Notes

HINDUSTAN ZINC

Time series :- A set of ordered quantitative variable taken at successive points in time is known as a 'Time series'. Or arrangment of Statistical data in chronological order i.e. in accordance with occurence of time is known as 'time series'.

A time series depicts the relation ship between two variables one of them being time e.g. the population (y_t) of a country in different years (t) & temperature (y_t) of a place in different days etc.

Mathematically a time series is defined by the functional relationship

$$y_t = f(t)$$

where, y_t is the value of the phenomenon (or variable) under consideration at time t. For ex. (i) the pop^n y_t for years t

(ii) the number of births and deaths (y_t) in different ~~months~~ months (t) of the year

(iii) the Sale of a departmental Store in different months.

thus if the value of a phenomenon (or variable) at times $t_1, t_2, \ldots t_n$ are $y_1, y_2 \ldots y_n$ respectively then the series

t : $t_1, t_2, t_3 \ldots, t_n$
y_t : $y_1, y_2, \ldots, y_n$

constitute a time series. Thus a time series invariably gives a bivariate distribution one is time & other is value.

Components of time series :-

The various forces at work affecting the values of a variable in time series can be broadly classified into the following categories :-

(i) secular trend or long term movement
(ii) periodic changes or short term fluctuations.
   (a) seasonal variations
   (b) cyclic variations.
(iii) Random or irregular movements.

(i) 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.

Thus this is true in most of series of business and economic statistics for ex. an upward tendency would be seen in data pertaining to population, agriculture production, currency in circulation etc. while a downward trend will be seen in data of births & deaths, epidamics etc. as a result of advancement in medical sciences, better medical facilities, literacy & higher standards of living

(1) It may be clearly noted that trend is the general, smooth, long term average tendency. It is not necessarily that the increase or decline should be in same direction throughout the given period. It may be possible that different tendencies of increase, decrease or stability are observed in different sections of time. However the overall tendency may be upward, downward or stable.

(2) It should not be inferred that all the series must show an upward or downward trend. We might come across certain series whose values fluctuate around a constant reading which doesnot change with time.

e.g. the series of barometric readings or the temperature of a particular place.

(3) Linear an Non linear (Curvi linear) trend :-

If the time series values plotted on graph cluster more or less, round a straight line, then the trend exhibited by the time series is termed as linear other wise Non linear trend.

(4) The term 'long period of time' is relative term and cannot be defined exactly. In some cases a period as small as a week may be fairly long while in some cases, a period as long as two years may not be enough. For eg. if the data of agriculture production for 24 months shows an increase it won't be termed as secular change over a period of 2 years whereas if the count of bacterial population

bacteria
of a culture every five minutes, your a week shows an increase, then we would regard it as a secular change.

(ii) Periodic changes :- These forces do not act continously but operate in a regular spasmodic manner. These can be classified as.

1. Seasonal Variations :- In time series that part of the movement which is assigned to the effect of the seasons on the years.
e.g. seasonal variation in rainfall.

These variations in time series are due to rythmic forces which operate in a regular and periodic manner over a span of less than a year. Seasonal variation in time series will be there if data are recorded quarterly, monthly, weekly, daily, hourly & so on.

Most of economic series are influenced by seasonal swings.
eg. prices, production & consumption of commodities; sales & profits in a departmental store, bank clearings & deposits etc. are affected by seasonal variations.

HINDUSTAN ZINC

Notes

These may be attributed to the following two Cases.

(a) Those resulting from natural forces :-

As the name suggests, the various seasons or weather conditions & climate changes play very important role in seasonal movements.

For instance, the sale of umbrellas pick up very fast in rainy season; the demand of electric fan goes up in summer; sale of ice & ice cream increases very much in summer; the sale of woolens go up in winter all being affected by natural forces.

Like wise the production of certain commodities like sugar, rice pulses eggs etc. depends on season.

(b) Resulting from manmade conventions :-

These variations in time series within a period of 12 months are due to habits, fashions customs & conventions of the people in the society. for eg. the sale of jewellery & ornaments goes up in marriages the sales & profit in departmental stores goes up considerably during marriages & festivals

like diwali, dussehra, christmas, eid, durga-puja.

The main objective of the measurement of seasonal variations is to isolate them from the trend and study their effect

free cyclic.

(iii) Irregular or Random movements :-

Apart from the regular variations, almost all the series contain another factor called random or irregular or residual fluctuations, which are not accounted for by secular trend and seasonal & cyclical variations. These fluctuations are purely random, erratic, unforseen, unpredictable and are due to numerous, non recurring and irregular circumstances which are beyond the control of human hand but at the same time are a part of our system such as earthquakes, wars, floods, famines, revolutions, epidamics etc.

(2) Cyclic variations :- The oscillatory movements in a time series with period of oscillation more than one year are termed a cyclic fluctuations.

HINDUSTAN ZINC

Notes

One complete period is called a 'cycle'.

In time series they are generally attributed to the so called business 'cycle' which may also be reffred to as the 'four-phase' 'cycle' composed of prosperity, recession, depression & recovery & normally lasts from seven to eleven years.

The upswings & downswings in business depend upon the cumulative nature of the economic forces and the ~~inter~~interaction between them.

Most of economic and comercial series eg. series relative to prices, production & wages etc. are effected by business cycles.

Notes

HINDUSTAN ZINC

* Measurement of Trend :-

1. Graphic Method :- A free hand smooth curve obtained on plotting the values of y_t against t enables us to form an idea about general 'trend' of series. smoothing of curve eliminates other components ie regular & irregular fluctuations.

This method does not involve any complex mathematical technique and can be used to describe all types of trend linear & non linear.

Thus simplicity & flexibility are strong points of this method.

Drawbacks :- The method is very subjective ie. the bias of the person handling the data plays a very important role and such diff. trend curves will be obtained by diff. persons for the same data.

(b) It does not enables us to measure trend.

(c) Method of Semi Averages :- In this method the whole data is divided into two parts with respect to time e.g. if we are given y_t for t

from 1991-2002 i.e our period of 12 years, the equal two parts will be the data from 91 to 96 & 97 to 2002.

In case of odd no. of years the two parts are obtained by omitting the value corresponding to the middle years eg. for data from 1991 to 2001 the value for year 1996 being omitted.

Next we compute the arithmetic mean for each part & plot these two means against the mid values of the respective time periods covered by each part. The line obtained on joining these two points is the required trend line and may be extended both ways to estimate intermediate or future values.

Advantage:- 1) As compared with graphic method the obvious advantage is its objectivity in the sense that every one who applies it would get the same result. We can also estimate trend values.

2) It is readily comprehensible as compared to the other two methods.

Limitations:- This method assumes linear relationship between the plotted points which may not exist. the limitations of arithmetic mean as an average also stand on its way.

(3) 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 extent (or period) m is a series of successive averages (arithmetic mean) of first m terms at a time.

Thus the first average is the mean of 1st m terms; 2nd is the mean of the m terms from 2nd to (m+1)th term and so on. & 3rd is the mean of m terms from 3rd to (m+2)th term.

If m is odd = (2k+1) (say) moving average is placed against the mid value of time interval it covers.

The graph obtained on plotting the moving average values against the corresponding time values gives the trend curve.

moving average method is very flexible in the sense that the addition of a few more figures to the data simply results in some more trend values the previous calculations are not affected at all.

Drawbacks:-

i) It does not provide trend values for all the terms eg. for a M.A. of period $2k+1$, we have to forgot the trend values for first k and the last k terms of the series.

ii) It cannot be used for forecasting or predicting future trend which is main objective of trend analysis.

M.A. method gives a correct picture of the long term trend of the series if

i) the trend is linear or approx. linear

ii) oscillatory movements affecting the data are regular in period and amplitude.

If trend is not linear, M.A. introduces bias in trends.

(4) Method of Curve fitting by principle of least squared.

→ It is the most popular & widely used method of fitting mathematical functions to a given set of data.

Curve fitting is the process of introducing mathematical relationship between dependent and independent variables in the form of a equation for a given data.

Method of least squares helps us to find the values of unknown a & b in such a way that the following conditions are satisfied.

1. The sum of deviations of observed values of y and corresponding expected values $\hat{y}$ of y will be zero

$$\Sigma(y-\hat{y})=0$$

the sum of squares of the of observed values y & corresponding estimated values should be atleast

i.e. $\Sigma(y-\hat{y})^2$ is at least

when the time series is found to be be increasing or decreasing by equal absolute amounts the straight line trend is used in this case the plotting of data will give a straight line graph.

(a) Fitting of Straight line by least square ~~not~~ method:-

The straight line trend between given time series values y_t and time t be given by the equations

$$y_t = a + bt \quad \text{— (1)}$$

principle of least square consists in minimising the sum of squares of the deviations between the given values of y_t and their estimates given by (1) or in other ~~words~~ sense we have to find a & b such that for given values of y_t corresponding to n different values of t

$$E = \sum_t (y_t - a - bt)^2$$

is minimum. For maxima & minima of E the variations in a & b we should have.

$$\frac{\delta E}{\delta a} = 0 = -2\Sigma(y_t - a - bt)$$

$$\frac{\delta E}{\delta b} = 0 = -2\Sigma t(y_t - a - bt)$$

$$\Rightarrow \Sigma y_t = na + b\Sigma t \quad \text{—} *$$

$$\Sigma y_t = a\Sigma t + b\Sigma t^2 \quad \text{—} **$$

which are normal equations for estimating a & b.

the values of $\Sigma y_t, \Sigma t, \Sigma t^2$ are obtained from data.

(b) fitting of second degree trend / parabolic curve:-

The equation of second degree parabola is

$$y_t = a + bt + ct^2$$

the curve can be fitted by method of least squares. following the usual procedure we get three normal equations.

HINDUSTAN ZINC

Notes

$$\Sigma y_t = na + b\Sigma t + c\Sigma t^2$$

$$\Sigma t y_t = a\Sigma t + b\Sigma t^2 + c\Sigma t^3$$

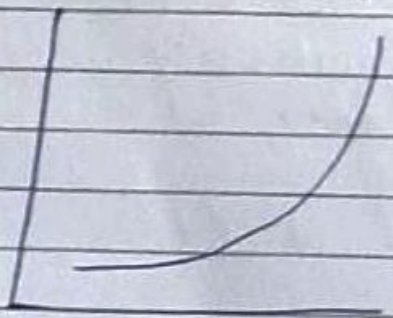
$$\Sigma t^2 y_t = a\Sigma t^2 + b\Sigma t^3 + c\Sigma t^4$$

Summation (Σ) runs over all pairs of Observations.

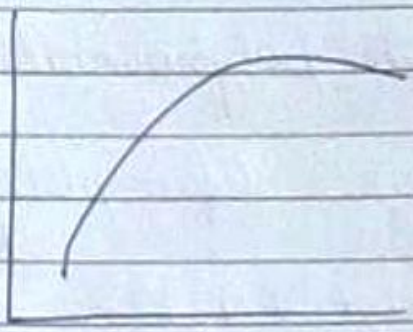
Solving the three equations we obtain the values of a, b & c substituting these values we get 2nd degree parabola.

The shape of second degree curve depends upon the values of b & c.

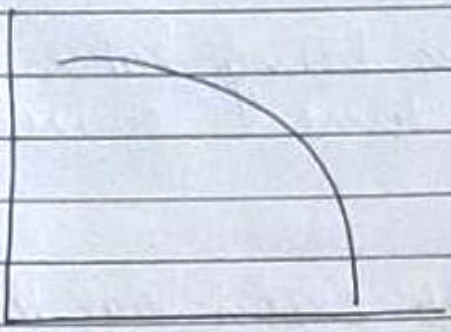
① If $b>0$, $c>0$ the curve is concave from left bottom to right top.



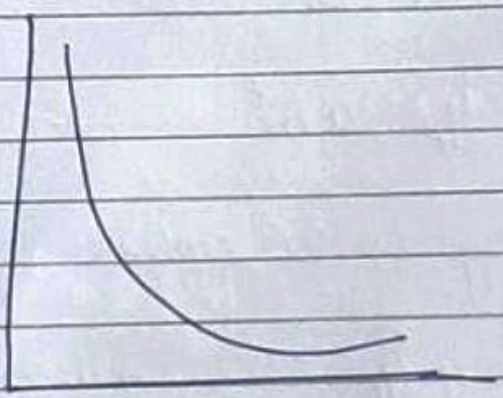
② If $b>0$ & $c<0$ the curve is convex from left bottom to right top.



③ If $b<0$, $c<0$ the curve is convex from left top to right bottom



④ If $b<0$, $c>0$, the curve is concave from left top to right bottom



HINDUSTAN ZINC

Notes

(c) fitting of k^{th} degree polynomial:-

The equation of k^{th} degree polynomial is

$$y_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \dots a_k t^k$$

this is an equation in $k+1$ unknowns
It has constant coefficients

the summation being taken over the values of the time series.

(d) fitting of an ~~exp~~ exponential curve:-

exponential curve is also called geometric trend. The equation of the curve is.

$$y_t = ab^t \quad \text{---(1)}$$

taking log both sides.

$$\log y_t = \log a + t \log b \quad \text{---(2)}$$

$$Y = A + Bt \quad (\text{say}) \quad \text{---(3)}$$

where $Y = \log y_t$, $A = \log a$ & $B = \log b$.

(3) is equation of straight line in t & Y thus the normal equations for estimating A & B are

$$\Sigma Y = nA + B\Sigma t$$

$$\& \quad \Sigma tY = A\Sigma t + B\Sigma t^2$$

these can be solved for A & B & finally we get

$$a = \text{Antilog } A \quad \& \quad b = \text{Antilog } B$$

(c) Second degree curve fitted to logarithms

:- Suppose the trend curve is

$$y_t = a\,b^t\,c^{t^2} \quad \text{— (1)}$$

taking log.

$$\log y_t = \log a + t\log b + t^2 \log c \quad \text{— (2)}$$

$$Y = A + Bt + Ct^2 \quad \text{— (3)}$$

(3) is second degree parabolic curve in t & can be fitted by the same technique of parabola fitting.

a = Antilog A , b = Antilog B , c = Antilog C

with these values of a, b & c the curve (1) becomes the best second degree curve fitted to algorithms.

when n = even then

$$x = \frac{t - \frac{1}{2}(\text{middle 2 years})}{\frac{1}{2}(\text{interval})}$$

elimination of trend is by subtracting trend values from given values.

* Most Plausible solution of a system of linear equation:-
(प्रशस्तीय) (उचित)

Method of least squares is helpful in finding the most plausible values of the variables satisfying a system of independent linear equations whose number is more than the number of variables under study.

Consider the following set of m equations in n variables $x, y, z, \ldots, t$:

$$a_1x + b_1y + c_1z + \cdots + k_1t = l_1$$

$$a_2x + b_2y + c_2z + \cdots + k_2t = l_2$$

$$\vdots$$

$$a_mx + b_my + c_mz + \cdots + k_mt = l_m$$

where $a_i, b_i, \ldots l_i$; $i = 1, 2, \ldots m$ are constants

If $m = n$, the system of equations can be solved uniquely with the help of algebra.

If $m > n$, it is not possible to determine a unique solution $x, y, z \ldots t$ which will satisfy the system. In this case we find the values of $x, y, z \ldots$, which will satisfy the system as nearly as possible.

Principle of least square consists in minimizing the sum of squares of the deviations or residuals or the errors. If

$$E_i = a_ix + b_iy + c_iz + \cdots + k_it - l_i \; ; \; i = 1, \ldots m$$

Is the residual for the i^{th} equation, then we have to determine $X, Y, Z, - - , T$ so that

$$U = \sum_{i=1}^{m} \epsilon_i^2 = \sum_{i=1}^{m} (a_i X + b_i Y + C_i Z + - - + k_i T - l_i)^2$$

is minimum. Using the principle of maxima & minima in differential calculas the partial derivatives of 'U' w.r.t. $X, Y, Z, - - T$ should vanish seperately. Thus.

$$\frac{\delta U}{\delta X} = 0 = \sum_{i=1}^{m} a_i (a_i X + b_i Y + C_i Z + - - k_i T - l_i)$$

$$\frac{\delta U}{\delta Y} = 0 = \sum_{i=1}^{m} b_i (a_i X + b_i Y + C_i Z + - + k_i T - l_i)$$

$$\frac{\delta U}{\delta T} = 0 = \sum_{i=1}^{m} k_i (a_i X + b_i Y + C_i Z + - + k_i T - l_i)$$

These are are known as knormal equations for $X, Y, Z, - - T$ respectively

Thus we have n normal equations in n unknowns $X, Y, Z - - T$ & their unique solution gives the best or the most plausible solution of the system.

Notes

HINDUSTAN ZINC

Note:

A System of linear equations means two or more linear equations that must all be true at the same time.

or

A system of equation is when we have two or more equations working together.

* Merits & Demerits of trend fitting by principle of Least Squares.

Merits.

Most popular & widely used method of fitting mathematical functions to given set of observations.

(1) Because of its mathematical or analytical character this method completely eliminates the element of subjective judgement or personal bias.

(2) unlike the method of M.A. this method enables us to compute the trend values for all given time periods of the given series.

(3) The trend equations can be used to estimate or predict the values of variable

HINDUSTAN ZINC

Notes

for any period t in future or even in intermediate periods of given series & the forecast values are quite reliable.

(4) This is only technique which enables us to obtain growth per annum for yearly data if linear trend is fitted.

Drawbacks:-

(1) Quite tendious & time consuming

(2) long term variation completely ignores cyclic, seasonal & irregular fluctuations.

(3) A new addition of single observation necessitates all calculations to be done afresh.

(4) serious limitation is the determination of type of trend curve to be fitted.

(5) Can not be used to fit growth curves to which most of economics & business series data conform.

Measurement of seasonal variations:-

The determination of seasonal effects is of paramount (सर्वोपरि) importance in planning
(i) business efficiency or
(ii) production program.

For example:-
the head of a departmental store would be interested to study the variations in the demands of different articles for different months in order to plan his future stocks to cater to the public demands due to seasonal swings.

(a) **Method of Simple Averages**:- This is the simplest method of measuring seasonal variation in a time series. and involves the following steps.

(i) Arrange the data by months, years or quarters)

(ii) compute the average $\bar{x}_i$ $(i=1,2,\ldots,12)$ for i^{th} month for all the years.

(iii) Compute the average $\bar{x}$ of the monthly / quarterly averages i.e.

$$\bar{x} = \frac{1}{12}\sum_{i=1}^{12}\bar{x}_i \quad \text{or} \quad \bar{x} = \frac{1}{4}\sum_{i=1}^{4}\bar{x}_i$$

iv) Seasonal indices for different months are ~~attained~~ obtained by expressing monthly (quarterly) averages as percentage of $\bar{x}$

Thus, seasonal index for i^{th} month or quarter $= \frac{\bar{x}_i}{\bar{x}} \times 100$

(Total of seasonal indices is $12 \times 100 = 1200$ for monthly ~~to~~ data

& $4 \times 100 = 400$ for quarterly data.

Merits & Demerits :- This method is based on the basic assumption that the data do not contain any trend & cyclic components and consists in eliminating irregular components by averaging the monthly or quarterly values over different years.

Year	I Qrt	II Qrt	III Qrt	IV Qrt
1995	30	40	36	34
1996	34	52	50	44
1997	40	58	54	48
1998	54	76	68	62
1999	80	92	86	82

(b) Ratio to trend method :- This is an improvement over the simple averages method and is based on assumption that seasonal variation for any given month or quarter is constant factor of trend.

(i) Compute the trend values by the principle of least squares by fitting ~~an exponential~~ appropriate mathematical curve.

(ii) Express the original data as the percentage of trend values. These values will contain the seasonal, cyclic & irregular components
cyclic & irregular

(iii) These components are then wiped out by averaging the percentage for different months / quarters, thus leaving us with indices of seasonal variations.

(iv) If the seasons are quarters the sum of seasonal indices / variation should be 400 & if month it should be 1200. but often the sum is not exactly what it should be. So it is adjusted by multiplying each of them by $\frac{400}{\text{sum of}}$ or $\frac{1200}{\text{S.I total}}$

Merits & Demerits:-

M → i) The Method is based on sound & logical footing.

(ii) It utilizes the complete information.

(iii) If periods are of short duration, it gives very good results.

iv) Easy to compute & understand.

D.M:→ If there are pronounced cyclic swings in the time series. a linear & curvilinear trend does not given good trend values. In such a situation ratio to moving average performs better than ratio to trend.

(c) Ratio to Moving Average Method:-

This is most popular & best on the ground that on taking 12 month M.A. in monthly data & 4-quarter M.A. in quarterly data, the trend & cyclic variations are removed by dividing the actual data by M.A. entered against it.

Notes

HINDUSTAN ZINC

(1) find 12-month or 4-quarter M.A. of data. these values will give estimates of combined effects of trend & cyclic variation (T×C).

(2) express the original data for each month of quarter as the percentages of the centred m.A. values. these values will present seasonal & irregular components.

(3) find the average for each period over the years to eliminate irregular variations. These average or median values for periods represent the seasonal indices.

(4) These indices must be 1200 or 400 for monthly or quarterly. if Not so then they are adjusted by multiplying S.I. by adjusted factor $k = \frac{1200/400}{\text{total of S.I.}}$

Merits: → i) This is based on most logical approach that yields most satisfactory results
(ii) better trend ratio-to-trend.
(iv) flexibility is more

HINDUSTAN ZINC

Notes

M.A Table

Years Quarter or month (1)	No of persons visiting (2)	12 pt no. total (3)	12 pt M.A (4) = (3) ÷ 12	12 Centred M.A. (5)	Ratio to Moving (6) = (2) ÷ (5) × 100

Demerits :- The only drawback is that S.I. for half of beginning period of the first year & half of the periods of last year is not calculatable.

Notes

HINDUSTAN ZINC

(d) Link Relative method:- Also known as pearson's method and is based on averaging the link relatives.

Link Relative is the value of one season expressed as a percentage of the value of the preceding season.

$$\text{Link Relative for any month} = \frac{\text{Current month/q value}}{\text{previous month/q value}} \times 100$$

Steps.

(i) Convert the original data into link relatives (L.R.) by the formula.
(This eliminates the effect of trend)

(ii) Average the link relatives for each month/quarter of the data. Mean or median may be used for averaging.

(iii) Convert the link Relatives to chain relatives. take C.R for the first season = 100

C.R. for 1^{st} period = 100.

$$\text{C.R. for } i^{th} \text{ period} = \frac{\text{L.R. for } i^{th} \text{ period} \times \text{C.R of } (i-1)^{th} \text{ period}}{100}$$

where i varies over all periods (month/quarters)

Or,

$$\text{C.R for any season} = \frac{\text{A.M of L.R of that season} \times \text{C.R. of previous season}}{100}$$

Now by taking last month/quarter value as base a New or second C.R. for first month is calculated.

$$\text{New/2nd C.R. for I season/period} = \frac{\text{Avg. L.R of I}^{st}\text{ period} \times \text{C.R of last period}}{100}$$

Correction factor

$$d = \frac{[\text{New C.R for I quarter} - 100]}{4}$$

if yearly $= \frac{1}{12}$

Adjusted C.R. for I quarter = C.R. of I quarter

A.C.R for 2nd quar. = C.R of II − d

3rd = C.R of III − 2d

4th = C.R of IV − 3d

Take Average of Adjusted (corrected) C.R. = Sum of A.C.R. / no. of prods.

Seasonal index for any period. = Adjusted C.R of that period / Avg of A.C.R. × 100

Merits & Demerits: (i) The link relatives averaged together contain both trend & cyclic components. Although the trend is subsequently eliminated by applying correction the method is effective only if the growth is of constant amount/rate.

(ii) Not so simple as the moving avg. method or so readily adaptable as others to the construction of some or more complex types of seasonal movements.

(iii) This method utilizes data more completely than M.A. method as there is loss by only one link relative i.e. for the first season.

(e) ~~Deseason~~ Deseasonalisation of data:-

As already point out, the objective of studying seasonal variations is (i) to measure them & (ii) to eliminate them from the given series.

Elimination of the seasonal effects from the the given values is termed as Deseasonalization of the data.

It helps ~~us~~ us to adjust the given time series for seasonal variations thus leaving us with trend, cyclic & irregular components.

Assuming multiplicative model of the time series the De-seasonalized,

Notes

HINDUSTAN ZINC

Values are obtained by dividing the given values by corresponding S.I's.

Deseasonalized data $= \frac{Y}{S} = \frac{TCSI}{S}$

It is specially needed for the study of cyclic component. It also helps businessmen & management executives for planning future production programs, for forecasting, and managerial control.

It also helps in proper interpretation of the data.