UNIT 10 FORECASTING AND TIME SERIES ANALYSIS

Structure

10.1 INTRODUCTION

In the previous unit you have seen how regression analysis was useful in decision making. If you carefully analyse most of our decisions and actions, be it by the Government, an institution, an industrial organisation or by an individual, they will greatly depend on the situation expected in future. For example, suppose the Government is thinking of a housing policy by which it will provide houses to all families of the fisher folk in Mumbai over the next five years. Then the Government must be able to assess what the number of families of fisher folk in Mumbai in the next five years would be. Similar assessments will be required in the case of designing an unemployment policy, and so on. An educational institution must assess the future needs so that it can start building up the required infrastructure, employ new teachers, and so on. A car manufacturer or a cement factory owner should be in a position to predict the future demands for her/his products so that she/he can plan the production schedules in the most profitable way. A person buying a real estate property will be benefitted by the knowledge on how the land value has been appreciating in the locality where she is planning to buy a piece of land.

Notice that in all the examples above, our interest lies in 'predicting' what the prevailing situation will be in future regarding various aspects. This kind of prediction is known as 'forecasting'. This is largely based on the past behaviour of the particular aspect being studied.

In this unit you will study the methods of forecasting using time series. Data related to a particular feature, arranged in chronological order, is known as time series. You will learn some simple methods of time series data analysis and how to use the analysis in forecasting. In Sec.10.2 and Sec.10.3, we shall give formal definitions of forecasting and time series

and look at some examples of their applications. In Sec. 10.4, you will be introduced to basic components of a time series data and to the additive and multiplicative models. Our main focus in the analysis of time series data will be on understanding the long-term trend. Some methods of such trend analysis will be discussed in Sec. 10.5.
Before studying this unit, we expect you to have gone through Unit 9 thoroughly. Many concepts and formulae used there will be applied here too.

Objectives

After studying this unit, you will be able to

- define forecasting and justify its need;
- define time series data and its four basic components;
- explain the additive and multiplicative models;
- fit the linear (and some simple nonlinear) trends using the method of least squares;
- fit the trend by the method of moving averages;
- build forecasting models using simple exponential smoothing;
- briefly describe the Holt-Winter double parameter exponential smoothing model.

10.2 FORECASTING

All of us read or listen to the weather forecast often enough, wondering whether rain or a storm is going to hit us just when we don't want it to. The forecast, as you know, is a prediction of future conditions based on an analysis of data received over a period of time.

There are other kinds of forecasting too. For example, the Government may be interested in predicting the extent of damage due to natural calamities in the year ahead so that sufficient stock of foodgrains and other necessary commodities may be procured and reserved for such emergencies.

Forecasting can be used for predicting qualitative as well as quantitative aspects of events such as those described above. However, in this unit, we shall confine our discussion to the prediction of quantitative characteristics such as extent of damage quantified in rupees, the amount of food-grains required, etc. We call the aspect or condition we want to forecast, the **characteristic of interest**. These characteristics are generally influenced by a variety of factors such as economic conditions, technological advancements, population, inflation rates, weather conditions, seasonal effects, and so on. In most cases, it may not be possible to identify all these factors which affect the characteristic of interest. So we can't use them in predicting the future values of the characteristic. Instead, we apply a process of studying the behaviour of the characteristic over time for forecasting the expected values at certain points of time in the future. Typically, the forecasting methods have 2 parts — observing the pattern of the data on the characteristic over a time period; and then predicting the future expected values by assuming that the same pattern will continue in the future. So, the accuracy of such forecasts depends heavily on the quality of the identification of the pattern and the validity of the assumption that the same pattern will continue in future. It is reasonable to assume in many cases that at least in the short run, the underlying pattern may not change. So this is a reasonable assumption for short-term forecasting. However, in the long run these assumptions can bring in errors. Such forecasting errors can be measured by computing the differences between actual values and the predicted values, and corrections be applied in future.

Let us examine how forecasting will be useful and how to compute the forecasting errors, through an example.

Example 1 : A grocery shop owner in a small colony gets 50 bread loaves every morning from a company and sells them to the residents of that colony. If he sells a loaf, he makes a profit of Rs.2/-. If a loaf is not sold on the same day, the shop owner returns it to the company the next morning but he loses Re.1/- on each loaf returned. In order to maximise his profits, the shop owner decided to study the pattern of the demand for the bread. (The daily demand for bread can be thought of as a random variable.) He collected the data shown in Table 1.

Table 1 : No. of Bread Loaves Demanded

Day	1st Week	2nd Week	3rd Week
Sunday	47	51	49
Monday	54	49	54
Tuesday	44	40	46
Wednesday	46	40	41
Thursday	63	59	57
Friday	51	56	50
Saturday	55	54	58

Let us now analyse this data. If we compute the average number of loaves that can be sold per day (that is, the demand per day) based on the three weeks' data, it comes to 50.66 loaves (= average of the 21 numbers in Table 1). In fact, from his past experience the shop owner found that he can sell approximately 50 loaves on an average. This is why he takes 50 loaves every morning from the company for selling.

So, based on the data gathered for 3 weeks, the owner's forecast for each day's sale is 50. Let us now compute the forecasting errors (for the owner's forecast) during the first 2 weeks (see Table 2).

Table 2 : Forecasting Errors For The First Two Weeks

Day	First Week			Second Week		
	Actual Demand	Predicted Demand	Forecasting Error	Actual Demand	Predicted Demand	Forecasting Error
Sun	47	50	–3	51	50	1
Mon	54	50	4	49	50	–1
Tue	44	50	–6	40	50	–10
Wed	46	50	–4	40	50	–10
Thurs	63	50	13	59	50	9
Fri	51	50	1	56	50	6
Sat	55	50	5	54	50	4

To check whether you have grasped the discussion so far, please try the following exercise now.

E1) Compute the forecasting errors for the third week.

Observe from Table 2, on Sunday of the first week the shop owner had to return 3 loaves, so he must have lost Rs.3/-. On Monday of the first week he fell short by 4 loaves. Had he had these loaves, he could have made an extra Rs.8/-profit. So this should also be considered as a loss resulting from his decision to get only 50 loaves every day. Therefore, the owner's profit on Sunday of the first week is equal to $(47 \times 2) - (3 \times 1) = 91$ rupees; his profit on Monday of the first week is equal to $(50 \times 2) - (4 \times 2) = 92$ rupees. In this way, if you work out the overall profit for the three weeks, it comes to Rs.1832/- (how?).

Do you think the owner can make a better decision so as to increase his profit? This may be possible if there is a way of making better forecasts of the sales. So, how can the owner improve upon his forecasting method so that the forecasting errors are smaller? To answer this, let us take another look at the data of Table 1. If we plot each week's data on a graph where the x-axis represents the days and the y-axis represents the demand, we get Fig. 1.

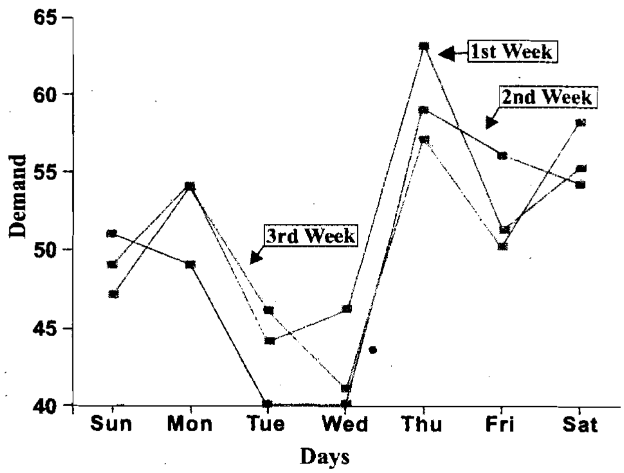


Fig. 1 : Demand of bread loaves vs. days.

From Fig. 1 we find the following trends :

i) the demand varies from day to day;
ii) on Tuesdays and Wednesdays the demand is lower than on other days of the week;
iii) the demand seems to be significantly higher on Thursdays than on other days.

Remark : Note that the daily demand is a random variable which is dependent on time, and its expected values in future depend on the observed values in the past.

Don't you think it will be wiser on the part of the owner, in the light of the observations above, to make separate forecasts for each day of the week and then make his decision accordingly? How would he do this? One way of doing so is to forecast each day's demand as the average of the three observations of that day. For example, the forecast for the demand on Sundays will be the average of 47, 51 and 49, which is equal to 49. Similarly, the forecast for the demand on Mondays will be the average of 54, 49 and 54, which is equal to 52.3. Since the number of loaves cannot be equal to 52.3, the owner may decide to get 52 loaves (rounded off to the nearest integer). In this way, the shop owner could decide to procure 49 loaves on Sundays, 52 on Mondays, etc.

Now, you may like to evaluate the benefits if the owner decides to make forecasts in this way, by doing the following exercise.

E2) Compute the forecasts for Tuesday to Saturday as explained above. Write down the forecast errors for the three weeks and compare them with those of Table 2. What is the profit over 3 weeks according to the new decision?

Since you have solved E2, you know that the revised forecasting method would lead to a profit of Rs.2015/-. You also calculated the profit got by applying the previous forecasting method as Rs.1832/-. So, the method just discussed is superior to the former one.

* * *

The example above was used to give you a flavour of forecasting — how forecasting helps in decision making, what forecasting errors are and how to compute them. Generally, in practice you may not find it easy to use models to produce very accurate forecasts. This is because of two reasons :

(i) the data may not follow a pattern that can be described by any mathematical model, and
(ii) the pattern of the data may change all of a sudden.

However, there are ways of making reasonable forecasts. In the next two sections you will see how forecasting models are built, and various elements involved in building such models.

10.3 TIME SERIES

In the previous section you have seen examples of how data collected over a period of time can help in forecasting. You have also seen that forecasting involves studying the behaviour of a characteristic over time and examining the data for any patterns. Then the forecasts are made by assuming that the characteristic will continue to behave according to the same pattern in future. The data gathered could be sales per week, units of output per day, the cost of running a company per month, and so on.

The data on any characteristic collected with respect to time over succeeding time periods is called a time series. For example, in Table 1 we had a day-by-day time series for the demand of bread.

Some time series cover a period of several years. For example, the Andhra Pradesh government wanted to study the changes in the cropping pattern, to be able to predict the future economic needs of the agricultural sector. For this purpose, they gathered the pertinent data, some of which are given in Table 3 below.

Table 3 : A Time Series for Crop Yield

Sl.No.	Year	Crop Area (Hectares)		Yield/Hectare (Tons)		Rainfall (Cms.)
		Rice	Sugar Cane	Rice	Sugar Cane	
1	1955	27.23	0.71	1137	7420	1064
2	1956	29.27	0.77	1163	8178	1128
3	1957	28.31	0.77	1180	8434	847
4	1958	30.10	0.74	1250	9304	1063
5	1959	30.81	0.83	1244	8605	1030
6	1960	29.61	0.91	1238	8888	851
7	1961	33.92	0.96	1239	8139	1017
8	1962	34.75	0.91	1220	9809	1134
9	1963	33.57	1.24	1292	8701	891
10	1964	34.60	1.45	1447	7477	920
11	1965	31.40	1.36	1262	8602	680
12	1966	33.23	1.08	1328	7778	948
13	1967	33.99	1.23	1375	8200	817
14	1968	28.49	1.56	1231	8180	787
15	1969	34.69	1.58	1248	7074	990
16	1970	35.21	1.20	1359	7923	956
17	1971	30.41	1.19	1551	9914	692
18	1972	29.28	1.34	1454	8245	727
19	1973	33.78	1.78	1653	8284	894
20	1974	35.53	1.95	1604	8570	848
21	1975	38.95	1.71	1657	7577	1104
22	1976	35.65	1.79	1410	7727	1024
23	1977	36.63	1.95	1565	7940	873
24	1978	39.79	1.62	1907	7067	1150
25	1979	34.69	1.39	1859	8222	743
26	1980	36.00	1.72	1991	7859	884
27	1981	38.24	2.21	2102	9142	945
28	1982	36.38	2.05	2156	7922	819
29	1983	41.63	1.72	2161	7332	1198
30	1984	34.98	1.70	2021	7322	734
31	1985	34.52	1.65	2264	7483	865
32	1986	34.59	1.68	1951	5754	868
33	1987	32.07	1.77	2258	6902	954
34	1988	42.18	1.96	2572	7690	1144
35	1989	42.06	2.10	2403	7699	1343
36	1990	40.36	2.28	2442	7281	982
37	1991	39.36	2.36	2400	7496	981
38	1992	36.04	1.71	2495	7107	837
39	1993	35.47	1.75	2759	7676	817
40	1994	36.37	2.09	2609	7150	1018
41	1995	36.92	3.58	2498	7303	971

Source : Directorate of Economics and Statistics, A.P.

Note that the successive time period in Table 3 is one year. Depending upon the characteristic of interest of the forecaster, this successive time period can be a day, a week, a month, a quarter, a year, etc.

The fluctuations you see in each time series is the net result of the effect of several forces acting on the characteristic of interest. For example, the yield of rice (or sugar cane) depends upon irrigation facilities, quality and availability of fertilizers, weather conditions, transportation facilities, etc. Scientists have classified the effects of such forces on a time series into four broad categories. They are

i) long-term (or secular) trend,
ii) seasonal variations,
iii) cyclic variations,
iv) irregular (or random) variations.

These are the four basic components of any time series on which the forecasting models are built. Let us try to understand each of these components one by one.

10.3.1 Long-term Trend

Look at the data on yield of rice in Table 3. The data are presented graphically in Fig. 2.

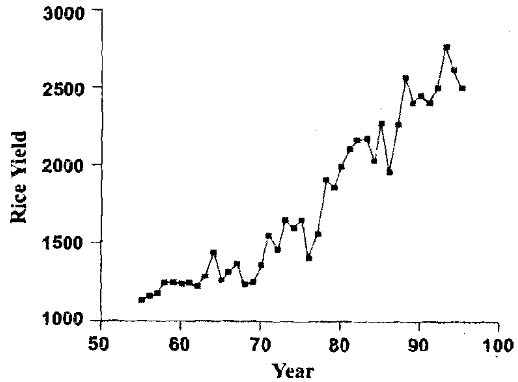
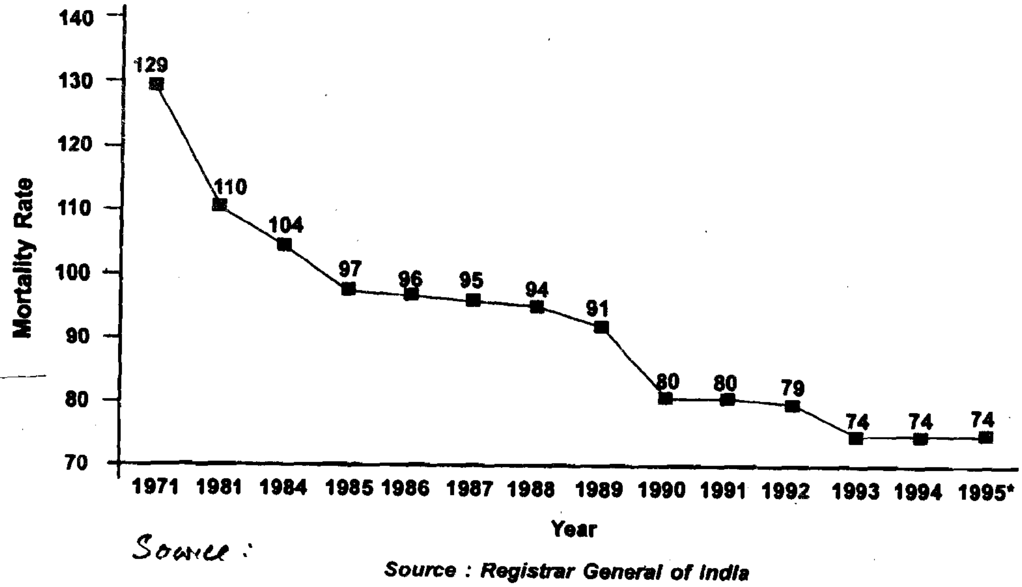


Fig. 2

From the graph, you can clearly see a general upward trend in the yield over a period of 40 years, though there are downward movements in between. The upward trend can possibly be attributed to better methods and facilities, use of new breeds of rice, etc. In fact, many business and economic statistics show upward trends over long periods.

Of course, there are series which show downward trends as well. For instance, the data on sugar cane yield presented in Table 3 exhibits a downward trend. Another example is the mortality rate of children below the age of 10 years in India (see Fig. 3). This also shows a very sharp downward trend over a long period of time.



Source : Registrar General of India

Fig. 3 : Infant mortality rate per 1,000 live births In India

There are characteristics which do not show any trend over a 15-year period or a longer period. For instance, in Fig. 4 we present the rainfall data of Table 3 graphically. Can you see any general upward or downward trend in it? In a time series where there is no trend, the long-term trend component will be absent in the models used for forecasting such time series.

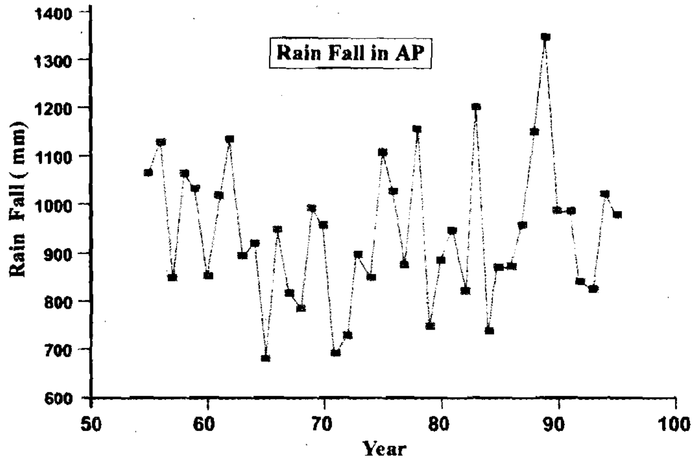


Fig.4 : Example of a time series with no trend

So, we have just seen examples of the long-term trend, which is an upward (or downward) movement in a time series over a **long period of time**, usually 15 years or a longer period. Here is an exercise about this now.

E3) Look at your newspaper or at events around you, and give one example each of a time series with a downward trend, an upward trend and with no trend.

In this sub-section we introduced you to a long-term effect on a characteristic of interest. Let us now look at the second component that we had listed as an effect on the characteristic.

10.3.2 Seasonal Variations

Suppose a readymade garment's manufacturer wants to forecast the sales of cotton shirts. He studies the data he has for the period 1995-2000. This data is of quarterly (i.e., 3-monthly) sales, which are given in Table 4.

Table 4 : Quarterly Sales of Cotton Shirts

Year	Shirt Sales			
	1st quarter	2nd quarter	3rd quarter	4th quarter
1995	100	243	250	120
1996	95	250	239	113
1997	111	227	241	110
1998	107	232	230	100
1999	110	230	237	97
2000	92	245	229	98

Let us draw the graph of these sales by the quarter (see Fig. 5) for 3 years.

From Fig. 5 you can see that for each year the sales are low in the first and fourth quarters, but high, and more or less the same, in the second and third quarters. So, within a year, the pattern is different in different quarters.

However, the same pattern repeats every year. We have shown the sales for only 3 years in Fig. 5. The sales for the other years follow the same pattern. This kind of repetition of a pattern within a time period (of a year in this case), and repeated every year is an example of a seasonal variation.

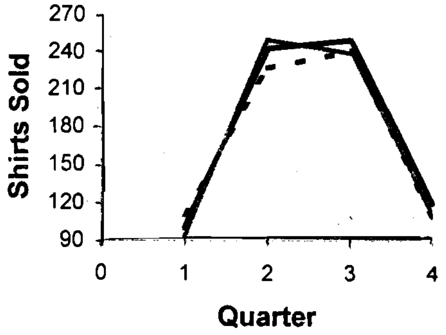


Fig. 5 : Quarterly sales of shirts

More generally, a **seasonal variation** of a characteristic is a pattern found in the data **within** a time period (a year in the example above) whose shape repeats in every successive period (in the example above, year after year). For instance, if you study the data of annual rainfall in India, but given month-wise, you will find high rainfall during the months of July and August, and almost no rainfall during April and May.

Seasonal variations need not only be due to changes in the natural weather conditions. They could be due to human-made conditions as well. For example, the number of STD phone calls made during a day will vary with the time slots specified by the telephone department (at present MTNL has two time slots with varying charges). Similarly, the number of passengers travelling in city buses is usually low on Saturdays and Sundays as compared to other (working) days of a week.

Now, try the following exercises.

E4) Give one example each from your own experience of a time series with seasonal variations
   i) due to natural weather conditions;
   ii) due to human-made conditions.

E5) Give an example of a time series which has no seasonal variation component.

As you may have realised, the seasonal variation plays an important role in planning or forecasting **only over a short period** like a year or less. However, there is a somewhat similar component of a time series that is, in a sense, linked with the long-term trend. Let us discuss this component now.

10.3.3 Cyclic Variations

When you were studying the data in Table 3, you may have noticed that the data on crop area of sugar cane presented there seems to increase and decrease repeatedly over the 41 years. In Fig.6, we present the same data graphically.

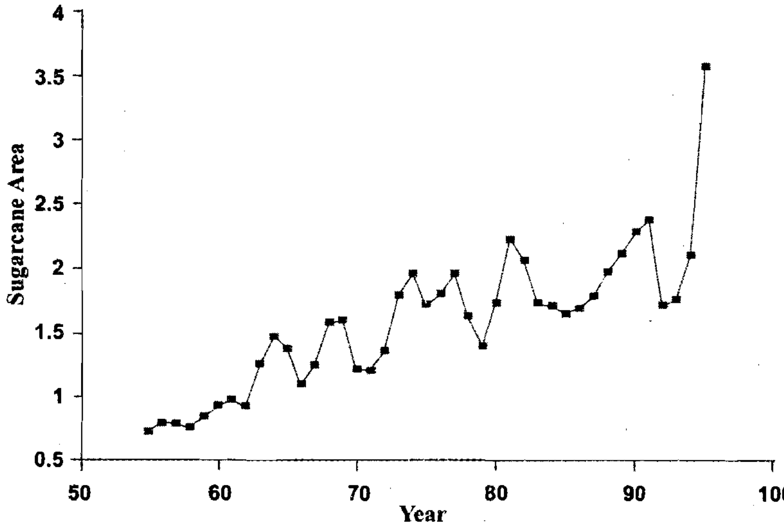


Fig. 6 : The crop area of sugar cane in Andhra Pradesh from 1955-1995

From the graph above, you can clearly see that the crop area is increasing upto a certain point and then dropping, then again increasing and dropping. This kind of movement in a time series is known as cyclic variation.

The time series from one peak to the next (or from one lowest turning point to the next lowest turning point) is known as a **cycle**. Unlike the seasonal variation, the length (or duration) of a cycle in a **cyclic variation** is not periodic, as you can see from Fig. 6. One cycle is from 1962-66, another is from 1966-71, etc., that is, they are of varying lengths.

Some time series may not exhibit any cyclic variations. Can you think of one such series? What about the time series of child mortality in India which is graphed in Fig. 3? There is no cyclic component visible here.

Cyclic variations are common among commercial and economic time series, in which the length of a cycle could vary from 2 to 10 years. From Fig. 5 and Fig. 6, you may say that seasonal variations also show peaks and troughs. But the duration of a seasonal variation is short, usually a year or less. The generally accepted convention is that **a cyclic variation is called a cycle only when its duration is more than a year and a seasonal variation otherwise.**

Why don't you try an exercise now?

E6) Consider the time series you presented while doing E3. Examine them for cyclic variations.

Now we shall briefly consider the fourth component of a time series.

10.3.4 Irregular Variations

Suppose I am studying the trends in the male vs. female birth rates in North India. While going through the annual figures per thousand from 1973 to 2000, I find that it is following a definite pattern — the number of female births is going down steadily. This pattern is disturbed suddenly at one point (in 1997) when the number of female births suddenly rises. Then the pattern continues again.

Such unexplained variation in a time series is called **random variation**, or **irregular variation**. It is the result of one or more chance causes which are purely random and unpredictable. So, this factor is represented by a random variable. The values it takes are, in fact, the estimates of the forecast errors. Therefore, it is expected that this is i.i.d normal with mean 0 (see Unit 2). Therefore, it is usually not taken into consideration while doing long-term planning and forecasting.

Here is an exercise about such variations.

E7) Give an example of a time series and an irregular variation in it.

So, you have seen the four basic components of any time series. Any given time series may or may not involve all four components. Here is an exercise based on this.

E8) Which of the four basic components have an effect on the characteristic under consideration in the following time series data? Explain your choice of components.

i) Number of cars produced monthwise by a leading car manufacturing company from April '91 toMarch '98 (there will be 84 observations in this time series).

ii) Number of fiction books issued by a lending library daywise from Jan. 1, 1999 to December 1, 1999.

iii) The yarn production data shown in Table 5.

Table 5 : Annual Yarn Production of Andhra Pradesh (in 1000 tons)

Year	Production
1971	29.9
1972	26.7
1973	25.5
1974	32.5
1975	30.1
1976	30.0
1977	28.3
1978	30.5
1979	31.6
1980	32.2

As you have seen, a time series helps us to see a pattern in the long-term behaviour of the particular characteristic we are studying. You have also seen that a time series, in general, has four basic components. The question is : how are these components put together so that a forecast analyst can be helped to understand the phenomena affecting the path followed by a time series? Let us try and answer this now.

10.4 FORECASTING MODELS

In this section, the key issue we shall discuss is that of appropriate models which explain the available time series data reasonably well. We shall restrict our discussion to two commonly used forecasting models — the additive model and the multiplicative model. While doing so, we shall use the notation y_t for the value of the time series at the time t. Since all time series will be in chronological order (i.e., in order of successive time periods), we can use serial numbering for time t. For example, for the data in E8 (iii) (Table 5), we can use t = 71, 72, …, 80, in which case the time series will be $y_{71} = 29.9$, $y_{72} = 26.7$, and so on. We can also serialise this data using t = 1, 2, …, 10, 1 standing for 71, 2 for 72, and so on. Then the time series will be $y_1 = 29.9$, $y_2 = 26.7$, and so on.

So let us look at the models, one-by-one.

10.4.1 The Additive Model

One of the most widely used models is the additive forecasting model. In this model it is assumed that at any time t, the time series is the sum of all the components. Symbolically, the model is

$$y_t = T_t + C_t + S_t + I_t$$

where T_t, C_t, S_t, I_t are the long-term trend, cyclic, seasonal and irregular variations, respectively. Furthermore, it is assumed that the effect of the cyclic component (C_t) remains the same for all cycles and that the effect of any seasonal variation (S_t) remains the same during any year (or corresponding period). Similarly, it is assumed that the irregular component (I_t) has the same effect throughout. (In other words, it is assumed that I_t is i.i.d normal (see Unit 2) with mean 0.)

As you have seen in Sec.10.3, it is not necessary that every time series must include all the four components. For instance, the model for the annual rice yield data does not have a seasonal component and the model for the annual rainfall data does not have a cyclical component.

Let us consider an example of the use of the additive model.

Example 3 : Let us revisit the situation in Example 1. To fit the model we use fresh data collected during five weeks in November and December, 1998, regarding these sales. These are presented in Table 6 below.

Table 6 : Bread Loaf Sales Data

Day	Week beginning :					Average
	Nov. 1	Nov.8.	Nov.15	Nov.22	Nov.29	
Sun	45	52	55	56	64	54.4
Mon	46	53	59	56	60	54.8
Tue	48	45	46	51	57	49.4
Wed	47	51	53	58	55	52.8
Thu	58	61	60	66	73	63.6
Fri	58	62	61	61	70	62.4
Sat	51	56	65	64	65	60.2
Average	50.4	54.3	57.0	58.9	63.4	56.8

Since the data are there only for 5 weeks, we shall assume that the cyclic component is absent in this time series. Therefore, our model will be

$$y_t = T_t + S_t + I_t, \; t = 1, 2, \ldots, 35.$$

Using a common sense (**ad hoc**) approach, let us identify and measure the trend and seasonal components. Here the seasonal variation is reflected by variations within each week, while the long-term trend is reflected by the movement of weekly average sales. For convenience, we have already computed the weekly and daily averages in Table 6. Notice that the weekly averages (last row) are showing an upward trend — from the first week to the second there is an increase of 4 units in the average, from the 2nd to the 3rd weeks 3 units, 3rd to 4th weeks 2 units, and from the 4th to 5th weeks 4 units. Therefore, we can conclude that every week there is an increase of 3.2 ($= \frac{4+3+2+4}{4}$) units on an average. If we extend this trend to future weeks, the sixth week average sales should be 66 loaves (again rounded off for simplicity), the seventh week's should be 69, and so on. So, looking at the average, we can now write down the trend component explicitly as follows :

All the averages are rounded off to the nearest integer for simplicity.

$$
\begin{aligned}
T_t &= 50 \text{ for } t = 1, 2, \ldots, 7 \text{ (1st week)} \\
&= 54 \text{ for } t = 8, \ldots, 14 \text{ (2nd week)} \\
&\vdots \quad \vdots \\
&= 66 \text{ for } t = 36, \ldots, 42 \text{ (6th week)}
\end{aligned}
$$

Let us now estimate the seasonal component. Since the seasonal variation is reflected by the variation within each week, we can estimate this by subtracting any week's average from that week's observations. For example, the seasonal variations from the first week's data are obtained by subtracting 50 (the rounded off first week's average) from that week's observations. The differences thus obtained are presented in Table 7.

Table 7 : Estimates of Seasonal Components For Data in Table 6

Day	Week					Average
	1st	2nd	3rd	4th	5th	
Sunday	–5	–2	–2	–3	1	–2
Monday	–4	–1	2	–3	–3	–2
Tuesday	–2	–9	–11	–8	–6	–7
Wednesday	–3	–3	–4	–1	–8	–4
Thursday	8	7	3	7	10	7
Friday	8	8	4	2	7	6
Saturday	1	2	8	5	2	4

Note : Averages are rounded off to the nearest integer.

Recall that we have mentioned that the effect of seasonal variations remains the same over all successive periods (the period is a week in this case). In other words, we would expect the differences in each week to be identical. However, this does not happen in practice. Therefore, the seasonal component will be estimated by averaging out the differences over the weeks. In the last column in Table 7 we have noted these averages, which give us the seasonal components. So, for example,

$S_t = -2$ for $t = 1, 8, 15, 22, 29, \ldots$ (corresponds to Sundays)
$S_t = -7$ for $t = 3, 10, 17, \ldots$ (corresponds to Tuesdays)

It is now your turn to write down the values of S_t for the other days of the week.

E9) Write down the values of S_t for $t = 2, 4, 6, 16$ and 26 in the example above.

What remains to be estimated is the irregular component, I_t. This component can be estimated by subtraction (i.e., using $I_t = y_t - T_t - S_t$) for the available data. However, estimating specific values of I_t usually does not interest the forecaster. What will interest her is analysing the distribution of the values of I_t because this will help in providing confidence intervals to the forecasts. In fact, the values of I_t are nothing but the estimates of forecast errors (recall that you have come across forecast errors in Table 2, E1 and E2). Such errors are expected to have a zero mean, that is, they cancel each other out in the long run. So, **future forecasts are made by adding only the non-random components**. For example, the forecast of y_{36} is given by

We use $\hat{y}$ to denote the predicted value of y.

$$\hat{y}_{36} = T_{36} + S_{36} = 66 + (-2) = 64.$$

* * *

Why don't you try a related exercise now?

E10) Using the additive model for finding the trend in the bread sales, forecast the 6th, 7th and 8th weeks sales. Also compare them with the actual sales given in Table 8 below.

Table 8 : Actual Bread Sales

Week \ Day	Sun	Mon	Tue	Wed	Thurs	Fri	Sat
6	59	65	58	62	79	66	69
7	64	66	60	57	81	70	69
8	69	75	67	69	78	74	73

Let us now consider another model used for forecasting.

10.4.2 The Multiplicative Model

We briefly outline this method, which is really a multiplicative version of the earlier one. In the additive model, we have assumed that the time series is **the sum** of the trend, cyclical, seasonal and random components. From practical experience, scientists have found that additive models are appropriate when the seasonal variations remain unchanged (that is, the seasonal variations do not depend on the trend of the time series). However, in practice, there are a number of situations where the seasonal variations change over time, as you will see in Example 4 below. When the seasonal variations exhibit an increasing or decreasing trend, we can try the **multiplicative model**. In the multiplicative model it is assumed that the time series is obtained as **a product** of the four time series components, that is,

$$y_t = T_t \cdot C_t \cdot S_t \cdot I_t .$$

Multiplicative models are found to be appropriate for many economic time series data such as data related to production of electricity, number of passengers going abroad, consumption of cold drinks, etc. In the following example, we will briefly describe the application of this model.

Example 4 : Examine the data on coconut sales in Hyderabad from 1995 given in Table 9. In each year, the number of coconuts sold are recorded for three seasons: (i) Season I – March to June, (ii) Season II – July to October, and (iii) Season III – November to February.

Table 9 : Coconut Sales Data

Period (t)	Year	Season	Coconuts sold (y_t) (in lakhs)	Period (t)	Year	Season	Coconuts sold (y_t) (in lakhs)
1	1994-95	I	14	12	1997-98	III	90
2	1994-95	II	35	13	1998-99	I	22
3	1994-95	III	65	14	1998-99	II	50
4	1995-96	I	14	15	1998-99	III	105
5	1995-96	II	43	16	1999-00	I	23
6	1995-96	III	77	17	1999-00	II	30
7	1996-97	I	16	18	1999-00	III	120
8	1996-97	II	40	19	2000-01	I	27
9	1996-97	III	84	20	2000-01	II	68
10	1997-98	I	17	21	2000-01	III	132
11	1997-98	II	46				

First let us observe the time series plot. This is given in Fig. 7.

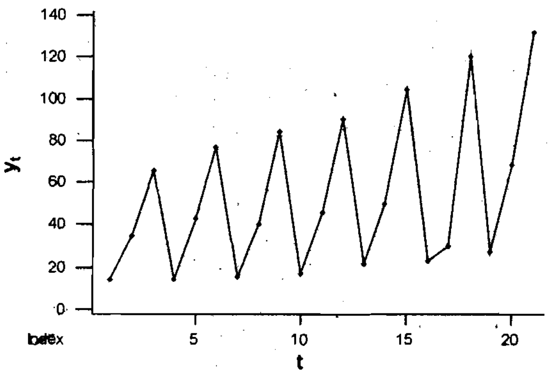


Fig. 7

From the plot you can easily see two things :

(i) the sales are gradually increasing (this indicates the increasing trend), and

(ii) the seasonal variation clearly exists, and, more importantly, it is increasing with the increasing trend.

Here is an exercise for you about this.

E11) Find out the **seasonal range** in each year. (e.g., the seasonal range for the year 1994-95 is 51 (= 65 – 14). Do you find any trend in the seasonal range over the years?

While doing E11, you would have realised that the seasonal variation is increasing with an increasing trend. So, as we said at the beginning of this sub-section, we should try the multiplicative model in this case. The cyclic variation that you see in Fig. 7 is actually seasonal variation. So, we drop

the cyclic component C_t, and include the seasonal component S_t in our model. Consequently, our model will be :
$y_t = T_tS_tI_t$.

We will see how we can estimate the time series components T_t, S_t and I_t, using formal methods, in the next section.

* * *

Note that in order to apply the multiplicative model the time series should have positive values. So, if we wish to use the multiplicative model to understand a time series with negative values, then we need to convert the time series to positive values by adding a suitable constant to each entry.

In this section you have seen examples of building forecasting models in some cases. Our approach was an ad hoc one, based on common sense. However, this is not the way analysts do it. In the next section you will learn some scientific methods of analysing time series data.

10.5 FORECASTING LONG-TERM TREND

Here we use the method of least squares, the moving average method and the method of exponential smoothing for finding the component T_t, mentioned in the models above. As you will see, identifying the trend in a time series requires elimination of other components from the time series.

Let us start with the simplest of these methods, which you studied in Unit-9.

10.5.1 The Least Squares Method

Let us once again look at the bread sales data presented in Table 6. The same data are reorganised in Table 10 day-wise.

Table 10 : Bread Sales Data

Day (t)	y_t	Day (t)	y_t
1	45	19	60
2	46	20	61
3	48	21	65
4	47	22	56
5	58	23	56
6	58	24	51
7	51	25	58
8	52	26	66
9	53	27	61
10	45	28	64
11	51	29	64
12	61	30	60
13	62	31	57
14	56	32	55
15	55	33	73
16	59	34	70
17	46	35	65
18	53		

Now, what do you expect its trend component to be like? Does the trend component of the time series increase (or decrease) at a constant rate? If it does, then the time series is said to have a **linear trend**, that is, it is a linear function of time. What this means algebraically is that if y_t has a linear trend, then we would expect $T_t = a + bt$, where a and b are constants. Recall (from Unit 9) how you fit linear equations. We can use the method of least squares. Here T_t is our dependent variable and t is our independent variable. Setting $y = T_t$ and $x = t$, we get

$$\bar{x} = 18,\ \bar{y} = 56.8,\ S_{xx} = 3570,\ S_{xy} = 1732,\ S_{yy} = 1715.6.$$

Therefore, the parameters a and b of the best fit linear equation are estimated as

$$\hat{b} = S_{xy}/S_{xx} = 0.485, \text{ and}$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = 48.06.$$

So, the regression equation in this case is given by

$$T_t = 48.06 + 0.485t. \qquad (1)$$

Notice that, x being a serial number, S_{xx} is easily computed.

This equation can now be used to obtain the trend component T_t.

$$R = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

If you work out the square of the correlations coefficient (see Sec. 9.3.2), you will find $R^2 = 0.49$. From Unit 9, you know that this means that this regression is reasonably reliable, but could be better.

Notice that in our earlier approach in Example 3, we had the same value for the trend component during any day of the week. In this regression approach, we get different values of the trend even within a week. In the following exercises we ask you to compare some of these different trend values.

E12) Compute T_t for our bread sales example for Wednesdays of each of the five weeks mentioned in Table 6, using Equation (1) above. Compare them with what we have got in Example 3.

E13) Fit the linear trend using part of the sugar cane crop area data presented in Table 11 below (extracted from Table 3).

Table 11 : Sugar Cane Crop Area in AP (in lakh hectares)

Year	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970
Area	0.96	0.91	1.24	1.45	1.36	1.08	1.23	1.56	1.58	1.20

Not all data exhibit a linear trend. Sometimes, by looking at the data points, it becomes reasonably clear that the time series data exhibit non-linear trends. For example, if you examine the rice yield data presented in Fig.2, there is a clear indication of a non-linear trend.

Let us look at another set of data, the population of Andhra Pradesh presented in Table 12 below.

Table 12 : Andhra Pradesh Population

Census	Population (in 10000s)	Census	Population (in 10000s)
1901	1906	1951	3111
1911	2144	1961	3598
1921	2142	1971	4350
1931	2420	1981	5355
1941	2729	1991	6651

Courtesy : Directorate of Economics and Statistics, AP

Look at the graph of the time series above given in Fig. 8 below. The points are certainly not lying around any line. Some non-linear curve may fit the data very well.

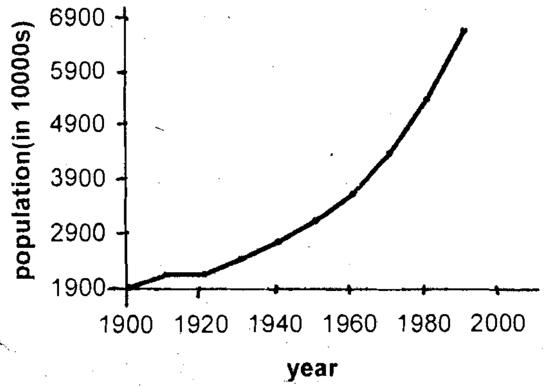


Fig. 8 : Graph of data in Table 12

A number of standard forms of curves have been found to be useful for fitting the data, in practice. These are polynomial curves, exponential curves and growth curves. The graphical plots of the time series is useful in identifying the form of the trend curve. Let us illustrate this with an example.

Notice that we are using year number for t rather than the serial number.

Example 5 : Let us try to fit a second degree equation to the trend of rice yield data of Table 3 using the method of least squares. We first plot the time series (see 'actual' curve in Fig. 9 below). Looking at it, it seems that a line will not be the best fit. Let us try a quadratic polynomial. The equation would be

$T_t = a_0 + a_1 t + a_2 t^2$, $t = 55, 56, \ldots, 95$.

The constants a_0, a_1 and a_2 are estimated by fitting a multiple linear regression model with two independent variables, $x_1 = t$ and $x_2 = t^2$, and the dependent variable $y = T_t$. Refer to Sec. 9.4.2 to recall the method of computing $\hat{a}_0$, $\hat{a}_1$ and $\hat{a}_2$, the estimates of a_0, a_1 and a_2, respectively.

The normal equations are derived from the least squares principle.

The first step in this is to tabulate the values of (x_1, x_2, y). There will be 41 observations of (x_1, x_2, y). From Table 3 we get the first one as $(55, 55^2, 1137)$, the second as $(56, 56^2, 1163)$, and so on. Next, we have to compute S_{11}, S_{22}, S_{12}, S_{1y} and S_{2y}. Then, $\hat{a}_1$ and $\hat{a}_2$ are obtained by solving the normal equations (10) and (11) of Sec. 9.4.2. Finally $\hat{a}_0$ is obtained from Equation (12) of Sec.9.4.2. We shall omit the details of this calculation and present the final results directly. The estimates are given by $\hat{a}_0 = 2660$, $\hat{a}_1 = -67.1$ and $\hat{a}_2 = 0.716$. Therefore, the trend curve is given by

$$T_t = 2660 - 67.1t + 0.716t^2 \qquad (2)$$

Here $R^2 = 0.94$ and the standard error $s = 126.8$.

The trend curve fitted in (2) and the actual time series values are plotted in Fig. 9. This may give you an idea about how good the fit is.

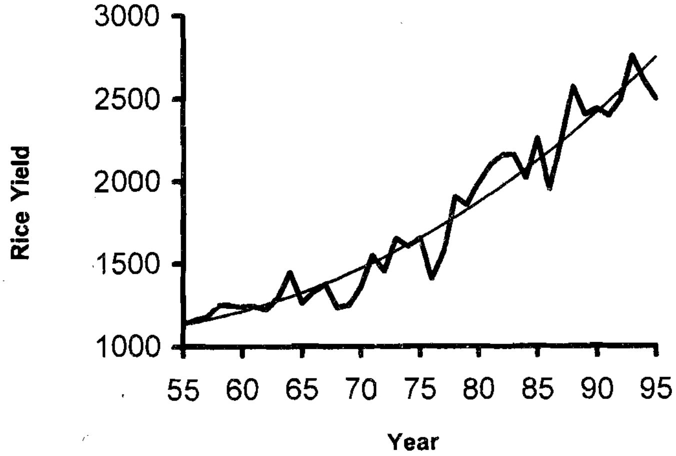


Fig. 9 : A quadratic fit to the trend curve

From Fig. 9 we see that the quadratic polynomial is a good fit for the trend of the rice yield data.

* * *

Now you should try to see how a second degree equation fits the population data of Table 12.

E14) Fit the trend curve to the population data of Table 12 using a second degree equation, i.e., $T_t = a_0 + a_1t + a_2t^2$. Use $t = 1, 2, \ldots, 10$. How good is your fit?

Let us now consider a situation where a non-linear trend is best fitted by an exponential curve.

Example 6 : Our country's population has been growing at a tremendous rate. In such a situation, when the growth rate is increasing fast, an exponential trend curve is probably the best fit to the data.

Examples of equations representing exponential curves are
$y = 2(3^t)$ and $y = (0.7)(2^t)$.
In general, it is of the form
$y_t = ab^t$, where a and b are positive constants.

Supposing we wish to fit this curve to the population data of Table 12. How do we go about it? Our equation for the trend curve would be

$$T_t = ab^t, \qquad (3)$$

where a and b are positive constants.

Notice that if we take the common logarithm (to the base 10) on both sides of Equation (3), we get

$$\log T_t = \log a + t.\log b \qquad (4)$$

If we put $z_t = \log y_t$, $Z_t = \log (T_t)$, $a_0 = \log a$ and $a_1 = \log b$, we can rewrite (4) as

$$Z_t = a_0 + a_1 t.$$

Now you know how to find out the values of a_0 and a_1 by the least squares method. Once you obtain the estimates of a_0 and a_1, you can calculate the values of a and b using antilogarithms. The calculations are done for you in Table 13 below.

Table 13 : AP Population Values For Curve Fitting

Census	t	Population (in 10,000s) (y_t)	$z_t = \log y_t$	Forecast for z_t (Z_t)	Forecast for y_t (T_t)
1901	1	1906	3.2801	3.2333	1711
1911	2	2144	3.3312	3.2925	1961
1921	3	2142	3.3308	3.3517	2248
1931	4	2420	3.3838	3.4109	2576
1941	5	2729	3.4360	3.4701	2952
1951	6	3111	3.4929	3.5293	3383
1961	7	3598	3.5561	3.5885	3877
1971	8	4350	3.6385	3.6477	4443
1981	9	5355	3.7288	3.7069	5092
1991	10	6651	3.8229	3.7661	5836

Z_t is the forecast for $\log(y_t)$ and is given by $Z_t = 3.1741 + 0.0592 \times t$, $R^2 = 0.96$, $S_e = 0.039$.

The original fit is obtained by taking antilogarithms. This is given by $T_t = \text{antilog}\,(3.1741)\,(\text{antilog}\,0.0592)^t = 1493 \times (1.1462)^t$.

* * *

Why don't you try an exercise now?

E15) (a) Compare the exponential fit with the quadratic fit you obtained in E14.

(b) The actual population of AP in 2001 is 7,57,27,541. Which of the two curves above gives a better forecast? What will the population be in 2011 according to both the curves fitted to the data?

E16) Estimate T_t using regression for the data in Example 4.

So far we have looked at two methods of forecasting trends. There is another method based on finding the averages of the data. Let us look at its strong and weak points.

10.5.2 The Method of Moving Averages

This method aims at identifying the long-term trend by eliminating seasonal variations. While doing this, the method also indicates the presence of seasonal and cyclic variations, if any. To appreciate this, let us apply this method to our bread sales data presented in Table 6.

Example 1 (Contd.) : To apply the method we need to present the data in a single column (in chronological order). This is done in Column 2 of Table 14 below.

Table 14 : Moving Averages for the Bread Sales Data

Day (t)	Sales(y_t)	Moving averages with length :				
		4	5	7	8	10
1	45					
2	46					
		46.50				
3	48		48.80			
		49.75				
4	47		51.40	50.43		
		52.75			50.63	
5	58		52.40	51.43		
		53.50			51.63	
6	58		53.20	52.43		50.30
		54.75			51.50	
7	51		54.40	52.00		50.90
		53.50			51.88	
8	52		51.80	52.57		52.40
		50.25			53.63	
9	53		50.40	53.00		53.80
		50.25			54.13	
10	45		52.40	53.57		54.70
		52.50			53.88	
11	51		54.40	54.29		54.40
		54.75			54.38	
12	61		55.00	54.71		54.50
		57.50			55.25	
13	62		57.00	55.57		54.00
		58.50			54.38	
14	56		58.60	55.71		54.10
		58.00			55.38	
15	55		55.60	56.00		54.80
		54.00			56.50	
16	59		53.80	55.86		56.40
		53.25			56.50	
17	46		54.60	55.71		57.80
		54.50			56.88	
18	53		55.80	57.00		57.30
		55.00			56.88	
19	60		57.00	57.14		56.70
		59.75			57.00	
20	61		59.00	56.71		56.20
		60.50			56.00	
21	65		59.60	57.43		56.50
		59.50			57.50	
22	56		57.80	58.14		57.20
		57.00			59.13	
23	56		57.20	59.00		58.70
		55.25			59.25	
24	51		57.40	59.00		59.80
		57.75			59.63	
25	58		58.40	58.86		60.20
		59.00			59.50	
26	66		60.00	60.00		60.10
		62.25			60.00	
27	61		62.60	60.57		59.30
		63.75			60.13	
28	64		63.00	61.43		59.20
		62.25			60.63	
29	64		61.20	61.00		60.90
		61.25			62.50	
30	60		60.00	62.00		62.80
		59.00			63.00	
31	57		61.80	63.29		63.50
		61.25			63.50	
32	55		63.00	63.43		
		63.75				
33	73		64.00	63.33		
		65.75				
34	70					
35	65					

For the time being ignore Columns (3), (4), (6) and (7) in the table. Let us see how we have obtained the entries in Column (5). We first compute the average of the first seven observations in Column (2), i.e., $\frac{45+\cdots+51}{7}$ = 50.43. We place this figure in Column (5) in line with Day 4, the mean of Days 1 to 7. Next we compute the average of the seven observations y_2, y_3, …, y_8, (that is, we drop y_1 and include y_8). This is 51.43. We place it in Column (5) in line with Day 5. In this way we compute the averages of seven consecutive observations, each time dropping the first observation and including the next observation. In this way, the last entry in Column (5), 63.33, is the average of the last seven observations of Column (2), and is placed in line with Day 32. These averages we just computed are called **the moving averages of length 7.**

The averages are called 'moving' averages because at each stage of calculating averages we move from one period to the next period.

Now consider Column (4). This consists of the moving averages of length 5. So, the first entry is $\frac{y_1+y_2+y_3+y_4+y_5}{5}$ = 48.80. This figure is written in line with Day 3. The next entry in Column (4) is $\frac{y_2 + y_3 + y_4 + y_5 + y_6}{5}$, and so on.

We have just seen how to find moving averages of odd length (7 and 5, respectively). Let us now see how to compute the moving averages of even length. Suppose we wish to compute the moving averages of length 4 for the bread sales data. We compute the average of the first 4 observations of Column (2) in Table 14. This average is equal to 46.5. What day does it correspond to? We imagine that it corresponds to 'Day 2.5' (i.e., $\frac{1+2+3+4}{4}$), and place it in Column (3) at the point between the rows corresponding to Day 2 and Day 3. Next, we compute the average of the 2^{nd}, 3^{rd}, 4^{th} and 5^{th} observations. This is 49.75. Correspond this with 'Day 3.5' in Column (3). Since 46.5 corresponds to Day 2.5 and 49.75 corresponds to Day 3.5, the average of 46.5 and 49.75 = (46.5 + 49.75)/2 = 48.125 should correspond to Day 3 (= average of 2.5 and 3.5).
In this way we continue, the last entry in Column (3) being $\frac{55+73+70+65}{4}$ = 65.75. This is placed in line with 'Day 33.5'.

You should calculate and check for yourself that all the entries shown in Columns (3) to (7) are correct.

To make sure that you have got the definition and computation of moving averages correctly, do the following exercises.

E17) Compute the first three moving averages of length 3 for the bread sales data and place them in line with the corresponding day numbers.

E18) For the bread sales data, compute the 2^{nd} and last moving averages of length 6. What days do these averages correspond to?

Now let us see the graphic representation of the moving averages. In Fig. 10, we have plotted the moving averages (of length 7) computed in Column (5), Table14, against the corresponding day numbers. The straight line trend obtained in Sec.10.5.1 (Equation (1)) is also drawn in the figure.

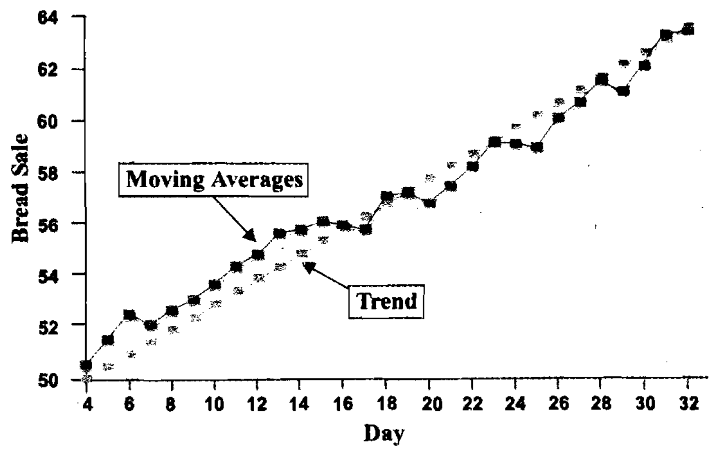


Fig. 10 : Plot of moving averages and linear trend for bread sales data

Now look at Fig.11, where the moving averages of lengths 4, 8 and 10 (shown in Table 14) are plotted along with the trend line.

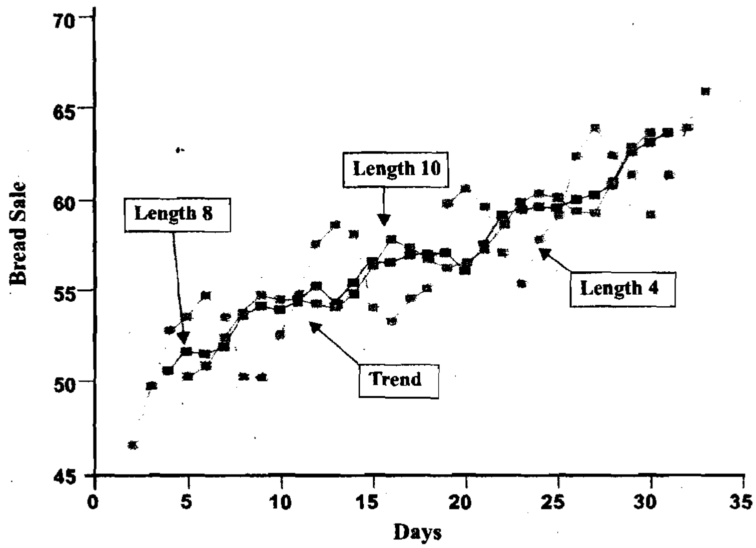


Fig. 11 : Moving averages of different lengths

You can see from Fig. 10 and Fig. 11 that fluctuations, i.e., peaks and troughs, are smallest when the length of the moving averages is 7. This is not surprising as the bread sales data have the seasonal component, the season being the **seven days** of the week. Observe from Fig. 10 that approximately half of the moving averages are above the trend line and the rest are below or on the trend line. This is to be expected whenever the linear trend is the best fit, since it represents the 'average' in a sense.

If you go back to Fig. 1, you will notice that the time series shown in it doesn't seem to show a linear, quadratic or exponential trend. In fact, it is a very jerky graph. But, related to the same data, Fig. 11 shows us that the moving averages of different lengths of the same data gives us smoother curves that show the long-term movements of the series. In this way the moving averages can be used to estimate the trends, if the time series does not help us in doing so.

* * *

From the example above, can you see what the basic idea is of using moving averages? If the time series of the data contains certain seasonal or cyclical variations, the effect of these variations can be eliminated by taking a moving average where the time period in the average equals the period of the season or the cycle. By smoothening the curve in this way, the trend doesn't get affected. For instance, in the example above the cycle is of length 7 and so the moving averages of length 7 will help us study the long-term trend, since the S_t-component of y_t gets removed.
Now, why don't you try the following exercise?

E19) Find the moving averages of lengths 3, 4 and 5 for the rice yield data given in Table 3. Using which of these averages can the effect of the cycles be eliminated? From this exercise, what conclusions can you draw regarding the C and S components of the time series?

E20) Using moving averages of length 3, estimate T_t for the data in Example 4.

Before we wind up this sub-section on the method of moving averages (MA), let us note the following points :

1) If a time series is a purely random sequence of numbers, you will find that a moving average of this time series will tend to show cyclical fluctuations. This is because a moving average is serially correlated. So, remember that many apparent cycles in moving averages may be spurious.

2) The peaks and troughs in the moving average may occur at different times than the peaks and troughs in the original time series (see Fig. 10 and Fig.11).

3) A moving average cannot be calculated for the latest or earliest years in a time series, since the average depends on numbers that precede or occur after these years in the time series.

4) The method of moving averages is useful when the trend in time series data is linear, or approximately so. This is because if you compute the moving averages for such a series, given by, say, $y_t = a + bt$, you will find that they will coincide with the time series values. In fact, this is the case irrespective of the length of the moving average. However, we can't claim this for general time series which is not linear. (You may like to check this for the quadratic time series $y_t = 2 + t^2$.)

5) It is effective when the fluctuations in time series are regular and periodic provided the length of the moving averages is that of the period.

6) The method is not very useful if it has to be used for forecasting future trend values. It is basically a tool to identify trend and cyclic components in a time series. For instance, examine the data in Table 14 on bread sales. We have 35 observations here. Let us say we want to predict the 36th observation using MA. To know the MA against 36, let us say of length 3, we need to have observations for the time points 36 and 37. As we don't have these observations when we are at time 35, we can't compute the required MA. For this reason MA is not useful for forecasting.

As you have seen, time series often have so many irregular fluctuations that we aren't able to see the general trend. The method of moving averages helps in removing such fluctuations to a large extent. Therefore, we call such a method a **smoothing technique**. We shall now consider another method for smoothening the long-term trend curves.

10.5.3 Exponential Smoothing

In this sub-section we shall introduce you to another smoothing technique in which **weighted averages** are calculated. In the method of moving averages we also attach weights, equal weights to each observation that is considered. In the method we shall now discuss, the weights assigned to past and current values of the time series are different fixed positive numbers. This method is called **exponential smoothing**. Let us see why through an example.

Example 7 : Consider the rainfall data presented in Table 3. From Fig.4, it appears that there is no trend in the time series. This is the same as saying that if $T_t = a + bt$ is the trend, then $b = 0$. Therefore, if we fit a linear regression to this time series data with $b = 0$, then the least square estimate of a is given by $\hat{a} = \frac{\sum_{t=1}^{41} y_t}{41}$, the simple average of the time series values.

Notice that the method of least squares gives equal weightage to all the observations in the time series.

Also, notice that since our model is $T_t = a +$ (forecast error), our forecast for rainfall is $\hat{a}$ for all t. This seems somewhat unreasonable because, even though there is no trend, the constant does not seem to be the same at all times t, at least going by Fig.4. The value of a seems to be higher during the period 1955 to 1965 compared to the value of a for the period 1965 to 1975. Therefore, it is logical to assume that the value of a is gradually changing over time. Therefore, we should denote it by a_t rather than a.

Accordingly we select a single weight w (called the **exponential smoothing constant**), where w lies between 0 and 1. Then we compute an exponentially smoothed series a_t as follows :

w is chosen between 0 and 1 because each a_i is a **convex combination** of y_i and a_{i-1}.

$$a_1 = y_1$$
$$a_2 = wy_2 + (1 - w)a_1$$
$$a_3 = wy_3 + (1 - w)a_2$$
$$\vdots$$
$$a_t = wy_t + (1 - w)a_{t-1} \qquad (6)$$

Two things needs to be specified here :

i) the quantity w is a constant known as **smoothing constant**, and there is a method of choosing this constant (which we shall discuss a little later); and

ii) we should know the value of a_0 (i.e., when $t = 1$, $a_{t-1} = a_0$) to initialise the computation of $a_1, a_2, \ldots$. We will take this initial value of a_0 as $\hat{a}$, the simple arithmetic mean of all the time series values.

Let us now compute some of these quantities for the rainfall data. Firstly,

$$a_0 = \hat{a} = \frac{\sum y_t}{41} = 939.95, \text{ rounded off to the 2}^{nd} \text{ decimal place.}$$

The value of the smoothing constant w can be anything between 0 and 1. However, from experience statisticians have found that one should choose an 'appropriate' value of w between 0.01 and 0.3. Let us take, to start with, $w = 0.02$. Then, from Equation (6) we get

$$\hat{a}_1 = 0.02y_1 + 0.98a_0 = (0.02 \times 1064) + (0.98 \times 939.95) = 942.43.$$

Our forecast for y_1 at time zero is $\hat{a}_0$. Therefore, the forecast error is

$$e_1 = y_1 - \hat{a}_0 = 1064 - 939.95 = 124.05.$$

The forecast error at $t = 2$ is

$$e_2 = y_2 - \hat{a}_1 = 1128 - 942.43 = 185.57.$$

Next,

$$\hat{a}_2 = 0.02y_2 + 0.98\hat{a}_1 = 0.02 \times 1128 + 0.98 \times 942.43 = 946.14,$$

and the forecast error

$$e_3 = y_3 - \hat{a}_2 = 847 - 946.14 = -99.14.$$

Now try the following exercise.

E21) Compute $\hat{a}_3$, $\hat{a}_4$, the corresponding forecasts and the forecasting errors.

In the table below we have given all the forecasts and errors.

Table 15 : Exponential Smoothing of Rainfall Data

Year	Rainfall y_t	Forecast a_t	Error $e_t = y_t - \hat{a}_t$
1955	1064	939.95	124.05
1956	1128	942.431	185.569
1957	847	946.1424	−99.14238
1958	1063	944.1595	118.8405
1959	1030	946.5363	83.46366
1960	851	948.2056	−97.20561
1961	1017	946.2615	70.7385
1962	1134	947.6763	186.3237
1963	891	951.4027	−60.40275
1964	920	950.1947	−30.19469
1965	680	949.5908	−269.5908
1966	948	944.199	3.801018
1967	817	944.275	−127.275
1968	787	941.7295	−154.7295
1969	990	938.6349	51.36509
1970	956	939.6622	16.33779
1971	692	939.989	−247.989
1972	727	935.0292	−208.0292
1973	894	930.8686	−36.86861
1974	848	930.1312	−82.13123
1975	1104	928.4886	175.5114
1976	1024	931.9988	92.00116
1977	873	933.8389	−60.83886
1978	1150	932.6221	217.3779
1979	743	936.9696	−193.9696
1980	884	933.0902	−49.09025
1981	945	932.1084	12.89156
1982	819	932.3663	−113.3663
1983	1198	930.0989	267.9011
1984	734	935.457	−201.457
1985	865	931.4278	−66.42783
1986	868	930.0993	−62.09927
1987	954	928.8573	25.14271
1988	1144	929.3601	214.6399
1989	1343	933.6529	409.3471
1990	982	941.8399	40.16012
1991	981	9942.6431	48.35692
1992	837	943.4102	−106.4102
1993	817	941.282	−124.282
1994	1018	938.7964	79.20362
1995	971	940.3805	30.61955

Now look at the graphs of the time series before and after smoothing, in Fig.12. (The 'before smoothing' graph is also given in Fig.4.) You can see that the peaks are barely there in the graph of the time series after smoothing, which is the dotted curve.

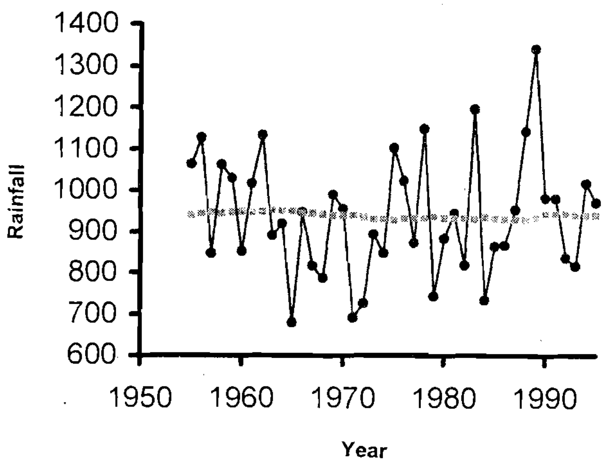


Fig.12

You can also try forecasts with other values of w, and see the curves you get.

We still haven't seen what the appropriate choice of w is. For this, we need to consider $S_E = \sqrt{\frac{e_1^2 + \cdots + e_n^2}{n}}$, that is, the square root of the average error sum of squares. Since this depends on the choice of w, we shall denote it by $S_E(w)$. If we compute this quantity for our rainfall data, we get

$$S_E(0.02) = \sqrt{\frac{(124.05)^2 + \cdots + (30.62)^2}{41}} = 146.36,$$ by using the calculations shown in Table 15.

Table 16: Values of $S_E(w)$ for the Rainfall Data

S.No.	w	$[E(w)]^2$
1	0.0005	8,62,707
2	0.001	8,62,710
3	0.005	8,62,795
4	0.01	8,63,036
5	0.05	8,68,546

We are now in a position to find out what value of w is most appropriate for our rainfall data — we should find the value of w for which $S_E(w)$ is minimum. How do we find this out? Note that if $S_E(w)$ is the minimum for $w = \alpha$, say, then $41(S_E(\alpha))^2$ will be the minimum of all values of $41(S_E(w))^2$ calculated for different values of w. So, to find the value of w that gives the least value of $S_E(w)$ it is enough to find the error sum of squares, say, $(E(w))^2$. In Table 16 we give $[E(w)]^2$ for different values of w. Note that $S_E(w)$ is small when $w = 0.0005$ (or $w = 0.001$). **The small value of w indicates that the average level of the time series is not changing much over time.**

Using the recursive equation (6) it can be shown that
$\hat{a}_t = wy_t + w(1-w)y_{t-1} + w(1-w)^2 y_{t-2} + \ldots + w(1-w)^{t-1}y_1 + (1-w)^t \hat{a}$.
So, for example, we get
$\hat{y}_{31} = \hat{a}_{30} = wy_{30} + w(1-w)y_{29} + \ldots + w(1-w)^{29} y_1 + (1-w)^{30} \hat{a}$.
Since $(1-w), (1-w)^2, \ldots$ are decreasing exponentially, the weightages given to more recent observations is more. So, **all** the observations are being given weightage here, and the latest observation is being given

greatest weightage. In this way, this method is a refinement of the method of moving averages.

Here is a related exercise.

E22) Apply the simple exponential smoothing procedure to the following data. Take $w = 0.1$.

Table 17 : No. of New Bank Branches Opened in a Town

Year	1981	1982	1983	1984	1985	1986	1987
No. of Branches	5	3	3	4	3	6	4

Let us now consider another **exponential smoothing forecasting method.** This is due to C.C. Holt, suggested by him in 1958.

Holt's method : We shall illustrate this procedure with the rice yield data. Suppose that the trend shown by the rice yield data is linear, namely, $y_t = a + bt + e_t$, where e_t is the error.

As before, we shall assume that a and b change gradually over time. Therefore, we denote the values of a and b at time t by a_t and b_t, respectively. As in simple exponential smoothing, a_t and b_t are smoothened using two smoothing constants w_1 and w_2 (both between 0 and 1). The recursive equations for computing these quantities and the forecasts are given by :

$$\hat{a}_t = w_1 y_t + (1 - w_1)[\hat{a}_{t-1} + \hat{b}_{t-1}], \tag{7}$$

$$\hat{b}_t = w_2[\hat{a}_t - \hat{a}_{t-1}] + (1 - w_2)\,\hat{b}_{t-1} \tag{8}$$

and the forecast for the immediate future y_{t+1} is given by

$$\hat{y}_{t+1} = \hat{a}_t + \hat{b}_t . \tag{9}$$

Here too, we need the initial values a_0 and b_0. These, together with w_1 and w_2, should be chosen so that the sum of the squares of the forecasting errors is minimised. One suggestion to obtain a_0, b_0, w_1 and w_2 is to first obtain a_0 and b_0 by fitting a linear regression to one half of the time series data (a_0 as the intercept and b_0 as coefficient). Then, using them as initial values, we should obtain the values of w_1 and w_2 which minimise S_E, the square root of the average of error sum of squares (see Table 18 below). Once w_1 and w_2 are obtained, then we can change the initial values a_0 and b_0 to the intercept and coefficient of the regression line fitted to the entire time series data. However, there is no guarantee that this would lead to the best choice of a_0, b_0, w_1 and w_2. In fact, this procedure of obtaining a_0, b_0, w_1 and w_2 in our example results in a very large value of S_E. A better choice is

$\hat{a}_0 = 1090$, $\hat{b}_0 = 32$, $w_1 = 0.4$, and $w_2 = 0.01$.

The resulting $S_E = 133.16$. The computations are shown in Table 18. The values in the columns have been rounded off to the nearest integer for simplification in calculations.

Table 18 : Holt's Exponential Smoothing of Rice Yield Data

Year	t	Rice Yield	$\hat{a}_t$	$\hat{b}_t$	Forecast	Error
1955	1	1137	1128	32	1122	15
1956	2	1163	1161	32	1160	3
1957	3	1180	1188	32	1193	–13
1958	4	1250	1232	32	1220	30
1959	5	1244	1256	32	1264	–20
1960	6	1238	1268	32	1288	–50
1961	7	1239	1276	32	1300	–61
1962	8	1220	1272	32	1307	–87
1963	9	1292	1299	31	1304	–12
1964	10	1447	1377	31	1330	117
1965	11	1262	1350	32	1409	–147
1966	12	1328	1360	31	1381	–53
1967	13	1375	1384	31	1391	–16
1968	14	1231	1342	31	1415	–184
1969	15	1248	1322	30	1372	–124
1970	16	1359	1355	30	1352	7
1971	17	1551	1451	30	1384	167
1972	18	1454	1470	30	1481	–27
1973	19	1653	1562	31	1501	152
1974	20	1604	1597	31	1592	12
1975	21	1657	1639	31	1628	29
1976	22	1410	1566	30	1670	–260
1977	23	1565	1584	30	1596	–31
1978	24	1907	1731	31	1613	294
1979	25	1859.	1801	31	1762	97
1980	26	1991	1896	32	1832	159
1981	27	2102	1997	33	1928	174
1982	28	2156	2080	33	2030	126
1983	29	2161	2133	33	2114	47
1984	30	2021	2108	33	2166	–145
1985	31	2264	2190	33	2141	123
1986	32	1951	2114	32	2223	–272
1987	33	2258	2191	33	2147	111
1988	34	2572	2363	34	2224	348
1989	35	2403	2399	34	2397	6
1990	36	2442	2437	34	2434	8

The forecasts and the actual time series values are plotted in Fig. 13.

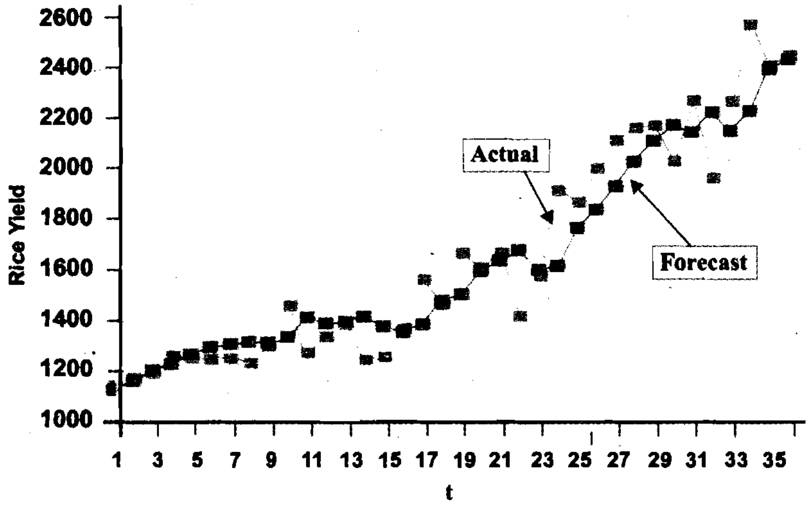


Fig. 13 : Forecast with Holt's model

Now, observe Equation (9), which gives forecasts of the immediate future. Assume that we have got the data only upto 1985, and that we wish to forecast the yields from 1986 to 1990. Note that t = 31 corresponds to year 1985. Equation (9) can now be generalised to make these forecasts as follows :

$$\hat{y}_{31+k} = \hat{a}_{31} + \hat{b}_{31}k, \; k = 1, 2, \ldots \qquad (10)$$

You may like to do the following exercise now.

E23) Assume that you have the rice yield data from 1955 to 1985 only. What will your forecasts be for 1986 to 1990 if you use Holt's exponential smoothing with $\hat{a}_0 = 1090$, $\hat{b}_0 = 32$, $w_1 = 0.4$, and $w_2 = 0.01$?

With this we come to the end of our discussion on finding the trends, and hence the predicted values. Actually, to forecast correctly, we need to estimate the other components too. We will do this in one case, just to give you a flavour.

Example 4 (contd.) : In Example 4, we observed that the model suitable for forecasting is $y_t = T_t \,.\, S_t \,.\, I_t$. While solving E20, we found that the method of moving averages gives us the estimate $T_t = \frac{1}{3}(y_{t-1} + y_t + y_{t+1})$.

Now, to estimate the seasonal variations, we first find the estimates of S_tI_t in the table below.

Table 19 : Coconut Sales Data

Period (t)	Year	Season	Coconuts Sold (y_t) (in lakhs)	Trend component (T_t)	Estimates of S_tI_t (= y_t/T_t)
1	1994-95	I	14	-	-
2	1994-95	II	35	38	0.921
3	1994-95	III	65	38	1.711
4	1995-96	I	14	41	0.341
5	1995-96	II	43	45	0.956
6	1995-96	III	77	45	1.711
7	1996-97	I	16	44	0.364
8	1996-97	II	40	47	0.851
9	1996-97	III	84	47	1.787
10	1997-98	I	17	49	0.347
11	1997-98	II	46	51	0.902
12	1997-98	III	90	53	1.698
13	1998-99	I	22	54	0.407
14	1998-99	II	50	59	0.847
15	1998-99	III	105	59	1.780
16	1999-00	I	23	53	0.434
17	1999-00	II	30	58	0.517
18	1999-00	III	120	59	2.034
19	2000-01	I	27	72	0.375
20	2000-01	II	68	76	0.895
21	2000-01	III	132	-	-

We are now in a position to estimate the seasonal effects, S_I, S_{II} and S_{III}. The estimate of S_I is the average of all those estimates of S_tI_t in which t corresponds to Season I. So, $S_I = \dfrac{S_4I_4 + S_7I_7 + \cdots + S_{19}I_{19}}{6} = 0.378$.

We can similarly estimate S_{II} and S_{III}.

Having estimated S_tI_t and S_t, we can now estimate I_t by using the equation $I_t = S_tI_t/S_t$.

* * *

You can wind up the example above by doing the following exercises.

E24) Estimate S_{II} and S_{III} in the example above. (Note that there are seven terms in S_{II} and six terms in S_{III}.)

E25) Estimate I_t for all values of t in Table 9.

E26) Using only the first 5 years of coconut sales data (see Table 9), build the multiplicative model and estimate the trend, seasonal and irregular components.
Use this model to forecast the sales for the next two years and compare them with actual sales.

We shall end our discussion on forecasting here. If you would like to go into greater depth, you could refer to the books listed under 'Further Reading' in the Course Introduction. You could also see the website: *www.bath.ac.uk/~masar/math0118/forecasting/node6.html.*

Now let us sum up the chief points made in this unit.

10.6 SUMMARY

In this unit we have discussed the following points.

1) Forecasting and its importance in future planning.
2) Time series and its four basic components — trend, seasonal variation, cyclic variation and random variation.
3) The additive and multiplicative models of time series data analysis.
4) The method of least squares for fitting linear and non-linear trends.
5) Time series data analysis by the method of moving averages.
6) The simple exponential smoothing procedure and Holt-Winters' exponential smoothing procedure.

10.7 SOLUTIONS AND ANSWERS

E1) The forecasting errors for Sunday to Saturday are –1, 4, –4, –9, 7, 0 and 8, respectively.

E2) The forecasts are 43, 42, 60, 52 and 56 for Tuesday to Saturday, respectively. The forecasting errors are given in the following table.

Table 20 : Forecasting Errors

Day	Forecast	Forecasting Errors		
		First Week	Second Week	Third Week
Sun	49	−2	2	0
Mon	52	2	−3	2
Tue	43	1	−3	3
Wed	42	4	−2	−1
Thu	60	3	−1	−3
Fri	52	−1	4	−2
Sat	56	−1	−2	2

The three weeks overall profit is equal to Rs.2015/-.

E3) The data regarding the death rate of India, the population of India, the summer temperatures in your town are examples. You can think of many other examples.

E4) i) Quarterly sales of woollen clothes, numbers of fans sold in a city month-by-month, etc.

ii) Number of people attending church day-wise. Here the variation is within a week, with a peak on Sunday.

E5) The monthly consumption of rice by a family, for instance.

E6) While doing this, remember that there are time series which may not exhibit cyclic variations. For instance, the data on death rates.

E7) One example could be a sudden warm wave in January due to weather irregularities.

E8) i) The number of cars produced monthwise is usually large in January to March every year. Therefore, there will be a seasonal variation. Any unexpected change in the economic or labour policy can bring in irregular variations in the long-term trend. Since the demand will be more in some months and little less in others over the years, there will be a cyclic variation also. So, all the four basic components are expected to be present.

ii) Again, all 4 components can be present : seasonal variation due to exams, irregular variation due to the shop closing for unexpected reasons.

iii) Trend, cyclic and irregular components could appear. The seasonal component may not be exhibited because the time series is annual, not monthly or quarterly.

E9) The S_t values for Mon, Wed, Thu and Fri are −2, −4, 7 and 6, respectively. So, $t_2 = -2$, $t_4 = -4$, $t_6 = 6$, $t_{16} = -2$ (Mon), $t_{26} = 7$ (Thurs).

E10) The trend components for the weeks 6, 7 and 8 are 66, 69 and 72, respectively. To get the 6th week's forecasts, we must add 66 to the average seasonal components (given in the last column of Table 7). Similarly, we must add 69 and 72 to the seasonal components to get the other two weeks' forecasts. The forecasts are tabulated below.

Table 21 : Forecasts (Actual Sales) for 6th, 7th and 8th Weeks

Day → / Week ↓	Sun	Mon	Tue	Wed	Thu	Fri	Sat
6	64(59)	64(65)	59(58)	62(62)	73(79)	72(66)	70(69)
7	67(64)	67(66)	62(60)	65(57)	76(81)	75(70)	73(69)
8	70(69)	70(75)	65(67)	68(69)	79(78)	78(74)	76(73)

E11) In 1995-96 it is 63, in 1996-97 it is 68, and so on till in 2000-01, it is 105. It is clearly increasing over the years.

E12) Wednesdays are represented by the day numbers t = 4, 11, 18, 25, 32, ... Substituting t = 4 in (1), we get $T_t = 50.0$. This is the trend value of Day 4. All the required trend values by the two methods, regression and the ad hoc approach of Example 3, are tabulated below. Remember that for the ad hoc method, the starting trend value is 50.4, and thereafter it increases by 3.2 every week.

Table 22 : Trend Values By Two Methods

Method	Day (t)					
	4	11	18	25	32	...
Regression	50.0	53.4	56.8	60.2	63.6	...
Ad hoc	50.4	53.6	56.8	60.0	63.2	...

E13) The trend equation is given by

$$T_t = -1.724 + 0.0455t, \; t = 61, 62, \ldots \qquad (11)$$

with $R^2 = 0.35$ and standard error $S_e = 0.1991$.

E14) Two equations are fitted for you to see, and compare. One is linear regression and the other is a second degree equation. Their graphs are shown in the figure below. Here t is taken as t = 1 for 1901, t = 2 for 1911, and so on.

$$Y_t = 769 + 485.68 \times t, \; t = 1, 2, \ldots, \quad R^2 = 0.87, S_e = 585.41, \qquad (12)$$

$$Y_t = 2305 - (282.398 \times t) + (69.825 \times t^2), \; t = 1, 2, \ldots, R^2 = 0.99, S_e = 154. \qquad (13)$$

From the values of R^2 in (12) and (13), you can see that the quadratic polynomial is a better fit than the linear one.

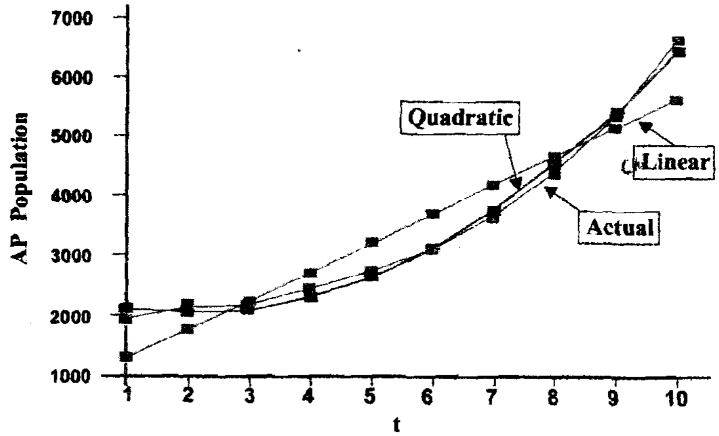


Fig. 14 : Linear and quadratic trend curves for population

E15) a) The graph for the exponential curve fitting and the actual time series values is given below. The quadratic fit is better in this case.

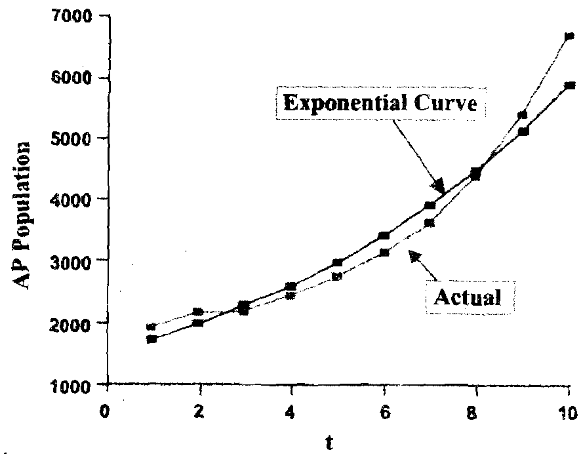


Fig.15 : Exponential curve for AP population

b) According to the exponential fit T_{11} = 6,68,80,580, and according to the quadratic fit T_{11} = 7,64,84,470. In this case the quadratic curve gives a better forecast.
According to the exponential fit, T_{12} = 7,66,47,850; and according to the quadratic fit T_{12} = 7,95,79,040.

E16) Take T_t as the dependent and t as the independent variable and fit a linear regression equation. If we do this for our data, we get the estimate of T_t as $T_t = 26.1 + 2.47t$.

E17) The first three moving averages of length 3 are 46.33, 47 and 51. These numbers should be placed against Days 2, 3 and 4, respectively.

E18) The first and last moving averages of length 6 are 50.33 and 63.33, and they correspond to Days 3.5 and 32.5, respectively.

E19) The moving averages are given in Columns (3), (4) and (5), respectively in Table 23 below.

Table 23

Year (t)	Crop Yield	Moving averages of length : 3	4	5
1955	27.23			
1956	29.27	28.27	28.7275	
1957	28.31	29.22667	29.6225	29.144
1958	30.1	29.74	29.7075	29.62
1959	30.81	30.17333	31.11	30.55
1960	29.61	31.44667	32.2725	31.838
1961	33.92	32.76	32.9625	32.532
1962	34.75	34.08	34.21	33.29
1963	33.57	34.30667	33.58	33.648
1964	34.6	33.19	33.2	33.51
1965	31.4	33.07667	33.305	33.358
1966	33.23	32.87333	31.7775	32.342
1967	33.99	31.90333	32.6	32.36
1968	28.49	32.39	33.095	33.122
1969	34.69	32.79667	32.2	32.558
1970	35.21	33.43667	32.3975	31.616
1971	30.41	31.63333	32.17	32.674
1972	29.28	31.15667	32.25	32.842
1973	33.78	32.86333	34.385	33.59
1974	35.53	36.08667	35.9775	34.638
1975	38.95	36.71	36.69	36.108
1976	35.65	37.07667	37.755	37.31
1977	36.63	37.35667	36.69	37.142
1978	39.79	37.03667	36.7775	36.552
1979	34.69	36.82667	37.18	37.07
1980	36	36.31	36.3275	37.02
1981	38.24	36.87333	38.0625	37.388
1982	36.38	38.75	37.8075	37.446
1983	41.63	37.66333	36.8775	37.15
1984	34.98	37.04333	36.43	36.42
1985	34.52	34.69667	34.04	35.558
1986	34.59	33.72667	35.84	35.668
1987	32.07	36.28	37.725	37.084
1988	42.18	38.77	39.1675	38.252
1989	42.06	41.53333	40.99	39.206
1990	40.36	40.59333	39.455	40
1991	39.36	38.58667	37.8075	38.658
1992	36.04	36.95667	36.81	37.52
1993	35.47	35.96	36.2	36.832
1994	36.37	36.25333		
1995	36.92			

Let us consider the curves given by these in Fig. 16 below.

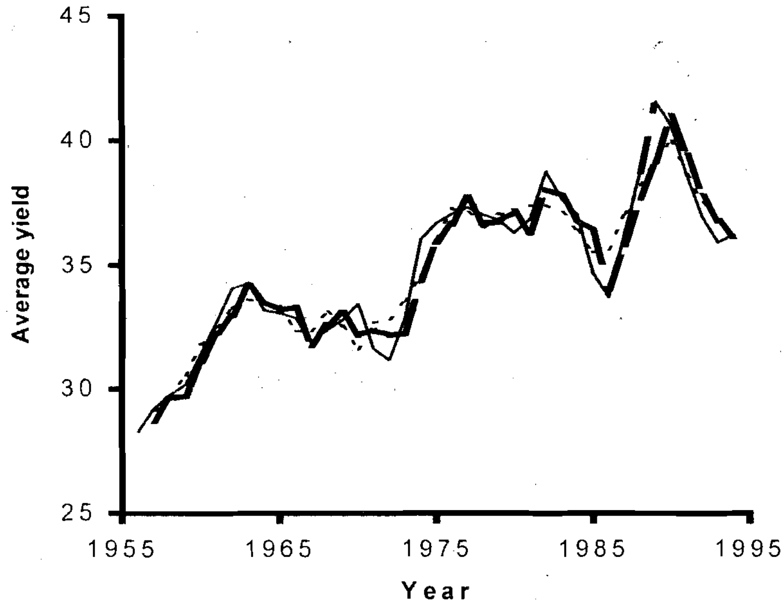


Fig.16

Looking at the curves, we see that the cycles are usually of 5-year periods because the moving averages of length 5 do not show the cyclical variation. The S-component, in any case, can be excluded because the data is annual and doesn't allow us to observe seasonal variation.

E20) Consider the average of y_1, y_2 and y_3. This is the average of sales belonging to three different seasons. Similarly, the average of y_2, y_3 and y_4 is also the average of sales belonging to three different seasons. This way, we can compute the average of y_{t-1}, y_t and y_{t+1} for t = 2, 3, …, 20 (rounded off to the nearest integer). Since these averages have the effect of all seasons and the random effect, they are treated as the estimates of T_t. We have obtained the values in the table below.

Table 24 : Coconut Sales Data

Period (t)	Year	Season	Coconuts Sold (y_t) (in lakhs)	Trend component (T_t)
1	1994-95	I	14	-
2	1994-95	II	35	38
3	1994-95	III	65	38
4	1995-96	I	14	41
5	1995-96	II	43	45
6	1995-96	III	77	45
7	1996-97	I	16	44
8	1996-97	II	40	47
9	1996-97	III	84	47
10	1997-98	I	17	49
11	1997-98	II	46	51
12	1997-98	III	90	53
13	1998-99	I	22	54
14	1998-99	II	50	59
15	1998-99	III	105	59
16	1999-00	I	23	53
17	1999-00	II	30	58
18	1999-00	III	120	59
19	2000-01	I	27	72
20	2000-01	II	68	76
21	2000-01	III	132	-

E21) $\hat{a}_3 = 0.02y_3 + 0.98\,\hat{a}_2 = 0.02 \times 847 + 0.98 \times 946.14 = 944.16.$

The corresponding forecast error,

$\hat{e}_4 = y_4 - \hat{a}_3 = 1063 - 944.16 = 118.84.$

$\hat{a}_4 = 0.02y_4 + 0.98\,\hat{a}_3 = (0.02 \times 1063) + (0.98 \times 944.16) = 946.54.$

The corresponding forecast error is

$\hat{e}_5 = y_5 - \hat{a}_4 = 1030 - 946.54 = 83.46.$

E22) **Table 25 : Exponential Smoothing of Banks Data**

Year	No. of Branches	Forecast	Error
81	5	4.00	1.00
82	3	4.01	–1.01
83	3	3.999	–0.99
84	4	3.99	0.01
85	3	3.99	–0.99
86	6	3.98	2.02
87	4	4.00	0.00

E23) **Table 26 : Forecasts of Rice Yield For 1985 to 1990**

Year	t	Rice Yield	Forecast	Error
1986	32	1951	2114	–163
1987	33	2258	2146	112
1988	34	2572	2178	394
1989	35	2403	2210	193
1990	36	2442	2242	200

E24) $$S_{II} = \frac{S_2 I_2 + \cdots + S_{20} I_{20}}{7} = 0.841$$

$S_{III} = 1.787$

E25) **Table 27**

Period (t)	Year	Season	Coconuts Sold (y_t) (in lakhs)	Trend component (T_t)	Estimates of S_tI_t ($= y_t/T_t$)	Estimates of S_t	Estimates of $I_t (= S_tI_t/S_t)$
1	1994-95	I	14	-	-	0.378	-
2	1994-95	II	35	38	0.921	0.841	1.095
3	1994-95	III	65	38	1.711	1.787	0.957
4	1995-96	I	14	41	0.341	0.378	0.903
5	1995-96	II	43	45	0.956	0.841	1.136
6	1995-96	III	77	45	1.711	1.787	0.958
7	1996-97	I	16	44	0.364	0.378	0.962
8	1996-97	II	40	47	0.851	0.841	1.012
9	1996-97	III	84	47	1.787	1.787	1.0
10	1997-98	I	17	49	0.347	0.378	0.918
11	1997-98	II	46	51	0.902	0.841	1.072
12	1997-98	III	90	53	1.698	1.787	0.95
13	1998-99	I	22	54	0.407	0.378	1.078
14	1998-99	II	50	59	0.847	0.841	1.007
15	1998-99	III	105	59	1.780	1.787	0.996
16	1999-00	I	23	53	0.434	0.378	1.148
17	1999-00	II	30	58	0.517	0.841	0.615
18	1999-00	III	120	59	2.034	1.781	1.138
19	2000-01	I	27	72	0.375	0.378	0.992
20	2000-01	II	68	76	0.895	0.841	1.064
21	2000-01	III	132	-	-	1.781	-

E26) Let us form a table, based on the first 15 readings of Table 9. Since we need to forecast the sales, we should use regression for finding the trend. (Remember, MA is not a useful method for forecasting the trend).

To estimate S_t, we use the same procedure used in E25. Let us use the ad hoc approach (see Example 3) for determining the long-term trend component.

Now, the average sales for the first year are $\frac{14+35+65}{3} = \frac{114}{3}$, for the 2nd year are $\frac{134}{3}$, $\frac{140}{3}$ for the 3rd year, $\frac{153}{3}$ for the 4th year, and $\frac{177}{3}$ for the 5th year. So, the average increase in trend is $\frac{1}{4}\left(\frac{20}{3}+\frac{6}{3}+\frac{13}{3}+\frac{24}{3}\right) = \frac{21}{4} = 5.25$ units. Therefore, the 5th, 6th, and 7th columns in the table will be as shown.

Table 28

Period (t)	Year	Season	Coconuts Sold (y_t) (in lakhs)	Estimates of T_t	Estimates of S_t	Estimates of I_t ($=y_t/S_t.T_t$)
1	1994-95	I	14	38	0.358	1.029
2	1994-95	II	35	38	0.895	1.029
3	1994-95	III	65	38	1.727	0.990
4	1995-96	I	14	43.25	0.358	0.904
5	1995-96	II	43	43.25	0.895	1.111
6	1995-96	III	77	43.25	1.727	1.031
7	1996-97	I	16	48.5	0.358	0.921
8	1996-97	II	40	48.5	0.895	0.921
9	1996-97	III	84	48.5	1.727	1.003
10	1997-98	I	17	53.75	0.358	0.883
11	1997-98	II	46	53.75	0.895	0.956
12	1997-98	III	90	53.75	1.727	0.97
13	1998-99	I	22	59	0.358	1.042
14	1998-99	II	50	59	0.895	0.947
15	1998-99	III	105	59	1.727	1.030

Using these estimates, the forecast for the three seasons in the next two years, ignoring the irregular component, are (64.25) (0.358), (64.25) (0.895), …, (69.5) (1.727), i.e., 23, 58, 111, 25, 62, 120 (rounded off to the nearest integer). The forecasting errors are 0, –28, 9, 2, 6, 12, respectively.