BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024

Tamil Nadu, India.

Programme: M.Sc. Statistics

Course Title: Stochastic Processes and Time Series Analysis

Course Code: 23ST02DEC

Unit-IV

Time Series Analysis

Dr. T. Jai Sankar
Associate Professor and Head
Department of Statistics

Ms. I. Angel Agnes Mary
Guest Faculty
Department of Statistics

UNIT – IV

TIME SERIES ANALYSIS

Time Series:

An arrangement of statistical data in accordance with time of occurrence or in a chronological order is called a 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.

Definition:

Time series is an ordered sequence of values of a variable at equally spaced time intervals.

For example, measuring the value of retail sales each month of the year would comprise a time series. This is because sales revenue is well defined, and consistently measured at equally spaced intervals.

Applications of time series:

The usage of time series models is twofold:

- Obtain an understanding of the underlying forces and structure that produced the observed data
- Fit a model and proceed to forecasting, monitoring or even feedback and feed forward control.

Time Series Analysis is useful in various fields such as:

- **Financial Analysis** − It includes sales forecasting, inventory analysis, stock market analysis, price estimation.
- **Weather Analysis** − It includes temperature estimation, climate change, seasonal shift recognition, weather forecasting.
- **Network Data Analysis** − It includes network usage prediction, anomaly or intrusion detection, predictive maintenance.
- **Healthcare Analysis** − It includes census prediction, insurance benefits prediction, patient monitoring and for many applications such as
- Economic and Sales Forecasting
- Stock Market Analysis
- Yield Projections
- Process and Quality Control
- Inventory Studies
- Census Analysis and many more . . .

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, y_t) represents a time series data.

Ex: Production of rice in Tamilnadu for the period from 2010-11 to 2016-17.

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

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 trend is the long term pattern of a time series. A trend can be positive or negative depending on whether the time series exhibits an increasing long term pattern or a decreasing long term pattern. If a time series does not show an increasing or decreasing pattern then the series is stationary in the mean.

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.

Cyclical Variations:

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

Seasonal Variations:

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.

Irregular variations:

This component is unpredictable. Every time series has some unpredictable component that makes it a random variable. In prediction, the objective is to "model" all the components to the point that the only Component that remains unexplained is the random component. 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.

Models of Time Series Analysis

The following are the two models which we generally use for the decomposition of time series into its four components. The objective is to estimate and separate the four types of variations and to bring out the relative effect of each on the overall behavior of the time series.

- Additive model, and
- Multiplicative model

Additive Model

In the additive model, we represent a particular observation in a time series as the sum of these four components.

$$\text{i.e.} \quad O = T + S + C + I$$

Where, O represents the original data, T represents the trend. S represents the seasonal variations, C represents the cyclical variations and I represent the irregular variations.

In another way, we can write

$$Y(t) = T(t) + S(t) + C(t) + I(t)$$

Multiplicative Model

In this model, four components have a multiplicative relationship. So, we represent a particular observation in a time series as the product of these four components:

$$\text{i.e.} \quad O = T \times S \times C \times I$$

Where, O represents the original data, T represents the trend. S represents the seasonal variations, C represents the cyclical variations and I represent the irregular variations.

In another way, we can write

$$Y(t) = T(t) \times S(t) \times C(t) \times I(t)$$

This model is the most used model in the decomposition of time series. To remove any doubt between the two models, it should be made clear that in Multiplicative model S, C, and I are indices expressed as decimal percentages whereas, in Additive model S, C and I are quantitative deviations about a trend that can be expressed as seasonal, cyclical and irregular in nature.

Example:

If in a multiplicative model:

$$T = 500,\ S = 1.4,\ C = 1.20 \text{ and } I = 0.7$$

then $\mathbf{O = T \times S \times C \times I}$

By substituting the values we get

O = 500 × 1.4 × 1.20 × 0.7 = 588

If in additive model,

T = 500, S = 100, C = 25, I = –60

then **O = T + S + C + I**

By substituting the values we get

O = 500 + 100 + 25 – 60 = 565

1. Measurement of Trend

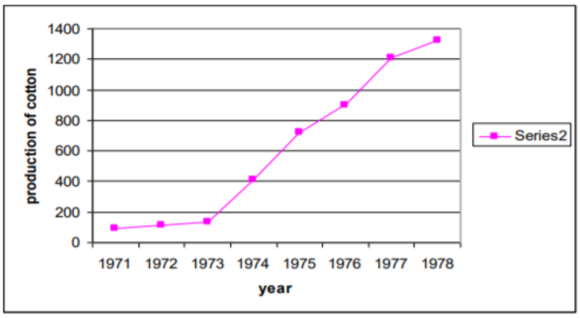
Trend is a long term movement in a time series. This component represents basic tendency of the series. The following methods are generally used to determine trend in any given time series. The following methods are generally used to determine trend in any given time series.

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

(i) Graphic method or eye inspection method

Graphic method is the easiest, simplest and the most flexible method of estimating secular trend. The procedure of obtaining a straight line trend by this method is given below:

- Plot the time series on a graph.
- Examine carefully the direction of the trend based on the plotted information (dots).
- Draw a straight line which will best fit to the data according to personal judgment. The line now shows the direction of the trend.



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.

(ii) Method of Moving Average

Moving average method is a simple device of reducing fluctuations and obtaining trend values with a fair degree of accuracy. In this method the average value of number of years (or months, weeks or days) is taken as the trend value for the middle point of the period of moving average. The process of averaging smoothes the curve and reduces the fluctuation.

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.

Advantages:

- This method is simple to understand 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.

Example

Determine the trend of the following by using 3-yearly and 4-yearly moving average method.

Year	2001	2002	2003	2004	2005	2006	2007	2008
Sales	40	45	40	42	46	52	56	61

Procedure

- Compute the value of first three years (1,2,3) and place the three year total against the middle year (i.e., 2nd year)
- Leave the first year's value and add up the values of the next three years i.e., 2,3,4 and place the three-year total against the middle year i.e., 3rd year
- This process must be continued until the last year's value is taken for calculating moving average.
- The three-yearly total must be divided by 3 and placed in the next column. This is the trend value of moving average.
- The formula calculating 3 yearly moving average is as follows:

$$(a+b+c)/3, (b+c+d)/3, (c+d+e)/3$$

Calculation

Calculation for 3-yearly moving average method

Year	Production	3-yearly moving average
2001	40	
2002	45	41.67
2003	40	42.33
2004	42	42.67
2005	46	46.67
2006	52	51.33
2007	56	56.33
2008	61	

Procedure

- Compute the values of the first four years and place the total in between the 2nd and the 3rd years.
- Leave the first year value and compute the value of the next four years and place the total in between the 3rd and 4th year.
- This process must be continued until the last year is taken into account.
- Compute the first two four-year totals and place it against the middle year (i.e., 3rd year).
- Leave the first four year total and compute the next four-year totals and place in the 4th year.
- This method must be continued until all the four-yearly totals are computed.
- Divide the above totals by 8 (because it is the total of the two four-yearly totals) and put in the next columns. This is the trend values.

Calculation

Calculation for 4-yearly moving average method

Year	production	4-yearly moving average	2-yearly moving average
2001	40		
2002	45		
		41.75	
2003	40		42.5
		43.15	
2004	42		44.12
		45	
2005	46		47
		49	
2006	52		51.38
		53.73	
2007	56		
2008	61		

Result

Thus we get the 3-yearly moving average trend values 41.67, 42.33, 42.67, 46.67, 51.33 and 56.33 which shall be plotted to the years 2002, 2003, 2004, 2005, 2006 and 2007 respectively.

Since, we get the 4-yearly moving average trend values 42.5, 44.12, 47 and 51.38 which is plotted to the years 2003, 2004, 2005 and 2006 respectively.

(iii) 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.

Merits

- It is simple and easier to understand than moving average and least square method
- As the line can be extended both ways, we can get the intermediate values and predict the future values.
- As it does not depend upon personal judgment, everyone who applies this method will get the same trend line unlike the former method.

Demerits

- Under this method, it has an assumption of linear trend whether such a relationship exists or not
- It is affected by the limitation of arithmetic mean
- This method is not enough for forecasting the future trend or for removing trend from original data

Example:

Determine the trend of the following by using semi-average method. Estimate the sales for the year 1984.

Year	1975	1976	1977	1978	1979	1980	1981	1982	1983
Sales	18	24	26	28	33	36	40	44	48

Procedure

- First step the whole data is divided in two equal parts.
- Second step omitting the middle year.
- Then calculate the averages by two parts.
- To calculate annual increment is,

  Annual Increment = $(\overline{x}_2 - \overline{x}_1) / 5$.

- And then calculate half yearly increment = (annual increment/2).
- Finally, calculate the trend values.

Calculation

Here the number of years is 9. The two middle parts will be 1975 to 1978 and 1980 to 1983. The value for the year 1979 will be ignored.

Year	Sales	**Semi averages**	**Trend Values**
1975	18		22.2 – 3.6 = 18.6
1976	24		24 – 1.8 = 22.2
		$\frac{18+24+26+28}{4} = 24$	Center = 24
1977	26		24 + 1.8 = 25.8
1978	28		25.8 + 3.6 = 29.4
1979	33		
1980	36		40.2 – 3.6 = 36.6
1981	40		42 – 1.8 = 40.2
		$\frac{36+40+44+48}{4} = 42$	Center = 42
1982	44		42 + 1.8 = 43.8
1983	48		43.8 + 3.6 = 47.4
1984			47.4 + 3.6 = 51

The value of 24 will be plotted in the middle of the years 1976 and 1977. Similarly the value 42 will be plotted against the middle of the years 1981 and 1982.

Annual Increment = $(\bar{x}_2 - \bar{x}_1) / 5 = (42 - 24)/5 = 3.6$

Again, 3.6/2 = 1.8 is the half yearly increment.

Result

- Since, we get two points 24 and 42 which shall be plotted corresponding to their middle years i.e. 1976 to 1977 and 1981 to 1982.
- The sale for the year 1984 is 51.

Example

Determine the trend of the following by using semi-average method.

Year	1982	1983	1984	1985	1986	1987
Sales	10	12	11	16	15	20

Procedure

- First step the whole data is divided in two equal parts.
- Then calculate the averages by two parts.
- To calculate annual increment is,

  Annual increment = $(\bar{x}_2 - \bar{x}_1) / 3$.

- Finally, calculate the trend values.

Calculation

Here the number of years is 8. The two middle parts will be 1982 to 1984 and 1985 to 1987.

Year	Sales	**Semi averages**	**Trend Values**
1982	10	$\frac{10+12+11}{3}=11$	11 – 2 = 9
1983	12		Center = 11
1984	11		11 + 2 = 13
1985	16	$\frac{16+15+20}{3}=17$	17 – 2 = 15
1986	15		Center = 17
1987	20		17 + 2 = 19

Annual Increment = ($\overline{x}_2$ - $\overline{x}_1$) / 3 = (17 - 11)/3 = 2

Result

Since, we get two points 11 and 17 which shall be plotted to the corresponding years 1983 and 1986.

(iv) 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,

$$Y_{tc} = a + bt \quad \ldots\ldots\ldots\ldots (1)$$

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 three equations we get the values of a and b.

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

$$Y_{tc} = a + bt$$

This is the required Straight line trend equation.

Example

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.

Procedure

- To find a and b using the normal equations.
- To calculate the equation of the straight line is $Y_c = a + bX$.
- Calculate the trend values.
- Finally draw the trend line in the graph.

Calculation

Fitting straight line trend by method of least square

Year	Sales	Deviation by 2003.5	Deviation × 2 (X)	XY	X^2	Y_c
2000	80	-3.5	-7	-560	49	83
2001	90	-2.5	-5	-450	25	85.5
2002	92	-1.5	-3	-276	9	88
2003	83	-0.5	-1	-83	1	90.5
2004	94	0.5	1	94	1	93
2005	99	1.5	3	297	9	95.5
2006	92	2.5	5	460	25	98
2007	104	3.5	7	728	49	100.5

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

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

The trend vales for various years are,

$Y_{2000} = 91.75 + 1.25(-7) = 83$
$Y_{2001} = 91.75 + 1.25(-5) = 85.5$
$Y_{2002} = 91.75 + 1.25(-3) = 88$
$Y_{2003} = 91.75 + 1.25(-1) = 90.5$
$Y_{2004} = 91.75 + 1.25(1) = 93$
$Y_{2005} = 91.75 + 1.25(3) = 95.5$
$Y_{2006} = 91.75 + 1.25(5) = 98$
$Y_{2007} = 91.75 + 1.25(7) = 100.5$

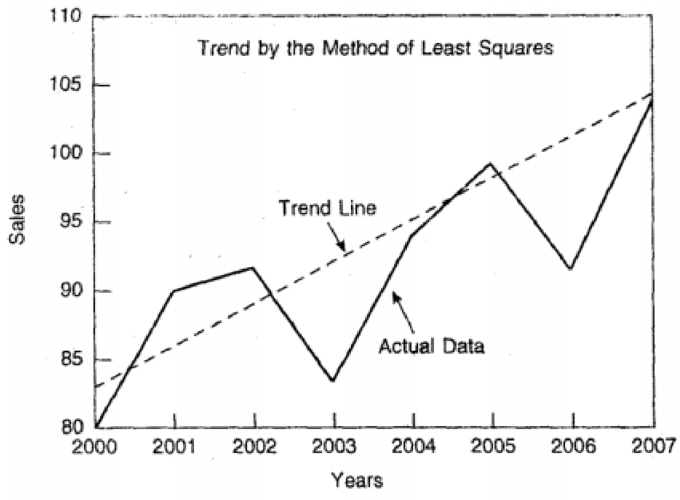
Result

The trend vales for various years are,

$Y_{2000} = 83$ $Y_{2001} = 85.5$ $Y_{2002} = 88$

$Y_{2003} = 90.5$ $Y_{2004} = 93$ $Y_{2005} = 95.5$

$Y_{2006} = 98$ $Y_{2007} = 100.5$



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

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

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

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

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

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

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

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

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

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

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

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

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

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

Solving these three equations we get the values of a, b and c.

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

$$Y_{tc} = a + bt + ct^2$$

This is the required parabolic trend equation.

Example

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 and plot the actual and trend values on the graph.

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

Procedure

- To find a, b and c using the normal equations.
- To calculate the equation of the straight line is $Y_c = a + bX + cX^2$.
- Calculate the trend values.
- To estimate the price of the commodity for the year 2008.
- Finally draw the trend line in the graph.

Calculation

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

$$\sum Y = na + b\sum X + c\sum X^2$$

$$\sum XY = a\sum X + b\sum X^2 + c\sum X^3$$

$$\sum X^2Y = a\sum X^2 + b\sum X^3 + c\sum X^4$$

Year	Prices	X	X^2	X^3	X^4	XY	X^2Y	Trend values
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	**848**	**3**	**19**	**27**	**115**	**771**	**3099**	

848 = 6 a + 3 b + 19 c (1)

771 = 3 a + 19 b + 27 c (2)

3099 = 19 a + 27 b + 115 c (3)

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

848 = 6 a + 3 b + 19 c

Eqn. (2)×2 => 1542 = 6 a+ 38 b + 54 c

\- - - -

-694 = -35 b – 35 c (4)

=> 35 b + 35 c = 694

Multiplying Eqn. (2) by 19 and Eqn. (3) by 3, we get

Eqn. (2)×19 => 14649 = 57 a + 361 b + 513 c

Eqn. (3)×3 => 9297 = 57 a + 81 b + 345 c

\- - - -

5352 = 280 b + 168 c (5)

Multiplying Eqn. (4) by 8 and solving equations (4) and (5), we get

Eqn. (4)×8 => 280 b + 280 c = 5552

Eqn. (5) => 280 b + 168 c = 5352

\- - -

112 c = 200 => **c = 1.786**

Substituting the value of c in Eqn. (4),

35 b + (35×1.786) = 694

35 b = 394 – 62.5 = 631.5

b = 18.042

Substituting the values of b and c in Eqn. (1),

$$848 = 6a + (3\times18.042) + (19\times1.786)$$
$$848 = 6a + 54.126 + 33.934$$
$$6a = 759.94$$
$$a = 126.657$$

Substituting the values of a, b and c in equation,

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

The trend vales for various years are,

$$Y_{2002} = 126.657 + 18.042(-2) + 1.786(-2)^2 = 97.717$$
$$Y_{2003} = 126.657 + 18.042(-1) + 1.786(-1)^2 = 110.401$$
$$Y_{2004} = 126.657 + 18.042(0) + 1.786(0)^2 = 126.657$$
$$Y_{2005} = 126.657 + 18.042(1) + 1.786(1)^2 = 146.485$$
$$Y_{2006} = 126.657 + 18.042(2) + 1.786(2)^2 = 169.885$$
$$Y_{2007} = 126.657 + 18.042(3) + 1.786(3)^2 = 196.857$$

Price for the year 2008 is,

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

$$Y_{2008} = 126.657 + 18.042(4) + 1.786(4)^2 = 227.401.$$

Thus the likely price of the commodity for the year 2008 s rs. 227.401 approx.

Result

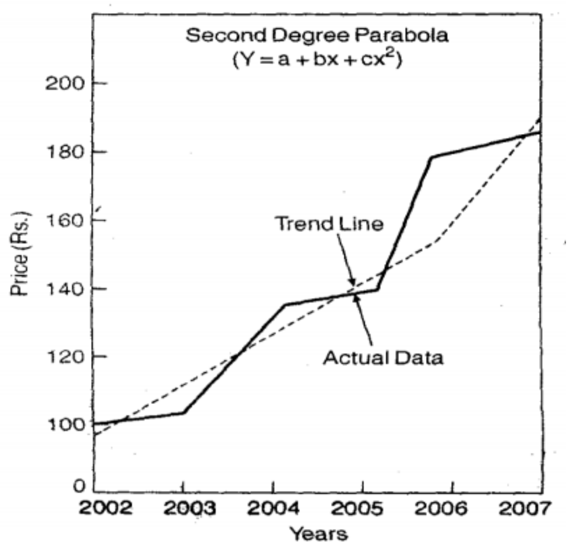
The trend vales for various years are,

$Y_{2002} = 97.717$	$Y_{2003} = 110.401$	$Y_{2004} = 126.657$
$Y_{2005} = 146.485$	$Y_{2006} = 169.885$	$Y_{2007} = 196.857$

Price for the year 2008 is,

$$Y_{2008} = 227.401.$$

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



Fitting of a Exponential 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 Exponential trend to be of the form $Y_{tc} = a\ b^t$ …………. (1)

Taking log on both sides, we get

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

$Y = A + Bt$ (say) …………. (2)

Where, $Y = \log Y_{tc}$, $A = \log a$, $B = \log b$ …………. (3)

Equation (2) is a straight line in t and Y and thus the normal equations for estimating A and B are

$$\sum Y = nA + B\sum t,$$
$$\sum tY = A\sum t + B\sum t^2,$$
…………. (4)

These equations can be solved for A and B and finally on using equation (3), we get

a = antilog (A) and b = antilog (B)

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

$$Y_{tc} = a\ b^t$$

This is the required Exponential trend equation.

Fitting of a Logarithms 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 Logarithms trend to be of the form $Y_{tc} = a\, b^t\, c^{t^2}$ …………. (1)

Taking log on both sides, we get

$$\log Y_{tc} = \log a + t \log b + t^2 \log c$$
$$Y = A + Bt + Ct^2$$
…………. (2)

Where, $Y = \log Y_{tc}$, $A = \log a$, $B = \log b$ and $C = \log c$ …………. (3)

Now, Equation (2) is a Logarithms curve in t and Y and thus the normal equations for estimating A, B and C are

$$\sum Y = nA + B\sum t + C\sum t^2,$$
$$\sum tY = A\sum t + B\sum t^2 + C\sum t^3,$$
$$\sum t^2Y = A\sum t^2 + B\sum t^3 + C\sum t^4,$$
…………. (4)

These equations can be solved for A, B and C and finally on using equation (3), we get

a = antilog (A), b = antilog (B) and c = antilog (C)

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

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

This is the required Logarithms trend equation.

2. Cyclical Variations

The various methods used for measuring cyclical variations are

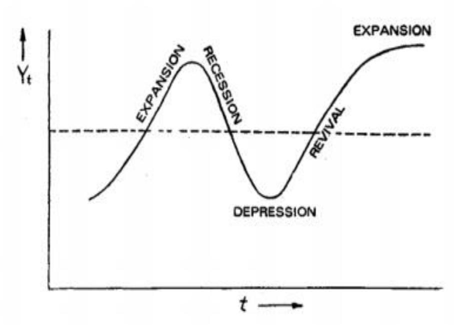
- Residual method
- Reference cycle analysis method
- Direct method
- Harmonic analysis method

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.

3. Seasonal Variations

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.

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

(i) 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 $\overline{x_i}$ (i = 1,2,…..12 for monthly and i=1,2,3,4 for quarterly) for the ith month or quarter for all the years.
- Compute the average $\bar{x}$ of the averages.
- i.e. $\bar{x} = \frac{1}{12}\sum_{i=1}^{12}\overline{x_i}$ for monthly and $\bar{x} = \frac{1}{4}\sum_{i=1}^{4}\overline{x_i}$ 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) $\frac{\overline{x_i}}{\bar{\bar{x}}} \times 100$

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

Calculate seasonal indices for various quarters by simple averages method from the following data:

Year	1st 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

Procedure

- Arrange the data by years and quarters.
- Compute the average $\overline{x_i}$ for all the quarters.
- Compute the average $\bar{x}$ of the averages.

$$\bar{x} = \frac{1}{4}\sum_{i=1}^{4}\overline{x_i} \text{ for quarterly}$$

- Seasonal indices for different quarters are obtained by expressing quarterly averages as percentages of $\bar{x}$. Thus seasonal indices for ith quarter is $\frac{\overline{x_i}}{\bar{x}} \times 100$

Calculation

Computation of seasonal indices

Year	1st 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 indices	**98.66**	**110.74**	**95.30**	**95.30**

The average $\bar{x}$ of the averages is,

$$\bar{x} = \frac{3.675 + 4.125 + 3.55 + 3.55}{4} = 3.725$$

$$Seasonal\ index = \frac{\overline{x_i}}{\bar{x}} \times 100$$

$$\text{Seasonal index for the first quarter} = \frac{3.756}{3.725} \times 100 = 98.66$$

$$\text{Seasonal index for the } 2^{nd} \text{ quarter} = \frac{4.125}{3.725} \times 100 = 110.74$$

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

Result

- The average of the average = 3.725
- Seasonal index for the first quarter = 98.66
- Seasonal index for the second quarter = 110.74
- Seasonal index for the third and fourth quarter = 95.30

(ii) 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 fluctuation 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 and last 6 months.

Example

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

Procedure

- Calculate the centered quarterly moving average of the given data.
- Express the original data as percentages of the centered moving average values.
- calculate the seasonal indices are obtained by eliminating the irregular or random components by averaging these percentages using A.M.

Calculation

Calculation of seasonal indices by Ratio to moving average method

Year	quarter	Production	4-quarter MA total	4-quarter MA	4-quarter MA Centered	Ratio to MA
2003	I	30				
	II	40				
			140	35		
	III	36			35.5	101.41
			144	36		
	IV	34			37.5	90.67
			156	39		
2004	I	34			46.25	73.51
			214	53.5		
	II	52			49.25	105.58
			180	45		
	III	50			45.75	109.29
			186	46.5		
	IV	44			47.25	93.12
			192	48		
2005	I	40			48.5	82.47
			196	49		
	II	58			49.5	117.17
			200	50		
	III	54			51.75	104.35
			214	53.5		
	IV	48			55.75	86.10
			232	58		
2006	I	54			59.75	90.38
			246	61.5		
	II	76			63.25	120.16
			260	65		
	III	68			68.25	99.63
			286	71.5		
	IV	62			73.5	84.35
			302	75.5		
2007	I	80			77.75	102.89
			320	80		
	II	92			82.5	111.52
			340	85		
	III	86				
	IV	82				

Calculate the Seasonal Indices

Year	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
2003	-	-	101.41	90.67
2004	73.51	105.58	109.29	93.12
2005	82.47	117.17	104.35	86.10
2006	90.38	120.16	99.63	84.35
2007	102.89	111.52	-	-
Total	349.25	454.43	414.68	354.24
Average	87.31	113.61	103.67	88.56
SI	88.83	115.58	105.47	90.10

Arithmetic Average of average = $\frac{393.15}{4} = 98.29$

$$\text{Seasonal index for the first quarter} = \frac{87.31}{98.29} \times 100 = 88.83$$

$$\text{Seasonal index for the second quarter} = \frac{113.61}{98.29} \times 100 = 115.58$$

$$\text{Seasonal index for the third quarter} = \frac{103.67}{98.29} \times 100 = 105.47$$

$$\text{Seasonal index for the fourth quarter} = \frac{88.56}{98.29} \times 100 = 90.10$$

Result

- The average of the average = 98.29
- Seasonal index for the first quarter = 88.83
- Seasonal index for the second quarter = 115.58
- Seasonal index for the third quarter = 105.47
- Seasonal index for the fourth quarter = 90.10

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 = $\frac{\text{Current periods figure}}{\text{Previous periods figure}} \times 100$

- 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

$$\frac{\text{Avg link relative for that period} \times \textit{chain} \text{ relative of the previous period}}{100}$$

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 $d = \frac{1}{N}[\text{New chain relative for first period} - 100]$ 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.

Example

Calculate seasonal indices by the method of link relative to the following data:

Year	1^{st} Quarter	2^{nd} Quarter	3^{rd} Quarter	4^{th} Quarter
2003	6.0	6.5	7.8	8.7
2004	504	7.9	8.4	7.3
2005	6.8	6.5	9.3	6.4
2006	7.2	5.8	7.5	8.5
2007	6.6	7.3	8.0	7.1

Procedure

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

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

- 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

$$\frac{\text{Avg link relative for that period} \times \textit{chain} \text{ relative of the previous period}}{100}$$

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 $d = \frac{1}{N}[\text{New chain relative for first period} - 100]$ where N denotes the number of periods i.e. 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.

Calculation

Calculate the seasonal indices by the method of link relatives

Year	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
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	88.6
Average	86.35	108.28	121.66	93.86
Chain relatives	100	108.28	131.73	123.64
Corrected Chain relatives	100	106.605	128.38	118.615
Seasonal Indices	88.18	94.01	113.21	104.60

Chain Relatives

Chain Relatives of the first quarter = 100

Chain Relatives of the second quarter = $\frac{100\times108.28}{100} = 108.28$

Chain Relatives of the third quarter = $\frac{121.66\times108.28}{100} = 131.73$

Chain Relatives of the fourth quarter = $\frac{93.86\times131.73}{100} = 123.64$

Adjusted Chain Relatives

Chain Relatives of the first quarter (on the basis of the last quarter) = $\frac{86.35\times123.64}{100} = 106.7$

The difference between these (first) chain relatives =106.7 - 100 = 6.7

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

Corrected chain relatives

Corrected chain relatives for first quarter = 100
Corrected chain relatives for second quarter = 108.28 - 1.675 = 106.605
Corrected chain relatives for third quarter = 131.73 - (2×1.675) = 128.38
Corrected chain relatives for fourth quarter = 123.64 - (3×1.675) = 118.615

Average of Average

$$\text{Average of corrected chain relatives } = \frac{100+106.605+128.38+118.615}{4} = \frac{453.6}{4} = 113.4$$

Seasonal Indices

$$\text{Seasonal index for the first quarter } = \frac{100}{113.4}\times 100 = 88.18$$

$$\text{Seasonal index for the sec ond quarter } = \frac{106.605}{113.4}\times 100 = 94.01$$

$$\text{Seasonal index for the third quarter } = \frac{128.38}{113.4}\times 100 = 113.21$$

$$\text{Seasonal index for the fourth quarter } = \frac{118.615}{113.4}\times 100 = 104.60$$

Result

- The average of the average = 113.4
- Seasonal index for the first quarter = 88.18
- Seasonal index for the second quarter = 94.01
- Seasonal index for the third quarter = 113.21
- Seasonal index for the fourth quarter = 104.60

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

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

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

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:

Calculate seasonal indices by Ratio to trend 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

Procedure

- To find a and b using the normal equations.
- To calculate the equation of the straight line is $Y_c = a + bX$.
- Calculate the yearly trend values.
- Calculate the quarterly trend values for each year.
- Calculate quarterly vales as percentage of trend values.
- To calculate average of average.
- finally calculate the seasonal indices.

Calculation

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

Calculating trend by method of least squares

Year	Totals	Average (Y)	X	XY	X^2	Trend value
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
		280		120	10	

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, $\sum X = 0$, $a = \frac{\sum Y}{N} = \frac{280}{5} = 56$ and $b = \frac{\sum XY}{\sum X^2} = \frac{120}{10} = 12$

The required line equation is $Y_c = 56 + 12X$

The trend vales for various years are,

$Y_{2003} = 56 + 12(-2) = 32$

$Y_{2004} = 56 + 12(-1) = 44$

$Y_{2005} = 56 + 12(0) = 56$

$Y_{2006} = 56 + 12(1) = 68$

$Y_{2007} = 56 + 12(2) = 80$

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

Half yearly increment = $\frac{3}{2} = 1.5$

Consider 2003, trend value for the middle of the quarter, i.e., half of 2^{nd} and half of 3^{rd} is 32. So that the trend value of 2^{nd} quarter is 32-1.5 = 30.5 and for 3^{rd} quarter is 32 + 1.5 = 33.5. the trend value for the 1^{st} quarter is 30.5 – 3 = 27.5 and for 4^{th} quarter is 33.5 – 3 = 36.5.5. thus we get quarterly trend values as 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	60.3
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 quarter of 2003 is (30/27.5)*100 = 109.09, for 2nd quarter is (40/30.5)*100 = 131.15, etc.

Quarterly vales as percentage 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
Seasonal indices	92.05	117.36	102.12	88.46

The average of the averages is,

$$\bar{x} = \frac{92.77 + 118.28 + 102.92 + 89.15}{4} = 100.78$$

$$Seasonal\ index = \frac{\bar{x}_i}{\bar{\bar{x}}} \times 100$$

$$Seasonal\ index\ for\ the\ first\ quarter = \frac{92.77}{100.78} \times 100 = 92.05$$

$$Seasonal\ index\ for\ the\ second\ quarter = \frac{118.28}{100.78} \times 100 = 117.36$$

$$Seasonal\ index\ for\ the\ third\ quarter = \frac{102.92}{100.78} \times 100 = 102.12$$

$$Seasonal\ index\ for\ the\ fourth\ quarter = \frac{89.15}{100.78} \times 100 = 88.46$$

Result

- The average of the average = 100.78
- Seasonal index for the first quarter = 92.05
- Seasonal index for the second quarter = 117.36
- Seasonal index for the third quarter = 102.12
- Seasonal index for the fourth quarter = 88.46

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. That is

$$Deseasonalized\ data = \frac{Original\ data}{Seasonal\ index} \times 100$$

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

In case of additive model

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

$$Deseasonalized\ data = \ Original\ data - \frac{Seasonal\ Index}{100} = Y_t - \frac{Seasonal\ Index}{100}$$

Index Numbers

An Index Number measures the relative change in price, quantity, value, or some other item of interest from one time period to another.

A simple index number measures the relative change in just one variable.

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.

Classification of Index Numbers

- **Price index number**

  It measures the changes in the prices of the commodities produced, consumed or sold in a given period with reference to the base period.

- **Quantity index numbers**

  These help to measure and compare the changes in the physical volume of goods produced, sold and purchased in a given period compared to some other given period.

- **Value index numbers**

  These indexes show changes in the value of any commodity in a given period in reference to the base period.

- **Consumer price index**

  These indexes measure the average over time in the prices paid by the consumers for a specific group of goods and services.

- **Special purpose index number**

  These indexes are framed for a special study relating to a particular variable or aspect.

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

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 without 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

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

Example

Calculate the base shift index number as Shift the base from 1998 to 2004 and recast the index numbers, with base year 1998 for the following data:

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

Calculation

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

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

$$\text{Index number for 2002} = \frac{400}{400} \times 100 = 100$$

$$\text{Index number for 2004} = \frac{400}{400} \times 100 = 100$$

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

Year	**Index Number (1998 as base)**	**Index Number (2004 as base)**	*Year*	**Index Number (1998 as base)**	**Index Number (2004 as base)**
1998	100	25	*2003*	410	102.5
1999	110	27.5	*2004*	400	100
2000	120	30	*2005*	380	95
2001	200	50	*2006*	370	92.5
2002	400	100	*2007*	340	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,

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

Example

Calculate Splicing of index number as the index A is started in 1993 and continued up to 2003 in which year another index B was started. The data is given below:

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

Calculation

Index B Splicing to Index A

Year	Index A	Index B	
1993	100		
1994	110		
1995	112		
-			
-			
2002	138		
2003	150	100	150
2004		120	180
2005		140	210
2006		130	195
2007		150	225

$$\text{Splicing Index number for } 2003 = \frac{150}{100} \times 100 = 150$$

$$\text{Splicing Index number for } 2004 = \frac{150}{100} \times 120 = 180$$

$$\text{Splicing Index number for } 2005 = \frac{150}{100} \times 140 = 210$$

$$\text{Splicing Index number for } 2006 = \frac{150}{100} \times 130 = 195$$

$$\text{Splicing Index number for } 2007 = \frac{150}{100} \times 150 = 225$$

This would be done by multiplying the old index by the ratio (100/150). Thus the spliced index for 1993 is (100/150)*100 = 66.7, for 1994 is (100/150)*110 = 73.3 and for 1995 is (100/150)*112 = 74.6, etc.

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.

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

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.

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

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

Example

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	3600	4200	5000	5500	6000	6400	6800	7200	7500
Price	100	120	145	160	250	320	450	530	600

Calculation

Index number showing changes in the real income of the worker

Year	**Income**	**Price**	**Real Income**	**Index**
1999	3600	100	3600.00	100.00
2000	4200	120	3500.00	97.22
2001	5000	145	3448.27	95.78
2002	5500	160	3437.50	95.49
2003	6000	250	2400.00	66.67
2004	6400	320	2000.00	55.56
2005	6800	450	1511.11	41.98
2006	7200	530	1358.49	37.74

2007	7500	600	1250.00	34.72

$$\text{Real Income for 1999} = \frac{3600}{100} \times 100 = 3600.00$$

$$\text{Real Income for 2000} = \frac{4200}{120} \times 100 = 3500.00$$

$$\text{Real Income for 2001} = \frac{5000}{145} \times 100 = 3448.27$$

$$\text{Real Income for 2002} = \frac{5500}{160} \times 100 = 3437.50$$

$$\text{Real Income for 2003} = \frac{6000}{250} \times 100 = 2400.00$$

$$\text{Real Income for 2004} = \frac{6400}{320} \times 100 = 2000.00$$

$$\text{Real Income for 2005} = \frac{6800}{450} \times 100 = 1511.11$$

$$\text{Real Income for 2006} = \frac{7200}{530} \times 100 = 1358.49$$

$$\text{Real Income for 2007} = \frac{7500}{600} \times 100 = 1250.00$$

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 52 rupee sales, rupee inventories of manufacturer's, wholesaler's and retailer's income, wages and the like.

Price Index Number

Measure changes in price over a specified period of time. It is basically the ratio of the price of a certain number of commodities at the present year as against base year.

In other words, a price index is a measure of price changes using a percentage scale. A price index can be based on the prices of a single item or a selected group of items, called a market basket.

Types of price indices

1. Index of retail prices,
2. Index of wholesale prices,
3. Cost of living index of industrial workers,
4. Export prices, and so on.

Cost of Living Index Number

The wholesale price index numbers measure the changes in the general level of prices and they fail to reflect the effect of the increase or decrease of prices on the cost of living of different classes or groups of people in a society.

Cost of living index numbers, also termed as ‘Consumer Price Index Numbers, or ‘Retail Price Index Numbers’ are designed to measure the effects of changes in the prices of a basket of goods and services on the purchasing power of a particular section or class of the society during any given (current) period with respect to some fixed (base) period.

Construction of Cost of Living Index Numbers:

The cost of living index numbers are constructed by the following formula:

$$\text{Cost of living index number} = \frac{\sum WP}{\sum W}$$

where, W is the weights and $P = \frac{p_1}{p_0}$

Example

Calculate the Cost of living index number from the following data:

Items	Price		Weights
	Base Year	Current Year	
Food	30	47	4
Fuel	8	12	1
Clothing	14	18	3
House Rent	22	15	2
Miscellaneous	25	30	1

Calculation

Calculations for cost of living index number

Items	Weights	Price		P = p_1/p_0	WP
		Base Year (p_0)	Current Year (p_1)		
Food	4	30	47	156.67	626.67
Fuel	1	8	12	150	150
Clothing	3	14	18	128.57	385.71
House Rent	2	22	15	68.18	136.36
Miscellaneous	1	25	30	120	120
	11				**1418.74**

$$\text{Cost of living index number} = \frac{\sum WP}{\sum W} = \frac{1418.74}{11} = 128.98.$$