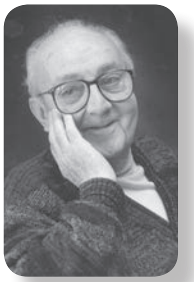
CHAPTER

7 TIME SERIES AND FORECASTING



G. E. P. Box

G. E. P. Box (1919-2013) was a Britsh Statistician was "one of the great statistical minds" of the 20th century, who received his Ph.D., from the University of London, under the supervision of E. S. Pearson. He served as President of Americal Statistical Association in 1978 and of the Institute of Mathematics in 1979. His name is associated with Box-Cox transformation in addition to Box-Jenkins models in time series.

G. M. Jenkins

G. M. Jenkins (1932-1982) was a British Statistician, earned his Ph.D. degree from University College, London under the supervision of F. N. David and N. L. Johnson. He served on the Research Section Committee and Council of Royal Statistical Society in 1960's. He was elected to the Institute of Mathematical Statitics

Both Box and Jenkins contributed to Auto regressive moving average models popularly known as Box-Jenkins Models.

LEARNING OBJECTIVES

The students will be able to

- ❖ understand the concept of time series
- ❖ know the upward and downward trends
- ❖ calculate the trend values using semi - average and moving average methods
- ❖ estimate the trend values using method of least squares
- ❖ compute seasonal indices
- ❖ understand cyclical and irregular variations
- ❖ understand the forecasting concept

Introduction

In modern times we see data all around. The urge to evaluate the past and to peep into the future has made the need for forecasting. There are many factors which change with the passage of time. Sometimes sets of observations which vary with the passage of time and whose measurements made at equidistant points may be regarded as time series data. Statistical data which are collected, observed or recorded at successive intervals of time constitute time series data. In the study of time series, comparison of the past and the present data is made. It also compares two or more series at a time. The purpose of time series is to measure chronological variations in the observed data.

12th_Statistics_EM_Unit_7.indd 182 12-12-2021 21:50:16

In an ever changing business and economic environment, it is necessary to have an idea about the probable future course of events. Analysis of relevant time series helps to achieve this, especially by facilitating future business forecasts. Such forecasts may serve as crucial inputs in deciding competitive strategies and planning growth initiatives.

7.1 DEFINITION

Time series refers to any group of statistical information collected at regular intervals of time. Time series analysis is used to detect the changes in patterns in these collected data.

7.1.1 Definition by Authors

According to Mooris Hamburg "A time series is a set of statistical observations arranged in chronological order".

Ya-Lun-Chou : "A time series may be defined as a collection of readings belonging to different time periods of some economic variable or composite of variables".

W.Z. Hirsch says "The main objective in analyzing time series is to understand, interpret and evaluate change in economic phenomena in the hope of more correctly anticipating the course of future events".

7.1.2 Uses of Time Series

- Time series is used to predict future values based on previously observed values.
- Time series analysis is used to identify the fluctuation in economics and business.
- It helps in the evaluation of current achievements.
- Time series is used in pattern recognition, signal processing, weather forecasting and earthquake prediction.

It can be said that time series analysis is a big tool in the hands of business executives to plan their sales, prices, policies and production.

7.2 COMPONENTS OF TIME SERIES

The factors that are responsible for bringing about changes in a time series are called the components of time series.

Components of Time Series

1. Secular trend
2. Seasonal variation
3. Cyclical variation
4. Irregular (random) variation

Approaches to time series

There are two approaches to the decomposition of time series data

(i) Additive approach

(ii) Multiplicative approach

The above two approaches are used in decomposition, depending on the nature of relationship among the four components.

The additive approach

The additive approach is used when the four components of a time series are visualized as independent of one another. Independence implies that the magnitude and pattern of movement of the components do not affect one another. Under this assumption the magnitudes of the time series are regarded as the sum of separate influences of its four components.

$$Y = T + C + S + R$$

where Y = magnitude of a time series

T = Trend,
C =Cyclical component,
S =Seasonal component, and
R = Random component

In additive approach, the unit of measurements remains the same for all the four components.

The Multiplicative approach

The multiplicative approach is used where the forces giving rise to the four types of variations are visualized as interdependent. Under this assumption, the magnitude of the time series is the product of its four components.

i.e. $Y = T \times C \times S \times R$

Difference between the two approaches

Multiplicative	Additive
(i) Four components of time series are interdependent	Four components of time series are independent
(ii) Logarithm of components are additive	Components are additive

7.3 MEASUREMENTS OF COMPONENTS

(i) Secular trend

It refers to the long term tendency of the data to move in an upward or downward direction. For example, changes in productivity, increase in the rate of capital formation, growth of population, *etc.*, follow secular trend which has upward direction, while deaths due to improved medical facilities and sanitations show downward trend. All these forces occur in slow process and influence the time series variable in a gradual manner.

12th_Statistics_EM_Unit_7.indd 184 12-12-2021 21:50:16

Methods of Measuring Trend

Trend is measured using by the following methods:

1. Graphical method
2. Semi averages method
3. Moving averages method
4. Method of least squares

7.3.1 Graphical Method

Under this method the values of a time series are plotted on a graph paper by taking time variable on the X-axis and the values variable on the Y-axis. After this, a smooth curve is drawn with free hand through the plotted points. The trend line drawn above can be extended to forecast the values. The following points must be kept in mind in drawing the freehand smooth curve.

(i) The curve should be smooth

(ii) The number of points above the line or curve should be approximately equal to the points below it

(iii) The sum of the squares of the vertical deviation of the points above the smoothed line is equal to the sum of the squares of the vertical deviation of the points below the line.

Merits

- It is simple method of estimating trend.
- It requires no mathematical calculations.
- This method can be used even if trend is not linear.

Demerits

- It is a subjective method
- The values of trend obtained by different statisticians would be different and hence not reliable.

Example 7.1

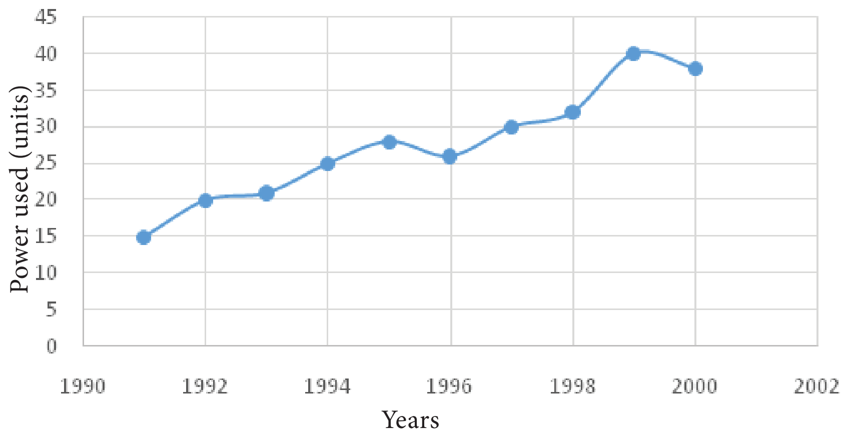
Annual power consumption per household in a certain locality was reported below.

Years	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Power used (units)	15	20	21	25	28	26	30	32	40	38

Draw a free hand curve for the above data.

12th_Statistics_EM_Unit_7.indd 185 12-12-2021 21:50:16

Solution:



7.3.2 Semi-Average Method

In this method, the series is divided into two equal parts and the average of each part is plotted at the mid-point of their time duration.

(i) In case the series consists of an even number of years, the series is divisible into two halves. Find the average of the two parts of the series and place these values in the mid-year of each of the respective durations.

(ii) In case the series consists of odd number of years, it is not possible to divide the series into two equal halves. The middle year will be omitted. After dividing the data into two parts, find the arithmetic mean of each part. Thus we get semi-averages.

(iii) The trend values for other years can be computed by successive addition or subtraction for each year ahead or behind any year.

Merits

- This method is very simple and easy to understand
- It does not require many calculations.

NOTE

In semi-average method if the difference between the semi-averages is negative then the trend values will be in decreasing order.

Demerits

- This method is used only when the trend is linear.
- It is used for calculation of averages and they are affected by extreme values.

Example 7.2

Calculate the trend values using semi-averages methods for the income from the forest department. Find the yearly increase.

Year	2008	2009	2010	2011	2012	2013
Income (in crores)	46.17	51.65	63.81	70.99	84.91	91.64

Source: The Principal Chief conservator of forests, Chennai-15. (pg. 231)

12th_Statistics_EM_Unit_7.indd 186 12-12-2021 21:50:16

Solution:

Year	Income	3-Year semi-total	Semi-average
2008	46.17		
2009	51.65	161.63	53.877
2010	63.81		
2011	70.99		
2012	84.91	247.54	82.513
2013	91.64		

Difference between the central years = 2012 – 2009 = 3

Difference between the semi-averages = 82.513 – 53.877 = 28.636

Increase in trend value for one year = $\dfrac{28.636}{3} = 9.545$

Trend values for the previous and successive years of the central years can be calculated by subtracting and adding respectively, the increase in annual trend value.

Example 7.3

Population of India for 7 successive census years are given below. Find the trend values using semi-averages method.

Census Year	1951	1961	1971	1981	1991	2001	2011
Population (in lakhs)	301.2	336.9	412.0	484.1	558.6	624.1	721.4

Solution:

Trend values using semi average method

Census Year	Population (in lakhs)	3-year semi-total	3-year semi-average	Trend values
1951	301.2			278.86
1961	336.9	1050.1	350.03	350.03
1971	412.0			421.2
1981	484.1			492.37
1991	558.6			563.54
2001	624.1	1904.1	634.7	634.71
2011	721.4			705.88

Difference between the years = 2001 – 1961 = 40

Difference between the semi-averages = 634.7 – 350.03 = 284.67

Increase in trend value for 10 year = $\dfrac{284.67}{4} = 71.17$

For example the trend value for the year 1951 = 350.03 – 71.17 = 278.86

The value for the year 2011 = 634.7 + 71.17 = 705.87

The trend values have been calculated by successively subtracting and adding the increase in trend for previous and following years respectively.

Example 7.4

Find the trend values by semi-average method for the following data.

Year	1965	1966	1967	1968	1969	1970	1971	1972
Production of bleaching powder (in tonnes)	7.4	10.8	9.2	10.5	15.5	13.7	16.7	15

Solution:

Trend values using semi averages method

Year	Production of bleaching powder	4 year semi-total	4 year semi-average	Trend
1965	7.4			7.315
1966	10.8	37.9	9.475	8.755
1967	9.2			10.195
1968	10.5			11.635
1969	15.5			13.075
1970	13.7	60.9	15.225	14.515
1971	16.7			15.955
1972	15			17.395

Difference between the years = 1970.5 – 1966.5 = 4

Difference between the semi-averages = 15.225 – 9.475 = 5.75

Increase in trend = $\frac{5.75}{4} = 1.44$

Half yearly increase in trend = $\frac{1.44}{2} = 0.72$

The trend value for 1967 = 9.475 + 0.72 = 10.195

The trend value for 1968 = 9.475 + 3 * 0.72 = 11.635

Similarly the trend values for the other years can be calculated.

7.3.3 Moving Averages Method

Moving averages is a series of arithmetic means of variate values of a sequence. This is another way of drawing a smooth curve for a time series data.

12th_Statistics_EM_Unit_7.indd 188 12-12-2021 21:50:16

Moving averages is more frequently used for eliminating the seasonal variations. Even when applied for estimating trend values, the moving average method helps to establish a trend line by eliminating the cyclical, seasonal and random variations present in the time series. The period of the moving average depends upon the length of the time series data.

The choice of the length of a moving average is an important decision in using this method. For a moving average, appropriate length plays a significant role in smoothening the variations. In general, if the number of years for the moving average is more then the curve becomes smooth.

Merits

- It can be easily applied
- It is useful in case of series with periodic fluctuations.
- It does not show different results when used by different persons
- It can be used to find the figures on either extremes; that is, for the past and future years.

Demerits

- In non-periodic data this method is less effective.
- Selection of proper 'period' or 'time interval' for computing moving average is difficult.
- Values for the first few years and as well as for the last few years cannot be found.

Moving averages odd number of years (3 years)

To find the trend values by the method of three yearly moving averages, the following steps have to be considered.

1. Add up the values of the first 3 years and place the yearly sum against the median year. [This sum is called moving total]
2. Leave the first year value, add up the values of the next three years and place it against its median year.
3. This process must be continued till all the values of the data are taken for calculation.
4. Each 3-yearly moving total must be divided by 3 to get the 3-year moving averages, which is our required trend values.

Example 7.5

Calculate the 3-year moving averages for the loans issued by co-operative banks for non-farm sector/small scale industries based on the values given below.

Year	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15
Loan by District Central Cooperative banks (Rupees in crores)	41.82	40.05	39.12	24.72	26.69	59.66	23.65	28.36	33.31	31.60	36.48

12th_Statistics_EM_Unit_7.indd 189 12-12-2021 21:50:17

Solution:

The three year moving averages are shown in the last column.

Year	Loan by District Central Cooperative Banks	3-year moving total	3-year moving average
2004-05	41.82	-	-
2005-06	40.05	120.99	40.33
2006-07	39.12	103.89	34.63
2007-08	24.72	90.53	30.18
2008-09	26.69	111.07	37.02
2009-10	59.66	110	36.67
2010-11	23.65	111.67	37.22
2011-12	28.36	85.32	28.44
2012-13	33.31	93.27	31.09
2013-14	31.60	101.39	33.80
2014-15	36.48	-	-

Moving averages - even number of years (4 years)

1. Add up the values of the first 4 years and place the sum against the middle of 2nd and 3rd year. (This sum is called 4 year moving total)
2. Leave the first year value and add next 4 values from the 2nd year onward and write the sum against its middle position.
3. This process must be continued till the value of the last item is taken into account.
4. Add the first two 4-years moving total and write the sum against 3rd year.
5. Leave the first 4-year moving total and add the next two 4-year moving total and place it against 4th year.
6. This process must be continued till all the 4-yearly moving totals are summed up and centered.
7. Divide the 4-years moving total by 8 to get the moving averages which are our required trend values.

Example 7.6

Compute the trends by the method of moving averages, assuming that 4-year cycle is present in the following series.

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Annual value	154.0	140.5	147.0	148.5	142.9	142.1	136.6	142.7	145.7	145.1	137.8

12th_Statistics_EM_Unit_7.indd 190 12-12-2021 21:50:17

Solution:

The four year moving averages are shown in the last column.

Year	Annual value	4-year moving total	Centered total	4-year moving average
1998	154.0			
		-		
1999	140.5		-	
		590.0		
2000	147.0		1168.9	146.11
		578.9		
2001	148.5		1159.4	144.93
		580.5		
2002	142.9		1150.6	143.83
		570.1		
2003	142.1		1134.4	141.8
		564.3		
2004	136.6		1131.4	141.43
		567.1		
2005	142.7		1137.2	142.15
		570.1		
2006	145.7		1141.4	142.68
		571.3		
2007	145.1		-	
		-		
2008	137.8			

7.3.4 Method of least squares

Among the four components of the time series, secular trend represents the long term direction of the series. One way of finding the trend values with the help of mathematical technique is the method of least squares. This method is most widely used in practice and in this method the sum of squares of deviations of the actual and computed values is least and hence the line obtained by this method is known as the line of best fit.

It helps for forecasting the future values. It plays an important role in finding the trend values of economic and business time series data.

Computation of Trend using Method of Least squares

Method of least squares is a device for finding the equation which best fits a given set of observations.

Suppose we are given n pairs of observations and it is required to fit a straight line to these data. The general equation of the straight line is:

12th_Statistics_EM_Unit_7.indd 191 12-12-2021 21:50:17

$$y = a + bx$$

where a and b are constants. Any value of a and b would give a straight line, and once these values are obtained an estimate of y can be obtained by substituting the observed values of y. In order that the equation $y = a + b\,x$ gives a good representation of the linear relationship between x and y, it is desirable that the estimated values of y_i, say $\hat{y}_i$ on the whole close enough to the observed values y_i, $i = 1, 2, \ldots, n$. According to the principle of least squares, the best fitting equation is obtained by minimizing the sum of squares of differences $\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$

That is, $\sum\left(y_i - \hat{y}_i\right)^2 = \sum_{i=1}^{n}\left(y_i - a - bx_i\right)^2$ is minimum. This leads us to two normal equations.

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i \tag{7.1}$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x + b\sum_{i=1}^{n} x_i^2 \tag{7.2}$$

Solving these two equations we get the vales for a and b and the fit of the trend equation (line of best):

$$y = a + bx \tag{7.3}$$

Substituting the observed values x_i in (7.3) we get the trend values y_i, $i = 1, 2, \ldots, n$.

Note: The time unit is usually of uniform duration and occurs in consecutive numbers. Thus, when the middle period is taken as the point of origin, it reduces the sum of the time variable x to zero $\left(\sum_{i=1}^{n} x_i = 0\right)$ and hence we get

$a = \dfrac{\sum_{i=1}^{n} y_i}{n}$ and $b = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$ by simplifying (7.1) and (7.2)

The number of time units may be even or odd, depending upon this, we follow the method of calculating trend values using least square method.

Merits

- The method of least squares completely eliminates personal bias.
- Trend values for all the given time periods can be obtained
- This method enables us to forecast future values.

Demerits

- The calculations for this method are difficult compared to the other methods.
- Addition of new observations requires recalculations.
- It ignores cyclical, seasonal and irregular fluctuations.
- The trend can be estimated only for immediate future and not for distant future.

12th_Statistics_EM_Unit_7.indd 192 12-12-2021 21:50:18

Steps for calculating trend values when n is odd:

i) Subtract the first year from all the years (x)

ii) Take the middle value (A)

iii) Find $u_i = x_i - A$

iv) Find u_i^2 and $u_i y_i$

Then use the normal equations:

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} u_i$$

$$\sum_{i=1}^{n} u_i y_i = a\sum_{i=1}^{n} u_i + b\sum_{i=1}^{n} u_i^2$$

Find $a = \dfrac{\sum_{i=1}^{n} y_i}{n}$ and $b = \dfrac{\sum_{i=1}^{n} u_i y_i}{\sum_{i=1}^{n} u_i^2}$

Then the estimated equation of straight line is:

$y = a + b\,u = a + b\,(x - A)$

Example 7.7

Fit a straight line trend by the method of least squares for the following consumer price index numbers of the industrial workers.

Year	2010	2011	2012	2013	2014
Index number	166	177	198	221	225

Solution:

Year	Index Number	$X = x_i - 2010$	$u_i = X - A$ $= X - 2$	u_i^2	$u_i y_i$	Trend
2010	166	0	–2	4	–332	165
2011	177	1	–1	1	–177	181.2
2012	198	2	0	0	0	197.4
2013	221	3	1	1	221	213.6
2014	225	4	2	4	450	229.8
	$\sum_{i=1}^{5} y_i = 987$		$\sum_{i=1}^{5} u_i = 0$	$\sum_{i=1}^{5} u_i^2 = 10$	$\sum_{i=1}^{5} u_i y_i = 162$	

The equation of the straight line is $y = a + bx$

$= a + bu$ where $u = X - 2$

The normal equations give:

$$a = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{987}{5} = 197.4$$

$$b = \frac{\sum_{i=1}^{n} u_i y_i}{\sum_{i=1}^{n} u_i^2} = \frac{162}{10} = 16.2$$

$y = 197.4 + 16.2\ (X - 2)$

$= 197.4 + 16.2\ X - 32.4$

$= 16.2\ X + 165$

That is, $y = 165 + 16.2X$

To get the required trend values, put $X = 0, 1, 2, 3, 4$ in the estimated equation.

$X = 0,\ y = 165 + 0 = 165$

$X = 1,\ y = 165 + 16.2 = 181.2$

$X = 2,\ y = 165 + 32.4 = 197.4$

$X = 3,\ y = 165 + 48.6 = 213.6$

$X = 4,\ y = 165 + 64.8 = 229.8$

Hence, the trend values for 2010, 2011, 2012, 2013 and 2014 are 165, 181.2, 197.4, 213.6 and 229.8 respectively.

Steps for calculating trend values when n is even:

i). Subtract the first year from all the years (x)

ii). Find $u_i = 2X - (n - 1)$

iii). Find u_i^2 and $u_i y_i$

Then follow the same procedure used in previous method for odd years

Example 7.8

Tourist arrivals (Foreigners) in Tamil Nadu for 6 consecutive years are given in the following table. Calculate the trend values by using the method of least squares.

Year	2005	2006	2007	2008	2009	2010
No. of arrivals (in lakhs)	12	13	18	20	24	28

12th_Statistics_EM_Unit_7.indd 194 12-12-2021 21:50:19

Solution:

Year x	No. of arrivals y_i	$X = x_i - 2005$	$u_i = 2X - 5$	u_i^2	$u_i y_i$
2005	12	0	–5	25	–60
2006	13	1	–3	9	–39
2007	18	2	–1	1	–18
2008	20	3	1	1	20
2009	24	4	3	9	72
2010	28	5	5	25	140
	$\sum_{i=1}^{6} y_i = 115$		$\sum_{i=1}^{6} u_i = 0$	$\sum_{i=1}^{6} u_i^2 = 70$	$\sum_{i=1}^{6} u_i y_i = 115$

The equation of the straight line is $y = a + bx$

$= a + bu$ where $u = 2X - 5$

Using the normal equation we have,

$$a = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{115}{6} = 19.17$$

$$b = \frac{\sum_{i=1}^{n} u_i y_i}{\sum_{i=1}^{n} u_i^2} = \frac{115}{70} = 1.64$$

$$\begin{aligned} y &= a + bu \\ &= 19.17 + 1.64\,(2X - 5) \\ &= 19.17 + 3.28X - 8.2 \\ &= 3.28X + 10.97 \end{aligned}$$

That is, $y = 10.97 + 3.28X$

To get the required trend values, put $X = 0, 1, 2, 3, 4, 5$ in the estimated equation. Thus,

$X = 0, y = 10.97 + 0 = 10.97$

$X = 1, y = 10.97 + 3.28 = 14.25$

$X = 2, y = 10.97 + 6.56 = 17.53$

$X = 3, y = 10.97 + 9.84 = 20.81$

$X = 4, y = 10.97 + 13.12 = 24.09$

$X = 5, y = 10.97 + 16.4 = 27.37$

Hence, the trend values for 2005, 2006, 2007, 2008, 2009 and 2010 are 10.97, 14.25, 17.53, 20.81, 24.09 and 27.37 respectively.

(ii) Seasonal variation

Seasonal variations are fluctuations within a year over different seasons.

Estimation of seasonal variations requires that the time series data are recorded at even intervals such as quarterly, monthly, weekly or daily, depending on the nature of the time series. Changes due to seasons, weather conditions and social customs are the primary causes of seasonal variations. The main objective of the measurement of seasonal variation is to study their effect and isolate them from the trend.

Methods of constructing seasonal indices

There are four methods of constructing seasonal indices.

1. Simple averages method
2. Ratio to trend method
3. Percentage moving average method
4. Link relatives method

Among these, we shall discuss the construction of seasonal index by the first method only.

7.3.5 Simple Averages Method

Under this method, the time series data for each of the 4 seasons (for quarterly data) of a particular year are expressed as percentages to the seasonal average for that year.

The percentages for different seasons are averaged over the years by using simple average. The resulting percentages for each of the 4 seasons then constitute the required seasonal indices.

Method of calculating seasonal indices

i) The data is arranged season-wise

ii) The data for all the 4 seasons are added first for all the years and the seasonal averages for each year is computed.

iii) The average of seasonal averages is calculated
(*i.e.*, Grand average = Total of seasonal averages /number of years).

iv) The seasonal average for each year is divided by the corresponding grand average and the results are expressed in percentages and these are called seasonal indices.

Example 7.9

Calculate the seasonal indices for the rain fall (in mm) data in Tamil Nadu given below by simple average method

Year	Season			
	I	II	III	IV
2001	118.4	260.0	379.4	70
2002	85.8	185.4	407.1	8.7
2003	129.8	336.5	403.1	12.0
2004	283.4	360.7	472.1	14.3
2005	231.7	308.5	828.8	15.9

12th_Statistics_EM_Unit_7.indd 196 12-12-2021 21:50:21

Solution:

Year	Season			
	I	II	III	IV
2001	118.4	260.0	379.4	70
2002	85.8	185.4	407.1	8.7
2003	129.8	336.5	403.1	12.0
2004	283.4	360.7	472.1	14.3
2005	231.7	308.5	828.8	15.9
Seasonal total	849.1	1451.1	2490.5	120.9
Seasonal average	169.82	290.22	498.1	24.18
Seasonal index	69	118	203	10

$$\text{Grand Average} = \frac{\text{Total of seasonal averages}}{4}$$

$$= \frac{169.82 + 290.22 + 498.1 + 24.18}{4}$$

$$= \frac{982.32}{4} = 245.58$$

$$\text{Seasonal Index} = \frac{\text{Seasonal average}}{\text{Grand average}} \times 100$$

$$\text{Seasonal Index for Season I} = \frac{169.82}{245.58} \times 100 = 69.15 \approx 69$$

$$\text{Seasonal Index for Season II} = \frac{290.22}{245.58} \times 100 = 118.18 \approx 118$$

$$\text{Seasonal Index for Season III} = \frac{498.1}{245.58} \times 100 = 202.83 \approx 203$$

$$\text{Seasonal Index for Season IV} = \frac{24.18}{245.58} \times 100 = 9.85 \approx 10$$

Example 7.10

Obtain the seasonal indices for the rain fall (in mm) data in India given in the following table.

Quarter \ Year	2009	2010	2011	2012
I	38.2	38.5	55	50.5
II	166.8	250.9	277.7	197
III	612.6	773.1	717.8	706.1
IV	72.2	153.1	65.8	101.1

12th_Statistics_EM_Unit_7.indd 197 12-12-2021 21:50:21

Solution:

Year	Quarter			
	I	II	III	IV
2009	38.2	166.8	612.6	72.2
2010	38.5	250.9	773.1	153.1
2011	55	277.7	717.8	65.8
2012	50.5	197	706.1	101.1
Seasonal total	182.2	892.4	2809.6	392.2
Seasonal average	45.55	223.1	702.4	98.05
Seasonal index	17	83	263	37

$$\text{Grand Average} = \frac{\text{Total of seasonal averages}}{4}$$

$$= \frac{45.55 + 223.1 + 702.4 + 98.05}{4}$$

$$= \frac{1069.10}{4} = 267.28$$

$$\text{Seasonal Index} = \frac{\text{Seasonal average}}{\text{Grand average}} \times 100$$

$$\text{Seasonal Index for Quarter I} = \frac{45.55}{267.28} \times 100 = 17.04 \approx 17$$

$$\text{Seasonal Index for Quarter II} = \frac{223.1}{267.28} \times 100 = 83.47 \approx 83$$

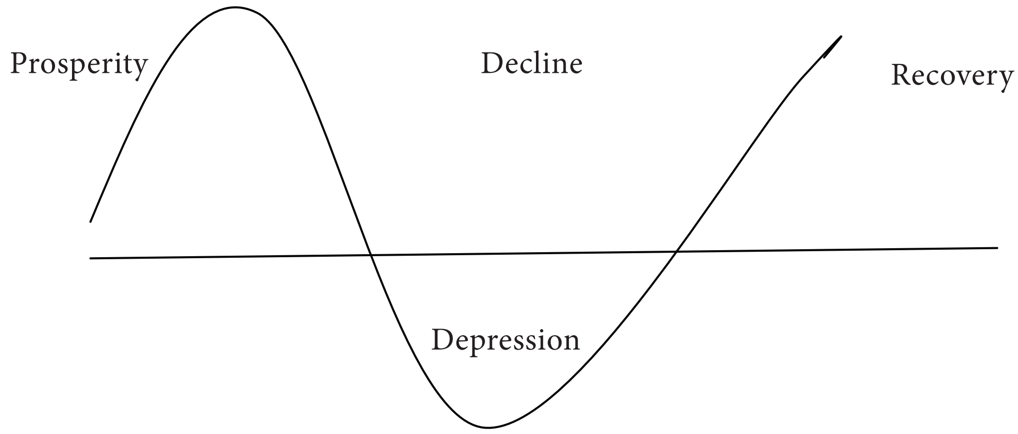
$$\text{Seasonal Index for Quarter III} = \frac{702.4}{267.28} \times 100 = 262.80 \approx 263$$

$$\text{Seasonal Index for Quarter IV} = \frac{98.05}{267.28} \times 100 = 36.69 \approx 37$$

12th_Statistics_EM_Unit_7.indd 198 12-12-2021 21:50:21

(iii) Cyclical variation

Cyclical variations refer to periodic movements in the time series about the trend line, described by upswings and downswings. They occur in a cyclical fashion over an extended period of time (more than a year). For example, the business cycle may be described as follows.



The cyclical pattern of any time series tells about the prosperity and recession, ups and downs, booms and depression of a business. In most of the businesses there are upward trend for some time followed by a downfall, touching its lowest level. Again a rise starts which touches its peak. This process of prosperity and recession continues and may be considered as a natural phenomenon.

YOU WILL KNOW

Cyclic movements are mainly due to Trade cycle.

(iv) Irregular variation

In practice, the changes in a time series that cannot be attributed to the influence of cyclic fluctuations or seasonal variations or those of the secular trend are classified as irregular variations.

NOTE

There is no statistical technique for measuring or isolating irregular fluctuations

In the words of Patterson, "Irregular variation in a time series is composed of non-recurring sporadic (rare) form which is not attributed to trend, cyclical or seasonal factors".

Nothing can be predicted about the occurrence of irregular influences and the magnitude of such effects. Hence, no standard method has been evolved to estimate the same. It is taken as the residual left in the time series, after accounting for the trend, seasonal and cyclic variations.

YOU WILL KNOW

Irregular variation is also called erratic fluctuations.

7.4 FORECASTING

The importance of statistics lies in the extent to which it serves as the basis for making reliable forecasts, against arbitrary forecast with no statistical background.

12th_Statistics_EM_Unit_7.indd 199 12-12-2021 21:50:21

Forecasting is a scientific process which aims at reducing the uncertainty of the future state of business and trade, not dependent merely on guess work, but with a sound scientific footing for the decision on the future course of action.

NOTE

1. Long term forecasting can be found using Trend.
2. Short term forecasting can be found using seasonal variations.

7.4.1 Definition

"Forecasting refers to the analysis of past and present conditions with a view of arriving at rough estimates about the future conditions.

According to T.S. Lewis and R.A. Fox "Forecasting is using the knowledge we have at one time to estimate what will happen at some future moment of time".

Forecasting is an important tool that serves many fields including business and industry, government, economics, environmental sciences, medicine, social science, politics and finance. Forecasting problems are often classified as short-term, medium-term, and long-term.

Short-term forecasting problems involve predicting events for a few time periods (days, weeks, months) into the future.

Medium-term forecast extends from one to two years into the future.

Long-term forecasting problems can extend beyond that by many years.

Short and medium-term forecasts are required for activities that range from operations management to budgeting and selecting new research and development projects. Long term forecasts impact issues relating to strategic planning.

POINTS TO REMEMBER

- Time series is a time oriented sequence of observations.
- Components of time series are secular trend, seasonal variations, cyclical variations and irregular (erratic) variations
- Methods of calculating trend values are graphical method, semi - averages method, moving averages method, and method of least squares.
- The line $y = a + b\,x$ found out using the method of least squares is called 'line of best fit'.
- Normal equations involved in the method of least squares are

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x + b\sum_{i=1}^{n} x_i^2$$

- Seasonal indices may be found out by using simple average method.
- Forecasting is the analysis of using past and present conditions to get rough estimates of the future conditions
- Forecasting methods can be short-term, medium-term and long-term.

12th_Statistics_EM_Unit_7.indd 200 12-12-2021 21:50:21

EXERCISE 7

I. Choose the best answer.

1. An overall tendency of rise or fall in a time series is called
   (a) seasonal variation (b) secular trend
   (c) cyclical variation (d) irregular variation

2. The component having primary use for short-term forecasting is
   (a) cyclical variation (b) irregular variation
   (c) seasonal variation (d) trend

3. Cyclical movements are due to
   (a) ratio to trend (b) seasonal
   (c) trend (d) trade cycle

4. Data on annual turnover of a company over a period of ten years can be represented by a
   (a) a time series (b) an index number
   (c) a parameter (d) a statistic

5. The component having primary use for long term forecasting is
   (a) cyclical variation (b) irregular variation
   (c) seasonal variation (d) trend

6. A time series is a set of data recorded
   (a) periodically (b) at equal time intervals
   (c) at successive points of time (d) all the above

7. A time series consists of
   (a) two components (b) three components
   (c) four components (d) five components

8. Irregular variation in a time series can be due to
   (a) trend variations (b) seasonal variations
   (c) cyclical variations (d) unpredictable causes

9. The terms prosperity, recession, depression and recovery are in particular attached to
   (a) secular trend (b) seasonal fluctuation
   (c) cyclical movements (d) irregular variation

10. An additive model of time series with components, T, S, C and I is
    (a) $Y = T \times S \times C \times I$ (b) $Y = T + S + C + I$
    (c) $Y = T \times S + C \times I$ (d) $Y = T \times S \times C + I$

12th_Statistics_EM_Unit_7.indd 201 12-12-2021 21:50:21

11. A decline in the sale of ice cream during November to March is associated with

(a) seasonal variation (b) cyclical variation

(c) irregular variation (d) secular trend

12. Business forecasts are made on the basis of

(a) future data (b) past data

(c) tax regulations (d) Government policies

13. The four components of time series in a multiplicative model are

(a) independent (b) interdependent

(c) constant (d) additive

14. In the least square theory the sum of squares of residuals is

(a) zero (b) minimum

(c) constant (d) maximum

15. No statistical techniques for measuring or isolating ________is available.

(a) cyclical variation (b) seasonal variation

(c) erratic fluctuations (d) secular trend

II. Give very short answers to the following questions

16. What is a time series?

17. What are the components of a time series?

18. Name different methods of estimating the trend?

19. Write short notes on irregular variation.

20. Mention the methods used to find seasonal indices?

21. What are the demerits of moving averages?

22. What are the merits of method of least squares?

23. Write the normal equations used in method of least squares?

24. Define forecasting.

25. What are the three types of forecasting?

26. What is a short-term forecast?

12th_Statistics_EM_Unit_7.indd 202 12-12-2021 21:50:21

III. Give short answers to the following questions

27. Write the uses of time series.
28. Explain semi-averages method
29. Write the merits of moving averages.
30. What is cyclical variation?
31. What is seasonal variation?
32. What are medium-term and long-term forecasts?
33. Describe the method of finding seasonal indices.
34. With what characteristic component of a time series should each of the following be associated.
    (i) An upturn in business activity
    (ii) Fire loss in a factory
    (iii) General increase in the sale of Television sets.

35. The number of units of a product exported during 1990-97 is given below. Draw the trend line using graphical method.

Year	1990	1991	1992	1993	1994	1995	1996	1997
No. of units exported (in '000)	12	13	13	16	19	23	21	23

36. Draw a time series graph relating to the following data and show the trend by free hand method

Year	1996	1997	1998	1999	2000	2001	2002	2003
Production in Million tonnes	40	44	42	48	51	54	50	56

37. Draw a trend line by the method semi averages

Year	1992	1993	1994	1995	1996	1997
Production of steel in Million tonnes	21	23	25	23	26	25

38. Yield of ground nut in Kharif season in India for the years 2003-04 to 2009-10 are given below. Calculate 3-year moving averages.

Year	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10
Yield (kg/hectare)	1320	909	1097	689	1386	1063	835

39. What do you understand by seasonal variations?

IV. Give detailed answers to the following questions

40. Explain the method of least squares

41. The following data states the number of ATM centers during 1995 to2001.

Year	1995	1996	1997	1998	1999	2000	2001
Number of ATM centres	50	63	75	100	109	120	135

Obtain the trend values using semi averages method

42. From the following data estimate the trend values using semi averages method

Year	2003	2004	2005	2006	2007	2008	2009	2010
Consumption of cotton (Thousands of bales)	677	696	747	755	766	777	785	836

43. Following data gives the yield of food grains in India for the years 2000-01 to 2009-10. Find the trend values using 4 year moving averages.

Year	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10
Yield (kg/ hectare)	1626	1734	1535	1727	1652	1715	1756	1860	1909	1798

44. Estimate the value of production for the year 1995 by using the method of least squares from the following data.

Year	1990	1991	1992	1993	1994
Production (1000s tons)	70	72	88	90	92

45. Find the following for the calculation of number of telephones for the year 2000.
    (1) Fit a straight line trend by the method of least squares.
    (2) Calculate the trend values.

Year	1990	1991	1992	1993	1994	1995
No. of telephones (in '00s)	20	21	23	25	27	29

46. The following data describes the export quantity of a company.

Year	1995	1996	1997	1998	1999	2000	2001
Export (in millions)	12	13	13	16	16	19	23

Fit a straight line trend and estimate the export for the year 2005.

12th_Statistics_EM_Unit_7.indd 204 12-12-2021 21:50:21

47. Calculate seasonal indices for the rainfall data of Tamil Nadu by using simple average method.

Year \ Quarter	I	II	III	IV
2000-01	314.5	335.6	16.8	118.4
2001-02	260.0	379.4	70.0	85.8
2002-03	185.4	407.1	8.7	129.8
2003-04	336.5	403.1	12.0	283.4
2004-05	360.7	472.1	14.3	231.7

48. Find seasonal Indices for the rainfall data in Tamil Nadu (in mm)

Year \ Quarter	2009	2010	2011	2012
I	38.2	38.5	55	50.5
II	166.8	250.9	277.7	197
III	612.6	773.1	717.8	706.1
IV	72.2	153.1	65.8	101.1

49. The following table gives quarterly expenditure over a number of years. Obtain seasonal correction for the data

Season \ Year	2000	2001	2002	2003
I	78	84	92	100
II	62	64	70	81
III	56	61	63	72
IV	71	82	83	96

50. Find the trend values using semi averages method. The following table shows the area covered for cultivation of Ragi in Tamil Nadu (in '000 hectares)

Year	2003	2004	2005	2006	2007	2008	2009	2010
Area (in '000 hectares)	118	109	100	95	94	90	82	76

(Hint: Decreasing trend)

12th_Statistics_EM_Unit_7.indd 205 12-12-2021 21:50:22

ANSWERS

I	**1.** (b)	**2.** (c)	**3.** (d)	**4.** (a)	**5.** (d)
	6. (d)	**7.** (c)	**8.** (d)	**9.** (c)	**10.** (b)
	11. (a)	**12.** (b)	**13.** (b)	**14.** (b)	**15.** (c)

III **34.** i). Cyclic variation ii). Irregular variation iii). Secular trend

37. 22.44, 23, 23.56, 24.11, 24.67, 25.22

38. –, 108.67, 898.33, 1057.33, 1046, 1094.67, –

IV **41.** 48.00, 62.67, 77.33, 92, 106.67, 121.33, 136

42. 691.66, 709.72, 727.78, 745.84, 763.91, 781.97, 800.03, 818.09

43. –, –, 1658.75, 1659.625, 1684.875, 1729.125, 1777.875, 1820.375, –, –

44. 70, 76.2, 82.4, 88.6, 94.8, 101

45. 19.52, 21.38, 23.24, 25.1, 26.95, 28.81

Number of telephones in the year 2000 is 3810

46. 10.86, 12.57, 14.29, 16, 17.71, 19.42, 21.14

Export for the year 2005 is 28 millions

47. 132, 181, 11, 77

48. 17, 83, 263, 37

49. 117, 91, 83, 109

50. 113, 108, 103, 98, 93, 88, 83, 78

12th_Statistics_EM_Unit_7.indd 206 12-12-2021 21:50:22