DJS3E - TIME SERIES and OFFICIAL STATISTICS

Unit-I (Time Series)
Components of time series – Additive and multiplicative models - Resolving components of a time series-measuring trend: Graphic, semi-averages, moving average and least squares methods.

Unit -II (Time Series)
Seasonal variation- measuring seasonal variation: method of simple averages, ratio-to-trend method, ratio-to-moving average method and link relative method- Cyclical and Random fluctuations- variate difference method.

Unit -III (Index Numbers)
Index numbers and their definitions - construction and uses of fixed and chain based index numbers - simple and weighted index numbers - Laspeyre's, Paasche's, Fisher's, and Marshall - Edgeworth index numbers – optimum tests for index numbers - Cost of living index numbers.

Unit -IV (Demographic Methods)
Demographic data – definition, sources and surveys –registration method. Fertility measurements – crude birth rate – general, specific, total fertility rates - gross and net reproduction rates. Mortality measurements – crude death rate – specific, standardized death rates – infant mortality rate – maternal mortality rate. Construction of Life table.

Unit -V (Official Statistics)
Present official statistics system in India – Ministry of statistics – NSSO, CSO and their functions - Registration of vital events – National Income Statistics – Agricultural Statistics – Industrial Statistics in India – Trade Statistics in India – Labour Statistics in India – Financial Statistics in India.

REFERENCE BOOKS:

1. Goon, A.M., M. K. Gupta and B. Das Gupta (2005) Fundamentals of Statistics- Vol. I, World press Ltd, Kolkata.
2. Gupta, S.C. and V.K. Kapoor (2007) Fundamentals of Applied Statistics, Sultan Chand & Sons, New Delhi.
3. Guilford, J. P. (1986) Fundamental Statistics in Psychology and Education, McGraw-Hill Book Company, New Delhi.
4. Srivastava, S. C. and S. Srivastava (2003) Fundamentals of Statistics, Anmol Publications Pvt. Ltd., New Delhi.

UNIT-1

TIME SERIES

1.1. Meaning

An arrangement of statistical data in accordance with time of occurrence or in a chronological order is called a time series. The numerical data which we get at different points of time-the set of observations-is known as time series.

In time series analysis, current data in a series may be compared with past data in the same series. We may also compare the development of two or more series over time. These comparisons may afford important guide lines for the individual firm. In Economics, statistics and commerce it plays an important role.

1.2 Definition and Examples

A time series is a set of observations made at specified times and arranged in a chronological order.

For example, if we observe agricultural production, sales, National Income etc., over a period of time, say over the last 3 or 5 years, the set of observations is called time series. Thus a time series is a set of time, quantitative readings of some various recorded at equal intervals of time. The interval may be an hour, a day, a week, a month, or a calendar year. Hourly temperature reading, daily sales in a shop, weekly sales in a shop, weekly sales in a market, monthly production in an industry, yearly agricultural production, population growth in ten years, are examples of time series.

From the comparison of past data with current data, we may seek to establish what development may be expected in future. The analysis of time series is done mainly for the purpose of forecasts and for evaluating the past performances. The chronological variations will be object of our study in time series analysis.

The essential requirements of a time series are:

- The time gap, between various values must be as far as possible, equal.

- It must consist of a homogeneous set of values.
- Data must be available for a long period.

symbolically if ‘t’ stands for time and ‘y_t’ represents the value at time t then the paired values (t, yt) represents a time series data.

Ex 1: Production of rice in Tamilnadu for the period from2010-11 to 2016-17.

Table 1.1. Production of rice in Tamilnadu (in ‘000 metric tons)

Year	**Production**
2010-11	400
2011-12	450
2012-13	440
2013-14	420
2014-15	460
2016-17	520

1.3 Uses of Time Series

The analysis of time series is of great significance not only to the economists and business man but also to the scientists, astronomists, geologists, sociologists, biologists, research worker etc. In the view of following reasons

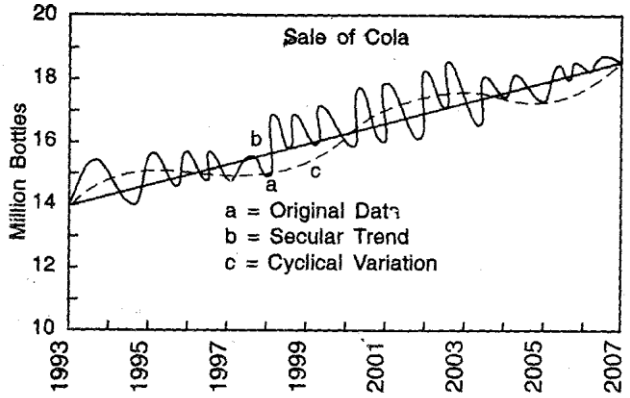
- It helps in understanding past behavior.
- It helps in planning future operations.
- It helps in evaluating current accomplishments.
- It facilitates comparison.

1.4 Components of Time Series

The values of a time series may be affected by the number of movements or fluctuations, which are its characteristics. The types of movements characterizing a time series are called components of time series or elements of a time series.

These are four types

- Secular Trend
- Seasonal Variations
- Cyclical Variations
- Irregular Variations



Secular Trend

Secular Trend is also called long term trend or simply trend. The secular trend refers to the general tendency of data to grow or decline over a long period of time. For example the population of India over years shows a definite rising tendency. The death rate in the country after independence shows a falling tendency because of advancement of literacy and medical facilities. Here long period of time does not mean as several years. Whether a particular period can be regarded as long period or not in the study of secular trend depends upon the nature of data. For example if we are studying the figures of sales of cloth store for 1996-1997 and we find that in 1997 the sales have gone up, this increase cannot be called as secular trend because it is too short period of time to conclude that the sales are showing the increasing tendency.

On the other hand, if we put strong germicide into a bacterial culture, and count the number of organisms still alive after each 10 seconds for 5 minutes, those 30 observations showing a general pattern would be called secular movement.

Mathematically the secular trend may be classified into two types

1. Linear Trend
2. Curvi-Linear Trend or Non-Linear Trend.

If one plots the trend values for the time series on a graph paper and if it gives a straight line then it is called a linear trend i.e. in linear trend the rate of change is constant where as in non-linear trend there is varying rate of change.

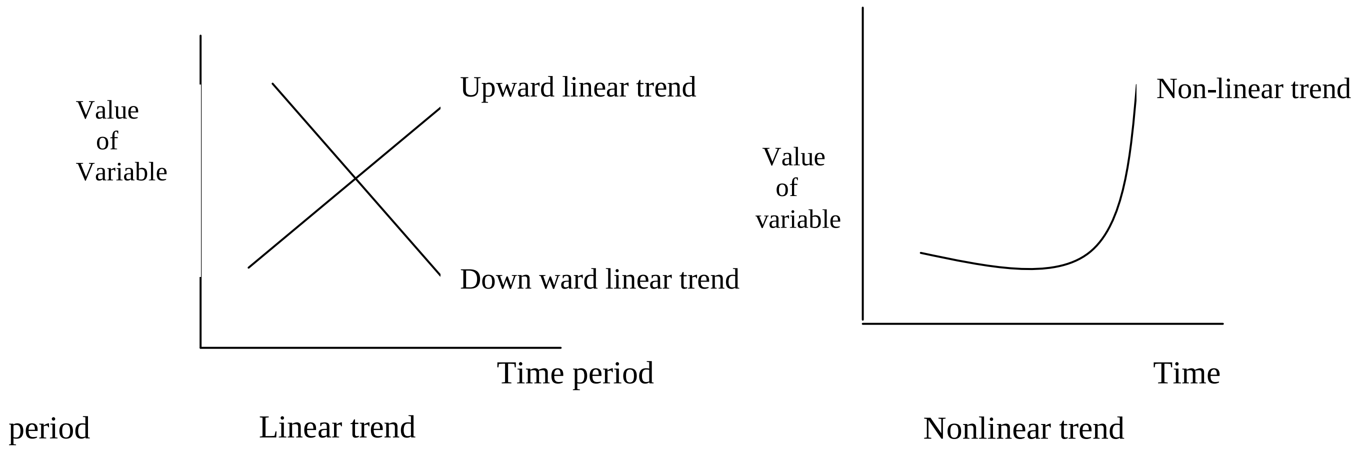


Figure 1.1. Linear Trend and Nonlinear Trend

Seasonal Variations

Seasonal variations occur in the time series due to the rhythmic forces which occurs in a regular and a periodic manner with in a period of less than one year. Seasonal variations occur during a period of one year and have the same pattern year after year. Here the period of time may be monthly, weekly or hourly. But if the figure is given in yearly terms then seasonal fluctuations does not exist. There occur seasonal fluctuations in a time series due to two factors.

- Due to natural forces
- Manmade convention.

The most important factor causing seasonal variations is the climate changes in the climate and weather conditions such as rain fall, humidity, heat etc. act on different products and industries differently. For example during winter there is greater demand for woolen clothes, hot drinks etc. Where as in summer cotton clothes, cold drinks have a greater sale and in rainy season umbrellas and rain coats have greater demand.

Though nature is primarily responsible for seasonal variation in time series, customs, traditions and habits also have their impact. For example on occasions like dipawali, dusserah, Christmas etc. there is a big demand for sweets and clothes etc., there is a large demand for books and stationary in the first few months of the opening of schools and colleges.

Cyclical Variations or Oscillatory Variation

This is a short term variation occurs for a period of more than one year. The rhythmic movements in a time series with a period of oscillation(repeated again and again in same manner) more than one year is called a cyclical variation and the period is called a cycle. The time series related to business and economics show some kind of cyclical variations.

One of the best examples for cyclical variations is 'Business Cycle'. In this cycle there are four well defined periods or phases.

- Boom
- Decline
- Depression
- Improvement

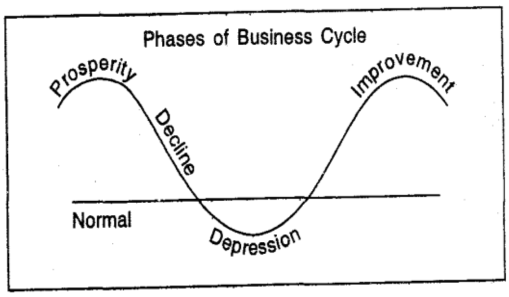


Figure 2. Phases of Business Cycle

Irregular Variation

It is also called Erratic, Accidental or Random Variations. The three variations trend, seasonal and cyclical variations are called as regular variations, but almost all the time series including the regular variation contain another variation called as random variation. This type of fluctuations occurs in random way or irregular ways which are unforeseen, unpredictable and due to some irregular circumstances which are beyond the control of human being such as earth quakes, wars, floods, famines, lockouts, etc. These factors affect the time series in the irregular ways. These irregular variations are not so significant like other fluctuations.

1.5 Mathematical Model

In classical analysis, it is assumed that some type of relationship exists among the four components of time series. Analysis of time series requires decomposition of a series, to decompose a series we must assume that some type of relationship exists among the four components contained in it.

The value Y_t of a time series at any time t can be expressed as the combinations of factors that can be attributed to the various components. These combinations are called as models and these are two types.

- Additive model
- Multiplicative model

Additive model

In additive model $Y_t = T_t + S_t + C_t + R_t$

Where T_t = Trend value at time t

S_t = Seasonal component

C_t = Cyclical component

R_t = Irregular component

But if the data is in the yearly form then seasonal variation does not exist, so in that situation $Y_t = T_t + C_t + R_t$

Generally the cyclical fluctuations have positive or negative value according to whether it is in above or below the normal phase of cycle.

Multiplicative model:

In multiplicative model $Y_t = T_t . S_t . C_t . R_t$.

The multiplicative model can be put in additive model by taking log both sides. However most business analysis uses the multiplicative model and finds it more appropriate to analyze business situations.

According to this model, the simple

One of the most important tasks before economists and businessmen these days is to make estimates for the future. For example, a businessman is interested in finding out his likely sales in the year 2016 or as a long-term planning in 2025 or the year 2030 so that he could adjust his production accordingly and avoid the possibility of either unsold stocks or inadequate production to meet the demand. Similarly, an economist is interested in estimating the likely population in the coming year so that proper planning can be carried out with regard to food supply, jobs for the people, etc. However, the first step in making estimates for the future consists of gathering information from the past. In this connection one usually deals with statistical data which are collected, observed or recorded at successive intervals of time. Such data are generally referred to as 'time series'. Thus when we observe numerical data at different points of time the set of observations is known as time series. For example if we observe production, sales, population, imports, exports, etc. at different points of time, say, over the last 5 or 10 years, the set of observations formed shall constitute time series. Hence, in the analysis of time series, time is the most important factor because the variable is related to time which may be either year, month, week, day and hour or even- minutes or seconds.

1.6 Measurement of Secular trend:

Secular trend is a long term movement in a time series. This component represents basic tendency of the series. The following methods are generally used to determine trend in any given time series. The following methods are generally used to determine trend in any given time series.

- Graphic method or eye inspection method
- Semi average method
- Method of moving average
- Method of least squares

Graphic method or eye inspection method

Graphic method is the simplest of all methods and easy to understand. The method is as follows. First plot the given time series data on a graph. Then a smooth free hand curve is drawn through the plotted points in such a way that it represents general tendency of the series. As the curve is drawn through eye inspection, this is also called as eye-inspection method. The graphic method removes the short term variations to show the basic tendency of the data. The trend line drawn through the graphic method can be extended further to predict or estimate values for the future time periods. As the method is subjective the prediction may not be reliable.

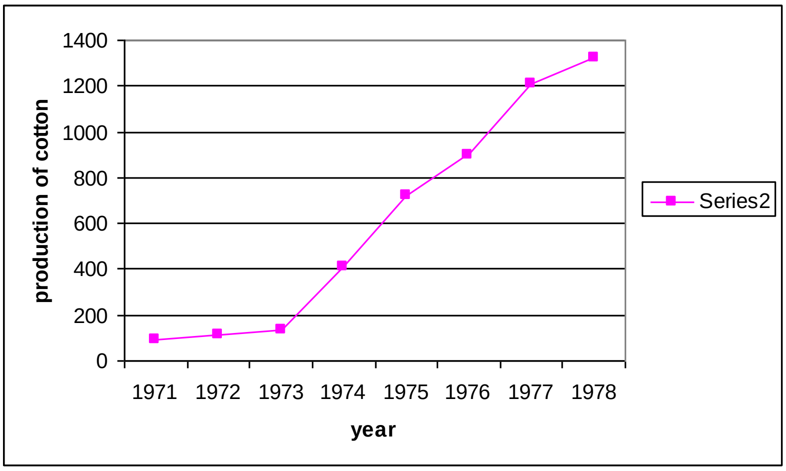


Figure 1.3. Graphic method for thee production of cotton base on year

Advantages

- It is very simplest method for study trend values and easy to draw trend.
- Sometimes the trend line drawn by the statistician experienced in computing trend may be considered better than a trend line fitted by the use of a mathematical formula.
- Although the free hand curves method is not recommended for beginners, it has considerable merits in the hands of experienced statisticians and widely used in applied situations.

Disadvantages:

- This method is highly subjective and curve varies from person to person who draws it.
- The work must be handled by skilled and experienced people.
- Since the method is subjective, the prediction may not be reliable.
- While drawing a trend line through this method a careful job has to be done.

Method of Semi Averages:

In this method the whole data is divided in two equal parts with respect to time. For example if we are given data from 1999 to 2016 i.e. over a period of 18 years the two equal parts will be first nine years i.e. from 1999 to 2007 and 2008 to 2016. In case of odd number of years like 9, 13, 17 etc. two equal parts can be made simply by omitting the middle year. For example if the data are given for 19 years from 1998 to 2016 the two equal parts would be from 1998 to 2006 and from 2008 to 2016, the middle year 2007 will be omitted. After the data have been divided into two parts, an average (arithmetic mean) of each part is obtained. We thus get two points. Each point is plotted against the mid year of the each part. Then these two points are joined by a straight line which gives us the trend line. The line can be extended downwards or upwards to get intermediate values or to predict future values.

Example:

Year	**Production**	**Semi averages**
2001	40	$\frac{40+45+40+42}{4} = 41.75$
2002	45	
2003	40	
2004	42	
2005	46	$\frac{46+52+56+61}{4} = 53.75$
2006	52	
2007	56	
2008	61	

Thus we get two points 41.75 and 53.75 which shall be plotted corresponding to their middle years i.e. 2002.5 and 2006.5. By joining these points we shall obtain the required trend line. This line can be extended and can be used either for prediction or for determining intermediate values.

Advantages:

- This method is simple to understand as compare to moving average method and method of least squares.
- This is an objective method of measuring trend as everyone who applies this method is bound to get the same result.

Disadvantages:

- The method assumes straight line relationship between the plotted points regardless of the fact whether that relationship exists or not.

- The main drawback of this method is if we add some more data to the original data then whole calculation is to be done again for the new data to get the trend values and the trend line also changes.
- As the Arithmetic Mean of each half is calculated, an extreme value in any half will greatly affect the points and hence trend calculated through these points may not be precise enough for forecasting the future.

Method of Moving Average:

It is a method for computing trend values in a time series which eliminates the short term and random fluctuations from the time series by means of moving average. Moving average of a period m is a series of successive arithmetic means of m terms at a time starting with 1^{st}, 2^{nd}, 3^{rd} and so on. The first average is the mean of first m terms; the second average is the mean of 2^{nd} term to $(m+1)^{th}$ term and 3^{rd} average is the mean of 3^{rd} term to $(m+2)^{th}$ term and so on.

If m is odd then the moving average is placed against the mid value of the time interval it covers. But if m is even then the moving average lies between the two middle periods which does not correspond to any time period. So further steps has to be taken to place the moving average to a particular period of time. For that we take 2-yearly moving average of the moving averages which correspond to a particular time period. The resultant moving averages are the trend values.

Ex:1 Calculate 3-yearly moving average for the following data.

Years	Production	3-yearly moving avg (trend values)
2001-02	40	
2002-03 →	45 →	(40+45+40)/3 = 41.67
2003-04 →	40 →	(45+40+42)/3 = 42.33
2004-05 →	42 →	(40+42+46)/3 = 42.67
2005-06 →	46 →	(42+46+52)/3 = 46.67
2006-07 →	52 →	(46+52+56)/3 = 51.33
2007-08 →	56 →	(52+56+61)/3 = 56.33
2008-09	61	

Ex :2 Calculate 4-yearly moving average for the following data.

Years	Production	4-yearly moving avg	2-yealry moving avg (trend values)
2001-02	40		
2002-03	45		
		→ (40+45+40+42)/3 = 41.75	
2003-04	40		→ 42.5
		→ (45+40+42+46)/3 = 43.15	
2004-05	42		→ 44.12
		→ (40+42+46+52)/3 = 45	
2005-06	46		→ 47
		→ (42+46+52+56)/3 = 49	
2006-07	52		→ 51.38
		→ (46+52+56+61)/3 = 53.75	
2007-08	56		
2008-09	61		

Advantages:

- This method is simple to under stand and easy to execute.
- It has the flexibility in application in the sense that if we add data for a few more time periods to the original data, the previous calculations are not affected and we get a few more trend values.
- It gives a correct picture of the long term trend if the trend is linear.
- If the period of moving average coincides with the period of oscillation (cycle), the periodic fluctuations are eliminated.
- The moving average has the advantage that it follows the general movements of the data and that its shape is determined by the data rather than the statistician's choice of mathematical function.

Disadvantages:

- For a moving average of 2m+1, one does not get trend values for first m and last m periods.

- As the trend path does not correspond to any mathematical; function, it cannot be used for forecasting or predicting values for future periods.
- If the trend is not linear, the trend values calculated through moving averages may not show the true tendency of data.
- The choice of the period is sometimes left to the human judgment and hence may carry the effect of human bias.

Method of Least Squares:

This method is most widely used in practice. It is mathematical method and with its help a trend line is fitted to the data in such a manner that the following two conditions are satisfied.

1. $\sum(Y - Y_c) = 0$ i.e. the sum of the deviations of the actual values of Y and the computed values of Y is zero.
2. $\sum(Y - Y_c)^2$ is least, i.e. the sum of the squares of the deviations of the actual values and the computed values is least.

The line obtained by this method is called as the "line of best fit".

This method of least squares may be used either to fit a straight line trend or a parabolic trend.

Fitting of a straight line trend by the method of least squares:

Let Y_t be the value of the time series at time t. Thus Y_t is the independent variable depending on t.

Assume a straight line trend to be of the form $Y_{tc} = a + bt$ ………… (1)

Where Y_{tc} is used to designate the trend values to distinguish from the actual Y_t values, a is the Y-intercept and b is the slope of the trend line.

Now the values of a and b to be estimated from the given time series data by the method of least squares.

In this method we have to find out a and b values such that the sum of the squares of the deviations of the actual values Y_t and the computed values Y_{tc} is least.

i.e. $S = \sum (Y_t - Y_{tc})^2$ should be least

i.e. $S = \sum (Y_t - a - bt)^2$ ………… (2) Should be least

Now differentiating partially (2) w.r.to a and equating to zero we get

$$\frac{\partial S}{\partial a} = 2\sum (Y_t - a - bt)(-1) = 0$$

$$\Rightarrow \sum (Y_t - a - bt) = 0$$
$$\Rightarrow \sum Y_t = \sum a + b\sum t$$

$$\Rightarrow \sum Y_t = na + b\sum t \quad \text{………………….. (3)}$$

Now differentiating partially (2) w.r.to b and equating to zero we get

$$\frac{\partial S}{\partial b} = 2\sum (Y_t - a - bt)(-t) = 0$$

$$\Rightarrow \sum t(Y_t - a - bt) = 0$$
$$\Rightarrow \sum tY_t = a\sum t + b\sum t^2 \quad \text{………………….. (4)}$$

The equations (3) and (4) are called 'normal equations'

Solving these two equations we get the values of a and b say $\hat{a}$ and $\hat{b}$.

Now putting these two values in the equation (1) we get

$$Y_{tc} = \hat{a} + \hat{b}t$$

which is the required straight line trend equation.

Note: The method for assessing the appropriateness of the straight line modal is the method of first differences. If the differences between successive observations of a series are constant (nearly constant) the straight line should be taken to be an appropriate representation of the trend component.

Illustration 10. Below are given the figures of production (in thousand quintals) of a sugar factory :

Year	2001	2002	2003	2004	2005	2006	2007
Production (in '000 qtls.)	80	90	92	83	94	99	92

(*i*) Fit a straight line trend to these figures.
(*ii*) Plot these figures on a graph and show the trend line.

(M. Com., Jiwaji Univ.; M. Com., Ajmer Univ.; B. Com., HPU; B.Com., Bangalore Univ.)

Solution. (*i*) FITTING THE STRAIGHT LINE TREND

Year	*Production ('000 qtls.) Y*	*X*	*X Y*	X^2	*Trend values* Y_c
2001	80	−3	−240	9	84
2002	90	−2	−180	4	86
2003	92	−1	−92	1	88
2004	83	0	0	0	90
2005	94	+1	+94	1	92
2006	99	+2	+198	4	94
2007	92	+3	+276	9	96
$N = 7$	$\Sigma Y = 630$	$\Sigma X = 0$	$\Sigma XY = 56$	$\Sigma X^2 = 28$	$\Sigma Y_c = 630$

The equation of the straight line is $Y_c = a + bX$.

To find a and b we have two normal equations

$$\sum Y = na + b\sum X$$

$$\sum XY = a\sum X + b\sum X^2$$

Since $\Sigma X = 0$; $a = \frac{\Sigma Y}{N}$, $b = \frac{\Sigma XY}{\Sigma X^2}$

$\Sigma Y = 630$, $N = 7$, $\Sigma XY = 56$, $\Sigma X^2 = 28$,

∴ $a = \frac{630}{7} = 90$; and $b = \frac{56}{28} = 2$

Hence the equation of the straight line trend is $Y_c = 90 + 2X$.

Origin, 2004 : *X* units, one year; *Y* units, production in thousand quintals.

For $X = -3$, $Y_c = 90 + 2(-3) = 84$

For $X = -2$, $y_c = 90 + 2(-2) = 86$

For $X = -1$, $y_c = 90 + 2(-1) = 88$.

Similarly, by putting $X = 0$, 1, 2, 3, we can obtain other trend values. However, since the value of *b* is constant, first trend value need be obtained and then if the value *b* is positive we may continue adding the value of *b* to every preceding value. For 2002 it will be 84 + 2 = 86, for 2003 it will be 86 + 2 = 88, and so on. If *b* is negative, then instead of adding we will deduct.

(*ii*) The graph of the above data is given below :

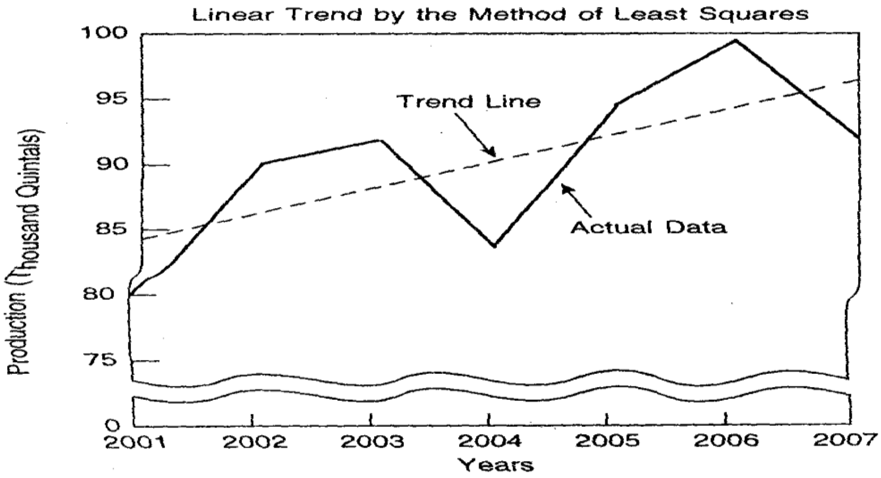


Illustration 32. Calculate trend values by the method of least-squares from the data given below :

Year	2000	2001	2002	2003	2004	2005	2006	2007
Sales	80	90	92	83	94	99	92	104

Plot the data showing also the trend line. *(B.Com.. HPU; M.A. Econ.. GNDU)*

Solution. FITTING STRAIGHT LINE TREND BY METHOD OF LEAST SQUARES

Years	*Sales Y*	*Deviations from 2003.5*	*Deviations multiplied by 2 X*	*XY*	X^2	Y_c
2000	80	− 3.5	− 7	− 560	49	83.0
2001	90	− 2.5	− 5	− 450	25	85.5
2002	92	− 1.5	− 3	− 276	9	88.0
2003	83	− .5	− 1	− 83	1	90.5
2004	94	+ .5	+ 1	+ 94	1	93.0
2005	99	+ 1.5	+ 3	+ 297	9	95.5
2006	92	+ 2.5	+ 5	+ 460	25	98.0
2007	104	+ 3.5	+ 7	+ 728	49	100.5
$N = 8$	$\Sigma y = 734$			$\Sigma XY = 210$	$\Sigma X^2 = 168$	$\Sigma Y_c = 734$

The equation of the straight line is $Y_c = a + bX$.

To find a and b we have two normal equations

$$\sum Y = na + b\sum X$$

$$\sum XY = a\sum X + b\sum X^2$$

Since $\Sigma X = 0,\ a = \frac{\Sigma Y}{N} = \frac{734}{8} = 91.75,\ b = \frac{\Sigma XY}{\Sigma X^2} = \frac{210}{168} = 1.25$

The required line equation is $Y = 91.75 + 1.25X$

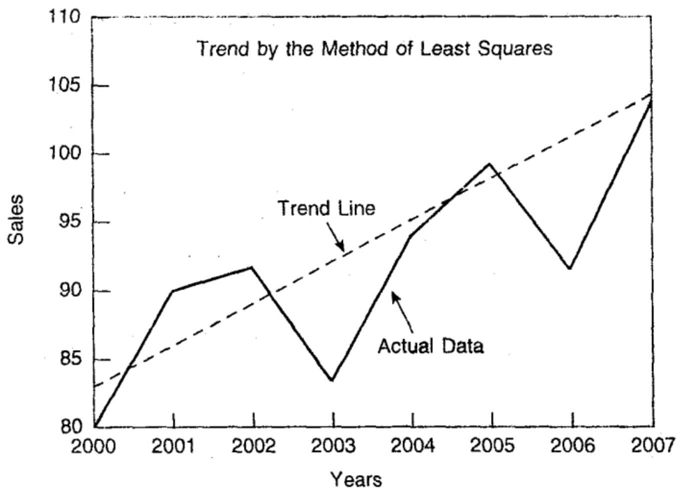
The trend values for various years are

$$Y_{2000} = 91.75 + 1.25\,(-7) = 91.75 - 8.75 = 83$$

For finding these trend values, double the value of b, *i.e.*, $1.25 \times 2 = 2.5$ and add to the preceding value :

$$Y_{2001} = 83 + 2.5 = 85.5$$

and so on



Fitting of a parabolic trend by the method of least squares

Let Y_t be the value of the time series at time t. Thus Y_t is the independent variable depending on t.

Assume a parabolic trend to be of the form $Y_{tc} = a + bt + ct^2$ ………… (1)

Now the values of a, b and c to be estimated from the given time series data by the method of least squares.

In this method we have to find out a, b and c values such that the sum of the squares of the deviations of the actual values Y_t and the computed values Y_{tc} is least.

i.e. $S = \sum (Y_t - Y_{tc})^2$ should be least

i.e. $S = \sum (Y_t - a - bt)^2$ ………… (2) Should be least

Now differentiating partially (2) w.r.to a and equating to zero we get

$$\frac{\partial S}{\partial a} = 2\sum (Y_t - a - bt - ct^2)(-1) = 0$$

$$\Rightarrow \sum (Y_t - a - bt - ct^2) = 0$$

$$\Rightarrow \sum Y_t = \sum a + b\sum t + c\sum t^2$$

$$\Rightarrow \sum Y_t = na + b\sum t + c\sum t^2 \quad \text{.................. (3)}$$

Now differentiating partially (2) w.r.to b and equating to zero we get

$$\frac{\partial S}{\partial b} = 2\sum (Y_t - a - bt - ct^2)(-t) = 0$$

$$\Rightarrow \sum t(Y_t - a - bt - ct^2) = 0$$

.... (4)

$$\Rightarrow \sum tY_t = a\sum t + b\sum t^2 + c\sum t^3$$

Now differentiating partially (2) w.r.to c and equating to zero we get

$$\frac{\partial S}{\partial c} = 2\sum (Y_t - a - bt - ct^2)(-t^2) = 0$$

$$\Rightarrow \sum t^2(Y_t - a - bt - ct^2) = 0$$

$$\Rightarrow \sum t^2 Y_t = a\sum t^2 + b\sum t^3 + c\sum t^4 \quad \text{.................. (5)}$$

The equations (3), (4) and (5) are called ‘normal equations’

Solving these three equations we get the values of a, b and c say $\hat{a}, \hat{b}$ and $\hat{c}$.

Now putting these three values in the equation (1) we get

$$Y_{tc} = \hat{a} + \hat{b}t + \hat{c}t^2$$

Which is the required parabolic trend equation

Note: The method for assessing the appropriateness of the second degree equation is the method of second differences. If the differences are taken of the first differences and the results are constant (nearly constant) the second degree equation be taken to be an appropriate representation of the trend component.

Illustration 14. The prices of a commodity during 2002-2007 are given below. Fit a parabola $Y = a + bX + cX^2$ to these data. Estimate the price of the commodity for the year 2008 :

Year	*Prices*	*Year*	*Prices*
2002	100	2005	140
2003	107	2006	181
2004	128	2007	192

Also plot the actual and trend values on the graph. *(B.Com. (H), DU; M. Com., M.D. Univ.)*

Solution : To determine the values of *a*, *b* and *c*, we solve the following normal equations :

$$\Sigma Y = Na + b\Sigma X + c\Sigma X^2 \quad ...(i)$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3 \quad ...(ii)$$

$$\Sigma X^2Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4 \quad ...(iii)$$

Year	*Prices (Rs.) Y*	X	X^2	X^3	X^4	XY	X^2Y	*Trend Values* (Y_C)
2002	100	−2	4	−8	16	−200	400	97.717
2003	107	−1	1	−1	1	−107	107	110.401
2004	128	0	0	0	0	0	0	126.657
2005	140	+1	1	+1	1	+140	140	146.485
2006	181	+2	4	+8	16	+362	724	169.885
2007	192	+3	9	+27	81	+576	1728	196.857
$N = 6$	$\Sigma Y = 848$	$\Sigma X = 3$	$\Sigma X^2 = 19$	$\Sigma X^3 = 27$	$\Sigma X^4 = 115$	$\Sigma XY = 771$	$\Sigma X^2Y = 3{,}099$	$\Sigma Y_C = 848.002$

$$848 = 6a + 3b + 19c \quad ...(i)$$

$$771 = 3a + 19b + 27c \quad ...(ii)$$

$$3{,}099 = 19a + 27b + 115c \quad ...(iii)$$

Multiplying the second equation by 2 and keeping the first as it is, we get .

$$848 = 6a + 3b + 19c$$

$$1{,}542 = 6a + 38b + 54c$$

$$-694 = -35b - 35c \quad ...(iv)$$

or $35b + 35c = 694$

Multiplying Eqn. (*ii*) by 19 and Eqn. (*iii*) by 3, we get

$$14{,}649 = 57a + 361b + 513c$$

$$9{,}297 = 57a + 81b + 345c$$

$$5{,}352 = 280b + 168c \quad ...(v)$$

Multiplying equation (iv) by 8, we have

$$280b + 280c = 5{,}552$$

Solving equations (*iv*) and (*v*)

$$280b + 280c = 5{,}552$$

$$280b + 168c = 5{,}352$$

$$112c = 200 \quad \text{or} \quad c = 1.786$$

Substituting the value of *c* in Eqn. (*iv*),

$$35b + (35 \times 1.786) = 694$$

$$35b = 694 - 62.5 = 631.5 \text{ or } b = 18.042$$

$$848 = 6a + 3(18.042) + 19(1.786) = 6a + 54.126 + 33.934$$

$$6a = 759.94 \quad \text{or} \quad a = 126.657$$

Thus $a = 126.657$, $b = 18.042$ and $c = 1.786$

Substituting these values in the equation,

$$Y = 126.657 + 18.042X + 1.786X^2$$

when $X = -2$

$$Y = 126.657 + 18.042(-2) + 1.786(-2)^2$$

$$= 126.657 - 36.084 + 7.144 = 97.717$$

when $X = -1$

$$Y = 126.657 + 18.042(-1) + 1.786(-1)^2$$

$$= 126.657 - 18.042 + 1.786 = 110.401$$

when $X = 1$,

$$Y = 126.657 + 18.042 + 1.786 = 146.485$$

when $X = 2$,

$$Y = 126.657 + 18.042\ (2) + 1.786\ (2)^2 = 169.885$$

when $X = 3$,

$$Y = 126.657 + 18.042\ (3) + 1.786\ (3)^2 = 196.857$$

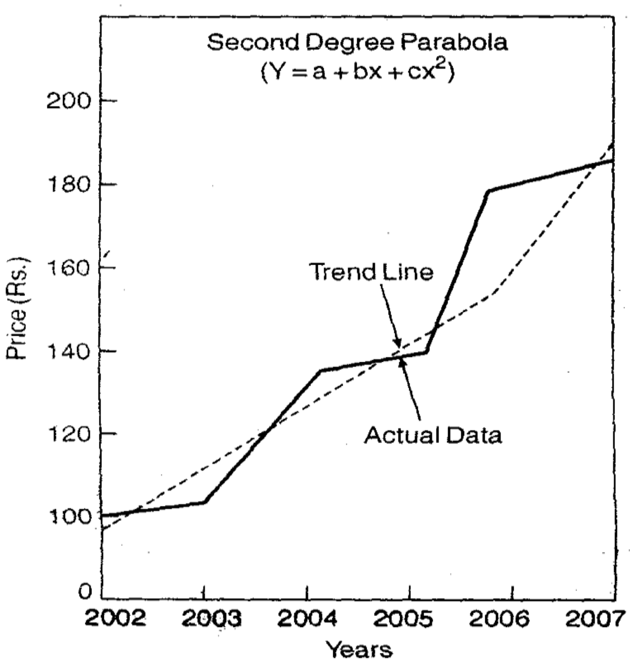
Price for the year 2008

For 2008 X would be equal to 4. Putting $X = 4$ in the equation,

$$Y = 126.657 + 18.042\ (4) + 1.786\ (4)^2$$
$$= 126.657 + 72.168 + 28.576 = 227.401.$$

Thus the likely price of the commodity for the year 2008 is Rs. 227.41 approx.

The graph of the actual and trend values is given below:



Advantages

- This is a mathematical method of measuring trend and as such there is no possibility of subjectiveness i.e. everyone who uses this method will get same trend line.
- The line obtained by this method is called the line of best fit.
- Trend values can be obtained for all the given time periods in the series.

Disadvantages

- Great care should be exercised in selecting the type of trend curve to be fitted i.e. linear, parabolic or some other type. Carelessness in this respect may lead to wrong results.
- The method is more tedious and time consuming.

- Predictions are based on only long term variations i.e trend and the impact of cyclical, seasonal and irregular variations is ignored.
- This method can not be used to fit the growth curves like Gompertz curve $\left(Y = K a^{b^X}\right)$, logistic curve $\left(Y = \frac{1}{K + ab^X}\right)$ etc.

Question bank:

1) Define a time series. Discuss its main components.
2) Define secular trend of a time series and explain methods that are used in isolating it.
3) Explain the method of moving average for the determination of trend. What are the advantages and disadvantages of this method?
4) What are the advantages and disadvantages of the graphic method and least square method in trend analysis?
5) Explain briefly the method of moving averages for calculating the trend.
6) How does analysis of time series help business forecasting?
7) Distinguish between secular trend, seasonal variations and cyclical fluctuations. Discuss various methods of measuring each.
8) Explain briefly the additive and multiplicative models of time series. Which of these models is more popular in practice and why?
9) Explain how you would determine seasonal variation by 12-monthly moving average.
10) What are the various methods for determining trend in a time series?
11) Describe in detail the method of least squares for determining trend.
12) The production data on steel in a factory in the past 10 years are given below:

Year	: 1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Production (tonnes)	: 75	86	98	90	96	108	124	140	150	165

Fit a staright line trend and tabulate the trend values. What is the expected production in the year 1997 on the basis of trend? $(Y = 113.2 + 5.08X;\ Y = 169.98)$

13) Fit a staright line trend of the following data by least square method. Also find an estimate for the year 1997;

Year	: 1989	1990	1991	1992	1993	1994	1995	1996
Production (tonnes)	: 12	13	13	16	19	23	21	23

$(Y = 17.5 + .893X;\ Y_{1997} = 25.54)$

14) Fit a staright line trend by the method of least square to the following data. Also find an estimate for the year 2000;

Year	: 1990	1991	1992	1993	1994	1995	1996	1997
Production (tonnes)	: 38	40	65	72	69	67	95	104

$(Y = 68.75 + 4.404X;\ )$

15) From the following data, calculate trend by 4 yearly moving average and find short-term oscillations:

Year	Production(tonnes)	Year	Production(tonnes)
1984	5	1990	9
1985	6	1991	10
1986	7	1992	9
1987	7	1993	10
1988	6	1994	11
1989	8	1995	11

UNIT-2

SEASONAL VARIATIONS

2.1 Introduction

Seasonal variations are regular and periodic variations having a period of one year duration. Some of the examples which show seasonal variations are production of cold drinks, which are high during summer months and low during winter season. Sales of sarees in a cloth store which are high during festival season and low during other periods.

The reason for determining seasonal variations in a time series is to isolate it and to study its effect on the size of the variable in the index form which is usually referred as seasonal index.

2.2 Measurement of seasonal variations:

The study of seasonal variation has great importance for business enterprises to plan the production schedule in an efficient way so as to enable them to supply to the public demands according to seasons.

There are different devices to measure the seasonal variations. These are

- Method of simple averages.
- Ratio to trend method
- Ratio to moving average method
- Link relative method.

Method of simple averages

This is the simplest of all the methods of measuring seasonality. This method is based on the additive modal of the time series. That is the observed values of the series is expressed by $Y_t = T_t + S_t + C_t + R_t$ and in this method we assume that the trend component and the cyclical component are absent.

The method consists of the following steps.

- Arrange the data by years and months (or quarters if quarterly data is given).

- Compute the average $\bar{x}_i$ (i = 1,2,…..12 for monthly and i=1,2,3,4 for quarterly) for the i th month or quarter for all the years.
- Compute the average $\bar{x}$ of the averages.
- i.e. $\bar{x} = \frac{1}{12}\sum_{i=1}^{12}\bar{x}_i$ for monthly and $\bar{x} = \frac{1}{4}\sum_{i=1}^{4}\bar{x}_i$ for quarterly
- Seasonal indices for different months (quarters) are obtained by expressing monthly (quarterly) averages as percentages of $\bar{x}$. Thus seasonal indices for i-th month (quarter) = $\frac{\bar{x}_i}{\bar{x}} \times 100$

Advantages and Disadvantages:

Method of simple average is easy and simple to execute.

This method is based on the basic assumption that the data do not contain any trend and cyclic components. Since most of the economic and business time series have trends and as such this method though simple is not of much practical utility.

Example: 1

Assuming that the trend is absent, determine if there is any seasonality in the data given below.

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
2004	3.7	4.1	3.3	3.5
2005	3.7	3.9	3.6	3.6
2006	4.0	4.1	3.3	3.1
2007	3.3	4.4	4.0	4.0

What are the seasonal indices for various quarters ? *(M. Com.. M.K. Univ.)*

Solution. COMPUTATION OF SEASONAL INDICES

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
2004	3.7	4.1	3.3	3.5
2005	3.7	3.9	3.6	3.6
2006	4.0	4.1	3.3	3.1
2007	3.3	4.4	4.0	4.0
Total	14.7	16.5	14.2	14.2
Average	3.675	4.125	3.55	3.55
Seasonal Index	98.66	110.74	95.30	95.30

Notes for calculating seasonal index

$$\text{The average of averages} = \frac{3.675 + 4.125 + 3.55 + 3.55}{4} = \frac{14.9}{4} = 3.725$$

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

$$\text{Seasonal Index for the first quarter} = \frac{3.675}{3.725} \times 100 = 98.66$$

$$\text{Seasonal Index for the second quarter} = \frac{4.125}{3.725} \times 100 = 110.74$$

$$\text{Seasonal Index for the third and fourth quarters} = \frac{3.55}{3.725} \times 100 = 95.30$$

Ratio to trend method:

This method is an improvement over the simple averages method and this method assumes a multiplicative model i.e

$$Y_t = T_t S_t C_t R_t$$

The measurement of seasonal indices by this method consists of the following steps.

- Obtain the trend values by the least square method by fitting a mathematical curve, either a straight line or second degree polynomial.
- Express the original data as the percentage of the trend values. Assuming the multiplicative model these percentages will contain the seasonal, cyclical and irregular components.
- The cyclical and irregular components are eliminated by averaging the percentages for different months (quarters) if the data are In monthly (quarterly), thus leaving us with indices of seasonal variations.
- Finally these indices obtained in step(3) are adjusted to a total of 1200 for monthly and 400 for quarterly data by multiplying them through out by a constant K which is

given by $K = \dfrac{1200}{Total\ of\ the\ indices}$ for monthly,

$K = \dfrac{400}{Total\ of\ the\ indices}$ for quarterly.

Advantages:

- It is easy to compute and easy to understand.
- Compared with the method of monthly averages this method is certainly a more logical procedure for measuring seasonal variations.
- It has an advantage over the ratio to moving average method that in this method we obtain ratio to trend values for each period for which data are available where as it is not possible in ratio to moving average method.

Disadvantages:

- The main defect of the ratio to trend method is that if there are cyclical swings in the series, the trend whether a straight line or a curve can never follow the actual data as closely as a 12- monthly moving average does. So a seasonal index computed by the ratio to moving average method may be less biased than the one calculated by the ratio to trend method.

Example 1:

Calculate seasonal indices by Ratio to moving average method from the following data.

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2003	30	40	36	34
2004	34	52	50	44
2005	40	58	54	48
2006	54	76	68	62
2007	80	92	86	82

Solution. For determining seasonal variation by ratio-to-trend method, first we will determine the trend for yearly data and then convert it to quarterly data.

CALCULATING TREND BY METHOD OF LEAST SQUARES

Year	*Yearly totals*	*Yearly average* *Y*	*Deviations from mid-year* *X*	*XY*	X^2	*Trend values*
2003	140	35	− 2	− 70	4	32
2004	180	45	− 1	− 45	1	44
2005	200	50	0	0	0	56
2006	260	65	+ 1	+ 65	1	68
2007	340	85	+ 2	+ 170	4	80
$N = 5$		$\Sigma Y = 280$		$\Sigma XY = 120$	$\Sigma X^2 = 10$	

The equation of the straight line trend is $Y = a + bX$.

$$a = \frac{\Sigma Y}{N} = \frac{280}{5} = 56 \qquad b = \frac{\Sigma XY}{\Sigma X^2} = \frac{120}{10} = 12$$

Quarterly increment = $\frac{12}{4} = 3$.

Calculation of Quarterly Trend Values. Consider 2003, trend value for the middle quarter, *i.e.*, half of 2nd and half of 3rd is 32. Quarterly increment is 3. So the trend value of 2nd quarter is $32 - \frac{3}{2}$, *i.e.*, 30.5 and for 3rd quarter is $32 + \frac{3}{2}$, *i.e.*, 33.5. Trend value for the 1st quarter is 30.5 − 3, *i.e.*, 27.5 and of 4th quarter is 33.5 + 3, *i.e.*, 36.5. We thus get quarterly trend values as shown below :

TREND VALUES

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2003	27.5	30.5	33.5	36.5
2004	39.5	42.5	45.5	48.5
2005	51.5	54.5	57.5	60.5
2006	63.5	66.5	69.5	72.5
2007	75.5	78.5	81.5	84.5

The given values are expressed as percentage of the corresponding trend values.

Thus for 1st Qtr. of 2003, the percentage shall be $(30/27.5) \times 100 = 109.09$, for 2nd Qtr. $(40/30.5) \times 100 = 131.15$, etc.

GIVEN QUARTERLY VALUES AS % OF TREND VALUES

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2003	109.09	131.15	107.46	93.15
2004	86.08	122.35	109.89	90.72
2005	77.67	106.42	93.91	79.34
2006	85.04	114.29	97.84	85.52
2007	105.96	117.20	105.52	97.04
Total	463.84	591.41	514.62	445.77
Average	92.77	118.28	102.92	89.15
S.I. Adjusted	92.05	117.36	102.12	88.46

Total of averages = 92.77 + 118.28 + 102.92 + 89.15 = 403.12.

Since the total is more than 400 an adjustment is made by multiplying each average by $\frac{400}{403.12}$ and final indices are obtained.

Ratio to moving average method:

The ratio to moving average method is also known as percentage of moving average method and is the most widely used method of measuring seasonal variations. The steps necessary for determining seasonal variations by this method are

- Calculate the centered 12-monthly moving average (or 4-quarterly moving average) of the given data. These moving averages values will eliminate S and I leaving us T and C components.
- Express the original data as percentages of the centered moving average values.
- The seasonal indices are now obtained by eliminating the irregular or random components by averaging these percentages using A.M or median.
- The sum of these indices will not in general be equal to 1200 (for monthly) or 400 (for quarterly). Finally the adjustment is done to make the sum of the indices to a total of 1200 for monthly and 400 for quarterly data by

multiplying them through out by a constant K which is given by

$$K = \frac{1200}{Total\ of\ the\ indices} \text{ for monthly}$$

$$K = \frac{400}{Total\ of\ the\ indices} \text{ for quarterly}$$

Advantages:

- Of all the methods of measuring seasonal variations, the ratio to moving average method is the most satisfactory, flexible and widely used method.
- The fluctuations of indices based on ratio to moving average method is less than based on other methods.

Disadvantages:

- This method does not completely utilize the data. For example in case of 12-monthly moving average seasonal indices cannot be obtained for the first 6 months and last 6 months.

Illustration 24. Calculate seasonal indices by the ratio to moving average method, from the following data :

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2005	68	62	61	63
2006	65	58	66	61
2007	68	63	63	67

Solution. CALCULATION OF SEASONAL INDICES BY 'RATIO TO MOVING AVERAGE' METHOD

Year	*Quarter*	*Given figures*	*4-figure moving totals*	*2-figure moving totals*	*4-figure moving average*	*Given figure as % of moving average*
2005	I	68				
	II	62				
			→ 254			
	III	61 →		505	63.186	96.54
			→ 251			
	IV	63 →		498	62.260	101.19
			→ 247			
2006	I	65 →		499	62.375	104.21
			→ 252			
	II	58 →		502	62.750	92.43
			→ 250			
	III	66 →		503	62.875	104.97
			→ 253			
	IV	61 →		511	63.875	95.50
			→ 258			
2007	I	68 →		513	64.125	106.04
			→ 255			
	II	63 →		516	64.500	97.67
			→ 261			
	III	63				
	IV	67				

CALCULATION OF SEASONAL INDEX

Year	Percentage to Moving Average			
	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
2005	—	—	96.63	101.20
2006	104.21	92.43	104.97	95.50
2007	106.04	97.67	—	—
Total	210.25	190.10	201.60	196.70
Average	105.125	95.05	100.80	98.35
Seasonal Index	105.30	95.21	100.97	98.52

Arithmetic average of averages $= \frac{399.32}{4} = 99.83$

By expressing each quarterly average as percentage of 99.83, we will obtain seasonal indices.

Seasonal index of 1st Quarter $= \frac{105.125}{99.83} \times 100 = 105.30$

Seasonal index of 2nd Quarter $= \frac{95.05}{99.83} \times 100 = 95.21$

Seasonal index of 3rd Quarter $= \frac{100.80}{99.83} \times 100 = 100.97$

Seasonal index of 4th Quarter $= \frac{98.35}{99.83} \times 100 = 98.52$

Link relative method:

This method is slightly more complicated than other methods. This method is also known as Pearson's method. This method consists in the following steps.

- The link relatives for each period are calculated by using the below formula

  $$Link\ relative\ for\ any\ period = \frac{Current\ periods\ figure}{Previous\ periods\ figure} \times 100$$

- Calculate the average of the link relatives for each period for all the years using mean or median.
- Convert the average link relatives into chain relatives on the basis of the first season. Chain relative for any period can be obtained by
- $$\frac{Avg\ link\ relative\ for\ that\ period \times Chain\ relative\ of\ the\ previous\ period}{100}$$

  the chain relative for the first period is assumed to be 100.
- Now the adjusted chain relatives are calculated by subtracting correction factor 'kd' from (k+1)th chain relative respectively.
- Where k = 1,2,…….11 for monthly and k = 1,2,3 for quarterly data.

  and $d = \frac{1}{N}\left[New\ chain\ relative\ for\ first\ period - 100\right]$ where N denotes the number of periods i.e. N = 12 for monthly N = 4 for quarterly
- Finally calculate the average of the corrected chain relatives and convert the corrected chain relatives as the percentages of this average. These percentages are seasonal indices calculated by the link relative method.

Advantages:

- As compared to the method of moving average the link relative method uses data more.

Disadvantages:

- The link relative method needs extensive calculations compared to other methods and is not as simple as the method of moving average.
- The average of link relatives contains both trend and cyclical components and these components are eliminated by applying correction.

Illustration 26. Apply the method of link relatives to the following data and calculate seasonal indices :

QUARTERLY FIGURES

Quarter	*2003*	*2004*	*2005*	*2006*	*2007*
I	6.0	5.4	6.8	7.2	6.6
II	6.5	7.9	6.5	5.8	7.3
III	7.8	8.4	9.3	7.5	8.0
IV	8.7	7.3	6.4	8.5	7.1

Solution. CALCULATION OF SEASONAL INDICES BY THE METHOD OF LINK RELATIVES

	Quarter			
Year	*I*	*II*	*III*	*IV*
2003	—	108.3	120.0	111.5
2004	62.1	146.3	106.3	86.9
2005	93.2	95.6	143.1	68.8
2006	112.5	80.6	129.3	113.3
2007	77.6	110.6	109.6	88.8
Arithmetic average	$\frac{345.4}{4}=86.35$	$\frac{541.4}{5}=108.28$	$\frac{608.3}{5}=121.66$	$\frac{469.3}{5}=93.86$
Chain relatives	100	$\frac{100\times 108.28}{100}$ $= 108.28$	$\frac{121.66\times 108.28}{100}$ $= 131.73$	$\frac{93.86\times 131.73}{100}$ $= 123.64$
Corrected chain relatives	100	$108.28 - 1.675$ $= 106.605$	$131.73 - 3.35$ $= 128.38$	$123.64 - 5.025$ $= 118.615$
Seasonal indices	$\frac{100 \times 100}{113.4}$ $= 88.18$	$\frac{106.605}{113.4}\times 100$ $= 94.01$	$\frac{128.38}{113.4}\times 100$ $= 113.21$	$\frac{118.615}{113.4}\times 100$ $= 104.60$

The calculations in the above table are explained below :

Chain relative of the first quarter (on the basis of first quarter) = 100

Chain relative of the first quarter (on the basis of the last quarter)

$$= \frac{86.35 \times 123.64}{100} = 106.7.$$

The difference between these chain relatives = 106.7 − 100 = 6.7.

Difference per quarter $= \frac{6.7}{4} = 1.675.$

Adjusted chain relatives are obtained by subtracting 1×1.675, 2×1.675, 3×1.675 from the chain relatives of the 2nd, 3rd and 4th quarters respectively.

Average of corrected chain relatives

$$= \frac{100 + 106.605 + 128.38 + 118.615}{4} = \frac{453.6}{4} = 113.4$$

$$\text{Seasonal variation index} = \frac{\text{Correct chain relatives}}{113.4} \times 100$$

Deseasonalisation

When the seasonal component is removed from the original data, the reduced data are free from seasonal variations and is called deseasonalised data. That is, under a multiplicative model

$$\frac{T \times S \times C \times I}{S} = T \times C \times I.$$

Deseasonalised data being free from the seasonal impact manifest only average value of data.

Seasonal adjustment can be made by dividing the original data by the seasonal index.

$$\text{That is, Deseasonalised data} = \frac{\text{Original data}}{\text{Seasonal index}} \times 100$$

where an adjustment-multiplier 100 is necessary because the seasonal indices are usually given in percentages.

In case of additive model

$$Y_t = T + S + C + I,$$

$$\text{Deseasonalised data} = \text{Original data} - \frac{\text{Seasonal index}}{100}$$

$$= Y_t - \frac{\text{Seasonal index}}{100}.$$

Uses and limitations of seasonal indices

Seasonal indices are indices of seasonal variation and provide a quantitative measure of typical seasonal behavior in the form of seasonal fluctuations.

Measurement of cyclical variations:

The various methods used for measuring cyclical variations are

- Residual method
- Reference cycle analysis method
- Direct method
- Harmonic analysis method

Business Cycle

According to Mitchell, "Business cycle are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises : a cycle consists of expansions occurring at about the same time in many activities, followed by general recessions, contractions and revivals which merge into the expansion phase of the next cycle; this sequence of changes is recurrent but not periodic; in duration business cycles vary from more than one year to ten or twelve years.

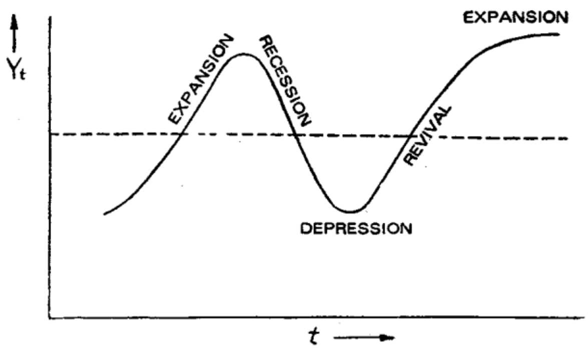
There are four phases of a business cycle, such as

(a) Expansion (prosperity)

(b) Recession

(c) Depression (contraction)

(d) Revival (recovery).



A cycle is measured either from trough-to-trough or from peak-to-peak. Recession and contraction are the result of cumulative downswing of a cycle whereas revival and expansion are the result of cumulative upswing of a cycle.

Question bank:

1. Distinguish between seasonal variations, and cyclical fluctuations. How would you measure secular trend in any given data?
2. Describe the method of link relatives for calculating the seasonal variation indices.
3. How would you determine seasonal variation in the absence of trend?
4. Briefly describe the relative merits and demerits of ratio to trend and ratio to moving average method.
5. What do you understand by cyclical fluctuations in time series?

6. What do you understand by random fluctuation in time series?
7. Explain the term 'Business cycle' and point out the necessity of its study in time series analysis.
8. Calculate seasonal variation for the following data of sale in thousands Rs. of a firm by the Ratio to trend method.

Year	**1st Quarter**	**2nd Quarter**	**3rd Quarter**	**4th Quarter**
1979	30	40	36	34
1980	34	52	50	44
1981	40	58	54	48
1982	52	76	68	62

9. Calculate seasonal indices by Ratio to moving average method from the following data.

Year	**1st Quarter**	**2nd Quarter**	**3rd Quarter**	**4th Quarter**
1980	75	60	54	59
1981	86	65	63	80
1982	90	72	66	85
1983	100	78	72	93

10. The data below gives the average quarterly prices of a commodity for five years. Calculate seasonal indices by method of link relatives.

Year	**1st Quarter**	**2nd Quarter**	**3rd Quarter**	**4th Quarter**
1979	30	26	22	31
1980	35	28	22	36
1981	31	29	28	32
1982	31	31	25	35
1983	34	36	26	33

UNIT-3

INDEX NUMBERS

3.1 Introduction

Historically, the first index was constructed in 1764 to compare the Italian price index in 1750 with the price level in 1500. Though originally developed for measuring the effect of change in prices, index numbers have today become one of the most widely used statistical devices and there is hardly any field where they are not used. Newspapers headline the fact that prices are going up or down, that industrial production is rising or falling, that imports are increasing or decreasing, that crimes are rising in a particular period compared to the previous period as disclosed by index numbers. They are used to feel the pulse of the economy and they have come to be used as indicators of inflationary or deflationary tendencies, In fact, they are described as 'barometers of economic activity', i.e., if one wants to get an idea as to what is happening to an economy, he should look to important indices like the index number of industrial production, agricultural production, business activity, etc.

Definition:

Index numbers are statistical devices designed to measure the relative change in the level of variable or group of variables with respect to time, geographical location etc.

In other words these are the numbers which express the value of a variable at any given period called "***current period*** "as a percentage of the value of that variable at some standard period called "***base period***".

Here the variables may be

1. The price of a particular commodity like silver, iron or group of commodities like consumer goods, food, stuffs etc.
2. The volume of trade, exports, imports, agricultural and industrial production, sales in departmental store.
3. Cost of living of persons belonging to particular income group or profession etc.

Ex: suppose rice sells at Rs.9/kg at BBSR in 1995 as compare to Rs. 4.50/Kg in 1985, the index number of price in 2015 compared to 2005.

Therefore the index number of price of rice in 2015 compared to 2005 is calculated as

$$\frac{Rs.9.00}{Rs.4.50}\times 100 = 200$$

***This means there is a net increase of 100% in the price of rice in 2015as compared to 2005* [the base year's index number is always treated as 100]**

Suppose, during the same period 1995 the rice sells at Rs. 12.00/kg in Delhi. There fore, the index number of price at Bhubaneswar compared to price at Delhi is

$$\frac{Rs.9.00}{Rs.12.00}\times 100 = 75$$

This means there is a net decrease of 25% in the price of rice in 2015as compared to 2005

The above index numbers are called **'*price index numbers*'**.

To take another example the production of rice in 2008 in Tamilnadu was 44, 01,780 metric c tons compare to 36, 19,500 metric tons in 2001. So the index number of the quantity produced in 2008 compared to 2001 is

$$\frac{4401780}{3619500}\times 100 = 121.61$$

That means there is a net increase of 21.61% in production of rice in 2008 as compared to 2001.

The above index number is called **'*quantity index number'***

3.2 Uses of index numbers:

Index numbers are indispensable tools of economics and business analysis. Following are the main uses of index numbers.

1) Index numbers are used as economic barometers:

Index number is a special type of averages which helps to measure the economic fluctuations on price level, money market, economic cycle like inflation, deflation etc. G.Simpson and F.Kafka say that index numbers are today one of the most widely used statistical devices. They are used to take the pulse of economy and they are used as indicators of inflation or deflation tendencies. So index numbers are called economic barometers.

2) Index numbers helps in formulating suitable economic policies and planning etc.

Many of the economic and business policies are guided by index numbers. For example while deciding the increase of DA of the employees; the employer's have to depend primarily on the cost of living index. If salaries or wages are not increased according to the cost of living it leads to strikes, lock outs etc. The index numbers provide some guide lines that one can use in making decisions.

3) They are used in studying trends and tendencies.

Since index numbers are most widely used for measuring changes over a period of time, the time series so formed enable us to study the general trend of the phenomenon under study. For example for last 8 to 10 years we can say that imports are showing upward tendency.

4) They are useful in forecasting future economic activity.

Index numbers are used not only in studying the past and present workings of our economy but also important in forecasting future economic activity.

5) Index numbers measure the purchasing power of money.

The cost of living index numbers determine whether the real wages are rising or falling or remain constant. The real wages can be obtained by dividing the money wages by the corresponding price index and multiplied by 100. Real wages helps us in determining the purchasing power of money.

6) Index numbers are used in deflating.

Index numbers are highly useful in deflating i.e. they are used to adjust the wages for cost of living changes and thus transform nominal wages into real wages, nominal income to real income, nominal sales to real sales etc. through appropriate index numbers.

3.3 Methods of constructing index numbers:

A large number of formulae have been derived for constructing index numbers. They can be

1) Unweighted indices
    a) Simple aggregative method
    b) Simple average of relatives.
2) Weighted indices
    a) Weighted aggregative method
        i) Lasperey's method
        ii) Paasche's method
        iii) Fisher's ideal method

iv) Marshal-Edgeworth method

b) Weighted average of relatives

Unweighted indices:

i) Simple aggregative method:

This is the simplest method of constructing index numbers. When this method is used to construct a price index number the total of current year prices for the various commodities in question is divided by the total of the base year prices and the quotient is multiplied by 100.

`Symbolically $P_{01} = \frac{\sum P_1}{\sum P_0} \times 100$

Where P_0 are the base year prices

P_1 are the current year prices

P_{01} is the price index number for the current year with reference to the base year.

Problem: 1.

Calculate the index number for 2005 taking 2001 as the base for the following data

Commodity	Unit	Prices 2001 (P_0)	Prices 2005 (P_1)
A	Kilogram	2.50	4.00
B	Dozen	5.40	7.20
C	Meter	6.00	7.00
D	Quintal	150.00	200.00
E	Liter	2.50	3.00
Total		166.40	221.20

Price index number = $P_{01} = \frac{\sum P_1}{\sum P_0} \times 100 = \frac{221.20}{166.40} \times 100 = 132.93$

$\therefore$ There is a net increase of 32.93% in 2005 as compared to 2001.

Limitations:

There are two main limitations of this method

1. The units used in the prices or quantity quotations have a great influence on the value of index.
2. No considerations are given to the relative importance of the commodities.

ii) Simple average of relatives

When this method is used to construct a price index number, first of all price relatives are obtained for the various items included in the index and then the average of these relatives is obtained using any one of the averages i.e. mean or median etc.

When A.M is used for averaging the relatives the formula for computing the index is

$$P_{01} = \frac{1}{n}\sum\left(\frac{P_1}{P_0}\times 100\right)$$

When G.M is used for averaging the relatives the formula for computing the index is

$$P_{01} = Anti\log\left[\frac{1}{n}\sum\log\left(\frac{P_1}{P_0}\times 100\right)\right]$$

Where n is the number of commodities

and price relative = $\frac{P_1}{P_0}\times 100$

Problem: 2.

Calculate the index number for 2005 taking 2001 as the base for the following data

Commodity	Unit	Prices 2001 (P_0)	Prices 2005 (P_1)	$\frac{P_1}{P_0}\times 100$
A	Kilogram	50	70	$\frac{70}{50}\times 100$ = 140
B	Dozen	40	60	150
C	Meter	80	90	112.5
D	Quintal	110	120	109.5
E	Liter	20	20	100
Total				

Price index number = $P_{01} = \frac{1}{n}\sum\left(\frac{P_1}{P_0}\times 100\right) = \frac{1}{5}\sum 612 = 122.4$

$\therefore$ There is a net increase of 22.4% in 2005 as compared to 2001.

Advantages:

- It is not affected by the units in which prices are quoted.
- It gives equal importance to all the items and extreme items don't affect the index number.
- The index number calculated by this method satisfies the unit test.

Disadvantages:

- Since it is an unweighted average the importance of all items are assumed to be the same.
- The index constructed by this method doesn't satisfy all the criteria of an ideal index number.
- In this method one can face difficulties to choose the average to be used.

Weighted indices: Weighted aggregative method:

These indices are same as simple aggregative method. The only difference is in this method, weights are assigned to the various items included in the index.

There are various methods of assigning weights and consequently a large number of formulae for constructing weighted index number have been designed.
Some important methods are

i. Lasperey's method:

This method is devised by Lasperey in year 1871.It is the most important of all the types of index numbers. In this method the base year quantities are taken weights. The formula for constructing Lasperey's price index number is

$$P_{01}^{La} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

ii. Paasche's method: In this method the current year quantities are taken as weights and the formula is given by

$$P_{01}^{Pa} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$$

iii. Fisher's ideal method: Fishers price index number is given by the G.M of the Lasperey's and Paasche's index numbers.

Symbolically

$$P_{01}^{F} = \sqrt{P_{01}^{La} P_{01}^{Pa}}$$

$$= \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100 \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100}$$

$$= \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$$

3.4 Quantity index numbers:

i. Lasperey's quantity index number: Base year prices are taken as weights

$$Q_{01}^{La} = \frac{\sum q_1 p_0}{\sum q_0 p_0} \times 100$$

ii. Paasche's quantity index number : Current year prices are taken as weights

$$Q_{01}^{Pa} = \frac{\sum q_1 p_1}{\sum q_0 p_1} \times 100$$

iii. Fisher's ideal method: $Q_{01}^{F} = \sqrt{Q_{01}^{La} Q_{01}^{Pa}} = \sqrt{\frac{\sum q_1 p_0}{\sum q_0 p_0} \times \frac{\sum q_1 p_1}{\sum q_0 p_1}} \times 100$

Fisher's index number is called ideal index number. Why?

The Fisher's index number is called ideal index number due to the following characteristics.

1) It is based on the G.M which is theoretically considered as the best average of constructing index numbers.

2) It takes into account both current and base year prices as quantities.

3) It satisfies both time reversal and factor reversal test which are suggested by Fisher.

4) The upward bias of Lasperey's index number and downward bias of Paasche's index number are balanced to a great extent.

Example:1. Compute price index numbers for the following data by

(*i*) *Laspeyre's method,*
(*ii*) *Paasche's method,*
(*iii*) *Fisher's ideal method,*
(*iv*) *Dorbish-Bowley's method,*
(*v*) *Marshall-Edgeworth's method.*

Year	*Commodity A*		*Commodity B*		*Commodity C*	
	Price	*Quantity*	*Price*	*Quantity*	*Price*	*Quantity*
1980	*4*	*50*	*3*	*10*	*2*	*5*
1985	*10*	*45*	*6*	*8*	*3*	*4*

Base year : 1980
Price and quantity given in arbitrary units.

Calculation of Indices

Commodities	*1980*		*1985*					
	Price p_0	*Quantity* q_0	*Price* p_1	*Quantity* q_1	p_1q_0	p_0q_0	p_1q_1	p_0q_1
A	4	50	10	45	500	200	450	180
B	3	10	6	8	60	30	48	24
C	2	5	3	4	15	10	12	8
Total	—	—	—	—	575	240	510	212

(*i*) Laspeyre's method :

$$L_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100 = \frac{575}{240} \times 100 = 239.58$$

(*ii*) Paasche's method :

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100 = \frac{510}{212} \times 100 = 240.57$$

(*iii*) Fisher's ideal method :

$$F_{01} = \sqrt{L_{01} \times P_{01}} = \sqrt{239.58 \times 240.57} = 240.07.$$

(*iv*) Dorbish-Bowley's method :

$$DB_{01} = \frac{L_{01} + P_{01}}{2}$$

$$= \frac{239.58 + 240.57}{2} \doteq 239.82.$$

3.5 Comparison of Lasperey's and Paasche's index numbers:-

In Lasperey's index number base year quantities are taken as the weights and in Paasche's index the current year quantities are taken as weights.

From the practical point of view Lasperey's index is often proffered to Paasche's for the simple reason that Lasperey's index weights are the base year quantities and do not change from the year to the next. On the other hand Paasche's index weights are the current year quantities, and in most cases these weights are difficult to obtain and expensive.

Lasperey's index number is said to be have upward bias because it tends to over estimate the price rise, where as the Paasche's index number is said to have downward bias, because it tends to under estimate the price rise.

When the prices increase, there is usually a reduction in the consumption of those items whose prices have increased. Hence using base year weights in the Lasperey's index, we will be giving too much weight to the prices that have increased the most and the numerator will be too large. Due to similar considerations, Paasche's index number using given year weights under estimates the rise in price and hence has down ward bias.

If changes in prices and quantities between the reference period and the base period are moderate, both Lasperey's and Paasche's indices give nearly the same values.

Demerit of Paasche's index number:

Paasche's index number, because of its dependence on given year's weight, has distinct disadvantage that the weights are required to be revised and computed for each period, adding extra cost towards the collection of data.

What are the desiderata of good index numbers?

Irving Fisher has considered two important properties which an index number should satisfy. These are tests of reversibility.

1. Time reversal test
2. Factor reversal test

If an index number satisfies these two tests it is said to be an ideal index number.

Weighted average of relatives:

Weighted average of relatives can be calculated by taking values of the base year (p_0q_0) as the weights. The formula is given by

When A.M is used $P_{01} = \frac{\sum PV}{\sum V}$

$$\text{When G.M is used } P_{01} = Anti\log \frac{\sum V \log P}{\sum V}$$

where $P = \frac{p_1}{p_0} \times 100$ and $V = p_0 q_0$ i.e. base year value

Illustration 8. From the following data compute price index by supplying weighted average of price method using :

(*a*) arithmetic mean, and
(*b*) geometric mean.

Commodity	p_0 *(Rs.)*	q_0	p_1 *(Rs.)*
Sugar	3.0	20 kg.	4.0
Flour	1.5	40 kg.	1.6
Milk	1.0	10 lt.	1.5

Solution. (*a*) INDEX NUMBER USING WEIGHTED ARITHMETIC MEAN OF PRICE RELATIVES

Commodity	p_0	q_0	p_1	p_0q_0 V	$\frac{p_1}{p_0} \times 100$ p	PV
Sugar	Rs. 3.0	20 kg.	Rs. 4.0	60	$\frac{4}{3} \times 100$	8,000
Flour	Rs. 1.5	40 kg.	Rs. 1.6	60	$\frac{1.6}{1.5} \times 100$	6,400
Milk	Re. 1.0	10 lt.	Rs. 1.5	10	$\frac{1.5}{1.0} \times 100$	1,500
				$\Sigma V = 130$		$\Sigma PV = 15{,}900$

$$p_{01} = \frac{\Sigma PV}{\Sigma V} = \frac{15{,}900}{130} = 122.31$$

This means that there has been a 22.3 per cent increase in prices over the base level.

(*b*) INDEX NUMBER USING GEOMETRIC MEAN OF PRICE RELATIVES

Commodity	p_0	q_0	p_1	V	p	*Log p*	*V Log p*
Sugar	Rs. 3.0	**20 kg.**	Rs. 4.0	60	133.3	2.1249	127.494
Flour	Rs. 1.5	40 kg.	Rs. 1.6	60	106.7	2.0282	121.692
Milk	Re. 1.0	10 lt.	Rs. 1.5	10	150.0	2.1761	21.761
				$\Sigma V = 130$			$\Sigma V \log p$ **= 270.947**

$$p_{01} = \text{Antilog}\left[\frac{\Sigma V \log p}{\Sigma V}\right] = \text{Antilog}\left[\frac{270.947}{130}\right] = \text{Antilog } 2.084 = 120.9$$

3.6 The Chain Index Numbers

In fixed base method the base remain constant through out i.e. the relatives for all the years are based on the price of that single year. On the other hand in chain base method, the relatives for each year is found from the prices of the immediately preceding year. Thus the base changes from year to year. Such index numbers are useful in comparing current year figures with the preceding year figures. The relatives which we found by this method are called link relatives.

$$\text{Thus link relative for current year} = \frac{Current\ years\ figure}{Previous\ years\ figure} \times 100$$

And by using these link relatives we can find the chain indices for each year by using the below formula

$$\text{Chain index for current year} = \frac{Linkrelative\ of\ current\ year \times Chain\ index\ of\ previous\ year}{100}$$

Note: The fixed base index number computed from the original data and chain index number computed from link relatives give the same value of the index provided that there is only one commodity, whose indices are being constructed.

Example: 3. From the following data of wholesale prices of wheat for ten years construct index number taking a) 1998 as base and b) by chain base method

Year	*Price of Wheat (Rs. per 40 kg.)*	*Year*	*Price of Wheat (Rs. per 40 kg.)*
1998	50	2003	78
1999	60	2004	82
2000	62	2005	84
2001	65	2006	88
2002	70	2007	90

Solution. (*a*) CONSTRUCTION OF INDEX NUMBERS TAKING 1998 AS BASE

Year	*Price of Wheat*	*Index Number (1998 = 100)*	*Year*	*Price of Wheat*	*Index Number (1998 = 100)*
1998	50	100	2003	78	$\frac{78}{50} \times 100 = 156$
1999	60	$\frac{60}{50} \times 100 = 120$	2004	82	$\frac{82}{50} \times 100 = 164$
2000	62	$\frac{62}{50} \times 100 = 124$	2005	84	$\frac{84}{50} \times 100 = 168$
2001	65	$\frac{65}{50} \times 100 = 130$	2006	88	$\frac{88}{50} \times 100 = 176$
2002	70	$\frac{70}{50} \times 100 = 140$	2007	90	$\frac{90}{50} \times 100 = 180$

This means that from 1998 to 1999 there is a 20 per cent increase; from 1999 to 2000 there is a 24 per cent increase; from 2000 to 2001 there is a 30 per cent increase. If we are interested in finding out increase from 1998 to 1999, from 1999 to 2000, from 2000 to 2001, we shall have to compute the chain indices.

(b) CONSTRUCTION OF CHAIN INDICES

Year	Price of Wheat	Link Relatives	Chain Indices (1998 = 100)
1998	50	100.00	100
1999	60	$\frac{60}{50} \times 100 = 120.00$	$\frac{120 \times 100}{100} = 120$
2000	62	$\frac{62}{60} \times 100 = 103.33$	$\frac{103.33 \times 120}{100} = 124$
2001	65	$\frac{65}{62} \times 100 = 104.84$	$\frac{104.84 \times 124}{100} = 130$
2002	70	$\frac{70}{65} \times 100 = 107.69$	$\frac{107.69 \times 130}{100} = 140$
2003	78	$\frac{78}{70} \times 100 = 111.43$	$\frac{111.43 \times 140}{100} = 156$
2004	82	$\frac{82}{78} \times 100 = 105.13$	$\frac{105.13 \times 156}{100} = 164$
2005	84	$\frac{84}{82} \times 100 = 102.44$	$\frac{102.44 \times 164}{100} = 168$
2006	88	$\frac{88}{84} \times 100 = 104.76$	$\frac{104.76 \times 168}{100} = 176$
2007	90	$\frac{90}{88} \times 100 = 102.27$	$\frac{102.27 \times 176}{100} = 180$

Note: the chain indices obtained in (b) are the same as the fixed base indices obtained in (a). in fact chain index figures will always be equal to fixed index figure if there is only one series.

Example-4: Compute the chain index number with 2003 prices as base from the following table giving the average wholesale prices of the commodities A, B and C for the year 2003 to 2007

Commodity	Average wholesale price (in Rs.) 2003	2004	2005	2006	2007
A	20	16	28	35	21
B	25	30	24	36	45
C	20	25	30	24	30

Solution. COMPUTATION OF CHAIN INDICES

Commodity	Relatives based on preceding year				
	2003	2004	2005	2006	2007
A	100	$\frac{16}{20} \times 100 = 80$	$\frac{28}{16} \times 100 = 175$	$\frac{35}{28} \times 100 = 125$	$\frac{21}{35} \times 100 = 60$
B	100	$\frac{30}{25} \times 100 = 120$	$\frac{24}{30} \times 100 = 80$	$\frac{36}{24} \times 100 = 150$	$\frac{45}{36} \times 100 = 125$
C	100	$\frac{25}{20} \times 100 = 125$	$\frac{30}{25} \times 100 = 120$	$\frac{24}{30} \times 100 = 80$	$\frac{30}{24} \times 100 = 125$

Total of Link Relatives	300	325	375	355	310
Average of Link Relatives	100	108.33	125	118.33	103.33
Chain Index (2003 = 100)	100	$\frac{108.33 \times 100}{100}$ = 108.33	$\frac{125 \times 108.33}{100}$ = 135.41	$\frac{118.33 \times 135.41}{100}$ = 160.23	$\frac{103.33 \times 160.23}{100}$ = 165.57

Conversion of fixed based index to chain based index

$$\text{Current year C.B.I} = \frac{Current\ \ years\ \ F.B.I}{\Pr evious\ \ years\ \ C.B.I} \times 100$$

Conversion of chain based index to fixed base index.

$$\text{Current year F.B.I} = \frac{Current\ \ years\ \ C.B.I \times \Pr evious\ \ years\ \ F.B.I}{100}$$

Example 5: Compute the chain base index numbers

Year	*1980*	*1981*	*1982*	*1983*	*1984*
Fixed base index	*100*	*120*	*150*	*130*	*160*

Solution. **Base year 1980 = 100**

Year	*Fixed base indices*	*Chain base index* $\left(\frac{I_1}{I_0} \times 100\right)$
1980	100	100
1981	120	$\frac{120 \times 100}{100} = 120$
1982	150	$\frac{150}{120} \times 100 = 125$
1983	130	$\frac{130}{150} \times 100 = 86.67$
1984	160	$\frac{160}{130} \times 100 = 123.08$

Example 6: Calculate fixed base index numbers from the following chain base index numbers

Year	*1978*	*1979*	*1980*	*1981*	*1982*
Chain base index numbers	*120*	*140*	*120*	*130*	*150*

Solution. Computation of fixed base index numbers

Year	*Chain Base Index Numbers*	*Fixed Base Index Numbers*
1978	120	120
1979	140	$\frac{140 \times 120}{100} = 168$
1980	120	$\frac{120 \times 168}{100} = 201.60$
1981	130	$\frac{130 \times 201.60}{100} = 262.08$
1982	150	$\frac{150 \times 262.08}{100} = 393.12$

Note: It may be remembered that the fixed base index for the first year is same as the chain base index for that year.

Merits of chain index numbers:

1. The chain base method has a great significance in practice, because in economic and business data we are often concerned with making comparison with the previous period.
2. Chain base method doesn't require the recalculation if some more items are introduced or deleted from the old data.
3. Index numbers calculated from the chain base method are free from seasonal and cyclical variations.

Demerits of chain index numbers:

1. This method is not useful for long term comparison.
2. If there is any abnormal year in the series it will effect the subsequent years also.

Differences between fixed base and chain base methods:

Chain base	**Fixed base**
1. Here the base year changes	1. Base year does not changes
2. Here link relative method is used	2. No such link relative method is used
3. Calculations are tedious	3. Calculations are simple
4. It can not be computed if any one year is missing	4. It can be computed if any year is missing
5. It is suitable for short period	5. It is suitable for long period
6. Index numbers will be wrong if an error is committed in the calculation of link relatives	6. The error is confined to the index of that year only.

Base shifting:

One of the most frequent operations necessary in the use of index numbers is changing the base of an index from one period to another with out recompiling the entire series. Such a change is referred to as *'base shifting'*. The reasons for shifting the base are

1. If the previous base has become too old and is almost useless for purposes of comparison.
2. If the comparison is to be made with another series of index numbers having different base.

The following formula must be used in this method of base shifting is

Index number based on new base year = $\frac{\text{current years old index number}}{\text{new base years old index number}} \times 100$

Example:

The following are the index numbers of prices with 1998 as base year

year	**Index**
1998	100
1999	110
2000	120
2001	200
2002	400
2003	410
2004	400
2005	380
2006	370
2007	340

Shift the base from 1998 to 2004 and recast the index numbers.

Solution:

Index number based on new base year = $\frac{\text{current years old index number}}{\text{new base years old index number}} \times 100$

Index number for 1998 = $\frac{100}{400} \times 100 = 25$

………………………………………….

Index number for 2007= $\frac{340}{400} \times 100 = 85$

Year	**Index number (1998as base)**	**Index number (2004 as base)**	**Year**	**Index number (1998as base)**	**Index number (2004 as base)**
1998	100	$\frac{100}{400} \times 100 = 25$	2003	410	$\frac{410}{400} \times 100 = 102.5$
1999	110	$\frac{110}{400} \times 100 = 27.5$	2004	400	$\frac{400}{400} \times 100 = 100$
2000	120	$\frac{120}{400} \times 100 = 30$	2005	380	$\frac{380}{400} \times 100 = 95$
2001	200	$\frac{200}{400} \times 100 = 50$	2006	370	$\frac{370}{400} \times 100 = 92.5$
2002	400	$\frac{400}{400} \times 100 = 100$	2007	340	$\frac{340}{400} \times 100 = 85$

Splicing of two series of index numbers:

The problem of combining two or more overlapping series of index numbers into one continuous series is called *splicing*. In other words, if we have a series of index numbers with some base year which is discontinued at some year and we have another series of index numbers with the year of discontinuation as the base, and connecting these two series to make a continuous series is called splicing.

The following formula must be used in this method of splicing,

Index number after splicing = $\frac{\text{index number to be spliced} \times \text{old index number of existing base}}{100}$

Example 7: The index A given was started in 1993 and continued up to 2003 in which year another index B was started. Splice the index B to index A so that a continuous series of index is made

Year	Index A	Index B	Year	Index A	Index B
1993	100		2002	138	
1994	110		2003	150	100
1995	112		2004		120
—			2005		140
—			2006		130
—			2007		150

Solution. INDEX *B* SPLICED TO INDEX *A*

Year	*Index A*	*Index B*	*Index B spliced to Index A 1993 as base*
1993	100		
1994	110		
1995	112		
—			
—			
2002	138		
2003	150	100	$\frac{150}{100} \times 100 = 150$
2004		120	$\frac{150}{100} \times 120 = 180$
2005		140	$\frac{150}{100} \times 140 = 210$
2006		130	$\frac{150}{100} \times 130 = 195$
2007		150	$\frac{150}{100} \times 150 = 225$

The spliced index now refers to 1993 as base and we can make a continuous comparison of index numbers from 1993 onwards.

In the above case it is also possible to splice the new index in such a manner that a comparison could be made with 2003 as base. This would be done by multiplying the old index by the ratio $\frac{100}{150}$. Thus the spliced index for 1993 would be $\frac{100}{150} \times 100 = 66.7$ for 1994, $\frac{100}{150} \times 110 = 73.3$, for 1995, $\frac{100}{150} \times 112 = 74.6$, etc. This process appears to be more useful because a recent year can be kept as a base. However, much would depend upon the object.

Deflating:

Deflating means correcting or adjusting a value which has inflated. It makes allowances for the effect of price changes. When prices rise, the purchasing power of money declines. If the money incomes of people remain constant between two periods and prices of commodities are doubled the purchasing power of money is reduced to half. For example if there is an increase in the price of rice from Rs10/kg in the year 1980 to Rs20/kg in the year 1982. then a person can buy only half kilo of rice with Rs10. so the purchasing power of a rupee is only 50paise in 1982 as compared to 1980.

Thus the purchasing power of money = $\frac{1}{price\ index}$

In times of rising prices the money wages should be deflated by the price index to get the figure of real wages. The real wages alone tells whether a wage earner is in better position or in worst position.

For calculating real wage, the money wages or income is divided by the corresponding price index and multiplied by 100.

$$\text{i.e. Real wages} = \frac{Money\ wages}{\Pr ice\, index} \times 100$$

$$\text{Thus Real Wage Index} = \frac{\operatorname{Re} al\ wage\ of\ current\ year}{\operatorname{Re} al\ wage\ of\ base\ year} \times 100$$

Example 8: The following table gives the annual income of a worker and the general Index Numbers of price during 1999-2007. Prepare Index Number to show the changes in the real income of the teacher and comment on price increase.

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007
income (Rs.)	3600	4200	5000	5500	6000	6400	6800	7200	7500
Price Index No.	100	120	145	160	250	320	450	530	600

Solution. INDEX NUMBER SHOWING CNANGES IN THE REAL INCOME OF THE WORKER

Year	*Income (Rs.)*	*Price Index No.*	*Real Income*	*Real income Index No.*
1999	3600	100	$\frac{3600}{100} \times 100 = 3600.00$	100.00
2000	4200	120	$\frac{4200}{120} \times 100 = 3500.00$	97.22
2001	5000	145	$\frac{5000}{145} \times 100 = 3448.27$	95.78
2002	5500	160	$\frac{5500}{160} \times 100 = 3437.50$	95.49
2003	6000	250	$\frac{6000}{250} \times 100 = 2400.00$	66.67
2004	6400	320	$\frac{6400}{320} \times 100 = 2000.00$	55.56
2005	6800	450	$\frac{6800}{450} \times 100 = 1511.11$	41.98
2006	7200	530	$\frac{7200}{530} \times 100 = 1358.49$	37.74
2007	7500	600	$\frac{7500}{600} \times 100 = 1250.00$	34.72

The method discussed above is frequently used to deflate individual values, value series or value indices. Its special use is in problems dealing with such diversified things as rupee sales, rupee inventories of manufacturer's, wholesaler's and retailer's income, wages and the like.

3.7 Cost of living index numbers (or) Consumer price index numbers:

The ***cost of living index numbers*** measures the changes in the level of prices of commodities which directly affects the cost of living of a specified group of persons at a specified place. The general index numbers fails to give an idea on cost of living of different classes of people at different places.

Different classes of people consume different types of commodities, people's consumption habit is also vary from man to man, place to place and class to class i.e. richer class, middle class and poor class. For example the cost of living of rickshaw pullers at BBSR is different from the rickshaw pullers at Kolkata. The consumer price index helps us in determining the effect of rise and fall in prices on different classes of consumers living in different areas.

Main steps or problems in construction of cost of living index numbers

The following are the main steps in constructing a cost of living index number.

1. **Decision about the class of people for whom the index is meant**

   It is absolutely essential to decide clearly the class of people for whom the index is meant i.e. whether it relates to industrial workers, teachers, officers, labors, etc. Along with the class of people it is also necessary to decide the geographical area covered by the index, such as a city, or an industrial area or a particular locality in a city.

2. **Conducting family budget enquiry**

   Once the scope of the index is clearly defined the next step is to conduct a sample family budget enquiry i.e. we select a sample of families from the class of people for whom the index is intended and scrutinize their budgets in detail. The enquiry should be conducted during a normal period i.e. a period free from economic booms or depressions. The purpose of the enquiry is to determine the amount; an average family spends on different items. The family budget enquiry gives information about the nature and quality of the commodities consumed by the people. The commodities are being classified under following heads i)Food

   ii) Clothing

   iii)Fuel and Lighting

iv)House rent

v) miscellaneous

3. Collecting retail prices of different commodities

The collection of retail prices is a very important and at the same time very difficult task, because such prices may vary from lace to place, shop to shop and person to person. Price quotations should be obtained from the local markets, where the class of people reside or from super bazaars or departmental stores from which they usually make their purchases.

Uses of cost of living index numbers:

1. Cost of living index numbers indicate whether the real wages are rising or falling. In other words they are used for calculating the real wages and to determine the change in the purchasing power of money.

$$\text{Purchasing power of money} = \frac{1}{\text{Cost of living index number}}$$

$$\text{R}eal\ Wages = \frac{\text{Money wages}}{\text{Cost of living index umbers}} \times 100$$

2. Cost of living indices are used for the regulation of D.A or the grant of bonus to the workers so as to enable them to meet the increased cost of living.
3. Cost of living index numbers are used widely in wage negotiations.
4. These index numbers also used for analyzing markets for particular kinds of goods.

Methods for construction of cost of living index numbers:

Cost of living index number can be constructed by the following formulae.

1) Aggregate expenditure method or weighted aggregative method
2) Family budget method or the method of weighted relatives

1) Aggregate expenditure method or weighted aggregative method

In this method the quantities of commodities consumed by the particular group in the base year are taken as weights. The formula is given by

$$\text{consumer price index} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

Steps:

i) The prices of commodities for various groups for the current year is multiplied by the quantities of the base year and their aggregate expenditure of current year is obtained .i.e. $\sum p_1q_0$.

ii) Similarly obtain $\sum p_0q_0$

iii) The aggregate expenditure of the current year is divided by the aggregate expenditure of the base year and the quotient is multiplied by 100.

Symbolically $\dfrac{\sum p_1q_0}{\sum p_0q_0}\times 100$

3) Family budget method or the method of weighted relatives

In this method cost of living index is obtained on taking the weighted average of price relatives, the weights are the values of quantities consumed in the base year i.e. $v = p_0q_0$. Thus the consumer price index number is given by

consumer price index = $\dfrac{\sum pv}{\sum v}$

Where $p = \dfrac{p_1}{p_o}\times 100$ for each item

$v = p_0q_0$, value on the base year

Note: It should be noted that the answer obtained by applying the aggregate expenditure method and family budget method shall be same.

Example 8: Construct the consumer price index number for 2007 on the basis of 2006 from the following data using (i) the aggregate expenditure method, and (ii) the family budget method.

Commodity	Quantity consumed in 2006	Units	Price in 2006		Price in 2007	
			Rs.	Paise	Rs.	Paise
A	6 Quintals	Quintal	5	75	6	0
B	6 Quintals	Quintal	5	0	8	0
C	1 Quintals	Quintal	6	0	9	0
D	6 Quintals	Quintal	8	0	10	0
E	4 Kg.	Kg.	2	0	1	50
F	1 Quintals	Quintal	20	0	15	0

Solution. COMPUTATION OF CONSUMER PRICE INDEX NUMBER FOR 2007
(Base 2006 = 100) BY THE AGGREGATE EXPENDITURE METHOD

Commodity	Quantities consumed	Unit	Price in 2006 p_0	Price in 2007 p_1	p_1q_0	p_0q_0
A	6 Qtl.	Qtl.	5.75	6.00	36.00	34.50
B	6 Qtl.	"	5.00	8.00	48.00	30.00
C	1 Qtl.	"	6.00	9.00	9.00	6.00
D	6 Qtl.	"	8.00	10.00	60.00	48.00
E	4 Kg.	Kg.	2.00	1.50	6.00	8.00
F	1 Qtl.	Qtl.	20.00	15.00	15.00	20.00
					$\Sigma p_1q_0 = 174$	$\Sigma p_0q_0 = 146.5$

$$\text{Consumer Price Index} = \frac{\Sigma p_1q_0}{\Sigma p_0q_0} \times 100 \; \frac{174}{146.5} \times 100 = 118.77$$

CONSTRUCTON OF CONSUMER PRICE INDEX NUMBER FOR 2007
(Base 2006 = 100) BY THE FAMILY BUDGET METHOD

Commodity	Quantities consumed q_0	Unit	Price in 2006 p_0	Price in 2007 p_1	$\frac{p_1}{p_0} \times 100$ P	p_0q_0 V	PV
A	6 Qtl:	Qtl.	5.75	6.0	104.34	34.5	3,600
B	6 Qtl.	Qtl.	5.00	8.0	160.00	30.0	4,800
C	1 Qtl.	Qtl.	6.00	9.0	150.00	6.0	900
D	6 Qtl.	Qtl.	8.00	10.0	125.00	48.0	6,000
E	4 Kg.	Kg.	2.00	1.5	75.00	8.0	600
F	1 Qtl.	Qtl.	20.00	15.0	75.00	20.0	1,500
						$\Sigma V = 146.5$	$\Sigma PV = 17{,}400$

$$\text{Consumer Price Index} = \frac{\Sigma PV}{\Sigma V} = \frac{17.400}{146.5} = 118.77$$

Thus, the answer is the same by both the methods. However, the reader should prefer the aggregate expenditure method because it is far more easier to apply compared to the family budget method.

Possible errors in construction of cost of living index numbers:

Cost of living index numbers or its recently popular name consumer price index numbers are not accurate due to various reasons.

1. Errors may occur in the construction because of inaccurate specification of groups for whom the index is meant.
2. Faulty selection of representative commodities resulting out of unscientific family budget enquiries.
3. Inadequate and unrepresentative nature of price quotations and use of inaccurate weights
4. Frequent changes in demand and prices of the commodity
5. The average family might not be always a representative one.

Question Bank:

1. Define fisher's ideal index number. Show that it satisfies both time reversal and factor reversal tests.
2. Define index number. Discuss various problems in the construction of index numbers.
3. Explain the need of weights in index numbers. Explain commonly used weighting schemes.
4. What are the tests of a good index number? Explain Paasche's, Lasperey's and fisher's index numbers. Verify if they satisfy these tests?
5. What is the chain base method of constructing of index numbers and how does it differ from fixed base method? Discuss the advantages and disadvantages of the two methods.
6. What is an index number? What does it measure?
7. Distinguish between weighted and unweighted index numbers.
8. Mention three important weighted index numbers.
9. Show that the mean of Lasperey's and Paasche's index numbers is always greater than fisher's index number.
10. Why fisher's index number is said to be an ideal index number?
11. Differentiate wholesale price index number with the cost of living index numbers.
12. Distinguish between Lasperey's and Paasche's index numbers.
13. What considerations are to be made for the selection of items for the construction of price index numbers?
14. Distinguish between simple aggregative index and simple average of relatives.
15. What you mean by average family for the purpose of construction of consumers price index numbers?
16. Differentiate between fixed base and chain base.
17. Describe briefly a method of constructing cost of living index numbers. Mention its uses.
18. Explain the meaning of upward bias and downward bias with reference to Lasperey's and Paasche's price indices.
19. For a data Lasperey's index number is 120 and fisher's index number is 125. Calculate Paasche's index number.
20. Explain simple aggregative method for the construction of index numbers. Explain its merits and demerits.

21. Explain simple average of relatives for the construction of index numbers. Explain its merits and demerits.

22. What are the merits and demerits of chain base method?

UNIT - 4

VITAL STATISTICS

4.1 Introduction:

The term vital statistics signifies either the data or the methods applied to the analysis of data which provide a description of vital event occurring in given communities. By vital events mean such events of human life as birth, death, sickness, marriage, divorce, adoption separation etc.; In short, all the events which have to do with an individuals entrance into or departure from life together with the changes in civil status which may occur to him during his life time.

4.2. Uses of Vital Statistics:

Vital Statistics are extremely useful and their significance can be judged from the following angles:

1. Use to individuals
2. Use to operating agencies
3. Use in research – demographic and medical
4. Use in public administration

Use to individuals: Records of births, deaths, marriages and divorce as well as those of legitimation, recognition, adoption, and so on. The basic registration document or a certified copy thereof has legal significance to the person concerned, which is equalled by a few of the other documents a man may acquire in his lifetime.

Use to operating agencies: Records of births, deaths, marriages are useful to government agencies for a variety of administrative purpose. For example, the control programmes for infectious disease within the family and within the community often depend on the death registration report for their initiation.

Uses in Research: Vital statistics are indispensable in demographic research. The study of population movement and the interrelationships of demographic with economic and social factors is of fundamental importance to society will become increasingly more so as the advances in technology and public health focus attention on demographic problems. The three directions which such an analysis takes are: 1. Population estimation 2. Population projection 3. Analytical study.

Uses in Public Administration: Statistics in general and vital statistics in particular are fundamental elements in public administration, which is the machinery and methods underlying all official programmes of economic and social development areas.

4.3. Measures of Mortality:

The following rates are used for measuring mortality Crude Death Rate, The annual crude death rate is defined as,

$$Crude\ \ Death\ Rate = \frac{Annual\ Deaths}{Annual\ Mean\ population} \times 1000$$

The crude death rate for a given year tells us what rate deaths have depleted the population over the course of the year. We can calculate crude death rate for males and females separately. The crude death rate usually lies between 8 and 30 per 1000. The female rate generally lower than the male rate.

4.4. Specific Death Rates:

Crude death rates are not enough for a detailed study of the mortality conditions in a community. We often need to know more about deaths occurring in different section of the population. For instance, people interested in infant or child welfare work study deaths taking place under 1 year of age or in such age groups as 1-4 years, 5-9 years etc. Those interested in maternal health, study how many deaths occurred among women of child – bearing age.

The formula for computing specific death rate is:

$$Annual\ Specific\ Death\ Rate\ \ for\ \ age = \frac{\begin{array}{l} Number\ of\ \ deaths\ which\ occured\ \ among \\ a\ \ specific\ age\ group\ of\ \ the\ \ population\ of\ \ a \\ given\ geograpic\ Area\ \ during\ given\ year \end{array}}{\begin{array}{l} Mid-year\ \ population\ of\ \ the\ \ specific\ age\ group \\ in\ \ given\ geograpic\ Area\ \ during\ same\ year \end{array}} \times 1000$$

These rates measures the risk of dying in each of the age groups selected for the computation. Usually such rates are computed for the entire span of years, and are further specified by sex, so that rates specific for age and sex are available.

4.5 Infant Mortality Rate

Infant mortality rates serve as one of the best indices to the general "healthiness" of the society. It is similar to age specific death rate for infants under 1 year of age. It is defined as,

$$Infant\ Mortality\ Rate = \frac{\begin{array}{c} Number\ of\ deaths\ under\ 1\ year\ of\ age\ which\ occured\ among\ the \\ population\ of\ a\ given\ geograpic\ area\ during\ a\ given\ year \end{array}}{\begin{array}{c} Number\ of\ live\ births\ which\ occured\ among\ the \\ population\ of\ the\ given\ geograpic\ Area\ during\ the\ same\ year \end{array}} \times 1000$$

IlLustrations:

1. Compute the crude death rate of the two populations A and B from the following data.

Age-group(years)	A		B	
	Population	**Deaths**	**Population**	**Deaths**
Below 5	15000	360	40000	1000
5-30	20000	400	52000	1040
Above 30	10000	280	8000	240

Solution:

$$Crude\ \ Death\ \ Rate = \frac{N}{P} \times 1000,\ where\ \ N - No.of\ \ deaths,\ \ P - Population\ ,$$

$$Crude\ \ Death\ \ Rate\ \ for\ \ town\ A = \frac{1040}{45000} \times 1000 = 23.11$$

$$Crude\ \ Death\ \ Rate\ \ for\ town\ \ B = \frac{2280}{100000} \times 1000 = 22.80$$

4.6. Measures of Fertility

In order to study the speed at which the population is increasing, fertility rates are used which are of various types.

Crude Birth Rate: The annual crude birth rate is defined as:

$$Crude\ \ Birth\ \ Rate = \frac{Annual\ Births}{Annual\ Mean\ population} \times 1000$$

In this measure the births are related to the mean population and not to the population at a particular date. The crude birth rate of a given year tells us at what rate births have augmented the population over the course of year. The crude birth rate usually lies between 10 1nd 55 per 1000.

4.7. Specific Fertility Rate

The concept of specific fertility arises out of the fact that fertility is affected by a number of factors such as age, marriage, state or region, urban-rural characteristic, etc. When fertility rate is calculated on the basis of age distribution, it is called age-specific fertility rate. While calculating age-specific fertility rate women of different ages in the child-bearing age are placed in a small age groups so as to put them at par with others of child-bearing capacity.

$$Specific\ Birth\ Rate = \frac{\begin{array}{c} Number\ of\ live\ births\ which\ occured\ to\ females \\ of\ a\ specified\ age\ group\ of\ the\ population\ of\ a \\ given\ geograpic\ Area\ during\ given\ year \end{array}}{\begin{array}{c} Mid-year\ female\ population\ of\ the\ specific\ age\ group \\ in\ given\ geograpic\ Area\ during\ same\ year \end{array}} \times 1000$$

4.8. General Fertility Rate:

This rate refers to the proportion of the number of children born per 1000 of females, the reproductive or child-bearing age. The formula for such a rate is

$$General\ Fertility\ Rate = \frac{\begin{array}{c} Number\ of\ live\ births\ which\ occured\ among\ the\ population \\ of\ a\ given\ geograpic\ Area\ during\ given\ year \end{array}}{\begin{array}{c} Mid-year\ female\ population\ of\ the\ specific\ age\ group \\ 15\ to\ 49\ in\ given\ geograpic\ Area\ during\ same\ year \end{array}} \times 1000$$

The general fertility rate shows how much the women in child-bearing ages have added to the existing population through births. It takes into account the sex composition of the population and also the age composition to a certain extent. Yet it is calculated without proper regard to the age composition of the female population in the child-bearing ages.

4.9. Total Fertility Rate:

The total fertility rate is the mean number of children which a female aged 15 can be expect to bear if she lives until at least the age of 50, provided she is subject to the given fertility conditions over the whole of her child-bearing period.

In order to calculate the total fertility rate we shall have to calculate specific fertility rates and then add them. Total fertility rate is thus the sum of the age-specific fertility rates from a given age to the last point of child-bearing age of female. We define a total fertility rate is,

$$Total\ Fertility\ Rate = t \times \sum \frac{Annual\ births\ to\ females\ aged\ x\ and\ under\ (x+5)}{Mean\ number\ of\ females\ aged\ x\ and\ under\ (x+5)} \times 1000$$

where t – the magnitude of the age class.

4.10. Life Tables

A life table composed of several sets of values showing how a group of infants all supposed to be born at the same time and experiencing unchanging mortality conditions would gradually die out. In other words, the life table is a convenient method for summarising the mortality experience of any population group that is, it provides concise measures of the longevity of that population. Such table are usually worked out after each decennial census to represent mortality conditions either during the previous decennium or during shorter periods covering the date of census. A life table can also be constructed to show how a group of babies would die if, hypothetically, one or more causes of death were eliminated.

Some of the Specific Uses of Life Table:

1. The preparation of population projections by age and sex
2. Analysis of effects of mortality on the age and sex composition of a population.
3. Comparisons of summarizing measures of mortality, as the life-table death rate to expectation of life at various ages, etc.
4. Computation of net reproduction rates; and
5. The appraisal of the accuracy of census enumerations and vital registration data.

In addition, life table techniques have been applied to the analysis of other types of demographic data; for example, in the computation of probability of marriage, specific age, and sex, on the basis census data classified by marital status.

4.11. Construction of Life Table:

Suppose we have the value of q_x for every x from 0 upwards. We can then start with cohort –say one of 100,000 (l_0)births. Multiplying l_0 by q_0, we get $l_0\, q_0 = d_0$. Then the $l_1 = l_0 - d_0$. Again $d_1 = l_1 d_1$, $l_2 = l_1 - d_1$, and so on. Having obtained the values in the

l_x column, we can then fill in the other columns viz. L_x, T_x (for which we start from the bottom of the table and get the values successively by using the relation $T_x = L_x + T_{x-1}$) and e_x^0, by means of the relation stated above.

If the probability that a person belonging to the age-group x to $x+1$ will die while in that age-group is denoted by $m_x^{'}$ then

$$m_x^{'} = \frac{d_x}{L_x} \approx \frac{d_x}{l_x - \frac{1}{2}d_x},$$

$$m_x^{'} = \frac{2q_x}{2-q_x},$$

$$where,\ q_x = \frac{2m_x^{'}}{2+m_x^{'}}$$

The probabilities $m_x^{'}$ are estimated by the observed age-specific death rate (m_x) for the community, where we now take

$$m_x = \frac{D_x}{P_x}(without\ the\ multiplier\ 1{,}000)$$

Hence the q_x values can be determined, if the m_x values are known,

$$q_x = \frac{2m_x^{'}}{2+m_x^{'}}$$

For the early years of life, the values of m_x are usually not so reliable owing to defects in census records. Besides, the assumption underlying that deaths are distributed uniformly over the years of age is not valid for the early ages, especially for age 0: mortality is generally very high in the first few weeks after birth and then it diminishes sharply. It is therefore, necessary to have alternative formulae for q_x for say $x = 0,1\ and\ 2$. We shall consider an alternative formulae for q_0 based on registration data alone. Here the assumption will be made that the effect of migration is negligible, which probably legitimate at age 0.

Note that in order to survive the first year of age, a child must survive till the end of the calendar year in which it is born and then live long enough in the next calendar year to attain exact age 1. Hence, denoting the probabilities of these two events by $p^{'}\ and\ p^{''}$, respectively, we have

$$p_0 = p' p''$$

The probabilities of p' *and* p'' are estimated by $(B_0 - D''')/B_0$

and $(B_{-1} - D' - D'')/(B_{-1} - D)$,

respectively, where

B_{-1} – number of children born in the preceding calendar year.

B_0 – number of children born in the current calendar year.

D' – number of children born and deceased in the preceding calendar year.

D'' – number of children born and deceased in the preceding calendar year and deceased in the

current calendar year before reaching age 1.

D''' – number of children born and deceased in the current calendar year.

If p_x persons of age x 1.b.d found alive in the community at the middle of a calendar period are the survivors out of, say, N_x persons alive at exact age x, then we should have, as the formula for estimating q_x,

$$q_x = \frac{D_x}{N}$$

However, N_x will commonly be an unknown quantity but related to P_x by the approximate relationship

$$P_x = N_x - (1 - \alpha_x) D_x$$

Hence, we have

$$q_x \approx \frac{D_x}{[P_x + (1 - \alpha_x) D_x]}$$

or

$$q_x \approx \frac{m_x}{[1 + (1 - \alpha_x) m_x]}$$

In case $\alpha_x = \frac{1}{2}$, that is, in case the deaths occurring in an age-interval may be assumed to be uniformly distributed over the interval.

Illustrations:)

1. Compute the specific fertility rate, general fertility rate and total fertility rate from the data given below

Age group (year)	**No. of women ('000)**	**No. of live births**
15-19	25	800
20-24	20	2400
25-29	18	2000
30-34	15	1500
35-39	12	500
40-44	6	120
45-49	4	10
Total	100	7330

It is given that out of 7330 the number of female birth was 4000.

SOLUTION:

Computation of specific fertility rate(S.F.R):

S.F.R (for age 15-19 years) = (800/25000)x1000 =32.00

S.F.R (for age 20-24 years) = (2400/20000)x1000 =120.00

S.F.R (for age 25-29 years) = (2000/18000)x1000 =111.11

S.F.R (for age 30-34 years) = (1500/15000)x1000 =100.00

S.F.R (for age 35-39 years) = (500/12000)x1000 =41.67

S.F.R (for age 40-44 years) = (120/6000)x1000 =20.00

S.F.R (for age 45-49 years) = (10/4000)x1000 =2.50

The computation of General Fertility Rate (G.F.R) is

$$General Fertility\ Rate = \frac{Number of\ live\ births}{No.\ of\ women of\ 15\ to\ 49\ years} \times 1000$$

$$General Fertility\ Rate = \frac{7330.00}{100000} \times 1000 = 73.3$$

It is clear from the above that for 15-19 age group S.F.R is 32. Accordingly 1000 womens exactly aged 15 would by the time they reached 20 have borne 32 x 20 =160 children. It is necessary to multiply by 5 since the specific fertility rate is a rate per annum and by the time the womens reach the age 20 they will have spent 5 years in the age group 15-19.

In the table below is shown the number of births which 1000 womens will have borne by the time they reach certain ages.

Exact Age (years)	**S.F.R x 5**	**Total births per 1000 Womens aged 15 by stated ages**
15	0	0
20	32 x 5 =160.00	160.00
25	120 x 5=600.00	760.00
30	111.11 x 5=555.55	1315.55
35	100 x 5=500.00	1815.55
40	41.67 x 5=208.35	2023.90
45	20 x 5=100	2123.00
50	2.5 x 5=12.5	2136.40

$$Total Fertility\ Rate = \frac{\sum S.F.R.}{1000} = \frac{21364}{1000} = 2.1364$$

Question Bank;

1. Explaining the purpose and procedure for standardizing death rates.
2. Explain the procedure for crude fertility rate and total fertility rate.
3. Explain crude and standardized death rates. In what way is standardized death rate is superior to crude death rate?

4. Define life table. Explain how would complete a life table given the death rate at each individual age.
5. Explain the procedure for construction of life table.
5. Compute the general fertility rate and total fertility rate from the data given below

Age group (year)	**No. of women ('000)**	**Specific fertility rate (per 1000)**
15-20	100	15
20-25	120	100
25-30	110	120
30-35	105	140
35-40	100	80
40-45	80	50
45-50	70	10

6. Calculate the general fertility rate, total fertility rate from the following data, assuming that every 100 girls 106 boys are born.

Age of Women	**Number of women**	**Age-specific fertility rate (per 1000)**
15-19	212619	98.0
20-24	198732	169.6
25-29	162800	158.2
30-34	145362	139.7
35-39	128109	98.6
40-44	106211	42.8
45-49	86753	16.9

7. Calculate the crude and standardized death rates for the local population from the following data.

Age-group(years)	A		B	
	Population	**Deaths**	**Population**	**Deaths**
0-10	600	18	400	16
10-20	1000	5	1500	6
20-60	3000	24	2400	24
60-100	400	20	700	21

UNIT - 5

OFFICIAL STATISTICS

5.1 What is official statistics?

Official statistics are statistics published by government agencies or other public bodies such as organizations. Official statistics make qualitative and quantitative information on economic and social development accessible to the public, allowing the impact of government policies to be assessed, thus improving accountability.

The Fundamental Principles of Official Statistics were adopted in 1992 by the United Nations Economic Commission for Europe (UNECE), and subsequently endorsed as a global standard by the United Nations Statistical Commission.

5.2 Historical perspective

In 1949, GOI established Central Statistical Unit in the Cabinet Secretariat and Prof. P C Mahalanobis was appointed its first Hon. Statistical Adviser. Later in May 1951 , it was Converted into Central Statistical Organisation and Prof. P V Sukhatme was appointed Statistical Adviser in the Ministry of Agriculture. A National Income Committee was appointed in the M/o Finance, in 1949 to work out a system for reliable estimation of national income. In 1954, National Income Unit was transferred to CSO. In April 1960, a separate Department of Statistics (DOS) was set up in the Cabinet Secretariat and the CSO and NSS became its part. In 1967, the Computer Centre was setup for data processing, as an attached office of DOS. In 1973, the Department of Statistics became a part of the newly created Ministry of Planning. In October 1999, the Department of Statistics (DOS) and the Department of Programme Implementation were merged and named as the Ministry of Statistics and Programme Implementation (MoSPI).

5.3 Present Statistical System In India

In India, Official statistics are collected and produced by National Statistics Office (NSO) under the Ministry of Statistics And Programme Implementation (MOSPI). The ministry is responsible for producing and disseminating official statistical information, providing the highest quality data. Quality in the context of official statistics is a multi-faceted concept, consisting of components such as relevance, completeness, timeliness, accuracy, accessibility, clarity, cost-efficiency, transparency, comparability and coherence.

At the centre, some ministries like agriculture, water,health, etc have full fledged statistical divisions. Large scale statistical operations like population census, Annual Survey Of Industries (ASI), Economic Census, etc are generally centralized. ISS and SSS officers perform the statistical activities in important ministries. CSO under MOSPI is the nodal agency for a planned development of the statistical system in india.

At the state level is the Directorate (formerly Bureau) of Economics and Statistics (DES), which is formally responsible for the coordination of statistical activities in the State.

5.4 Present Official Statistics System in India

Flow of data is upwards from village to block to district to State Government departments, and from there to the corresponding ministries at the Centre.

The Ministry of Statistics and Programme Implementation (MOSPI) came into existence as an Independent Ministry on 15.10.1999 after the merger of the Department of Statistics and the Department of Programme Implementation. The Ministry has two wings, one relating to Statistics and the other Programme Implementation.

The Statistics Wing called the National Statistical Office (NSO) consists of

- Central Statistical Office (CSO)
- Computer Centre
- National Sample Survey Office (NSSO).

The Programme Implementation Wing has three Divisions, namely

- Twenty Point Programme
- Infrastructure and Project Monitoring and
- Member of Parliament Local Area Development Scheme(MPLADS).

Besides these two wings, there is National Statistical Commission (NSC) created through a Resolution of Government of India (MOSPI) and one autonomous Institute, viz., Indian Statistical Institute declared as an institute of National importance by an Act of Parliament.

The Ministry of Statistics and Programme Implementation attaches considerable importance to coverage and quality aspects of statistics released in the country. The statistics released are based on administrative sources, surveys and censuses conducted by the Centre and State Governments and non-official sources and studies. The surveys conducted by the Ministry are based on scientific sampling methods. In line with the emphasis on the quality of statistics released by the Ministry, the methodological issues concerning the compilation of national accounts are overseen by Committees like Advisory Committee on

National Accounts, Standing Committee on Industrial Statistics, Technical Advisory Committee on Price Indices etc. The Ministry compiles datasets based on current data, after applying standard statistical techniques and extensive scrutiny and supervision.

5.5 Functions of NSO

NSO is mandated with the following responsibilities:-

- Acts as the nodal agency for planned development of the statistical system in the country, lays down and maintains norms and standards in the field of statistics, involving concepts and definitions, methodology of data collection, processing of data and dissemination of results;
- Coordinates the statistical work in respect of the Ministries/Departments of the Government of India and State Statistical Bureaus (SSBs), advises the Ministries/Departments of the Government of India on statistical methodology and on statistical analysis of data;
- Prepares national accounts as well as publishes annual estimates of national product, government and private consumption expenditure, capital formation, savings, estimates of capital stock and consumption of fixed capital, as also the state level gross capital formation of supra-regional sectors and prepares comparable estimates of State Domestic Product (SDP) at current prices;
- Maintains liaison with international statistical organisations, such as, the United Nations Statistical Division (UNSD), the Economic and Social Commission for Asia and the Pacific (ESCAP), the Statistical Institute for Asia and the Pacific (SIAP), the International Monetary Fund (IMF), the Asian Development Bank (ADB), the Food and Agriculture Organisation (FAO), the International Labour Organisation (ILO), etc.
- Compiles and releases the Index of Industrial Production (IIP) every month in the form of 'quick estimates'; conducts the Annual Survey of Industries (ASI); and provides statistical information to assess and evaluate the changes in the growth, composition and structure of the organised manufacturing sector;
- Organizes and conducts periodic all-India Economic Censuses and follow-up enterprise surveys, provides an in-house facility to process the data collected through various socio-economic surveys and follow-up enterprise surveys of Economic Censuses;

- Conducts large scale all-India sample surveys for creating the database needed for studying the impact of specific problems for the benefit of different population groups in diverse socio-economic areas, such as employment, consumer expenditure, housing conditions and environment, literacy levels, health, nutrition, family welfare, etc;
- Examines the survey reports from the technical angle and evaluates the sampling design including survey feasibility studies in respect of surveys conducted by the National Sample Survey Organisation and other Central Ministries and Departments;
- Dissemination of statistical information on various aspects through a number of publications distributed to Government, semi-Government, or private data users/ agencies; and disseminates data, on request, to the United Nations agencies like the UNSD, the ESCAP, the ILO and other international agencies;
- Releases grants-in-aid to registered Non-Governmental Organizations and research institutions of repute for undertaking special studies or surveys, printing of statistical reports, and financing seminars, workshops and conferences relating to different subject areas of official statistics.

5.6 Central Statistics Office (CSO)

The Central Statistical Office which is one of the two wings of the National Statistical Organisation (NSO) is responsible for coordination of statistical activities in the country and for evolving and maintaining statistical standards.

Type of Activities

- compilation of National Accounts;
- conduct of Annual Survey of Industries(ASI) and Economic Censuses,
- compilation of Index of Industrial Production(IIP), as well as Consumer Price Indices(CPI).
- deals with various social statistics, training, international cooperation, Industrial Classification etc.

The CSO is headed by a Director-General who is assisted by 5 Additional Director-Generals looking after the National Accounts Division, Social Statistics Division, Economic Statistics Division, Training Division and the Coordination and Publication Division.

Divisions of CSO

- National Accounts Division(NAD)
- Social Statistics Division(SSD)
- Economic Statistics Division(ESD)
- Coordiantion and Publication Wing(cap)
- Training Division

National Accounts Division (NAD) Including Prices and Cost of Living Units Activities:

- Compilation and release of national accounts including national income, government /private final consumption expenditure, capital formation ,savings and input-output tables.
- Compilation and release of quarterly and advanced estimates of national income and its expenditure components.
- Compilation and release of CPI(Consumer Price Index)and CFPI(consumer food price index) for rural, urban and combined.
- Conduct of training/workshops for the officers of the state directorates of economics and statistics and provide the necessary support on the matters related to state income and regional accounts.

Social Statistics Division (SSD) Activities

- Statistical monitoring of the millenium development goals(MDG), SAARC development goals, environmental economic accounting, National/International awards for statisticians,National Data Bank(NDB) on socio religious categories, ,time use survey
- DEVELOPMENT of data bases, evolving new statistical tools / surveys to bridge data gaps
- Monitoring of Basic Statistics For Local Level Development(BSLLD) pilot scheme
- Release of regular and adhoc statistical publications,and grant in aid for research,workshop/ seminars/conferences in official/applied statistics.

Economic Statistics Division(ESD) Activities

- Conduct of Annual Survey of Industries and Periodic Economic Censuses
- Compiling All-India Index of Industrial Production (IIP)
- Guidance to States on Compiling State-level IIP
- Compiling Indices of Services Production
- Bringing out Industrial and Product Classifications
- Release of Infrastructure Statistics
- Release of Energy Statistics

7.7 National Statistical Systems Training Academy (NSSTA)

This division is primarily responsible for the training of statistical personnel in official statistics and data management to tackle the emerging challenges of data collection, collation, analysis and dissemination required for evidence based policy making as also for planning, monitoring and evaluation. National Statistical Systems Training Academy (NSSTA) under training division (CSO) of the Ministry, is a premier Institute fostering human resource development in Official Statistics, came into existence on 13th February, 2009.NSSTA is a Centre of Excellence in imparting training on Official Statistics and undertaking research activities thereto.

Activities of NSSTA:

- To train statistical as well as non-statistical manpower in undertaking monitoring and evaluation of large scale programmes/projects, through specialized short/medium term training programmes.
- To create a pool of trainers and develop training material in consultation and collaboration with academicians, researchers and professionals from universities, external professional institutions, UN/bilateral agencies so as to decentralize training to the state and sub-district level.
- To facilitate and organize International training programmes sponsored by Statistical Institute for Asia and Pacific(SIAP), Economic And Social Commission For Asia And Pacific(ESCAP), United Nations Statistics Division(UNSD),FAQ and other multilateral and bilateral donor development agencies.

- To support the central and State Governments with necessary consultation in matters relating to training of its statistical personnel, designing and conducting surveys & research and for building capacity of their statistical institutes and entities.
- To introduce in near future certificate/diploma courses duly recognized through distant-cum-in house efforts in the field of theoretical and official statistics.

5.8. National Sample Survey Office (NSSO) Activities

- Socio economic surveys viz. consumer expenditure, employment-unemployment, debt and investment etc by conducting nation wide household surveys to assess the socio-economic condition of the country.
- Annual Survey Of Industries(ASI) to assess the industrial growth of the organized manufacturing sector of the country.
- Enterprise surveys to fill the data gaps in respect of the unorganized sector of the economy.
- Collection of data on urban and rural prices for price monitoring and generating various price indices.
- Supervision of the area enumeration and crop estimation surveys of the state agencies for improvement of crop statistics.
- Bringing out the technical journal of NSSO entitled "SARVEKSHANA".

NSSO Divisions:

1. Survey Design and Research Division (SDRD)
2. Field Operations Division (FOD)
3. Data Processing Division (DPD)
4. Co-Ordination and Publication Division (CPD)

Survey Design and Research Division (SDRD) is a professional organ of NSSO, mandated to do the job of:

- Planning of the survey
- Formulation of sample design
- Drawing up of schedules of inquiry
- Formulation of concepts and definitions
- Preparation of instruction manual for survey field work

- Training of field and data processing personnel on survey methodology
- Formulation of scrutiny check points
- Drawing up of tabulation programme
- Preparation of survey reports
- Analysis and presentation of survey results and
- Undertaking studies for the improvement of survey methodology

SDRD, NSSO is located at Mahalanobis Bhavan, Kolkata and is headed by an Additional Director General - a Higher Administrative Grade (HAG) level officer, and has sanctioned strength of three SAG (Senior Administrative Grade), fifteen JAG (Junior Administrative Grade), eight STS (Senior Time Scale) and four JTS (Junior Time Scale) level officers of Indian Statistical Service besides one Deputy Director (Administration) and the supporting staff members.

The **Field Operations Division (FOD)**, one of the four Divisions of the National Sample Survey Office, is responsible for conducting surveys in the field of Socio- Economic, Industrial Statistics, Agricultural Statistics, Prices, etc. as per the approved programmes.

Agricultural Statistics: It provides technical assistance to the states in adopting statistical techniques to obtain reliable estimates of crop yield, uniformity in the concepts and definitions and provide assistance to the state field personnel. It also consolidates the survey results of General Crop Estimation Survey conducted by the state governments for the yield estimation.

Urban Frame Survey: It consists of preparation and updation of mutually exclusive and identifiable urban blocks with a population strength of 600-800 and 120-180 households in each block. These blocks form the first stage units of the socio economis surveys of NSSO.The frame is updated after every five years.So the formation of boundaries of such blocks is the unique work of FOD.

Industrial Statistics: FOD conducts ASI wherein it collects data on various aspects of the registered manufacturing sector of the economy such as employment,wages,input,output,capital formation etc on annual basis for use in the estimation of national income.

Price Statistics: FOD conducts market survey in almost 59 urban centres to collect prices of the commodities consumed by the urban population for calculation of CPI(urban) .

5.9 National Income Statistics

National income may be defined as the value of commodities and services produced by the nationals of a country during a given period, counted without duplication. It consists of (a) the net domestic product (NDP) and (b) the net income earned from abroad, The NDP is the unduplicated without originating within the country. This can be obtained in three ways:

(i) Adding up the value of gross output of all producers and deducting from the total the purchases made by these producers from other producers (i.e., the value of intermediate products) and the depreciation of equipment used up in the process of production. A net figure of the kind can be obtained from each producer separately and represents the value added by him to the value of intermediate product which be production.

(ii) Adding up the wages, profits and other forms of income that accrue in productive activity. The sum-total value of the commodities and services is then obtained by adding up various incomes accrued.

(iii) Aggregating all final products available for consumption or for investment and adding up the corresponding values leading again to the same total.

These three approaches to the estimation of national income are called, respectively, the products approach, the income approach and the expenditure approach. Any one of these approaches or a combination may be applied for estimating national income.

5.10. Agriculture Statistics:

The statistics of crop yields are obtained in respect of principle crops by two methods.

1) Traditional Method

2) Random sampling method.

1) Traditional Method:- The Traditional method depends on eye estimates either directly or by multiplying normal yield with condition factor. The normal yield or standard yield is defined by the government as average yield on an average soil in a year of average character. In other words, the normal yield is the yield which a farmer has right to expect in the normal conditions. The condition factor is a factor less than or equal to unity and represents the condition of the crop compared to the normal one. Normally the condition factor is estimated twice in a season, once during the growth of the crop and other at the time of harvesting.

2) Random sampling method:-Under the method of Random sampling the yield per hector or acre is estimated objectively on the basis of crop cutting experiments conducted on randomly selected plots in a random sample of fields. This is at present most popular method in all the stages of the country and is free from subjective bias as in traditional method.

Crop forecast Uses:

Forecast of crop production assumes great importance for a country for ensuring sufficiency of food grains and their equal distribution in different areas of the country.

Crop forecast gives the advance warning in case of a bad crop and enable the government to have future planning to meet the possible shortages.

In case of commercial crops like cotton, jute and sugarcane forecasts are useful for trade and industry, to make availability of raw materials during the season.

Method of crop forecast:

Forecast of the crop production are made while crop is standing in the field. There are so many methods of crop forecasting.

One method is the production of a crop is determined by multiplying the area under the crop with the avg yield per hectare.

Where Avg yield = Normal yield x Condition factor

Where normal yield is the yield which a farmer has right to expect in the normal conditions and condition factor is a factor less than or equal to unity depending upon the abnormal conditions or normal conditions.

5.11 Industrial statistics

Statistics of industrial production in India may be considered under two heads.(a) statistics relating to the factory sector and (b) statistics relating to the non-factory sector. The factory sector covers industrial units registered under the Factories Act, 1948. The non-factory sector covers household and non-household units which are not registered under the said Act. The factory and the non-factory sector are also designated as, respectively, the organised and the unorganised sector.

Organized sector

The principal sources of data relating to this sector are the Annual Survey of Industries (ASI), the monthly returns received by the Directorate-General of Technical Development (DGTD) from units registered with its various Directorates and the data-collection systems of organisations like the Directorate of Sugar, the office of the Textile Commissioner, the office of the Jute Commissioner, the Ministry of petroleum, the SAIL etc.

In order to collect industrial data, a census used to be taken every year beginning from 1944. For the purpose of the census, called the Census of Manufacturing industries (CMI),the manufacturing industries were divided into 63 groups, but data were collected for 29 of these groups only. Again, the census was confined to factories employing 20 of more workers and using power. Factories under the control of the Ministry of Defence were excluded form the purview of the census. The census related to the calender year (1 January to 31 December), except in the case of the sugar industry for which the year July-June was used. For each industry, the quantities, and values of different products and by-products ,different kinds of fuel and materials consumed in value terms as well as in quantity terms, details of employment together with wages and salaries, component-wise fixed and working capital employed, etc., were recorded. The data were available in published form for the whole of India and also for each State separately. From 1950 onwards, a sample survey of manufacturing industries (SSMI)also used to be conducted every year on the recommendation of the National Income Committee. This survey covered all the 63 groups of industries and in each group data were initially collected for factories employing 20 or more workers on any day during the year and using power. The items of information considered were value of output, capital employed, total value of inputs, employment, wages and salaries etc.

Since 1959, the Annual Survey of industries (ASI) has replaced both the CMI and the SSMI. Conducted under the statutory provisions of the Collection of Statistics Act, 1953 and the Collection of Statistics, (central) Rules. 1959 in all the States (except Jammu and Kashmir where it is conducted under the J. & K. Collection of Statistics Rules, 1964).The ASI covers factories engaged in manufacturing, units engaged in the production and distribution of gas and water as well as those engaged in sanitary services, motion picture production, laundering, Job-dyeing and biri and cigar making are covered by the ASI. Factories under the control of Ministry of Defence as well as those engaged in oil storage and distribution,technical training institutes,hotels and cafes are, however,kept out of its purview. Large industrial establishments contributing substantial part of value added and other economic factors (exact frame changing form time to time) are completely enumerated. The rest of the establishments coming under Factory Act, for each industry is sampled, the sampling fraction and design changing from time to time. However, fhe final estimate for all the ultimate industries is based on both the census and sample data. Electricity undertakings, irrespective of the number of employees, are, however, completely enumerated. Data collected in any year relate to the operation of the units during the previous year. Data are collected on the principal characteristics, viz. Number of factories, fixed capital, invested capital, outstanding loans, number of workers, man-days worked, wages and salaries, fuels consumed, total inputs, products, total output, depreciation,value added and net income. Summary results of the survey are made available within six months after the completion of field work; but detailed results are published with a considerable time lag.

The DGTD covers units registered with its Directories in all the industries except iron and steel, sugar, tea, coffee , vanaspati, cotton textiles, jute textiles and petroleum, etc. While the entire factory sector is covered by the ASI, the number of units covered by the DGTD is about 6,000, which represent large- and small-scale units each having a minimum investment of Rs. 10 lakhs in plant and machinery (excluding land and buildings) for general and Rs.15 lakhs for ancillary industries. The data are collected on a monthly basis and consolidated statements on production, stocks, etc., in respect of some 400 items are sent to the CSO every month. Similar statements are also received from the Textile commissioner, Jute Commissioner, |Ministry of Petroleum, Steel Authority of India, etc.

The CSO issues a monthly index of industrial production along with the production data for the index items.

Unorganized sector

The position in respect of this sector is unsatisfactory, for there is no provision for data-collection on a regular basis. One has to depend on the data thrown up by some and hoc surveys for information on this sector.

An attempt was made in 1971 to conduct a comprehensive survey, but is was confined to urban areas. A complete listing of units in the non-factory sector was done and detailed information collected from units employing five or more workers. The survey was conducted more or less on a census basis, information was collected on number of factories using power, gross fixed assets owned, net assets owned, working capital, employment, emoluments, inputs consumed, products and by-products, total output, gross value added,loan due at the end of year, owned capital stock at the end of the year. The state Industries Development Organisation (SIDO) conducted quick surveys of units in a few selected industries in 1971 and also in 1972 covering both the factory and the non-factory sector. Information was collected an investment, employment, material consumption, capacity, production and exports.

A census of small-scale industrial units with 1972 as reference year was conducted in 1973-74 by the Development Commissioner, Small Scale Industries (DCSSI). It was restricted to small-scale units registered with the Directorate of Industries that came under the purview of the DCSSI and those under the modern small-scale sector (which, by definition, excluded small-scale units falling within the jurisdiction of specialised Boards and Agencies) . The DCSSI started collection of data on small-scale units from 1976 on a two percent sampling basis.

During the population census of 1961,1971,and 1981, data were collected in respect of census houses used as factories and workshops on registration particulars, description of the product, employment size and type of power-separately for rural and urban areas. The census thus provided data on registered household industries functioning in both rural and urban areas for each district.

The NSSO also survey the unregistered sector at the national level and collect data from household enterprises as a part of their multipurpose surveys in some of the rounds. In the 7^{th} to 10^{th} rounds (1953-56) data were collected in respect of small manufacturing units. In the 23^{rd} round (1968-69), all the household and non-household manufacturing units that were not registered under the Factories Act were covered, so that the data on the principal characteristics for the unregistered sector could be aggregated to the data for the registered sector, In the 29^{th} round, which covered the whole of the unorganised sector of manufacturing excluding the factory sector covered by the ASI , the NSSO conducted a household enquiry on self-employment in on-agricultural enterprises.

The responsibility for data-collection on certain segments of the unorganised sector rests with the All-India Handicrafts Board, All-India Handloom Board, Central silk Board Coir Board and Khadi and villages Industries Commission. These bodies, conduct industry and area surveys to meet their immediate data needs . An indirect method is used by the Handloom Board and Coir Board in estimating production. The production of the handloom industry is estimated on the basis of mill yarn supplied of weavers. On the other hand, the annual output of coir and coir goods is estimated from the movement of coir by rail and coastal steamers.and allied data.

In 1977, the CSO conducted a census throughout t the country, however, excluding Sikkim, Lakshadweep, Nagaland and some pockets of Jammu and Kashmir, of all establishments engaged in non-agricultural activities.

The important publications relating to industrial statistics are:

(i) Annual Survey of industries (Census Sector), Vols . I-X

(ii) Annual Survey of Industries –Summary Results for Factory Sector.

(iii)Annual Survey of Industries-Summary Results for Census Sector.

(iv)Monthly statistics of the Production of Selected Industries in India.

All these are brought out by the Industrial Statistics Wing (ISW) CSO. We should also mention the occasional publication Statistics for Iron and Steel Industry in India, Which provides comprehensive data on various aspects of the iron and steel industry and related information. While part A provides data on the production, consumption number of employees, etc., in respect of the Industry within the country Part B deals with details of world production, consumption, employment, etc., in the industry. Started by the Hindusthan Steel Ltd. In 1964, it is now brought out by the Steel Authority of India Ltd (SAIL).

The Indian Minerals Year-Book (annual), brought out by the Indian Bureau of Mines (IBM), supplies comprehensive data on the Indian minerals industry,including all-India and State-wise indices of mineral production, and also describes the major developments during the year in mineral producing countries of the world. Mineral Statistics of India (quarterly), also issued by the IBM, presents basic data on mineral production and value, metal production, external trade in and prices of minerals and metals, etc., The annual publication Statistics of Mines in India.Vol.I (coal and Vol.II (non-coal), brought out by the Director-General of Mines and the Indian Coal Statistics. For data on the coal Industry, there is also the Monthly Coal Bulletin, published by the Coal Controller.

5.12 Trade statistics

Trade statistics are collected as a by-product of industrial activity in India. These may be classified into two groups corresponding to (a) external trade and (b) inland trade-wholesale and retail (including coastal trade).

External trade

There are three types of administrative activity in the case of external trade,viz,licensing, actual shipment/arrival of goods and receipt/remittance of payments. These three give rise to three corresponding types of statistics of external trade,viz.licensing statistics,balance of trade statistics and balance of payment statistics. The agencies responsible for these three types are, respectively, the office of Chief Controller of Imports and Exports (CCIE), the Directorate General of Intelligence and Statistics (DGCIS) and the Reserve Bank of India (RBI) .

Licensing Statistics cover all the items falling under the purview of the important and export trade control. The import and export licenses issued by the CCIE and his regional offices in accordance with the import and export trade control policy of the Government of India are used to compile such statistics. The important publications of licensing statistics are:

(i) Weekly Bulletin of Important Licenses, Export Licences and Industrial Licenses

(ii) Annual Administration Report for the Import and Export Trade Control Organisation.

The first publication has six sections relating to important licences issued, release orders issued, from receiving export licences and licences issued under the Industries Development and Regulation, Act, 1951. The second publication reviews the work done by the impact and Export Trade Control Organisation during the relevant year. Part I describes the various aspects of the import and export trade polices and the working the Export Promotion councils. Part II contains statistics for the current year as well as the previous year on: the number of and value of import licences and release orders issue: category-wise value of important licenses according to office of issue: agency-wise break-up of release orders issued by the various offices; value of export licences issued for selected commodities; imports, value (including re-exports) and balance of trade; overall balance of payments; actual imports of items placed under OGL; class-wise distribution of exports; receipts, disposal and pendency of export licence applications and other receipts.

The primary sources of statistics of India's balance of trade are; (i) the customs authorities at the various sea ports and air ports, land customs stations and inland waterways in respect of trade with Bangaladesh: (ii) the foreign postal authorities: and (iii) border check-posts along the Indo-Nepal border. According to the Sea Customs Act, 1962, movement of all merchandise from or into the country has to be with the prior written permission of the customs authorities. On the basis on the declarations made by traders at the time of seeking such approach, the customs authorities prepare daily trade returns showing full particulars of each consignment exported or re-exported from or important into the country. Again, according to the present rules, each letter or parcel sent by foreign post which contains merchandise for export must be accompanied by the 'customs declaration slips. A Duplicate copy of the slip is sent to the DGCIS. Besides, the foreign postal authorities also furnish a monthly statement showing imports by foreign post for compiling data on imports. Data on the overland trade between India and Nepal are furnished in the form of monthly returns by border check-posts along the Indo-Nepal border.

Statistics relating to balance of trade are available in the following DGCIS publications:

(i) Monthly Staistics of Foreign Trade of India-Vol. I (Exports and Re-exports) and Vol. II (imports) , providing monthly and cumulative (from April to that month of the financial year) export import figures from / to India by commodity, and country at the 8 digit level of ITC codes.

(ii) Statistics of the Foreign Trade of India by Countries (monthly). Like the first publications, this too appear in two volumes. Vol.I contains details for exports and re-exports, while Vol.II gives similar information for imports. They provide monthly and cumulative export/import figures from / to India by commodity and for three consecutive months in one issue. Each volume present data in three separate tables. Table I shows the share of each of the 16 economic regions in the exports (including re-exports) form and imports into India. Table 2 gives the shares of different countries of the world in exports (including re-exports) from and imports into India. Table 3 provides the details of commodity-wise exports (including re-exports) from and imports into India from various countries.

(iii) Foreign Trade Statistics of India (monthly) providing monthly and cumulative figures for India's export/import of principal commodities by countries along with the corresponding figures of the previous year.

(iv) Monthly Brochure on Foreign Trade Statistics of India (Principal Commodities and Countries). This monthly brochure is published in order to meet immediate requirement of the Ministry of Commerce as well as order organisations. The brochure gives Indian exports/imports of principal commodities by countries and by principal ports.

(v) Indian Trade Journal (weekly). It presents, among other data relating to trade, commerce and industry, statistics of overland trade with Nepal. Tibet and Bhutan. It also contains periodical reports from Indian trade representative abroad.

For dissemination of the export/import data by economic zones these are provided at present in magnetic tapes/floppies or in computer hard copies only in respect of a specified and selected number of items for some specified periods for which data are finalised.

The primary sources of statistics on balance of payments are the GR forms which the traders are required to submit to the banks for the purpose of remitting money to or receiving money from abroad for imports made from or exports made to those countries.

These statistics are compiled by the RBI and published in the following.

(i) Reserve bank of India Bulletin (monthly) and

(ii) Report on Currency and Finance (annual).

5.13 Statistics of Labour and Employment

The decennial population census constitutes the main source of information on the economically active population of the country. A large mass of data on such items as age and sex-composition of workers, their rural-urban distribution, and their industrial and occupational classification flow out of the census. An indication of the magnitude of unemployment is also available from the census data.

The NSSO has been conducting surveys of employment regularly once in every five years 1972-73. The latest surveywas conducted in the 37^{th} round (1982-83) of the NSS.

The Employment Market Information (EMI) programme of the Directorate-General of Employment & T raining (DGET),in terms of the provisions of the Employment Exchange (Complsory Notification of vacancies) Act, 1959,makes available data on the organised sectors of the economy. The programme covers all public sector establishments (except for the defence establishments na armed forc es) and those private sector establishments which employ at least 25 persons on any day during the given-quarter. Beginning from 1966, private sector establishments employing 10 to 24 persons also being covered on a voluntary basis. The date collected under the programme are presents in the Quaretly Emplloyment Review (DGET, the Em,plloyment Review (DGET,annual) and the Occupatinal Educational Pattern in India (DGET,biennial),and the last with one series for the Public Sector and another for the Private Sector. The DGET also brings out the Census of Central Government Employees (annual),giving detailed data on gazetted and non-gazetted employees of the Central Government. The DGET data,it is to be noted, do not cover self-employed people,part-time employees, agricultural and allied occupations, household establishments and establishments employing fewer than 10 workers in the private sector. Besides, the coverage of employment in construction, particularly private,is inadequate.

The National Employment service, with nearly 440 employment exchanges, is another source of data on employment. The data relate to job seekers registered with the employment exchanges. But these data suffer from some obvious defects: (1) the registration being voluntary,not all unemployed people register themselves with the exchanges: (ii) a registrant need not necessarily be unemployed;(iii)there is the possibility of multiple registration; and (iv) as the employment exchanges are locates mostly in urban areas, the data do not reflect the magnitude of unemploymne in the rural areas.

The labour Bureau(LB) is the other important source of labour statistics. It collects, compiles and publishes statistics of employment in respect of factories, mines, plantations, shops and commercial establishments, etc., on an all-India basis. Most of these data are obtained as a by-product of the administration of the various Labour Laws operating in the relevant sectors. Information on employment and unemployment of agricultural labour is collected through the Agricultural/ Rural Labour Enquiries conducted at intervals of six years or more. The data so collected appear in the following publications.

(i) Indian Labour Statistics (LB, annual). It presents principal statistics relating to population census economic data, wages and earnings in different sectors, levels of living, industrial disputes, trade unions, industrial injuries, absenteeism, labour turnover and social security.

(ii) Indian Labour Journal (LB, monthly). The first part of the journal contains reports and studies on labour, labour activities in states, labour laws etc. The second part gives monthly statistics relating to prices and price indices, number of manshifts worked and employment in coal mines and cotton textiles mills, employment exchange statistics and statistics of absenteeism in certain industries, and also time series of such data.

(iii) Indian Labour Year-Book (LB). It provides in compact from reviews of labour problems and also the principal statistics on important aspects of labour that are currently available from various sources.

The CSO, through its publications, viz, the Statistical Abstract, Monthly Abstract of Statistics, etc., also gives statistics on number of persons employed by economic activity (in the organised sector), unemployment by occupational group, wages, industrial accidents and disputes.

For the mining sector, data on employment, hours of work, labour productivity, wages, index numbers of wages, industrial accidents, disputes and absenteeism are complied by the Directorate-General of Mines Safety (DGMS). These are published in the Statistics of Mines in India. Vol. I (coal) and Vol. II (non-coal) The Monthly Coal Bulletin (DGMS) presents data relating to labourers in coal mines.

The publication Agricultural Wages in India (DES, annual) presents statistics of wages received by different types of agricultural labour at selected centres during the year.

5.14 Financial and Banking Statistics

Financial statistics can be divided into two groups: (i) statistics relating to banking and insurance, and (ii) statistics relating to public finance.

Banking statistics are compiled and published by the Reserve Bank of India (RBI). The RBI is the note-issuing authority and controls the country's foreign exchange. It is the bank of bankers. To discharge these duties, the RBI collects a large mass of data. These are presented by the RBI in the following publications.

(1) Statement of the affairs of the Reserve Bank of India (weekly)- It gives the data at the close of Friday on the assets and liabilities of the banking and issue departments of the RBI separately, loans and advances made to scheduled banks and state co-operative banks, transactions in foreign currency, clearing house statistics and money rates.

(2) Report of Currency and Finance (annual)-Part |I of the report gives and over-all review of the Indian economy. Part 2 deals in detail with developments in various sectors of the economy. Part 3 contains a wealth of statistical materials on various sectors, including the sector of banking.

(3) Trend and Progress of Banking in India (annual)- It is annual report giving a review of important events in the field of banking during the year. The statistics given here are also available in other RBI publications.

The main publications giving insurance statistics are:

(i) The Indian Insurance Year-Book (Controller of Insurance, Ministry of Finance

(ii) Annual Report of the Life Insurance Corporation of India (LIC).

Public finance statistics are available in the annual budgets of the Central and State Governments. These give a complete account of the respective Governments, Public finance statistics regarding the Railways are separately available from the Railway Budget of the Central Government. The important publications are:

(i) Budget of the Central Government (annual)

(ii) Economic Survey (annual)

(iii) Report on Currency and Finance (annual).

****