	13	14	17	20

Year	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**
Y_t	19	17	15	12	11	9	7	4	3	2

Year	**2002**	**2003**	**2004**
Y_t	1	2	3



Random Shocks
This component, also called random error, random noise or irregular component of time series is composed of fluctuations that are caused by abrupt changes in events, such as, earth quick, wars, weather catastrophe . This component, random in nature, is written by the symbol, E_t. It is the most difficult component of the time series to understand, model and estimate accurately.

19.4.2 Ensemble of Time Series
A time series (Y_t) may be considered as an ensemble of all the four or fewer components. This ensemble appears in one of the three possible forms.

i) Addition Form
In this form the components of time series are added to each other . That is.
$Y_t = T_t + S_t + C_t + E_t$

ii) **Multiplication Form**
In the multiplication form all components multiply each other. That is,
$Y_t = T_t \text{ x } S_t \text{ x } C_t \text{ x } E_t$

A time series of this form can be written in to an addition form using Logarithm operator(Log). In this case, the time series becomes:
$\text{Log } Y_t = \text{Log } T_t + \text{Log } S_t + \text{Log } C_t + \text{Log } E_t$

Where the logarithmic operator “Log “ may be a common logarithmic (a logarithmic with base 10) or a natural (or Naperian) logarithmic having base e, the value of which is approximately 2.7183 .

iii) **Mixed form**
In this form , a time series appears as a mixture of addition and multiplication of components. For example,
$Y_t = T_t \text{ x } S_t \text{ x } C_t + E_t$
and so on.

19.4.3 Characteristics of Time Series

#1. Discrete and Continuous Time Series
A time series observed at equally spaced points in time is called a discrete time series; whereas, observations taken on continuous points in time are continuous time series.

#2. Linearity of Time Series
Time series are linear in nature if the observation exhibit a straight line (or first degree equation) pattern.

#3. Stationary and Non-stationary Time Series
A time series is called stationary if its underlying generating process is based on a constant mean and a constant variance in a statistical sense. It means that in case of stationary time series the statistical properties in general and means and variances in particular of the series are independent of time period during which it was observed. In case these properties are not satisfied, the time series are non-stationary.

A time series is strictly stationary if the underlying process of generating time series is in a particular state of equilibrium, otherwise, the stationary status is of weak nature. In any case, a stationary time series has no or insignificant trend.

For testing whether a time series is stationary or not various testing procedures are available. These procedures are based on parametric and non parametric test statistics. Some of these which are commonly used are Auto correlation Test Statistic, Von Neumann's Ratio Test Statistic, Modified Box Pierce Test Statistic, Daniel's Test Statistic and Kendal's Tau Statistic.

#4. Random Walk
A time series is known to be of random walk nature if its data points wander in an entirely unpredictable manner. In other words, there is no relationship between its consecutive values; i.e., its autocorrelation is zero. Some economic and financial phenomena are known to follow random walk, such as, prices of certain commodities and stocks.

#5. Gaussian (or Normal) Nature of Time Series.

A time series is Gaussian (or normal) in nature if the joint distribution of its data points is a multivariate normal. It will be strictly stationary if the means and the variances completely specify the joint Probability distributions.

#5. White Noise
A noise process is called a white noise process if its terms are independently and identically distributed (i.e. uncorrelated) with mean zero and some constant variance.

#19.4.4 Forecast Accuracy Measures

To measure accuracy of forecasts, generated by the application models (methods or techniques) , many types of criteria are available; such as, Bias (B) , Mean Absolute Error (MAE), Tracking Signal (TS), , Whiteness of Residuals (WOR), Akram Test Statistic (ATS) and Durban Watson Statistic (DWS).

For a time series $\{y_t\}_{t=1,\ldots,n}$ having one step ahead forecasts $\{ f_{t+1}\}_{t=1,\ldots,}\, n_e$ and $\{ e_t\}_{t=1,\ldots,}\, n_e$, the one step ahead forecasts residuals or errors; such that at time t, $e_t = y_t - f_{t+1}$; these criteria are defined as follows.

19.4.4.1 Bias

The bias of forecasts is simply an average of the residuals or errors. It is defined as: $\mathbf{B = \Sigma\, e_t / n_e}$
Where, $\mathbf{n_e}$ is the number of one step ahead residuals or error at time t.

It, that appears systematically in error terms due to drift of forecasts from the actual observations, arises from a number of sources, such as, use of inappropriate or incapable model and forecaster's preconceived notions, specially where priors are fed in to operating models.

For good forecasts the value of Bias should be minimum, preferably in neighborhood of zero.
A positive value of Bias indicates that the model (method or technique) considered for forecasting purpose is generating one step ahead forecasts lower than the actual observations. On the other hand a negative value of Bias leads us to the conclusion that the model (method or technique) is generating forecasts higher than the actual observations.

19.4.4 2 Mean Absolute Error (MAE)
The mean absolute error (some time called the Mean Absolute Deviation) is an average of absolute residuals or errors. It is obtained by adding all error terms without regard to sign. (the negatives are considered as positive) and dividing it by the number of errors summed up. That is,

$$\mathbf{MAE = \Sigma\, |e_t| \;/\; n_e}$$

It is used to compare the forecasts generated by various models (methods or techniques).

19.4.4.3 Mean Absolute Percentage Error (MAPE)
This criterion of forecast accuracy is an average of the ratios of the absolute errors and the actual observations, multiplied by 100.

$$\mathbf{MAPE = = [\, \Sigma\, (|\, e_t\,|\, /\, y_t\,) \;/\; n_e\,]\; 100}$$

Like all other criteria MAE, is also used to compare the forecasts generated by various models (methods or techniques). However, it alone lets us know about the percentage of errors or residuals generated by the model being used. For a suitable model, its value must be quite low, such as, less than 5%.

19.4.4.4 Mean Square Error (MSE) and Root Mean Square Error (RMSE)
The mean square error of forecasts is an average of the squares of errors. It is computed as:

$$\mathbf{MSE = \Sigma\;\, e_t^2 \;/\; n_e}$$

Its value is interpreted like a variance that gives more weight to large residuals or errors. To ensure forecasts of good quality the MSE should be minimum possible.

For better interpretation, its squared root, called Root Mean Square Error :

$$\mathbf{RMSE = \sqrt{MSE}}$$

is preferred as its value, considered like the standard deviation, is easy to interpret
.

19.4.4.5 Tracking Signal (TS)
To periodically monitor and control the forecasting process the tracking signal statistic is commonly used. In its simplest form at time t , it may be computed as follows:

$$(TS)_t \;= \Sigma\, e_t \;/\; (MAE)_t \;\;= n_e\, B_t\, /(MAE)_t$$

Where B_t is a bias at time t.

A forecasting process may be considered satisfactory if the values of TS are within the lower control limit (LCL) and the upper control limit (UCL).

As a rule of thumb for simple forecasting models these limits LCL= 3 and UCL= 8 may be considered.

In case of a simple (or single) exponential smoothing models, these limits may approximately be computed as UCL= $\{0.884/\sqrt{\alpha}\}$LCL; where $0< \alpha = 1-\beta <1$.

For example, in case of α=0.25 and LCL=2, the value of UCL=3.54; whereas, in case of α=0.10 and LCL=2, the value of UCL=5.59.

In general, the performance of a model employed for analysis and forecasting of time series may be considered unsatisfactory of the values of TS_t values lie outside the control limits for two or more consecutive terms.

19.4.4.6 Whiteness of Residuals

There are many ways to check the whiteness of residuals or errors. Some of these, which provide meaningful insight in to the behaviour of residuals are defined as follows.

19.4.4.6.1 **Auto Correlation of Forecast Residuals**

The auto correlation of one step ahead forecast residuals or errors is the correlation of successive error terms et within themselves, lagged by 0, 1, 2 or more periods. Its coefficient at time lag period L is defined as:

$\boldsymbol{\rho_L = \mathrm{Cov}(e_t, e_{t+L}) / \mathrm{Var}(e_t)}$ for a positive integer value of L, say L=1,2,3,…

Where $\boldsymbol{\mathrm{Cov}(e_t, e_{t+L}) = {}_{t=1}\sum^{n-L} (e_t - e^-)(e_{t+L} - e^-)/ n_e}$ and $\boldsymbol{\mathrm{Var}(e_t) = {}_{t=1}\sum^{n} (e_t - e^-)^2 / n_e}$

A value of ρ_L not significantly different from zero (specially at L=1) ensures that the residual terms are uncorrelated or white. It in turn implies that the forecasts are good and the model employed is doing its job in a suitable manner.

19.4.4.6.2 **DW (Durban Watson) Test Statistic**

It is defined as:

$$DW= {}_{t=2}\sum^{ne} (e_t - e_{t-1})^2 / {}_{t=1}\sum^{ne} e_t^2$$

19.4.4.6.3 **ATS (Akram Test Statistic)**

A model used to analyze time series and generate forecasts is considered suitable or appropriate if it yields optimum forecasts. That is forecasts with residuals having whiteness within its terms. That is, the residual with independent terms. If it does not happen, then it means that the model being used is not capable of filtering color of noise (if present) and therefore is not a suitable.

To assist the forecasters and professional in search of suitable model the following stepwise identification procedure is suggested.

Step #1

Analyze discrete time series using a candidate model with a white noise component and capable of accommodating AR(p) noise processes, such as, Linear Dynamic System and State Space models of Akram (1992) and Harrison –Akram (1983). Estimate the parameters of the model using an optimum estimation technique; such as, Least Square Error or Maximum Likelihood, generate one step ahead forecasts and residuals. Standardize the residuals, using standard normal variate

$$Z_t = (e_t - e_t^-) / S_e$$

Where for n-number of observations and $n_{p\,-}$ number of parameters of the model,

$e_t^- = \Sigma e_t / n_e$ is the mean of the residuals for n_e number of residuals.

and $S_e = \Sigma e_t^2 / (n- n_p)$ is the standard error of estimate.

Compute the ASL of these standardized residuals and move to the second step.

Step#2 (Formulation of Hypotheses)
The three different scenario for formulating the null hypothesis (H_0) and the alternative hypothesis (H_1). These are described as follows.

Scenario#1	**Scenario#2**	**Scenario#3**
H_0: $\mu_{ASL} \geq 2$	H_0: $\mu_{ASL} = 2$	H_0: $\mu_{ASL} \leq 2$
H_1: $\mu_{ASL} < 2$	H_1: $\mu_{ASL} \neq 2$	H_1: $\mu_{ASL} > 2$

Where μ_{ASL} is the parent population's Average String Length of residuals or error terms .

One of these three scenario is selected in the light of the sampled information, purpose of study.

Step#3 (Test Statistic)
For the standardized residuals, compute the test statistic

ATS : $\tau = \{ 2 (n_e -1) / (n_{-\to}n_+ + n_{+\to}n_-) \}$ for $n \geq 30$

Where n_e are number of error or residual terms, $n_{-\to}n_+$ are number od sign changes from negative to positive signs and $n_{+\to}n_-$ are number of sign changes from positive to the negative signs.

Step#4 (Decision Regions)
For making decision on the fate of null hypothesis, the acceptance region (AR) and the rejection regions (RR) at some level of significance α are defined as follows.

a) For a two tailed test of significance under Scenario #2

AR: $\{2n/ (n-1) + Z\alpha_{/2} \sqrt{}(n-1)\} \leq$ Region $\leq \{2n/ (n-1) - Z\alpha_{/2} \sqrt{}(n-1)\}$

RR: Region $< 2n/ (n-1) + Z\alpha_{/2} \sqrt{}(n-1)$ or Region $> \{2n/ (n-1) - Z\alpha_{/2} \sqrt{}(n-1)\}$

b) For a one tailed test of significance under Scenario#1
AR: Region $\geq \{2n/ (n-1) + Z\alpha \sqrt{}(n-1)\}$
RR: Region $< 2n/ (n-1) + Z\alpha \sqrt{}(n-1)$

c) For a one test of significance under Scenario#3
AR: Region $\leq \{2n/ (n-1) - Z\alpha \sqrt{}(n-1)\}$
RR: Region $> 2n/ (n-1) - Z\alpha \sqrt{}(n-1)$

Step #5 (Decision Rule)
Accept H_0 of the value of test statistic $\boldsymbol{\tau}$ lies in the acceptance region AR; otherwise, reject H_0 and accept H_1. The acceptance of H_0 implies that the residuals are white; where rejection of H_0 leads us to the conclusion that the residuals are not white, but colored.

Step #6 (Determination of AR Order of Noise Process)
In case of colored noise the question is of what type is this noise AR type , MA or ARMA type and the order p, q or (p,q) of the noise process. Here, discussion is confined to AR(p) processes as Andel (1981) and others have shown that an ARMA process can be approximated and adequately represented by an AR process.

For the empirical value of ASL computed from the data a corresponding value **Φ** of AR(1) coefficient is determined. Using this **Φ** value the initially used candidate model is restructured or updated and again apply to data. Generate one step ahead forecasts and residuals and move through the above testing procedure again. If H_0 is accepted then it will lead to the conclusion that the model with AR(1) noise process is suitable. Its one step ahead forecasts, therefore would be optimum.

If H_0 is rejected again then it means that the model with AR(1) colored noise component, having failed to filter the noise and therefore is not suitable for analysis of data. In this case determine **Φ** again from the table of theoretical of ASL. The candidate model with AR(1) process would therefore be considered inappropriate and model with AR(1) process will be reconstructed considering AR(2) noise process, using $\boldsymbol{\Phi_1}$ (the first Φ) and $\boldsymbol{\Phi_2}$ (the second **Φ)** coefficients.

The above procedure is repeated until the color is filtered out i.e. , H_0 is accepted. For more discussion see Akram (2001)

19.4.4.6.4 **Akaike Information Criterion**
This criterion, developed by Hirotogo Akaike (1971) is defined as:

$$\mathbf{AIC = n \ln (SSE) + 2 n_p}$$

Where n = number of observations
Ln = Log_e is Naperion or natural logrothim
SSE = Sum of Squares of Error
n_p = Number of parameters of model being used.

19.4.4.6.5 **Schwarz Bayesian Criterion**
This Bayesian criterion, developed by Schwarz in 1978 is defined as:

$$\mathbf{BIC = n \log_e (SSE) + n_p \ln (n)}$$

Where n = number of observations
Ln = Log_e is Naperion or natural logrothim
SSE = Sum of Squares of Error
n_p = Number of parameters of model being used.

19.5 Statistical Methods of Forecasting

Various methods of forecasting which make use of Statistics concepts are available. These methods depend upon statistical models, naïve to the sophisticated ones, for analysing the time series and generating the forecasts. Before going through these models let us first define the symbols that we are going to use in methods and models.

t= time (hour, days, weeks, months, year, etc.)
y_t = an observation at time t
m_t = an estimate of an observation at time t
f_t = one step ahead forecast of an observation at time t
e_t= one step ahead forecast error at time t

19.5.1 Simple Average Method

In this method the simple average of observations , at time t, is obtained by adding the observations or the data values up to the time t and dividing the sum by the number of values added. For n number of observations the simple averages (m_t) are defined as:

$$\mathbf{m_t = \{ {}_{j=1}\Sigma^t \ y_j \} \div t} \qquad \text{for } t=1,\ldots,n$$

The averages of the observations, obtained in such a manner, are then used to generate one step ahead forecasts as:

$$\mathbf{f_t = m_{t-1}}$$

The residuals of these forecasts, called one step ahead forecast errors, (e_t) are computed using the relation:

$$\mathbf{e_t = y_t - f_t = y_t - m_{t-1}}$$

On basis of these errors the forecast accuracy identifiers are computed to evaluate the forecasts generated by the method in the light of the forecast accuracy criteria.

Example (Simple Average Method)

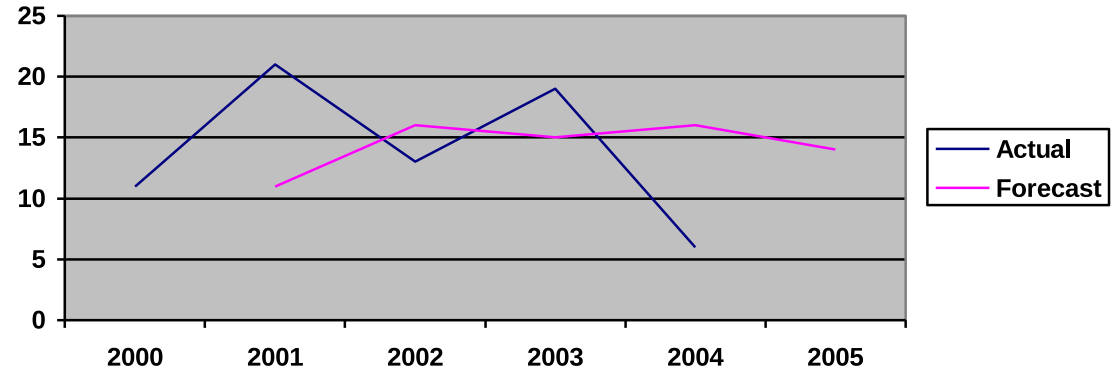
Consider the price (in $) of certain commodity for the last 5 years .

Year:	2000	2001	2002	2003	2004
Price:	11	21	13	19	6

Let us analyze these price data by a simple average method and forecast the price of the commodity for 1999.

Year	**t**	$\mathbf{y_t}$	$\mathbf{m_t}$	$\mathbf{f_t}$	$\mathbf{e_t}$	$\mathbf{\lvert e_t \rvert}$	$\mathbf{\lvert e_t \rvert / y_t}$	$\mathbf{e_t^2}$
2000	1	11	11	-	-	-	-	-
2001	2	21	16	11	10	10	0.4762	100
2002	3	13	15	16	-3	3	0.2308	9
2003	4	19	16	15	4	4	0.2105	16
2004	5	6	14	16	-10	10	1.6667	100
2005	6	forecast		**14**	-	-	-	
Total	-	-	-	-	1	27	2.5842	225

The original time series along with the one-step ahead forecast are shown below :



The forecast accuracy identifiers for this method are:

1) Bias : $B = \Sigma e_t / ς = 1/4 = 0.25$
2) Mean Absolute Error: $MAE = \Sigma |e_t| / ς = 27/4 = 6.75$
3) $MAPE = = \Sigma (| e_t | / y_t) / ς = 2.5842/4 = 0.6460$ (or 64.6%)
4) Mean Square Error: $MSE = \Sigma e_t^2 / ς = 225/4 = 56.25$
5) Root Mean Square Error : $RMSE = \sqrt{MSE} = \sqrt{56.25} = 7.5$
6) Tracking Signal : $TS = \Sigma e_t / MAE = ς B / MAE = 1/6.75 = 0.1482$

To see whether this method is better than any other method available to generate forecasts of the above price data we need to compare these values with the ones obtained from the competitive or rival method.

19.5.2 Moving Average Method

In this method the observations of data being analysed are added over a predetermined period (p) and then divided by the length of this time period to obtain the average. This procedure is carried on from the very first observation to the last observation, moving in the shape of blocks (consisting of p number of observations).

$$\mathbf{m_t = (1/p) \sum_{j=0}^{p-1} y_{t-j}} \qquad \text{for } t= p,\ldots, n$$

The averages obtained so are written corresponding to the p-th value position and so on.
It will be noticed that for the moving average process with moving period p equal weights (1/p) are assigned for each average.

For example if the moving period for annual data is 2 years, then first two observations are added, divided by two to obtain average and move to the second block of two observation by discarding the first observation and adding the third observation. We carrying on this process until the last observation is taken care of. For this process each average is obtained by assigning weight ½ = 0.5.

On the basis of these averages which are estimates of the observations one step ahead forecasts are obtained by considering $\mathbf{f_t = m_{t-1}}$

The residuals of these forecasts, called one step ahead forecast errors, (e_t) are defined as :

$$\mathbf{e_t = y_t - f_t = y_t - m_{t-1}}$$

These errors are then used to evaluate forecasts by computing the forecast accuracy identifiers and using forecast accuracy criteria.

For better insight let us go through the following numerical example.

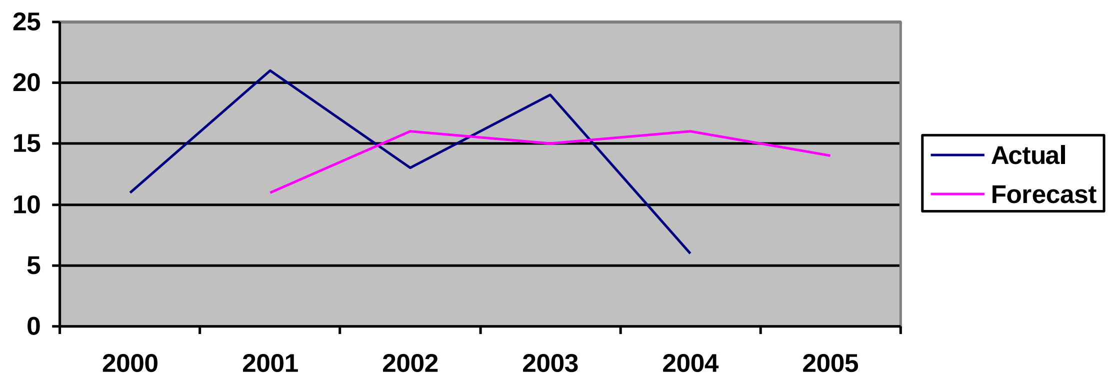
Example #1 (Moving Average Method)

Let us reconsider the above data (Example#1) on the price of commodity for the last 5 years and analyse these data by a moving average method.

Assuming, 2 years moving period (p=2) we obtain the following results.

Year	**t**	**y_t**	**m_t**	**f_t**	**e_t**	**$\|e_t\|$**	**$\|e_t\|/y_t$**	**e_t^2**
2000	1	11	-	-	-	-	-	-
2001	2	21	16	-	-	-	-	-
2002	3	13	17	16	-3	3	0.2308	9
2003	4	19	16	17	2	2	0.1053	4
2004	5	6	12.5	16	-10	10	1.6667	100
2005	6	forecast	-	**12.5**				
Total	-	-	-	-	-11	15	2.0028	113

The original time series along with the one-step ahead forecast are shown below :



To measure the accuracy of forecast the identifiers are:

1. Bias : $B = \Sigma e_t / \varsigma = -11/3 = 3.6667$
2. Mean Absolute Error: $MAE = \Sigma |e_t| / \varsigma = 15/3 = 5$
3. MAPE = = $\Sigma (| e_t | / y_t) / \varsigma = 2.0028/3 = 0.6676$ (or 66.67%)
4. Mean Square Error: $MSE = \Sigma e_t^2 / \varsigma = 113/3 = 37.6667$
5. Root Mean Square Error: $RMSE = \sqrt{37.6667} = 6.1373$
6. Tracking Signal : $TS = \Sigma e_t / MAE = \varsigma B / MAE = -11/5 = -2.2$

Comments:

Comparing these values with the similar values obtained from the simple average method the following conclusions are made.

i) In terms of Bias the Simple Average Method gives us better forecasts than the Moving Average Method as Bias of the first method (0.25) is less than the Bias of the second method(3.67) .

ii) Looking at the Mean Absolute Error (MAE) we notice that the SimpleAverage Method yields forecasts worst than the Moving Average Method as MAE of the first method (6.75) is more than the MAE of the second method (5.00).

iii) Comparing the Mean Square Error(MSE) of the Simple Average and the Moving Average methods we notice that MSE of the Moving Average method (37.67) is less than the MSE of the Simple Average method (56.25). Therefore, the Moving Average method yields better forecasts than the Simple Average method.

19.5.3 Simple Exponential Smoothing Method

In simple (or single) exponential smoothing the estimates of the observations (y - values) are computed as follows.

$$\mathbf{m_t = \beta\, m_{t-1} + (1 - \beta)\, y_t}$$

or

$$\mathbf{m_t = m_{t-1} + (1 - \beta)\, e_t}$$

where at time t,

e_t is one step ahead forecast residual or error. It is defined as:

$$\mathbf{e_t = y_t - m_{t-1}}$$

β is a smoothing coefficient having value between zero and one.
(i.e. $0 \le \beta \le 1$). It may or may not be known. If unknown, the n it may be approximately estimated as:

$$\boldsymbol{\beta \cong [\,(n-1) / (n+1)\,]}$$

$\mathbf{m_0}$ = prior estimate or information on the observation.

The above equations of m_t yield the first estimate m_1 of the observation y_1 on receipt of Information on the first observation y_1 , the smoothing coefficient $\boldsymbol{\beta}$ and the prior m_0 . After yielding the first estimate the process of updating the estimates starts as soon as new observation is received. This is explained through the following example.

Example (Simple Exponential Smoothing)

Let us revisit the price data (example#1 &2) and analyze these data by a simple exponential smoothing method. In this method, we estimate (m_t) of the price values (y_t) as:

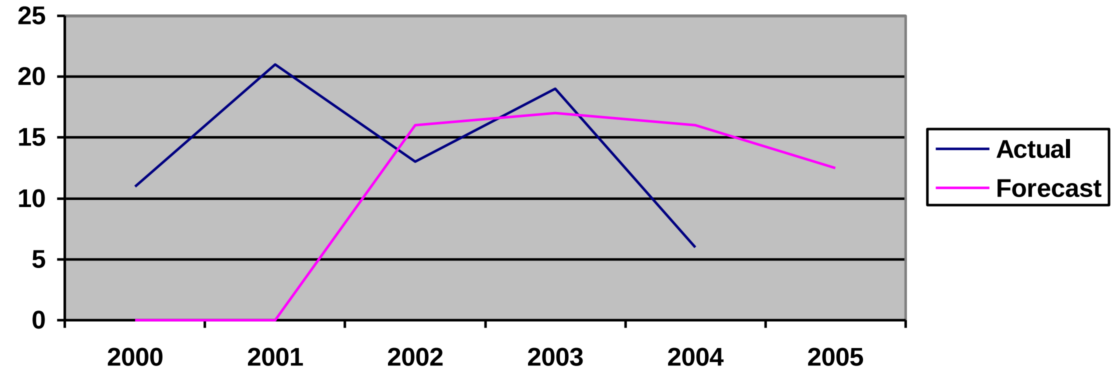
$$\mathbf{m_t = \beta\ m_{t-1} + (1-\beta)\, y_t}$$

by considering $m_0 = 15$ and $\beta \cong [\,(5-1) / (5+1)\,] \cong 0.7$

Using these prior values we obtain the following results.

Year	**t**	$\mathbf{y_t}$	$\mathbf{m_t}$	$\mathbf{f_t}$	$\mathbf{e_t}$	$\mathbf{\lvert e_t\rvert}$	$\mathbf{\lvert e_t\rvert / y_t}$	$\mathbf{e_t^2}$
2000	1	11	13.80	15.00	-4.00	4.00	0.3636	16.00
2001	2	21	15.96	13.80	7.20	7.20	0.3429	51.84
2002	3	13	15.07	15.96	-2.96	2.96	0.2277	8.76
2003	4	19	16.25	15.07	3.93	3.93	0.2068	15.44
2004	5	6	13.18	16.25	-10.25	10.25	1.7083	105.06
2005	6	forecast		**13.18**				
Total	-	-	-	-	-6.08	28.34	2.8493	197.1

The original time series along with the one-step ahead forecast are shown below



The forecast accuracy Identifiers

1. Bias : $B = \Sigma\, e_t \;/\; ς = -6.08/5 = -1.216$

2. Mean Absolute Error: $MAE = \Sigma\, |e_t| \;/\; ς = 28.34/5 = 5.668$

3. $MAPE = = \Sigma\, (|\, e_t\, |\, /\, y_t\,) \;/\; ς = 2.8493/5 = 0.5699$ (or 57 %)

4. Mean Square Error: $MSE = \Sigma\; e_t^2 \;/\; ς = 197.1/5 = 39.42$

5. Root Mean Square Error: $RMSE = \sqrt{39.42} = 6.2785$

6. Tracking Signal : $TS = \Sigma\, e_t \;/\; MAE = ς\, B \,/\, MAE = -6.08/5.668 = -1.073$

Notice that, unlike the Moving Average method (where equal weights are used for the averaging process) the Exponential Smoothing method relies on unequal weighting system whereby, more weight is given to the most recent observation and less weight is considered for past observations. These weights which are assigned in an exponential manner decay monotonically over the age of the observations. In other words, in Exponential Smoothing method more importance is given to the most recent observation and less to the older ones and so on.

Exponential Weights

For the single exponential smoothing model: $Y_t = a+E_t$ (I)
the parameter a, the level of the underlying process, at time t is estimated by the m_t as follows.
$m_t = \beta m_{t-1} +(1-\beta)y_t$ (II)

Defining B as a backward shift operator, such that $Bm_t=m_{t-1}$, we can write the equation (II) as:
$m_t = \beta Bm_t +(1-\beta)y_t$

or $(1- \beta B) m_t = (1-\beta)y_t$

or $m_t = (1-\beta) (1- \beta B)^{-1} y_t$ (III)

Writing $1-\beta = \alpha$, a smoothing coefficient, such that $0<\alpha<0$ we can write equation (III) as:

$m_t = \alpha (1- \beta B)^{-1} y_t$

$= \alpha \{ {}_{j=0}\sum^{\infty} (\beta B)^j y_t$

$= \alpha \{ {}_{j=0}\sum^{\infty} \beta^j y_{t-j} \}$ (IV)

It can be proved that the sum of exponential weights $\{w_j\}_{j=0,1,\ldots,\infty} = \{\alpha\beta^j\}_{j=0,1,\ldots,\infty} =1$ (V)

Example.

For n=9 observations, $\beta \cong (n-1)/(n+1) = 8/10 = 0.8 \rightarrow \alpha =1-\beta =1-0.8= 0.2$ the weights $\{w_j\}_{j=0,1,\ldots,\infty}$ would be:

$w_1= \alpha\beta^0 = 0.2(1) =0.20$
$w_2= \alpha\beta^1 = 0.2(0.8)=0.16$
$w_3= \alpha\beta^2 = 0.2(0.64)=0.128$
$w_4= \alpha\beta^3 = 0.2(0.512)=0.1024$

and so on.

The sum of these weights $\{w_j\}_{j=0,1,\ldots,\infty} = 0.2\{0.8^j\}_{j=0,1,\ldots,\infty}$
$= 0.20 + 0.16 + 0.128 + 0.1024 +\ldots\ldots+ 0.0000=1.$

19.6 Linear Regression Based Methods

In linear regression method we assume that the observations of a time series depend upon time. The dependence on time is expressed through the linear regression model which may be simple or multiple in nature. Both of these types of methods are explained as follows.

19.6.1 Simple Linear Regression Method

In this method we assume that the observations of a time series, being studied, can be adequately modelled as:

$$\mathbf{Y_t = a + b\,t + E_t \quad ; \qquad E_t \approx N(0,V)}$$

Where at time t,

a= trend or level of the process underlying the time series being studied
b= growth or decline coefficient of the trend.
E= random error or irregular component of time series. This error is assumed to be independently, identically and Normally distributed with mean zero and some constant variance, say V.

To estimate the values of observations at time t we first need to estimate this model and its parameters. This estimate of the model is:

$$\mathbf{m_t = a^{\wedge} + b^{\wedge}\, t}$$

Where for $\mathbf{\bar{t} = \Sigma t / n}$ **and** $\mathbf{\bar{y} = \Sigma y / n}$

$$\mathbf{b^{\wedge} = Covariance\,(t,y) / Variance\,(t) = S_{t,y} / S^2_t}$$

$$\mathbf{S_{t,y} = \Sigma\,(t - \bar{t})(y - \bar{y}) / (n-1)}$$

$$\mathbf{S^2_t = \Sigma\,(t - \bar{t})^2 / (n-1)}$$

$$\mathbf{a^{\wedge} = Average(y_t) - b^{\wedge}\; Average\,(t)}$$
$$= \mathbf{\bar{y} - b^{\wedge}\,\bar{t}}$$

Example (Simple Linear Regression)

Consider the above data (example #1,2,3) on the price of certain commodity again.

Year:	2000	2001	2002	2003	2004
Price:	11	21	13	19	6

For these data a simple linear regression model is written as:

Price = Level of process underlying the price + Growth (or decline) coefficient (time) + Random Error

$$\mathbf{Y_t = a + b\,t + E_t}$$

Assuming that $E_t \approx N(0,V)$ the estimate of the model is:

$$\mathbf{m_t = a^{\wedge} + b^{\wedge}\, t}$$

To compute this estimate let us complete the following table.

Year	**t**	$\mathbf{y_t}$	$\mathbf{(t - \bar{t})}$	$\mathbf{(t_t - \bar{t})^2}$	$\mathbf{(y_t - \bar{y})}$	$\mathbf{(t - \bar{t})(y_t - \bar{y})}$
2000	1	11	-2	4	-3	6
2001	2	21	-1	1	7	-7
2002	3	13	0	0	-1	0
2003	4	19	1	1	5	5
2004	5	6	2	4	-8	-16
Total	-	70	0	10	0	-12

From these results the parameters of the regression model are as follows.

$$\bar{t} = \Sigma t / n = 15/5 = 3$$
$$\bar{y} = \Sigma y / n = 70/5 = 14$$
$$S_{t,y} = \Sigma (t - \bar{t})(y - \bar{y}) / (n-1) = -12/4 = -3$$
$$S^2_t = \Sigma (t - \bar{t})^2 / (n-1) = 10/4 = 2.5$$
$$\hat{b} = S_{t,y} / S^2_t = -3/ 2.5 = - 1.2$$
$$\hat{a} = \bar{y} - \hat{b}\ \bar{t} = 14 – (-1.2)(3) = 17.6$$

Using these estimates of the parameters of the model the estimate of the model is:

$$m_t = \hat{a} + \hat{b}\ t = 17.6 – 1.2\ t$$

On the basis of this estimate of the model the m_t values, one step ahead forecasts are as follows.

Year	**t**	$\mathbf{y_t}$	$\mathbf{m_t}$	$\mathbf{f_t}$	$\mathbf{e_t}$	$\mathbf{\lvert e_t \rvert}$	$\mathbf{\lvert e_t \rvert / y_t}$	$\mathbf{e_t^2}$
2000	1	11	16.4	16.4	-5.4	5.4	0.4909	29.16
2001	2	21	15.2	15.2	5.8	5.8	0.2762	33.64
2002	3	13	14.0	14.0	-1.0	1.0	0.0769	1.00
2003	4	19	12.8	12.8	6.2	6.2	0.3263	38.44
2004	5	6	11.6	11.6	-5.6	5.6	0.9333	31.36
2005	6		forecast	10.4				
Total	-	70	70.0	-	0	24.0	2.1036	133.60

The original time series along with the one-step ahead forecast are shown below :

The forecast accuracy Identifiers

1. Bias : B = $\Sigma\, e_t$ / ς = 0/5 = 0
2. Mean Absolute Error: MAE = $\Sigma\, |e_t|$ / ς = 24/5 = 4.8
3. MAPE = = $\Sigma\, (|\, e_t\, |\, /\, y_t\,)$ / ς = 2.1036/5 = 0.42072 = 42.072%
4. Mean Square Error: MSE = $\Sigma\, e_t^2$ / ς = 133.6/5 = 26.72
5. Root Mean Square Error: RMSE = $\sqrt{26.72}$ = 5.1691
6. Tracking Signal : TS = $(\Sigma\, e_t)$ / MAE = ς B / MAE = 0/4.8 = 0

Comments:
This linear regression model is used assuming that the random errors are uncorrelated and independent random variables forming a white noise sequence.

This assumption is rarely true in real life situations as in many cases some sort of color of relationship appears in the error terms. In such cases, the estimates of parameters of the model are not optimal. These sub optimal estimates, in turn, yield bad forecasts. To cope with such situations we need models having provision of the colored noise processes, such as, Linear Dynamic System and State Space Models.

19.7 Seasonal Forecasting

For analysis and forecasting of seasonal time series, numerous models, from very simple to sophisticated are available. Some of these commonly used models are:

- Moving Average Seasonal Models with Trend in:
  - Additive form, for time series: $Y_t = T_t + S_t + E_t$
  - Multiplicative form, for time series: $Y_t = T_t \times S_t \times E_t$
  - Additive - Multiplicative – form, for time series: $Y_t = T_t \times S_t + E_t$
- Holt – Winter Seasonal Models with Trend in:
  - Additive form for time series: $Y_t = T_t + S_t + E_t$, called HWA model.
  - Multiplicative form for time series: $Y_t = T_t \times S_t \times E_t$, called HWM model.
- Seasonal Auto-Regressive and Moving Average Models.
- Seasonal Linear Dynamic System Models
- Seasonal State Space Models

To have better insight in to some of the above models, let us go through the following examples.

Example #1 (Moving Average Method using Additive Model)

Consider the following quarterly sale (in million $) of **Nolege**, a business organization, for the last three years.

Quarters

Year	Q_1	Q_2	Q_3	Q_4
2000	1	3	2	6
2001	5	7	6	10
2002	9	11	10	14

These sales data are analyzed by the Moving Average Method, using 4 quarter (seasonal) moving average and considering , at time t, the trend - seasonal model in additive form. That is,

$$Y_t = T_t + S_t + E_t$$

Where at time t :

Y_t is an observation, T_t is a Trend component, S_t is a seasonal variation component and E_t is a random noise, which is assumed independently and identically distributed with ,mean zero and some constant variance, say V_E.

An estimate of this model at time t is written as:

$$Y_t^{\wedge} = T_t^{\wedge} + S_t^{\wedge}$$

Where at time t:
$f_t = Y_t^\wedge$ is full or seasonally adjusted forecast , $T_t^\wedge = m_t$ is an estimate of trend T_t and $S_t^\wedge$ is an estimate of the seasonal variations S_t .

These values are computed through the following working table #1.

Working Table #1.1

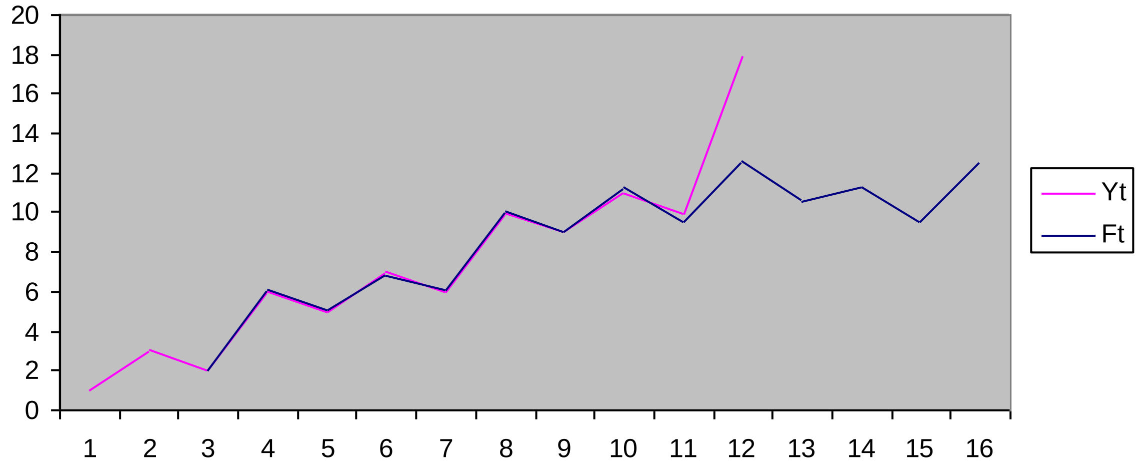
Year	**Q**	**T**	**Y_t**	**4Q-Moving Average**	**4Q-Centred Moving Average $m_t = T_t^\wedge$**	**Δt = Yt $-T_t^\wedge$**	**$S_t^\wedge$**	**$f_t = T_t^\wedge + S_t^\wedge$**	**$e_t = y_t - f_t$**
2000	(I)	1	1	-	-		-0.4375	-	
	(II)	2	3	-	-		0.3125	-	
				3					
	(III)	3	2		3.5	-1.5	-1.4375	2.0625	-0.0625
				4					
	(IV)	4	6		4.5	1.5	1.5625	6.0625	-0.0625
				5					
2001	(I)	5	5		5.5	-0.5	-0.4375	5.0625	-0.0625
				6					
	(II)	6	7		6.5	0.5	0.3125	6.8125	0.1875
				7					
	(III)	7	6		7.5	-1.5	-1.4375	6.0625	-0.0625
				8					
	(IV)	8	10		8.5	1.5	1.5625	10.0625	-0.0625
				9					
2002	(I)	9	9		9.5	-0.5	-0.4375	9.0625	-0.0625
				10					
	(II)	10	11		11.0	0.0	0.3125	11.3125	-0.3125
				12					
	(III)	11	10	-	11.0	-	-1.4375	9.5625	-
	(IV)	12	18	-	11.0	-	1.5625	12.5625	-
2003	**(I)**	**13**	**forecast**	**-**	**11.0**		**-0.4375**	**10.5625**	**-**
	(II)	**14**		**-**	**11.0**		**0.3125**	**11.3125**	**-**
	(III)	**15**		**-**	**11.0**		**-1.4375**	**9.5625**	**-**
	(IV)	**16**		**-**	**11.0**		**1.5625**	**12.5625**	**-**

The seasonal component of the model is computed through the following working table #2.

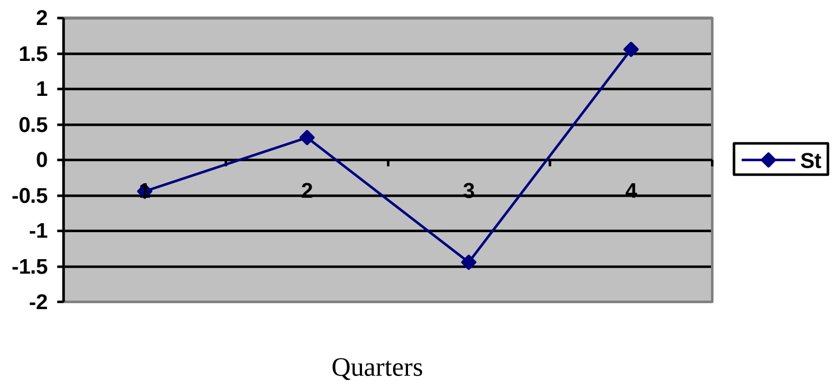
Working Table #1.2
For Estimation of Seasonal Variations
Quarters

Year	I	II	III	IV	Total
2000	-	-	-1.5	1.5	
2001	-0.5	0.5	-1.5	1.5	
2002	-0.5	0.0	-	-	
Total	-1.0	0.5	-3.0	3.0	
Average (Av)	-0.5	0.25	-1.5	1.5	-0.25
$S_t^\wedge$ = Av + AF = A_v + .0625	-0.4375	0.3125	-1.4375	1.5625	0.0000

Where the Adjustment factor : AF = –0.25 / 4 = -0.0625 (to be added into the average values)



Plot #1.1 (The observations (actual sales figures) and their forecasts)



Plot #1.2 (Seasonal Variations)

Example #2 (Moving Average Method with Additive - Multiplicative Model)

Reconsider the quarterly sale (in million $) of Nolege , for the last three years. That is,

	Quarters			
Year	**Q_1**	**Q_2**	**Q_3**	**Q_4**
2000	1	3	2	6
2001	5	7	6	10
2002	9	11	10	14

Now, these data are analyzed by assuming an Additive - multiplicative model:

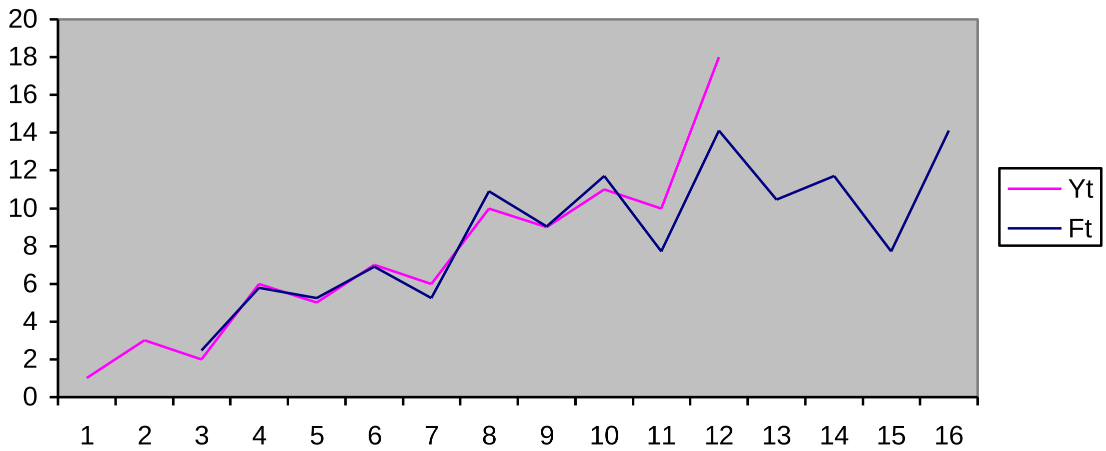
$$Y_t = T_t \times S_t + E_t$$

Where the components of this model are as defined earlier.

An estimate of this model is:

$$\hat{Y}_t = \hat{T}_t \times \hat{S}_t$$

These estimated components of the model are computed through the following working table.

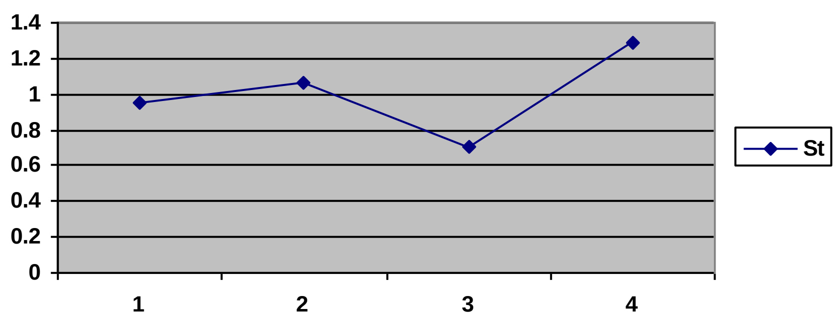


Plot #2.1 (Observations and Seasonally Adjusted Forecasts)

Working Table #2.2
For Estimation of Seasonal Indices
Quarters

Year	I	II	III	IV	Total
2000	-	-	0.57	1.33	
2001	0.91	1.08	0.80	1.18	
2002	0.95	1.00	-	-	
Total	1.86	2.08	1.37	2.51	
Average (Av)	0.930	1.040	0.685	1.255	3.910
$\hat{S}_t$ = Av/AF =Av /0.9775	0.9514	1.0639	0.7008	1.2839	4.0000

Where, Adjustment factor : AF = 3.910 / 4 = 0.9775 (the average values to be divided by 0.9775)



Quarters

Plot #2.2 (Seasonal Indices)

Comments:
Looking at the plots of seasonally adjusted forecasts, generated by the additive moving average model (example #1) and by the additive-multiplicative moving average model (example #2), it is to be noticed that the first (additive) model performed better than the second (additive-multiplicative) model. This is owing to the fact that in the first case, the model generated. Relatively, better forecasts of the sale of Nolege.

19.8 Simulation of Forecasts

For generating Forecasts for any phenomenon of interest using Monte-Carlo simulation technique, consider the price (Y_t in \$) of certain commodity over the last 25 days.

Day:	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Price:	20	30	20	40	30	50	30	40	50	40
Day:	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20
Price:	60	50	40	50	60	40	50	30	70	60
Day:	#21	#22	#23	#24	#25					
Price:	70	60	70	80	70					

Organizing these prices in to frequency distribution we get the following table.

Table #1

Observed Price at time t Y_t	Frequency F	Probability P	100 p	Intervals
20	2	0.08	8	00<08
30	4	0.16	16	08<24
40	5	0.20	20	24<44
50	5	0.20	20	44<64
60	4	0.16	16	64<80
70	4	0.16	16	80<96
80	1	0.04	4	96<100
Total	25	1.00	100	

Based on this information let us simulate 5 sets of forecasts for next 10 days of a month, using the following random numbers.

RAN # (Set #1) 06, 12, 97, 20, 06, 32, 85, 38, 19, 20
RAN # (Set#2) 24, 45, 40, 40, 71, 75, 22, 35, 52, 42
RAN# (Set#3) 53, 54, 45, 75, 49, 40, 75, 80, 63, 34
RAN# (Set#4) 62, 38, 41, 41, 24, 36, 28, 68, 77, 31
RAN# (Set#5) 36, 21, 88, 34, 92, 33, 64, 34, 33, 63

Table #2

Day	RAN #1	Y_1^*	RAN #2	Y_2^*	RAN #3	Y_3^*	RAN #4	Y_4^*	RAN #5	Y_5^*	Averages Or $F^{'}$ Casts Y^*
D_1	06	20	24	40	53	50	62	50	36	40	40
D_2	12	33	45	50	54	50	38	40	21	30	40
D_3	97	80	40	40	45	50	41	40	88	70	56
D_4	20	30	40	40	75	60	41	40	34	40	42
D_5	06	20	71	60	49	50	24	40	92	70	48
D_6	32	40	75	60	40	40	36	40	33	40	44
D_7	85	70	22	30	75	60	28	40	64	60	52
D_8	38	40	35	40	80	70	68	60	34	40	50
D_9	19	30	52	50	63	50	77	60	33	40	46

D_{10}	20	30	42	40	34	40	31	40	63	50	50

Comments:

#1. In table #1 the frequencies are written by considering just the repetition of Y- values. However, instead of just Y- values, intervals, groups or classes may be considered to construct frequency distribution.

#2. In table #1, the probabilities are computed by simply using the relative frequency approach. However, these probabilities may be generated by employing some standard or known probability functions, such as, Binomial, Poisson and Normal.

#3. In table #2, 5 sets of simulated forecasts are generated to show the working for forecasts. In practice, more or less than 5 sets may be generated, using random numbers and averages taken. However, in case of more sets the accuracy of forecasts is expected to increase.

#4. Based on 5 sets average simulated forecasts, the mean of means comes to 468/10 = 46.8. This values is not far away from the average of the original 25 data values, which is 48.5.

#5. The Simulated forecasts may be generated for 25 days, instead of just 10 days, and compared with the original time series by computing various forecast accuracy measuring criteria, such as, Bias, MAE and MSE .

#6. The forecasts are generated using a very simple procedure, whereby, the probabilities are computed using frequencies of occurrences of Events (or X values). In other situations the probabilities may be generated using statistical probability models.

Exercise # 19

#19.1

a) A time series may be decomposed in to four components. What are these components?. Explain by giving examples.

b) The following data show demand (in million tons) of certain commodity for the last 8 years.

Year	**Demand (Y_t)**
1994	15
1995	25
1996	20
1997	30
1998	10
1999	30
2000	22
2001	40

i) Analyze these data by Moving Average Method, considering 2 years as a moving period.

ii) Forecast the demand of commodity for the years 2002 and 2003.

iii) Do we obtain more accurate one step ahead forecasts, in the sense of Mean Square Error, if instead of two years, four years moving period is considered ?
If yes! Why? And if not! Why not?

#19.2

a) A time series consists of four components. What are these components ?
Describe these components by giving examples and stating the models that can be used to estimate these components in an optimum manner.

b) The following data show price (in$) of certain commodity for the last three years, recorded four times a year.

		Quarter			
		1	**2**	**3**	**4**
Year	**2000**	3	2	5	3
	2001	6	4	7	5
	2002	10	7	11	9

i) Analyze these quarterly data by the **Moving Average Method**

ii) Estimate the **Seasonal Component** assuming that it follows an additive model.

iii) Using your estimated model **Forecast** the seasonally adjusted price of the commodity for all quarters of 1998.

In all spheres of life we, as individuals or as a group, always desire to foresee future events of interest for better planning and optimizing our decisions.

For example :

iv) an investor would like to know future outcome (dividend) of a stock before purchasing its shares.

v) a business organization would like to know possible sales of its new product in the near to distant future for planning purpose.

vi) both, individuals and organizations would be interested to know shapes of new things to appear in future, such as, types of computers, cars and cameras.

#19.3(June 2006)

At time t, consider a discrete time series model: $Y_t = T_t + S_t + E_t$

a) Explain the components of this model.

b) As an example on this model, consider the following tri-annual data on sale (in million $) of Shahzad Enterprise.

Year	Period #1	Period #2	Period #3
2003	12	14	13
2004	15	17	13
2005	18	20	19

i)Analyze these sale data by Moving Average method, assuming the above model and estimate the trend and the seasonal components of this model.
ii)Generate seasonally adjusted forecasts of sale for all three periods of next year.

Solution #19.3

b)

Year	Period	Y_t	$T_t^{\wedge}$= MA(3)	$D_t = Y_t - T_t^{\wedge}$	$S_t^{\wedge}$	F_{t+1}
2003	1	12	-	-		
	2	14	13	1		
	3	13	14	-1		
2004	1	15	15	0		
	2	17	15	2		
	3	13	16	-3		
2005	1	18	17	1		
	2	20	19	1		
	3	19	-	-		
Total		141				
2006	1	-	19		0.56	19.56
	2	-	19		1.39	20.39
	3	-	19		-1.95	17.05

Year/Period	1	2	3	Total
2003	-	1	-1	
2004	0	2	-3	
2005	1	1	-	
Total	1	4	-4	
Average	0.5	1.33	-2	-0.17
$S_t^{\wedge+}$	0.56	1.39	-1.95	0.00

#19.4 (June 2007)
At time t, consider a discrete time series model: $Y_t = T_t + S_t + E_t$
a) Explain all the components of this model. (3 marks)

b) As an example on this time series analysis model, consider the following data on sale (in million \$) of XYZ business, recorded every 4^{th} month of a year for four consecutive years.

Year	Period #1	Period #2	Period #3
2007	3	8	4
2008	6	11	7
2009	9	17	13
2010	15	20	16

i)Analyze these sale data by Moving Average method, using the above time series model and estimate the trend and the seasonal component of this model.
ii)Generate seasonally adjusted forecasts of sale of XYZ business for next year.

Solution Question #4
a) $Y_t = T_t + S_t + E_t$; Where at time t, Y_t is an observation on sale of XYZ business, T_t is a trend component, S_t is a seasonal component, representing seasonal variations in sale within a year and E_t is random error.

Additive Model

Year	Period	y_t	$T_t^\wedge$MA(3)	$d_t=y_t- T_t^\wedge$	$S_t^\wedge$	f_t
2006	#1	3	-	-		
	#2	8	5	3		
	#3	4	6	-2		
2007	#1	6	7	-1		
	#2	11	8	3		
	#3	7	9	-2		
2008	#1	9	11	-2		
	#2	17	13	4		
	#3	13	15	-2		
2009	#1	15	16	-1		

	#2	20	17	3		
	#3	16	-	-		
2010	#1	-	17	-	-1.303	15.697
	#2	-	17	-	3.277	20.277
	#3	-	17	-	-1.973	15.027

Estimation of Seasonal Variations

d_t	#1	#2	#3	∑	
2007	-	3	-2		
2008	-1	3	-2		
2009	-2	4	-2		
2010	-1	3	-		
∑	-4	13	-6		
Av.	-1.33	3.25	-2	-0.08	AF= -0.08/3= -0.027
$\hat{S}_t$	-1.303	3.277	-1.973	0	

QM353 Class Lecture Notes #16 **Dr. M. Akram** (12th December, 2011)

Time Series Analysis and Forecasting

#1. Time Series and its Components

A time series is a sequence of observations or instant records on an event of interest, recorded over equal passage of time, such as, hours, days, weeks, months and years in a chronological order. When analyzed, it reveals valuable information about phenomenon being studied and leads analysts to foresee future in a meaningful manner. It may be considered as an ensemble of four or few of the components: Trend, Seasonal , Cyclical and Random Shocks. These components are explained as follows.

Trend (T_t) is a long term movement of time series that exhibits increase or decrease over a passage of time in a phenomenon being studied. It is linear if increase or decrease, on the average, remains constant; otherwise, the trend is non linear. As an example, consider the following sale (in million $) of two products: ice cream and woollen clothes.

Table #1 (Sale of Ice Cream) Table #2 (Sale of Woollen Clothes)

Year	**2006**	**07**	**08**	**09**	**10**	**Σ**		**2006**	**07**	**08**	**09**	**10**	**Σ**
Annual Sale	5	6	8	10	16	**45**		15	11	9	8	7	**50**

These figures reflect positive growth in trend of the sale of ice cream and decline in trend of the sale of the woollen clothes over the passage of 5 years time period.

Seasonal Component (S_t) is a **short term wavelike pattern** of time series **within a year.** It is typically found in summer and winter seasons of a year. However, quarterly and monthly data, quite often, also exhibit seasonal variations. As an example, consider the following sale of ice cream and the woollen clothes..

Table#1 (Sale of Ice Cream) **Table #2** (Sale of Woollen Clothes)

Season	**2006**	**07**	**08**	**09**	**10**	**Σ**		**2006**	**07**	**08**	**09**	**10**	**Σ**
Summer	3	5	6	7	12	**36**		6	2	2	3	2	**15**
Winter	2	1	2	3	4	**12**		9	9	7	5	5	**35**
Σ	**5**	**6**	**8**	**10**	**16**	**45**		**15**	**11**	**9**	**8**	**7**	**50**

These figures clearly show that in summer the sale of ice cream goes up and the sale of woolen clothes goes down.

Cyclical Component (C_t) are **long term wavelike upward and downward fluctuations** (or cycles) over a passage of many years. These fluctuations arise due to changes in business activities or/and economic conditions. For example, a fashion cycle is believed to be repeated in 15 to 20 years, a business cycle in 20 to 30 years and an economic cycle in 25 to 35 years and.

Random Error (Shock or Noise : E_t) is an irregular component of time series caused by unexpected abrupt changes in events, such as, earth quick, war and weather catastrophe. It is the most difficult component of the time series to understand, model and estimate accurately. Usually, the mean of E(t) is assumed zero.

Time Series Models

There are three possible ways to model time series (Y_t),containing all the four or fewer components.

i)**Additive Model:** $Y_t = T_t + S_t + C_t + E_t$ (Linear Form), an estimate of which is: $Y_t^{\wedge} = T_t^{\wedge} + S_t^{\wedge} + C_t^{\wedge}$

ii)**Multiplicative Model**: $Y_t = T_t \text{ x } S_t \text{ x } C_t \text{ x } E_t$ (Non linear form).

In linear form this model is written as: Log Y_t = Log T_t + Log S_t + Log C_t + Log E_t , an estimate of which is: Log $Y_t^{\wedge}$=Log $T_t^{\wedge}$+ Log $S_t^{\wedge}$+ Log $C_t^{\wedge}$

iii)**Mixed (Multiplicative-Additive)** Model: $Y_t = T_t \times S_t \times C_t + E_t$, an estimate of which is: $Y_t^{\wedge} = T_t^{\wedge} \times S_t^{\wedge} \times C_t^{\wedge}$

#2. Forecasting is an art of foreseeing future through the scientific perceptions of Statistics or otherwise. There are three types of forecasts: Short Term(up to one year), Medium Term (2 to 4 years) and Long Term forecasts (5 or more years). These forecasts may be qualitative or/and quantitative in nature. Here, the discussion is confined to the quantitative forecasting only.

#3. Estimation and Projection or Forecasting of Trend (T_t)

For estimation of trend the model considered is $Y_t = T_t + E_t$ an estimate of which is $Y_t^{\wedge} = T_t^{\wedge}$ This estimate may be computed by using numerous techniques, such as, Linear Regression, Moving Average and Exponential Smoothing. Some of these techniques are discussed here.

#3.1 Simple Linear Regression Method

Model: $Y_t = T_t + E_t = A + Bt + E_t$

Where the parameter **A** is the level of the underlying process, the parameter **B** is the regression coefficient and E is a random error term (assumed to be independently, identically and normally distributed with mean zero and a constant variance $\sigma_E^{\ 2}$ {i.e. **$E_t \sim$ IIN(0, $\sigma_E^{\ 2}$)}.**

Example #1 (SLR Model)

Given below are the sale data (Y_t in million $) of ice cream during the last five years: These data are analyzed using simple linear regression model **$Y_t = T_t + E_t = A + Bt + E_t$**, the **Least Square Error** (LSE) estimate of which is **T_t^ = 1.2+2.6 t.** On basis of this equation estimates of trend values, Bias, MAE and MSE are computed as follows.

Year	**Time t**	**Y_t**	**$T_t^{\wedge}$ = $Y_t^{\wedge}$**	**$e_t = Y_t - Y_t^{\wedge}$**	**$\lvert e_t \rvert$**	**e_t^2**
2006	1	5	3.8	1.2	1.2	1.44
2007	2	6	6.4	-0.4	0.4	0.16
2008	3	8	9.0	-1.0	1.0	1.00
2009	4	10	11.6	-1.6	1.6	2.56
2010	5	16	14.2	1.8	1.8	3.24
$\sum$		**45**		**0.0 = $\sum e_t$**	**6.0 = $\Sigma \lvert e_t \rvert$**	**8.4= Σe_t^2**
2011	6	---	**16.8**	-	-	-
Accuracy Measures	---	---	---	**Bias: B** =Σe_t/(n-2)= **0**	**MAE=** $\Sigma \lvert e_t \rvert$ /(n-2)= **2**	**MSE=** Σe_t^2/(n-2)= **2.8**
Standard Error	---	---	---	**---**	**---**	**$S_e = \sqrt{MSE}$=1.6733**

Where df= n-2=3 as there are two parameters (A, B) of the linear regression model.
The projection (or forecast) of trend of sale of ice cream in the year 2011 is $16.8 million.

Comments:

i) By Bias B=0 means that the simple linear regression model is unbiased. It is neither over estimating nor under estimating the actual sale of ice cream.
ii) MAE and MSE are used for identification of an appropriate model from a family of available models.

QM353 Class Lecture Notes #17 **Dr. M. Akram** **(14^{th} December, 2011)**

Time Series Analysis and Forecasting

Moving Average (MA) Method. In this method we use the **Model:** $\mathbf{Y_t = T_t + E_t}$ **;** where $\mathbf{T_t = A}$, the level of the underlying process, and E is a random error (assumed to be independently and identically distributed with mean zero and a constant variance $\sigma_E{}^2$ {i.e. $\mathbf{E_t \sim II(0, \sigma_E{}^2)}$**}.** The parameter A, at time t, is estimated by the Moving averages having the moving period P ≥2 (an integer).

Example #1. Given below are sales (in million $) of a certain commodity for the last 6 years.

Year	2006	2007	2008	2009	2010	2011
Sales	4	8	6	10	5	9

Let us analyze these data by the Moving Average (MA) model $\mathbf{Y_t = T_t + E_t}$**,** using the moving periods: P= 2 and P=3, generate one step ahead forecasts (f_{t+1}), obtain one step ahead forecast errors (e_t) and compute the forecast accuracy measures (Bias, Mean Absolute Error (MAE) and Mean Square Error (MSE).

#1.1 MA(2)

Year	**Time (t)**	$\mathbf{Y_t}$	$\mathbf{\hat{T}_t}$ **MA(2)**	$\mathbf{f_{t+1}}$	$\mathbf{e_t=Y_t-f_{t+1}}$	$\mathbf{\lvert e_t \rvert}$	$\mathbf{e_t^2}$
2006	1	4	-		-	-	-
2007	2	8	6	-	-	-	-
2008	3	6	7	6	0	0	0
2009	4	10	8	7	3	3	9
2010	5	5	7.5	8	-3	3	9
2011	6	9	7	7.5	1.5	1.5	2.25
∑		**42**	**-**		$\mathbf{1.5 = \sum e_t}$	$\mathbf{7.5=\Sigma\lvert e_t\rvert}$	$\mathbf{20.25= \Sigma e_t^2}$
2012 F'cast	7	-		**7**	-	-	-
Accuracy Measures					**Bias: B** $=\Sigma e_t/3$ **=0.5**	**MAE=**$\Sigma\lvert e_t\rvert/3$ **=2.5**	**MSE=** $\Sigma e_t^2/3$ **=6.75**
Standard Error							$\mathbf{S_e=\sqrt{MSE}}$ **=2.5981**

Where, df =(n-No. of Parameters –Moving Av. period). In our case, there is **one** parameter **A** only amd P=2, therefore, df**=** (6-1-2)=3 and forecast of sale during the year 2012 is $7 million .

#1.2 MA(3)

Year	**Time (t)**	$\mathbf{Y_t}$	$\mathbf{\hat{T}_t}$ **MA(3)**	$\mathbf{f_{t+1}}$	$\mathbf{e_t=Y_t-f_{t+1}}$	$\mathbf{\lvert e_t \rvert}$	$\mathbf{e_t^2}$
2006	1	4	-		-	-	-
2007	2	8	-	-	-	-	-
2008	3	6	6	-	-	-	-
2009	4	**10**	8	**6**	**4**	**4**	**16**
2010	5	**5**	7	**8**	**-3**	**3**	**9**
2011	6	**9**	8	**7**	**2**	**2**	**4**
∑		**42**	**-**		$\mathbf{3 = \sum e_t}$	$\mathbf{\Sigma\lvert e_t\rvert =9}$	$\mathbf{\Sigma e_t^2 =29}$
2012 (F'Cast)	7	-		**8**	-	-	-
Accuracy Measures					**Bias: B** $=\Sigma e_t/2$ =1.5	**MAE=** $\Sigma\lvert e_t\rvert/2$=**4.5**	**MSE=** $\Sigma e_t^2/2$=**14.5**
Standard Error							$\mathbf{S_e=\sqrt{MSE}}$ **=3.8079**

Where, df =(n- Number of Parameters –Moving Av. Period). In our case, there is one parameter **A** only and P=3, therefore, df**=** (6-1-3)=2 and forecast of sale during the year 2012 is $8 million .

Identification of Moving Period that generates the best one step ahead forecasts.

Forecast Accuracy Measure Criteria

Moving Average	Bias	MAE	MSE
MA(2)	0.5	2.5	6.75
MA(3)	1.5	4.5	14.5

Looking at above table it is noticed that: MA(2) performs better than MA(3), because Bias and MSE of MA(2) are less than MA(3). Therefore, MA(2) is selected to yield one step ahead forecasts.

The Bias itself is interpreted as follows.

i) Bias: B=0 $\rightarrow$ the model (method or technique) of forecasting being used is neither underestimating nor overestimating the observations of phenomenon being studied . That is the forecasts, on the average, are statistically equal to the observations.

ii Bias: B> 0 $\rightarrow$ the model (method or technique) of forecasting being used is underestimating the observations of phenomenon being studied . That is the forecasts, on the average, are statistically less than the observations.

iii) Bias: B< 0 $\rightarrow$ the model (method or technique) of forecasting being used is overestimating the observations of phenomenon being studied . That is the forecasts, on the average, are statistically more than the observations.

Exercise #17

<table>
<tr><th>Questions</th><th>Answers</th></tr>
<tr><td>

#1. Given below are daily prices (Yt) of a certain commodity during the last seven days of a week.

Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Y_t	3	9	6	15	12	21	18

a) Analyze these prices by the Moving Average model using the moving periods: i) P= 2, ii) P= 3, iii)P=4
b)Forecast price of the commodity for Monday of the next week.
c) Compute the Bias (B), Mean Absolute Error (MAE) and Mean Square Error (MSE) of the one step ahead forecasts.
d) Identify the moving period (P=2, P=3 or P=4) which generates better one step ahead forecasts.

</td><td>

#1.
a)
b)
c)
d)

Model/Criteria	Bias	MAE	MSE
M_1			
M_2			
M_3			

</td></tr>
<tr><td>

#2. To identify the best model out of three available models the forecast accuracy measures : Bias, MAE and MSE are computed as follows.

Model/Criteria	Bias	MAE	MSE
M_1	-2.5	10	16
M_2	1.4	8	12
M_3	0.0	14	18

a)For each model interpret the Bias.
b)Which model, in your opinion, is better than the others to generate one step ahead forecasts.

</td><td>

#2.
a)The model:
i) M_1 is overestimating as Bias is negative
ii)M_2 is underestimating as Bias is positive
iii)M_3 is neither overestimating nor underestimating as Bias is zero.
b)The Model M_2 performs better than the other models as MAE=8 and MSE=12 of this model are lower than the MAE and MSE of the other models.

</td></tr>
</table>

QM353 Class Lecture Notes #18 **Dr. M. Akram** **(20th December, 2011)**

Analysis and Forecasting of Seasonal Time Series

For analysis and forecasting of time series containing trend and seasonal variations a multiplicative-additive model considered is:

$\mathbf{Y_t = T_t \times S_t + E_t}$ { where $E_t \sim IIN(0, \sigma_E^2)$ }.

An estimate of model is:

$\mathbf{Y_t^{\wedge} = T_t^{\wedge} \times S_t^{\wedge}}$.

The trend component T_t of this model is estimated by the simple linear regression model and the seasonal component S_t is estimated by using the Moving Average Model by taking the moving average period P equal to the seasonal period (the number of observations in a year) .

Example #1 (Analysis of Trivial time series using Moving Average Model)
Given below are time series recorded three times a year. These series are analyzed by using the above model . The trend values of this model are obtained using the regression line equation: $\mathbf{T_t^{\wedge}=27- 0.8t}$ and the seasonal variations are estimated through the three period moving average method.

Table #1.

Year	**Period**	**Time(t)**	$\mathbf{Y_t}$	**Tt^ =MA(3)**	$\mathbf{I_t=Y_t/T_t^{\wedge}}$		
2009	P_1	1	21	-----	----		
	P_2	2	30	75/3=25	1.20		
	P_3	3	24	72/3=24	1.00		
2010	P_1	4	18	72/3=24	0.75		
	P_2	5	30	75/3=25	1.20		
	P_3	6	27	72/3=24	1.13		
2011	P_1	7	15	75/3=25	0.60		
	P_2	8	33	57/3=19	1.74		**Seasonally Adjusted**
	P_3	9	9	-	-		**Forecasts**
				$\mathbf{T_t^{\wedge}}$		$\mathbf{S_t^{\wedge}}$	$\mathbf{f_{t+1}=T_t^{\wedge}xS_t^{\wedge}}$
2012	P_1	10	-	**19.0**	-	**0.649**	**19.0x0.649=12.3310**
	P_2	11	-	**18.2**	-	**1.327**	**18.2x1.327=24.1514**
	P_3	12	-	**17.4**	-	**1.024**	**17.4x1.024=17.8176**

Table #2 (**Estimation of seasonal variations (or Indices)** $\mathbf{S_t^{\wedge}}$ **using** I_t –values)

Year/Period	P_1	P_2	P_3	∑
2009	----	1.20	1.00	
2010	0.75	1.20	1.13	
2011	0.60	1.74	-----	
∑	1.35	4.14	2.13	
Averages: AVs	0.675	1.38	1.065	∑AVs= 3.12 → Adjustment Factor: AF= AVs/3 = **1.04**
$S_t^{\wedge}$=AVs/AF	**0.649**	**1.327**	**1.024**	∑ $\mathbf{S_t^{\wedge}}$**=3.00** (=Seasonal/Moving Average Period)

Example #2 (Biannual Time Series)
Given below are data on sale of ice cream , recorded two times a year over the time period of 5 years. Let us analyze these data by considering the model: $\mathbf{Y_t=T_t \, xS_t + E_t}$ and using **two period moving average** (= seasonal period) and project the trend by using the regression line equation: $\mathbf{T_t^{\wedge}= 3.4+0.7429t.}$

Table #1

Year	Season	Time (t)	Y_t	MA (P=2)	Trend Centred MA=Tt^	$I_t=Y_t/T_t^{\wedge}$	Seasonal Indices	Seasonally Adjusted f'casts
2009	S	1	**6**	----	-----	------		
	W	2	**2**	8/2=4	(4+4.5)/2=4.25	0.4706	---	---
2010	S	3	**7**	9/2=4.5	(4.5+5)/2=4.75	1.4737	---	---
	W	4	**3**	10/2=5.0	(5+7.5)/2=6.25	0.4800	---	---
2011	S	5	**12**	15/2=7.5	(7.5+9)/2=8.25	1.4545	---	---
	W	6	**6**	18/2=9	--------	-----	---	---
			----	-----	$\mathbf{T_t^{\wedge}}$	---	$\mathbf{S_t^{\wedge}}$	$\mathbf{f_{t+1}=T_t^{\wedge}xS_t^{\wedge}}$
2012	**S**	**7**	---	-----	**8.6003**	**---**	1.5098	**12.9847**
	W	**8**	---	-----	**9.3432**	**---**	0.4902	**4.5600**

Table #2: Estimation of seasonal variations (or Indices) $\mathbf{S_t^{\wedge}}$ using I_t -values

Year	Summer (S)	Winter (W)	∑
2009	---------	0.4706	
2010	1.4737	0.4800	
2011	1.4545	---------	
∑	2. 9282	0.9506	
Averages (Av)	1.4641	0.4753	∑AVs=1.9394→ AF = AVs /2= **0.9697**
$\mathbf{S_t^{\wedge}=AVs/AF}$	1.5098	0.4902	$\sum S_t^{\wedge}$=2.0 (Seasonal/Moving Av. period)

Note that:
i)In seasonal forecasting the **Moving Average period P = Seasonal Period** (the number of times data are observed in a year). In our case, P=2 and sale figures are recorded two times a year. The moving /seasonal period in case of quarterly data is 4 and monthly data is 12.

ii) The **centralization** of the moving averages for alignment to the seasonal periods is required in case of even (P= 2,4,…) seasonal period only (not in case of odd seasonal/moving period; e.g. P=3,5,..) as the moving averages are already corresponding to the seasonal periods and not in between the periods.

Exercise #18

Questions	Answers
#1. Analyze the following quarterly time series by considering the Model: $\mathbf{Y_t=T_t \, xS_t + E_t}$ and using the Moving Average Method, project the trend by using the regression equation $\mathbf{T_t^{\wedge}=5.97 + 1.40 \, t}$ **,** estimate the seasonal indices and generate seasonally adjusted forecasts for all quarters of next year.	**#1.**

Year Period	2008 I	II	III	IV	2009 I	II	III	IV	2010 I	II	III	IV	2011 I	II	III	IV
$\mathbf{Y_t}$	**6**	**3**	**9**	**15**	**6**	**9**	**21**	**30**	**24**	**18**	**30**	**27**	**15**	**33**	**12**	**22**

QM353 Class Lecture Notes #19 **Dr. M. Akram** **(27th December, 2011)**

Analysis and Forecasting of Time Series

Single (or simple) Exponential Smoothing Method

For analysis and forecasting of time series $\{y_t\}_{t=1,2,\ldots,n}$ using single exponential smoothing one step ahead forecasts are obtained as follows.

$$\mathbf{F_{t+1} = F_t + \alpha (y_t - F_t)} \qquad \{ \text{ Or } \quad \mathbf{F_{t+1} = \alpha y_t + (1-\alpha) F_t} \}$$

Where:
F_t is forecast at time t
F_{t+1} is forecast at time t+1
α is a damping or smoothing factor between zero and one ; i.e**.** $\mathbf{0 \le \alpha \le 1}$**. Sometime** $\boldsymbol{\beta = 1-\alpha}$ is considered as a damping factor which lies between zero and one; i.e. $\mathbf{0 \le \beta \le 1}$**.**

At time t=0, $\mathbf{F_0}$ is the **prior information** on phenomenon being studied . If it does not exist then its value is taken as y_1.
One step ahead forecast error is defined as: $\mathbf{e_t = y_t - Ft+1}$

Example (Simple Exponential Smoothing)

Consider the following average prices (in $) of a certain commodity for the last five years.

Year	2007	2008	2009	2010	2011
Price ($): y_t	11	21	13	19	6

Let us analyze these data by simple exponential smoothing, using the smoothing (or damping) factor α=0.3 and Prior forecast at time t=0, F_1=y_1=11. On basis of this damping factor the simple exponential smoothing equation is: $\mathbf{F_{t+1} = F_t + 0.3 (y_t - F_t)}$

On basis of this equation at time:
$t=1 \rightarrow F_2 = F_1+0.3(y_1-F_1) = 11+0.3(11-11) = 11$
$t=2 \rightarrow F_3 = F_2+0.3(y_2-F_2) = 11+0.3(21-11) = 14$
$t=3 \rightarrow F_4 = F_3+0.3(y_3-F_3) = 14+0.3(13-14) = 13.7$
$t=4 \rightarrow F_5 = F_4+0.3(y_4-F_4) = 13.7+0.3(19-13.7) = 15.29$
$t=5 \rightarrow F_6 = F_5+0.3(y_5-F_5) = 15.29+0.3(6-15.29) = 12.503$

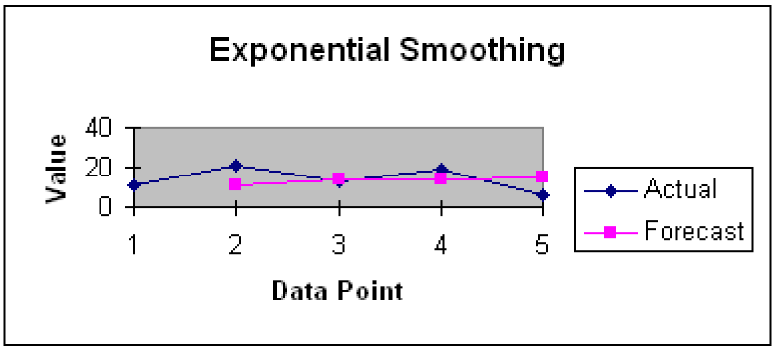
Writing these values in the table, given below, we obtain Bias, MAE and MSE.

Year	**t**	$\mathbf{y_t}$	$\mathbf{F_{t+1}}$	$\mathbf{e_t}$	$\mathbf{\lvert e_t \rvert}$	$\mathbf{e_t^2}$
2007	1	11	F_2=11	0		
2008	2	21	F_3=14	7		
2009	3	13	F_4=13.7	-0.7		
2010	4	19	F_5=15.29	3.71		
2011	5	6	F_6=12.503	-6.503		
$\sum$		**-**				
2012(F'Cast)	6	-	**12.503**			
				B = $\Sigma\, e_t$ /4 =	MAE = $\Sigma\, \lvert e_t \rvert$ /4 =	MSE = $\Sigma\, e_t^2$ / 4 =

Note that, unlike the Moving Average method (where equal weights are used for the averaging process) the Exponential Smoothing method relies on unequal weighting system ($\alpha, \alpha^2, \alpha^3, \ldots$) whereby, more weight is given to the most recent observation and less weight is considered for past observations. These weights which are assigned in an exponential manner decay monotonically over the age of the observations. In other words, in Exponential Smoothing method more importance is given to the most recent observation and less to the older ones and so on.

Excel Output
Simple or Single Exponential Smoothing
With damping factor **β=1-α=0.7**

Year	**t**	**Y(t)**	**F(t+1)**
2007	1	11	11
2008	2	21	14
2009	3	13	13.7
2010	4	19	15.29
2011	5	6	12.503
2012	**6F'Cast**		**12.503**



Exercise #19

Questions	Answers
#1. a) Analyze the exampled data using the damping factor : i) α=0.1 or β=0.9 ii) α=0.2 or β=0.8 iii) α=0.4 or β=0.6 b) Suggest one step ahead forecasts in all three cases. c)Compute the Bias, MAE and MSE of one step ahead forecasts d) Identify the damping factor that yields better one step ahead forecasts.	

Exercise: (Solution) Quarterly Data Using Multiplicative Model)

Let us analyze the following quarterly (recorded four times a year) time series, by the moving average method, estimate its seasonal indices, project the trend in to future, the four periods of the next year, and generate full (seasonally adjusted) forecasts for all time periods of next year. For this purpose, the time series model considered is:
$Y_t = T_t \, x S_t + E_t$, where $E_t \sim (0, \ \sigma_E{}^2)$

The estimate of model is:
$Y_t{}^{\wedge} = T_t{}^{\wedge} \text{ x } S_t{}^{\wedge}$
Where an estimate of trend is computed through the simple linear regression line equation:
$T_t{}^{\wedge} = 6.775 \ + 1.262 \text{ t.}$

Year	Period	Time(t)	Y_t	MA(4)	Centred MA=Tt^	$d_t = Y_t / T_t{}^{\wedge}$		
2006	#1	1	6	-				
	#2	2	3					
	#3	3	9					
	#4	4	15					
2007	#1	5	6					
	#2	6	9					
	#3	7	21					
	#4	8	30					
2008	#1	9	24					
	#2	10	18					
	#3	11	30					
	#4	12	27					
2009	#1	13	15					
	#2	14	33					
	#3	15	12					
	#4	16	22					Forecasts
			$\sum Y_t$= 280		$T_t{}^{\wedge}$		$S_t{}^{\wedge}$	$f_{t+1} = T_t{}^{\wedge} x S_t{}^{\wedge}$
2010	#1	17	-			-		
	#2	18	-			-		
	#3	19	-			-		
	#4	20				-		

Estimation of seasonal Variations (Indices)
d_t - values

Year	Period #1	Period #2	Period #3	Period #4	∑
2006	-				
2007					
2008					
2009					
∑					
Averages (AVs)					∑AVs=
$S_t{}^{\wedge}$=Adjusted Avs.					$\sum S_t{}^{\wedge}$=4.0

Adjustment Factor: AF= AVs /4=

t	Y(t)
1	5
2	6
3	8
4	10
5	16

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.943119
R Square	0.889474
Adjusted R Square	0.852632
Standard Error	1.67332
Observations	5

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	67.6	67.6	24.14286	0.016145
Residual	3	8.4	2.8		
Total	4	76			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 9*
Intercept	1.2	1.754993	0.683763	0.543204	-4.38517	6.785171	-4.38517	6.78
t	2.6	0.52915	4.913538	0.016145	0.916008	4.283992	0.916008	4.28

RESIDUAL OUTPUT

Observation	*Predicted Y(t)*	*Residuals*
1	3.8	1.2
2	6.4	-0.4
3	9	-1
4	11.6	-1.6
5	14.2	1.8

Example #1 Dr. Akram

Estimation of Trend

Year	Time(t)	Y(t)
2006	1	6
2007	2	12

2008	3	9
2009	4	15
2010	5	12

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.693375
R Square	0.480769
Adjusted R Square	0.307692
Standard Error	2.84605
Observations	5

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	22.5	22.5	2.777778	0.194171
Residual	3	24.3	8.1		
Total	4	46.8			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	6.3	2.984962	2.110579	0.125298	-3.19948	15.79948
Time(t)	1.5	0.9	1.666667	0.194171	-1.3642	4.364202

RESIDUAL OUTPUT

Observation	*Predicted Y(t)*	*Residuals*
1	7.8	-1.8
2	9.3	2.7
3	10.8	-1.8
4	12.3	2.7
5	13.8	-1.8

Estimation of Trend using SLR

Example#2

Year	Time(t)	Y(t)
2006	1	6
2007	2	9
2008	3	15
2009	4	12
2010	5	18
Total		
2011	6	

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9
R Square	0.81
Adjusted R Square	0.746667
Standard Error	2.387467
Observations	5

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	72.9	72.9	12.78947	0.037386
Residual	3	17.1	5.7		
Total	4	90			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	3.9	2.503997	1.55751	0.217224	-4.06884	11.86884	-4.06884	11.86884
Time(t)	2.7	0.754983	3.576237	0.037386	0.297306	5.102694	0.297306	5.102694

RESIDUAL OUTPUT

Observation	*Predicted Y(t)*	*Residuals*
1	6.6	-0.6
2	9.3	-0.3
3	12	3

4	14.7	-2.7
5	17.4	0.6

Example #3

Time(t)	Y_t
1	21
2	30
3	24
4	18
5	30
6	27
7	15
8	33
9	9

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.276026
R Square	0.07619
Adjusted R Square	-0.05578
Standard Error	8.155629
Observations	9

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	38.4	38.4	0.57732	0.472174
Residual	7	465.6	66.51429		
Total	8	504			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Up 95.*
Intercept	27	5.924927	4.557018	0.002614	12.98977	41.01023	12.98977	41.0
Time(t)	-0.8	1.052887	-0.75982	0.472174	-3.28968	1.689683	-3.28968	1.68

RESIDUAL OUTPUT

Observation	*Predicted Yt*	*Residuals*

1	26.2	-5.2
2	25.4	4.6
3	24.6	-0.6
4	23.8	-5.8
5	23	7
6	22.2	4.8
7	21.4	-6.4
8	20.6	12.4
9	19.8	-10.8

Qm353 Class Lecture Notes **#16** **M. Akram (12th December, 2011)**

Analysis and Forecasting of Seasonal Data

Example #3 (Trivial Data)
Let us analyze the following trivial(recorded three times a year) time series, by the moving average method, estimate its seasonal indices, project the trend in to future, the three periods of the next year, and generate full (seasonally adjusted) forecasts for all time periods of next year. For this purpose, the time series model considered is:

$Y_t = T_t \, x S_t + E_t$, where $E_t \sim (0, \ \sigma_E^{\ 2})$

The estimate of model is:

$Y_t^{\wedge} = T_t^{\wedge} \text{ x } S_t^{\wedge}$

Where an estimate of trend is computed through the simple linear regression line equation:

$T_t^{\wedge} = 5.97 + 1.40 \ t.$

Year	Period	Time(t)	Y_t	MA(3)	$I_t = Y_t/MA(3)$		
2005	#1	1	6	-	-		
	#2	2	3	18/3=6	0.5		
	#3	3	9	27/3=9	1.0		
2006	#1	4	15	10	1.5		
	#2	5	6	10	0.6		
	#3	6	9	12	0.75		
2007	#1	7	21	20	1.05		
	#2	8	30	25	1.2		
	#3	9	24	24	1.0		
2008	#1	10	18	24	0.75		
	#2	11	30	25	1.2		
	#3	12	27	24	1.125		
2009	#1	13	15	25	0.6		
	#2	14	33	60/3=20	1.65		
	#3	15	12	-	-		Forecasts
			$\sum Y_t = 258$	$T_t^{\wedge}$		$S_t^{\wedge}$	$f_{t+1} = T_t^{\wedge} x S_t^{\wedge}$
2010	#1	16	-		-		
	#2	17	-		-		
	#3	18	-		-		

Estimation of seasonal Variations (Indices)
d_t - values

Year	Period #1	Period #2	Period #3	$\sum$
2005	-	0.500	1.000	
2006	1.500	0.600	0.750	
2007	1.050	1.200	1.000	
2008	0.750	1.200	1.125	
2009	0.600	1.650	-	
$\sum$				
Averages (AVs)				$\sum$AVs=
$S_t^{\wedge}$=Adjusted Avs.				$\sum S_t^{\wedge} = 3.00$

Adjustment Factor: AF= AVs/3 =

Example : Three Period Seasonal Data

Year	Period	Time(t)	Y(t)	MA(3)	d(t)
2005	1	1	6	NA	NA
	2	2	3	6	0.5
	3	3	9	9	1
2006	1	4	15	10	1.5
	2	5	6	10	0.6
	3	6	9	12	0.75
2007	1	7	21	20	1.05
	2	8	30	25	1.2
	3	9	24	24	1
2008	1	10	18	24	0.75
	2	11	30	25	1.2
	3	12	27	24	1.125
2009	1	13	15	25	0.6
	2	14	33	20	1.65
	3	15	12		NA
Total			258		
2010	1				
	2				
	3				

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.637235
R Square	0.406069
Adjusted R Square	0.360382
Standard Error	7.877897
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	551.6036	551.6036	8.888049	0.010614
Residual	13	806.7964	62.06126		
Total	14	1358.4			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Up 95.*

Intercept	5.971429	4.28052	1.395024	0.186386	-3.27607	15.21893	-3.27607	15.2
Time(t)	1.403571	0.470794	2.981283	0.010614	0.386482	2.420661	0.386482	2.42

RESIDUAL OUTPUT

Observation	*Predicted Y(t)*	*Residuals*
1	7.375	-1.375
2	8.778571	-5.77857
3	10.18214	-1.18214
4	11.58571	3.414286
5	12.98929	-6.98929
6	14.39286	-5.39286
7	15.79643	5.203571
8	17.2	12.8
9	18.60357	5.396429
10	20.00714	-2.00714
11	21.41071	8.589286
12	22.81429	4.185714
13	24.21786	-9.21786
14	25.62143	7.378571
15	27.025	-15.025

Example #4 Quarterly Time Series Analysis and Forecasting

Year	Period	Time(t)	Y(t)	MA(4)	Centred	I(t)
2006	1	1	6		Nil	Nil
	2	2	3	8.25	Nil	Nil
	3	3	9	8.25	8.25	1.0909
	4	4	15	9.75	9	1.6667
2007	1	5	6	12.75	11.25	0.5333
	2	6	9	16.5	14.625	0.6154
	3	7	21	21	18.75	1.12
	4	8	30	23.25	22.25	1.3483
2008	1	9	24	25.5	24.375	0.9846
	2	10	18	24.75	25.125	0.7164
	3	11	30	22.5	23.625	1.2698
	4	12	27	26.25	24.375	1.1077
2009	1	13	15	21.75	24	0.625
	2	14	33	20.5	21.125	1.5621

3	15	12	Nil	Nil
4	16	22	Nil	Nil

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.626294
R Square	0.392244
Adjusted R Square	0.348833
Standard Error	7.739984
Observations	16

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	541.2971	541.2971	9.03557	0.00944
Residual	14	838.7029	59.90735		
Total	15	1380			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Up 95.*
Intercept	6.775	4.058882	1.669179	0.117281	-1.93044	15.48044	-1.93044	15.4
Time(t)	1.261765	0.41976	3.005922	0.00944	0.36147	2.162059	0.36147	2.16

RESIDUAL OUTPUT

Observation	*Predicted Y(t)*	*Residuals*
1	8.036765	-2.03676
2	9.298529	-6.29853
3	10.56029	-1.56029
4	11.82206	3.177941
5	13.08382	-7.08382
6	14.34559	-5.34559
7	15.60735	5.392647
8	16.86912	13.13088
9	18.13088	5.869118
10	19.39265	-1.39265
11	20.65441	9.345588
12	21.91618	5.083824
13	23.17794	-8.17794
14	24.43971	8.560294
15	25.70147	-13.7015
16	26.96324	-4.96324

Example # (Quarterly Data)

Year	Period	t	T_t	S_t	C_t	Y_t
2008	#1					
	#2					
	#3					
	#4					
2009	#1					
	#2					
	#3					
	#4					
2010	#1					
	#2					
	#3					
	#4					
2010						
Σ						

Theoretical Values of ASL

{ AR(1) Colored Noise Process }

Φ	ASL	Φ	ASL
-.99	1.047380	-.49	1.508438
-.98	1.068312	-.48	1.516762
-.97	1.084994	-.47	1.525126
-.96	1.099504	-.46	1.533532
-.95	1.112648	-.45	1.541981
-.94	1.124836	-.44	1.550476
-.93	1.136311	-.43	1.559018
-.92	1.147234	-.42	1.567608
-.91	1.157713	-.41	1.576250
-.90	1.167828	-.40	1.584944
-.89	1.177641	-.39	1.593692
-.88	1.187198	-.38	1.602497
-.87	1.196536	-.37	1.611360
-.86	1.205685	-.36	1.620283
-.85	1.214671	-.35	1.629268
-.84	1.223515	-.34	1.638317
-.83	1.232234	-.33	1.647431
-.82	1.240843	-.32	1.656613
-.81	1.249357	-.31	1.665865
-.80	1.257785	-.30	1.675189
-.79	1.266139	-.29	1.684586
-.78	1.274427	-.28	1.694059
-.77	1.282658	-.27	1.703610
-.76	1.290838	-.26	1.713242
-.75	1.298975	-.25	1.722955
-.74	1.307074	-.24	1.732753
-.73	1.315140	-.23	1.742637
-.72	1.323180	-.22	1.752611
-.71	1.331196	-.21	1.762676
-.70	1.339195	-.20	1.772835
-.69	1.347179	-.19	1.783090
-.68	1.355152	-.18	1.793443
-.67	1.363119	-.17	1.803899
-.66	1.371081	-.16	1.814458
-.65	1.379043	-.15	1.825124
-.64	1.387007	-.14	1.835899
-.63	1.394977	-.13	1.846787
-.62	1.402954	-.12	1.857790
-.61	1.410942	-.11	1.868911
-.60	1.418942	-.10	1.880154
-.59	1.426958	-.09	1.891521
-.58	1.434991	-.08	1.903015
-.57	1.443045	-.07	1.914641
-.56	1.451120	-.06	1.926402
-.55	1.459220	-.05	1.938301
-.54	1.467346	-.04	1.950342
-.53	1.475500	-.03	1.962528
-.52	1.483684	-.02	1.974864
-.51	1.491900	-.01	1.987353
-.50	1.500151	0.00	2.000000

Φ	ASL	Φ	ASL
.01	2.012809	.51	3.032932
.02	2.025785	.52	3.067466
.03	2.038931	.53	3.103051
.04	2.052253	.54	3.139744
.05	2.065756	.55	3.177606
.06	2.079445	.56	3.216704
.07	2.093324	.57	3.257108
.08	2.107401	.58	3.298896
.09	2.121679	.59	3.342151
.10	2.136166	.60	3.386964
.11	2.150866	.61	3.433436
.12	2.165787	.62	3.481673
.13	2.180935	.63	3.531795
.14	2.196317	.64	3.583930
.15	2.211940	.65	3.638223
.16	2.227811	.66	3.694828
.17	2.243938	.67	3.753921
.18	2.260329	.68	3.815694
.19	2.276993	.69	3.880361
.20	2.293938	.70	3.948159
.21	2.311173	.71	4.019357
.22	2.328708	.72	4.094255
.23	2.346552	.73	4.173190
.24	2.364717	.74	4.256549
.25	2.383212	.75	4.344765
.26	2.402050	.76	4.438340
.27	2.421241	.77	4.537850
.28	2.440799	.78	4.643957
.29	2.460737	.79	4.757439
.30	2.481068	.80	4.879203
.31	2.501806	.81	5.010323
.32	2.522967	.82	5.152078
.33	2.544566	.83	5.306005
.34	2.566621	.84	5.473975
.35	2.589149	.85	5.658286
.36	2.612167	.86	5.861803
.37	2.635697	.87	6.088140
.38	2.659759	.88	6.341950
.39	2.684374	.89	6.629334
.40	2.709566	.90	6.958474
.41	2.735359	.91	7.340647
.42	2.761778	.92	7.791923
.43	2.788852	.93	8.336141
.44	2.816610	.94	9.010532
.45	2.845082	.95	9.877212
.46	2.874301	.96	11.04981
.47	2.904304	.97	12.76558
.48	2.935127	.98	15.63866
.49	2.966810	.99	22.10603
.50	2.999397	.999	69.29176

#3.3 Single Exponential Smoothing Method
For this method the model considered is: $\mathbf{Y_t = T_t + E_t}$; where $\mathbf{T_t = A}$, the level of the underlying process and E is a random error term (assumed to be independently and identically distributed with mean zero and a constant variance $\sigma_E{}^2$ {i.e. $\mathbf{E_t \sim II(0, \sigma_E{}^2)}$}. This model is exponentially estimated and forecasted , at time t, by $\mathbf{f_{t+1}}$ through the equation: $\mathbf{f_{t+1} = (1- \alpha)\, f_t + \alpha\, y_t}$ **;** where $0< \alpha < 1$ is a smoothing coefficient**.**
For better understanding let us go through the following example.

Example #3 (Single Exponential Smoothing)
Let us analyze the data on sale of ice cream by using the smoothing value $\mathbf{\alpha=0.2}$ and prior $\mathbf{f_1 = 10.}$
On basis of these values $\mathbf{f_{t+1} = 0.8f_t + 0.2\, y_t}$ **.** which generates one step ahead forecasts:
At time t=1, $f_2 = 0.8\, f_1 + 0.2\, y_1 = 0.8(10) + 0.2(5) = \mathbf{9.0}$
At time t=2, $f_3 = 0.8\, f_2 + 0.2\, y_2 = 0.8(9) + 0.2(6) = \mathbf{8.4}$ and so on.
On basis of these one step ahead forecasts let us compute the **B**ias, MAE and MSE.
Table #1

Year	**Time t**	$\mathbf{Y_t}$	$\mathbf{f_{t+1}}$	$\mathbf{e_t = Y_t - f_{t+1}}$	$\mathbf{\lvert e_t \rvert}$	$\mathbf{e_t^2}$
2006	1	5	f_1= 10	-5.000	5.000	25.000
2007	2	6	f_2 =9	-3.000	3.000	9.000
2008	3	8	f_3 = 8.4	-0.400	0.400	0.160
2009	4	10	f_4=8.32	1.680	1.680	2.822
2010	5	16	f_5=8.656	7.435	7.435	54.279
$\sum$		**45**	**-**	$\mathbf{0.715 = \sum e_t}$	$\mathbf{17.515 = \Sigma \lvert e_t \rvert}$	$\mathbf{92.261 = \Sigma\, e_t^2}$
2011	6	-	$\mathbf{f_6=10.1248}$	-	-	-
Accuracy Measures	-	-	-	**Bias**$=\Sigma e_t/4$ **= 0.18**	**MAE=** $\Sigma \lvert e_t \rvert /4$= 4.38	**MSE=** $\Sigma e_t^2/4$=**23.07**
Standard Error	-	-	-	**-**	**-**	$\mathbf{S_e=\sqrt{MSE}= 4.8}$

Where, df =(n-number of model parameters). In our model there is one parameter A only, therefore, df**=** (5-1)=4 and forecast of sale of ice cream during the year 2011 is million .

Comments:
i)By Bias B= + 0.156 means that the MA(2) model is slightly underestimating the sale of ice cream. That is, the forecasts are slightly lower than the actual observed sale of ice cream.
ii) Comparing the Exponential Smoothing (ES) method with the Moving Average (MA) method it is noticed that ES performs better than the MA as both the MAE and MSE of ES are lower than MA.