VIVEKANANDA COLLEGE
THAKURPUKUR
KOLKATA-700063

NAAC ACCREDITED 'A' GRADE

Topic: Time_Series_Analysis

Course Title: Introduction & Trend Analysis

Paper: CC10

Unit: 2

Semester: 4

Name of the Teacher: Sutapa Biswas

Name of the Department: Statistics

Time Series Analysis

A series of value of some business or economic variable given according to their time of occurence is called a time series.

Since the values of the variable in a time series occur chronologically, so it is clear that the time variable in a time series is not taken haphazardly, but systematically at regular intervals from past to future.

Mathematically, a time series is defined by the functional relationship,

$$y_t = f(t)$$

where y_t is the value of the variable at time t.

Some examples of the time series are —

i) The annual production of wheat in a country over the last 20 years.

ii) The daily closing price of a share on a stock exchange for a week.

iii) Monthly electric bills for 12 months.

iv) Hourly temperature recorded by the meterological office in a day.

Objective of Time Series Analysis :-

Time series are analysed for several reasons —

i) It helps to detect the type and nature of variations in the data.

ii) It helps us to understand past behaviour of the variable under study.

iii) Analysis of the past behaviour of the variable under study enables one to forecast the future behaviour and thereby facilitates in proper planning for future operations. Business executives may judiciously formulate their policy decisions.

iv) It helps to segregate and study the effects of different components.

Components of a Time Series :-

When time series data are graphically exhibited, the graph shows an overall picture of haphazard movement. In reality it is observed that atleast a part of the changes, known as the systematic part, can be accounted for and the remaining part is irregular. Among four components, the three main factors constituting the systematic part are,

(i) Secular Trend

(ii) Seasonal Variation

(iii) Cyclical Variation

The fourth component is 'Random' or 'Irregular' movement.

The value of y_t of a time series at any time (t) is considered as the resultant of the combined effect of secular trend (T_t), Seasonal Variation (S_t), Cyclical variation (C_t) and Irregular movements (I_t).

Here two models, Multiplicative model and Additive model are used for decomposition of a time series into its components which are

$$y_t = T_t \times S_t \times C_t \times I_t \quad , \text{ multiplicative or product model}$$

and $$y_t = T_t + S_t + C_t + I_t \quad , \text{ additive or sum model}$$

In product model, T_t has same unit as that of Y_t, where S_t, C_t and I_t are unit free.

In sum model, all components have same unit as that of y_t.

*> Multiplicative model is the most commonly assumption, as compared to additive model —

In additive model it is assumed that all the four components of the time series operate independently of one another, so that none of these components has any effect on the remaining three. But practically it cannot happen always that four components operate independently. So in multiplicative model it is assumed that the four components are due to different causes but are not independent and may affect one another. The multiplicative model cannot be applied unless the time series is translated by adding a suitable positive value. So multiplicative model is more often employed in practice.

Secular Trend :- By secular trend or simply trend we mean the general tendency of the data to increase or decrease during a long period of time. So it is the general, smooth, long-term, average tendency.
It is not necessary that the increase or decline should be in the same direction throughout the given period. Different tendencies of increase, decrease or stability may be observed in different sections of time. Such tendencies are more or less constant for a long time or which change very gradually and continuously over a long period of time as the change in the population, tastes, habits and customs of the people in a society and so on. For example, an upward tendency would be seen in population growth, agricultural production, currency in circulation etc; while a downward tendency will be noticed in data of births and deaths, epidemics etc., and a constant reading is the series of barometric readings.

Trend can be studied or measured by following methods —

(i) Graphic (or Free hand Curve fitting) Method.
(ii) Method of Semi-Averages,
(iii) Method of Curve Fitting by Principle of Least Square
(iv) Method of Moving Averages.

Among four method we mainly use Method of moving averages and Method of curve fitting by Principle of Least Square.

Seasonal Variations :- By seasonal variations we mean a regular and periodic movement (i.e., a movement that repeats itself at regular intervals of time), where the period is not longer than one year.
Seasonal variations will be there where the data are recorded quarterly, monthly, weekly, daily, hourly and so on. For example, Monthly expenditure of a family etc. Most of economic time series are influenced by seasonal swings, e.g., prices, production and consumption of commodities, bank clearings and bank deposits etc., are all affected by seasonal variations.

These variations in economic time series may be attributed to two broad factors namely, (i) climatic changes of various seasons (e.g., the sale of ice-cream, umbrella etc.) and (ii) habits, customs and conventions followed by people at different times (e.g., sale of consumer goods increases during festival months).

Seasonal variation can be measured by following method —

(i) Method of monthly averages
(ii) Ratio to moving average method
(iii) Ratio to trend method
(iv) Method of Link Relatives.

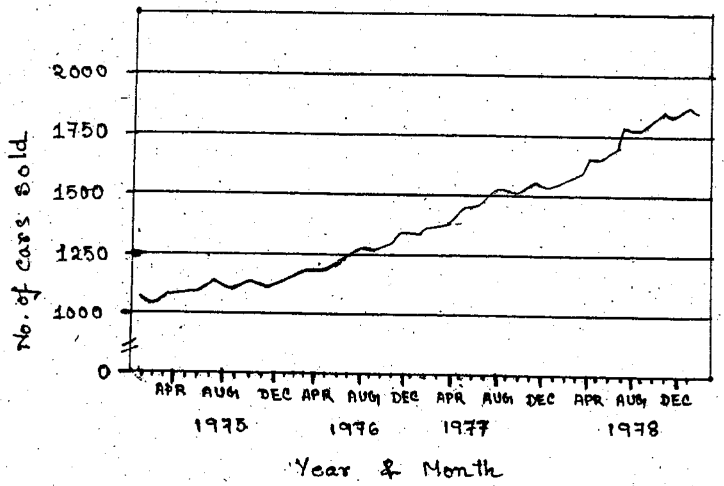
Cyclic Variations :– The oscillatory movements in a time series with period of oscillation more than one year are termed as cyclic fluctuations. One complete period is called a 'cycle'. These variations, though more or less regular, are not necessarily periodic. Such fluctuations in a time series are usually attributed to a 'business cycle' comprising four successive phases viz. properity (or boom), recession, depression and recovery and normally lasts from seven to eleven years. The swing from boom to recovery and back again to boom is found to vary in time span. Most of the economic and commercial series, e.g., series relating to prices, production and wages, etc, are affected by business cycles.

Irregular (or Random) Component :– Irregular or residual fluctuations is inherent in every time series and caused by unforeseen and unpredictable forces which are purely erratic and random in character. For example, war, earthquakes, floods, famines, revolutions, epidemics etc. cause irregular fluctuations. Random component of a time series includes all those variations which are not accounted for by trend, seasonal and cyclical movements. So these variations may be regarded as residual component of a time series.

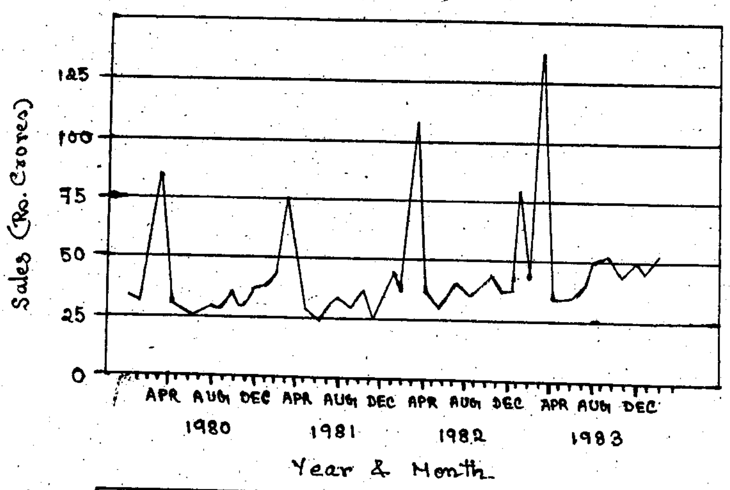
Graph of various components

1) Secular Trend

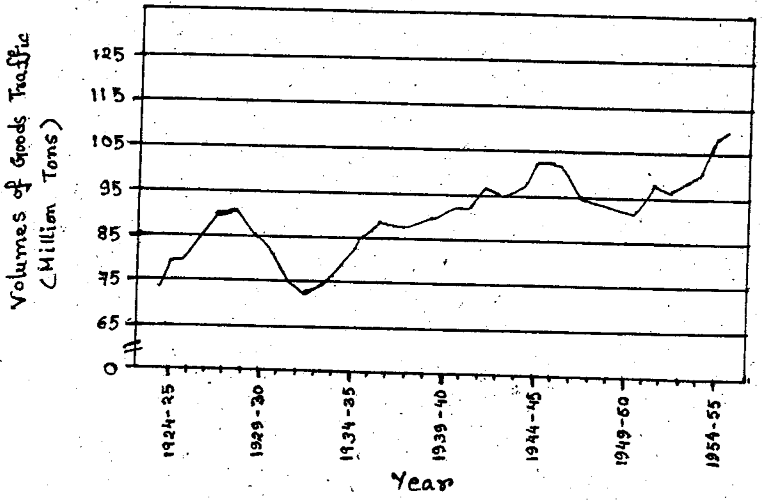
[Figure Shows upward Trend]



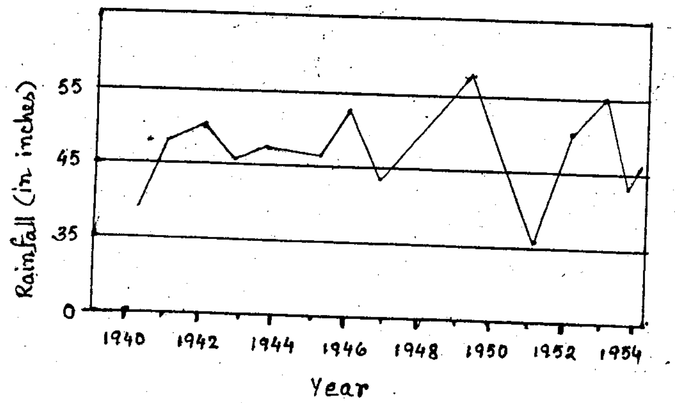
2) Seasonal Variations



3) Cyclical Fluctuations



4) Irregular Fluctuations



Method of Moving Averages :-

It consists in measurement of trend by smoothing out the fluctuations of the data by means of a moving average. Moving average of period m is a series of arithmetic means of m terms at a time, starting with 1st, 2nd, 3rd term etc. Thus, the first average is the mean of the 1st m terms. The 2nd is the mean of the m terms from 2nd to $(m+1)$th term, the 3rd is the mean of the m terms from 3rd to $(m+2)$th term and so on.

If m is odd equal to $(2k+1)$ say, moving average is placed against the mid-value of the time interval it covers i.e. against $t=k+1$ and if m is even equals to $(2k$ say$)$, it is placed between the two middle values of the time intervals it covers, i.e. between $t=k$ and $t=k+1$. In the latter case the moving averages does not coincide with an original time period so here a subsequent two-item moving average is computed to make the resulting moving average values corresponding to given times (i.e. to obtain the centered moving averages). The graph obtained on plotting the moving average values against the corresponding time values gives trend curve.

Interpretation of Moving Average :-

The interpretation of moving average is very simple. A k-point moving average may be interpreted as the estimated value for the middle of the period covered from successive linear curves fitted through the first k points, through the 2nd to the $(k+1)$st values and so on, and lastly through the last k points.

Consider the first k points $y_1, y_2, ---, y_k$. Let the origin be shifted to the middle of the period so that $\sum_{i=1}^{k} t_i = 0$.

The normal equations for fitting a curve $y = a+bt$ through $y_1, y_2, --- y_k$ are

$$\Sigma y_i = ka + b\,\Sigma t_i$$

$$\Sigma t_i y_i = a\,\Sigma t_i + b\,\Sigma t_i^2$$

so that $\hat{a} = \dfrac{\sum_i y_i}{k} = \bar{y}$ and $\hat{b} = \dfrac{\sum_i t_i y_i}{\sum_i t_i^2}$

Hence, for the estimated value for the middle of the period covered i.e. for $t=0$, from the curve $y = \hat{a} + \hat{b}t$ is $\hat{a}$, which is the first moving average value. Similarly it can be shown that the estimated value from the fitted linear curve through $y_2, y_3, ---, y_{k+1}$, the second moving average value and so on.

Merits of Moving Average Method

i) Moving Average method is flexible in the sense that if some more observations are added to the original series, the previous calculations remain unaffected and we get some more averages.

ii) The application of moving average method is simple and does not involve complicated calculations.

iii) If the time series contain regular cyclical fluctuations, these fluctuations are automatically removed provided an appropriate period is chosen. Even when the fluctuations are not completely eliminated, moving average process reduces their intensity.

iv) Moving Averages adapt themselves to the general movements of data. Any change in the trend is faithfully reflected by the moving averages.

Demerits of Moving Average Method

i) In moving average method some trend values at each end of the series cannot be estimated, their number increasing with increase in the period of moving average.

ii) The period of moving average has to be chosen very carefully. There are no hard and fast rules for the purpose.

iii) This process cannot be used for forecasting as it assumes no law of change.

iv) The method is satisfactory when the trend is more or less linear. But if the trend is non-linear then moving averages may deviate considerably from the trend.

Method of Mathematical curves :-

It is the most important and rational method of determining trend. Here we first choose a suitable trend equation from the graphical representation of the data and then estimate the constants associated with the equation on the basis of the available data.

Choose a polynomial of suitable degree for the original variables or for a transformed variable and the constants are obtained by the method of least squares. Graphical representation of time series data enables in selecting the appropriate polynomial.

Let us consider that a polynomial of degree 'r' in 't' is selected for exhibiting the trend as,

$$T_t = a_0 + a_1 t + a_2 t^2 + \dots + a_r t^r \quad \text{——— (1)}$$

The constants $a_0, a_1, a_2, \dots, a_r$ are estimated by the method of least squares, which requires $\Sigma(y_t - T_t)^2$, a minimum with respect to $a_0, a_1, a_2, \dots, a_r$. Then the estimates of these unknown constants will be the solutions of the following normal equations.

$$\Sigma y_t = n a_0 + a_1 \Sigma t + a_2 \Sigma t^2 + \dots + a_r \Sigma t^r$$

$$\Sigma t y_t = a_0 \Sigma t + a_1 \Sigma t^2 + a_2 \Sigma t^3 + \dots + a_r \Sigma t^{r+1}$$

$$\Sigma t^2 y_t = a_0 \Sigma t^2 + a_1 \Sigma t^3 + a_2 \Sigma t^4 + \dots + a_r \Sigma t^{r+2}$$

$$\dots \quad \dots \quad \dots$$

$$\Sigma t^r y_t = a_0 \Sigma t^r + a_1 \Sigma t^{r+1} + a_2 \Sigma t^{r+2} + \dots + a_r \Sigma t^{2r}$$

where the values of y_t are given for n values of t. Using the estimates we can obtain the trend value for any given time t by putting the value of t in eq.(1).

Fitting of Straight Line:

If we put $r=1$ in equation (1) and the trend equation is

$$T_t = a_0 + a_1 t$$

Here the least-square estimates of the constants a_0 and a_1 will be the solutions of the normal equations

$$\Sigma y_t = n a_0 + a_1 \Sigma t$$

$$\Sigma t y_t = a_0 \Sigma t + a_1 \Sigma t^2$$

Fitting a Second Degree Polynomial:

If we put $r=2$ in equation (1) and the trend equation becomes

$$T_t = a_0 + a_1 t + a_2 t^2$$

The normal equations for obtaining the constants a_0, a_1 and a_2 by the method of least squares are

$$\sum y_t = n a_0 + a_1 \sum t + a_2 \sum t^2$$
$$\sum t y_t = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3$$
$$\sum t^2 y_t = a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4$$

<u>Note</u> :- Generally the successive time points will be equidistant. Consider the mid-point of the entire time span as origin. Then sum of odd powers of t is zero. If d be the common difference between two successive time values, then for odd or even number of values, one may take d or $d/2$ as the new unit ~~of~~ for t for each of computation.

<u>Exponential Curves</u>:

The trend equation may be,

$$T_t = a b^t$$

Then through logarithmic transformation, it is reduced to the linear form

$$\log T_t = A + Bt \quad , \text{ where } A = \log a, \; B = \log b$$

Here the normal equations for determining A and B will be

$$\sum \log y_t = nA + B \sum t \quad \text{and}$$
$$\sum t \log y_t = A \sum t + B \sum t^2$$

The antilogarithms of the estimates of A and B give the values of a and b and hence the exponential trend equation is fitted to the time series.

<u>Note</u>: Exponential trend is appropriate when time series data in successive points of time change by more or less a constant ratio.

<u>Merits</u> :- 1) This method is most objective, since the appropriate ~~when time series data in successive points of time change by more or less a constant ratio.~~ curve to be fitted can be objectively determined through graphical representation of the data or otherwise.

2) Unlike the method of moving averages, this method enables us to compute the trend values for all the given time periods in the series.

3) The trend equation can be used to estimate or predict the values of the variable for any period 't' in future or even in the intermediate periods of the given series and the forecast values are also quite reliable.

Demerits :- 1) The method is quite tedious and time-consuming as compared with other methods. It is rather difficult for a non-mathematical person to understand and use.

2) The addition of even a single observation to the given data necessitates all calculations to be done afresh.

3) Future prediction or forecasts based on this method are based only on the long term variations, i.e., trend and completely ignore the cyclical, seasonal and irregular fluctuation.

4) The most serious limitations of this method is the determination of the type of the trend curve to be fitted, viz, whether we should fit a linear or a parabolic trend or some other more complicated trend curve.

Reduction of yearly trend equation to monthly, quarterly and half yearly trend equation

Case I: Suppose the yearly trend equation based on yearly total is

$$T_t = a + bt \quad \text{—— (1)}$$ with origin of t at 1990 (for example) and unit 1 year.

To obtain monthly trend equation, we are to divide the constants by 12 and write $t/12$ for t. Then the trend equation for monthly values is

$$T_t = \frac{a}{12} + \frac{b}{12} \cdot \frac{t}{12} \quad \text{—— (2)}$$

Now, the origin of the above equation is middle of 1990. But monthly trend values should correspond to the midpoints of months. For proper centering, the origin should be shifted half a month to the right or to the left. If we want to shift to the right, i.e., to the middle of July, 1990, we have to write $(t+\frac{1}{2})$ for t and for left i.e. to the middle of June, 1990, $(t-\frac{1}{2})$ for t in (2). Thus, the monthly trend equation will be

$T_t = \frac{a}{12} + \frac{b}{12} \cdot \frac{1}{12}(t+\frac{1}{2})$, with origin of t at July, 1990 and unit 1 month

or, $T_t = \frac{a}{12} + \frac{b}{12} \cdot \frac{1}{12}(t-\frac{1}{2})$, with origin of t at June, 1990 and unit 1 month.

Proceeding in the same way, the quarterly trend equation reduced from (1) will be

$T_t = \frac{a}{4} + \frac{b}{4} \cdot \frac{1}{4}(t+\frac{1}{2})$, with origin of t at 3rd quarter of 1990 and unit 1 quarter

or, $T_t = \frac{a}{4} + \frac{b}{4} \cdot \frac{1}{4}(t-\frac{1}{2})$, with origin at the 2nd quarter of 1990.

Again, the half yearly trend equation will be

$T_t = \frac{a}{2} + \frac{b}{2} \cdot \frac{1}{2}(t+\frac{1}{2})$, origin at 2nd half of 1990 and unit $\frac{1}{2}$ year.

or, $T_t = \frac{a}{2} + \frac{b}{2} \cdot \frac{1}{2}(t-\frac{1}{2})$ when origin is at first half of 1990.

Case II: Suppose the yearly trend equation based on yearly totals for an even number of years, with 1995 and 1996 (for example) in the middle is

$$T_t = a + bt \quad \text{—— (3)}$$, with origin of at at the middle of 1995 and 1996 and unit $\frac{1}{2}$ year.

Preceeding as before, but keeping in keeping in mind that unit of t is now 6 months and origin is end of December, 1995 the monthly trend equation will be

$T_t = \frac{a}{12} + \frac{b}{12} \cdot \frac{1}{6}(t+\frac{1}{2})$, with origin of t at January, 1996 and unit 1 month

or, $T_t = \frac{a}{12} + \frac{b}{12} \cdot \frac{1}{6}(t-\frac{1}{2})$, with origin at December, 1995.

Similarly, quarterly trend equation will be

$T_t = \frac{a}{4} + \frac{b}{4} \cdot \frac{1}{2}\left(t+\frac{1}{2}\right)$ or $T_t = \frac{a}{4} + \frac{b}{4} \cdot \frac{1}{2}\left(t-\frac{1}{2}\right)$

according as origin is at first quarter of 1996 or last quarter of 1995, and unit 1 quarter.

Again, equation for half yearly trend will be

$T_t = \frac{a}{2} + \frac{b}{2}\left(t+\frac{1}{2}\right)$ or,

$T_t = \frac{a}{2} + \frac{b}{2}\left(t-1/2\right)$

according as origin is at first half of 1996 or 2nd half of 1995, and unit 1/2 year.

Second Degree Curve Fitted to Logarithms

Suppose the trend curve is

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

Then through logarithmic transformation, it is reduced to the quadratic form

$$\log T_t = \log a + t \log b + t^2 \log c$$
$$= A + Bt + Ct^2$$

where $A = \log a$, $B = \log b$, $C = \log c$

Hence the normal equations for determining A, B and C will be

$$\sum \log T_t = nA + B\sum t + C\sum t^2$$
$$\sum t \log T_t = A\sum t + B\sum t^2 + C\sum t^3$$
$$\sum t^2 \log T_t = A\sum t^2 + B\sum t^3 + C\sum t^4$$

The antilogarithms of the estimates of A, B and C give the values of a, b and c and hence the trend curve becomes best second degree curve fitted to logarithms.

Remark:

(i) When the time series is found to be increasing or decreasing by equal absolute amounts, the straight line trend is used. The plotting of the data will give a straight line graph.

(ii) When the series is increasing or decreasing by a constant percentage rather than a constant absolute amount, the exponential curve ($T_t = ab^t$) is used. The data plotted on a semi-logarithmic scale will give a straight line graph.

(iii) If the data plotted on a semi-logarithmic scale is not a straight line graph but shows curvature, being concave upward or downward then Second degree curve fitted to logarithms is used.

Advantages and Disadvantages of Trend Fitting by the Principle of Least Square Method :—

Advantages :-

(i) This method completely eliminates the elements of subjective judgement or personal bias on the part of the investigator because of its mathematical nature.

(ii) This method enables us to compute the trend values for all the given time periods in the series.

(iii) The trend equation can be used to estimate or predict the values of the variable for any period 't' in future. It is also reliable for finding the intermediate periods of the given series and the forecast values.

(iv) The least square method is the only technique to obtain the rate of growth per annum, for yearly data, if linear trend is fitted.

Disadvantages :-

(i) This method is quite tedious and time-consuming as compared with other methods. It is also difficult for layman to understand and use.

(ii) All calculations have to be done afresh for the addition of a single new observation.

(iii) Future predictions or forecasts based on this method are based only on the long term variation i.e. trend and completely ignore the cyclical, seasonal and irregular fluctuations.

(iv) The most serious limitations of this method is the determination of the type of the trend curve to be fitted, viz., whether we should fit a linear or a parabolic trend or some other more complicated trend value.

(v) It cannot be used to fit growth curves like Modified Exponential curve, Gompertz curve and Logistic curve.

Growth Curves and Their Fitting

The various growth curves, viz. the modified exponential, Gompertz and Logistic curves cannot be determined by the principle of least squares. Here the curve to be fitted can be made by use of the following theorem based on finite difference.

(i) Modified Exponential Curve and its Fitting :-

Modified exponential curve is given as

$$T_t = a + bc^t \quad , \quad a > 0 \qquad \text{——— (1)}$$

where T_t is the trend equation at time t and a, b, c are constants, called its parameters.

Practically we start computing with time series values y_t.

Taking first difference of equation (1) w.r.t y_t

$$\Delta y_t = y_{t+h} - y_t = (a + bc^{t+h}) - (a + bc^t)$$

$$= bc^t (c^h - 1) \quad \text{where 'h' is the interval of differencing}$$

Similarly,

$$\Delta y_{t-h} = y_t - y_{t-h} = bc^{t-h}(c^h - 1)$$

$$\therefore \quad \frac{\Delta y_t}{\Delta y_{t-h}} = c^h \text{ (constant)}$$

Here we see that the first differences of the consecutive value of y_t corresponding to equivalent values of t change by a constant ratio. So if we plot the first differences of y_t on a semi-logarithmic graph paper, it gives straight line.

Here $y_t = a$ is the only asymptote of the curve.

There are two methods of fitting modified exponential curve.

A) Method of Three Selected Points :-

We take 3 ordinates y_1, y_2, y_3 corresponding to 3 equidistant values of t which are t_1, t_2, t_3 such that

$$t_2 - t_1 = t_3 - t_2$$

Then, $y_1 = a + bc^{t_1}$, $y_2 = a + bc^{t_2}$, $y_3 = a + bc^{t_3}$ ——— (*)

$$\Rightarrow \quad y_2 - y_1 = b(c^{t_2} - c^{t_1}) = bc^{t_1}(c^{t_2 - t_1} - 1) \qquad \text{— (2)}$$

and $$y_3 - y_2 = b(c^{t_3} - c^{t_2}) = bc^{t_2}(c^{t_3 - t_2} - 1) \qquad \text{— (3)}$$

Dividing (3) by (2) we get, $$\frac{y_3 - y_2}{y_2 - y_1} = c^{t_2 - t_1} \Rightarrow c = \left[\frac{y_3 - y_2}{y_2 - y_1}\right]^{\frac{1}{t_2 - t_1}}$$

$[\because t_2 - t_1 = t_3 - t_2]$

Substituting the value of c in (2) we get

$$y_2 - y_1 = b\left[\frac{y_3-y_2}{y_2-y_1}\right]^{\frac{t_1}{t_2-t_1}}\left[\frac{y_3-y_2}{y_2-y_1} - 1\right]$$

or $$y_2 - y_1 = b\left[\frac{y_3-y_2}{y_2-y_1}\right]^{\frac{t_1}{t_2-t_1}}\left[\frac{y_3-2y_2+y_1}{y_2-y_1}\right]$$

or, $$b = \frac{(y_2-y_1)^2}{y_3-2y_2+y_1}\left[\frac{y_2-y_1}{y_3-y_2}\right]^{\frac{t_1}{t_2-t_1}}$$

Now putting the value of b and c in (*) we get the value of a,

$$a = y_1 - bc^{t_1}$$
$$= y_1 - \frac{(y_2-y_1)^2}{y_3-2y_2+y_1}\left[\frac{y_2-y_1}{y_3-y_2}\right]^{\frac{t_1}{t_2-t_1}}\cdot\left[\frac{y_3-y_2}{y_2-y_1}\right]^{\frac{t_1}{t_2-t_1}}$$
$$= \frac{y_1y_3 - 2y_2y_1 + y_1^2 - y_2^2 + 2y_2y_1 - y_1^2}{y_3-2y_2+y_1}$$
$$= \frac{y_1y_3 - y_2^2}{y_3-2y_2+y_1}$$

Now putting the value of a, b and c in (1) we get the trend equation,

$$T_t = \frac{y_1y_3-y_2^2}{y_3-2y_2+y_1} + \frac{(y_2-y_1)^2}{y_3-2y_2+y_1}\left[\frac{y_2-y_1}{y_3-y_2}\right]^{\frac{t_1}{t_2-t_1}}\left[\frac{y_3-y_2}{y_2-y_1}\right]^{\frac{t}{t_2-t_1}}$$

$$= \frac{y_1y_3-y_2^2}{y_3-2y_2+y_1} + \frac{(y_2-y_1)^2}{y_3-2y_2+y_1}\left[\frac{y_2-y_1}{y_3-y_2}\right]^{\frac{t_1}{t_2-t_1}}\left[\frac{y_3-y_2}{y_2-y_1}\right]^{\frac{t_1}{t_2-t_1}}\cdot\left[\frac{y_3-y_2}{y_2-y_1}\right]^{\frac{t}{t_1}}$$

$$= \frac{y_1y_3-y_2^2}{y_3-2y_2+y_1} + \frac{(y_2-y_1)^2}{y_3-2y_2+y_1}\left[\frac{y_3-y_2}{y_2-y_1}\right]^{\frac{t}{t_1}}$$

B) **Method of Partial Sums :-**

Here the given time series data are split up into 3 equal parts. Each part contains n consecutive values of y_t corresponding to $t = 1, 2, \ldots, n$; $t = n+1, n+2, \ldots, 2n$ and $t = 2n+1, 2n+2, \ldots, 3n$.

Let S_1, S_2 and S_3 are the partial sums of the 3 parts respectively so that

$$S_1 = \sum_{t=1}^{n} y_t, \quad S_2 = \sum_{t=n+1}^{2n} y_t, \quad S_3 = \sum_{t=2n+1}^{3n} y_t$$

Now $$S_1 = \sum_{t=1}^{n} y_t = \sum_{t=1}^{n}(a+bc^t) = na + b(c+c^2+c^3+\cdots+c^n)$$
$$= na + bc(1+c+c^2+\cdots+c^{n-1}) \quad \text{—— (1)}$$

$$= na + bc\left[\frac{c^n-1}{c-1}\right] \quad \text{——} \quad (2)$$

Similarly we get, $S_2 = na + bc^{n+1}\left[\frac{c^n-1}{c-1}\right]$

and $S_3 = na + bc^{2n+1}\left[\frac{c^n-1}{c-1}\right]$

Now $S_2 - S_1 = bc\left[\frac{c^n-1}{c-1}\right]\left[c^n-1\right] = bc\frac{(c^n-1)^2}{c-1}$ —— $(*)$

and $S_3 - S_2 = bc^{n+1}\left[\frac{c^n-1}{c-1}\right]\left[c^n-1\right] = bc^{n+1}\frac{(c^n-1)^2}{c-1}$

Now, $\frac{S_3-S_2}{S_2-S_1} = c^n \Rightarrow c = \left[\frac{S_3-S_2}{S_2-S_1}\right]^{1/n}$

Substituting for c in $(*)$ we have

$$S_2 - S_1 = b\left[\frac{S_3-S_2}{S_2-S_1}\right]^{1/n}\frac{1}{\left[\left(\frac{S_3-S_2}{S_2-S_1}\right)^{1/n}-1\right]}\left[\frac{S_3-S_2}{S_2-S_1}-1\right]^2$$

$$= b\left[\frac{S_3-S_2}{S_2-S_1}\right]^{1/n}\frac{1}{\left[\left(\frac{S_3-S_2}{S_2-S_1}\right)^{1/n}-1\right]}\left[\frac{S_3-2S_2+S_1}{S_2-S_1}\right]^2$$

$$\therefore \quad b = \frac{\left(\frac{S_3-S_2}{S_2-S_1}\right)^{1/n}-1}{\left(\frac{S_3-S_2}{S_2-S_1}\right)^{1/n}}\;\frac{(S_2-S_1)^3}{(S_3-2S_2+S_1)^2}$$

Putting the values of b and c in (2) we have

$$S_1 = na + bc\left[\frac{c^n-1}{c-1}\right]$$

$$\text{or} \quad a = \frac{1}{n}\left[S_1 - \frac{bc}{c-1}\left(c^n-1\right)\right]$$

$$= \frac{1}{n}\left[S_1 - \frac{(S_2-S_1)^3}{(S_3-2S_2+S_1)^2}(c^n-1)\right] \quad \left[\because \frac{bc}{c-1}\left[\frac{S_3-S_2}{S_2-S_1}-1\right]^2 = S_2-S_1\right]$$

$$= \frac{1}{n}\left[S_1 - \frac{(S_2-S_1)^3}{(S_3-2S_2+S_1)^2}\left\{\frac{S_3-S_2}{S_2-S_1}-1\right\}\right]$$

$$= \frac{1}{n}\left[S_1 - \frac{(S_2-S_1)^2}{S_3-2S_2+S_1}\right]$$

$$= \frac{1}{n}\left[\frac{S_1S_3 - 2S_2S_1 + S_1^2 - S_2^2 + 2S_2S_1 - S_1^2}{S_3-2S_2+S_1}\right]$$

$$= \frac{1}{n}\left[\frac{S_1S_3 - S_2^2}{S_3-2S_2+S_1}\right]$$

Now the trend equation is

$$T_t = \frac{1}{n}\left[\frac{S_1 S_3 - S_2^2}{S_3 - 2S_2 + S_1}\right] + \frac{\left(\frac{S_3 - S_2}{S_2 - S_1}\right)^{1/n} - 1}{\left(\frac{S_3 - S_2}{S_2 - S_1}\right)^{1/n}} \; \frac{(S_2 - S_1)^3}{(S_3 - 2S_2 + S_1)^2} \cdot \left(\frac{S_3 - S_2}{S_2 - S_1}\right)^{t/n}$$

$$= \frac{1}{n}\left[\frac{S_1 S_3 - S_2^2}{S_3 - 2S_2 + S_1}\right] + \left[\left(\frac{S_3 - S_2}{S_2 - S_1}\right)^{1/n} - 1\right] \frac{(S_2 - S_1)^3}{(S_3 - 2S_2 + S_1)^2} \left(\frac{S_3 - S_2}{S_2 - S_1}\right)^{\frac{t-1}{n}}$$