23✓

Time Series Analysis

07.08.17

Ref: 1. Gupta Kapoor / Parimal Mukherje (Applied Statistics)
2. Jackfield Chatfield

Data recorded at different time points are known as time series data. For example, rainfall in West Bengal over the last five years.

If we observe or record numerical features of an individual or a population for different time points or intervals, the set of observations forms a time series. [Time sets are of equal length; No missing values]

Let, if Y_t be the value of the time series at time t. Then we can write $Y_t = f(t)$. In case of period data t is considered as the mid-point of the t^{th} period.

NOTE: Time series data are analysed to detect the nature of variations in the data and subsequently to enable one to plan the future judiciously. This analysis is mainly important for business forecasting.

□ Components of a time series

When time series data are graphically exhibited, the variations can be readily observed. Apparently the graph represents an overall picture of 'haphazard' movement.

But in reality it is not so. It is observed that atleast a part of the changes (known as the systematic part) can be accounted for while the remaining part is irregular. The factors ~~consisting~~ constituting the systematic part are:

(1) Secular Trend / Trend
(2) Seasonal Variation
(3) Cyclical Fluctuations

Thus the value of time series at time t, Y_t is the resultant of the combined effect of trend ~~(S_t)~~ (T_t), seasonal variation (S_t), cyclical fluctuations (C_t) and irregular variations (I_t)

□ Two Models / Approaches

Ans. 1

There are two approaches by which we can express Y_t

(1) Additive Model or Sum Model

Here Y_t is expressed as,

$$Y_t = T_t + S_t + C_t + I_t$$

Moreover, here all the components ~~[illegible]~~ has the same unit as that of Y_t.

(2) Multiplicative Model (Product Model)

Here Y_t is expressed as,

$$Y_t = T_t \cdot S_t \cdot C_t \cdot I_t$$

T_t has the same unit as that of Y_t whereas S_t, C_t, I_t are unit free.

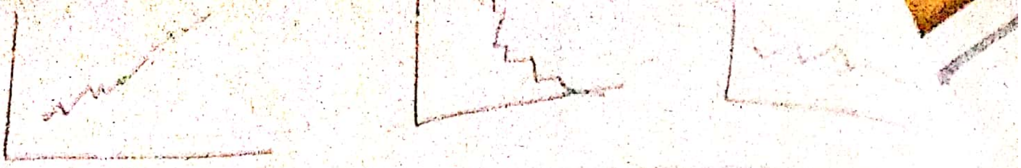
that

ile the

~~resisting~~

NOTE: In additive model, all the components are assumed to be independent but in multiplicative model, no such assumption is

However



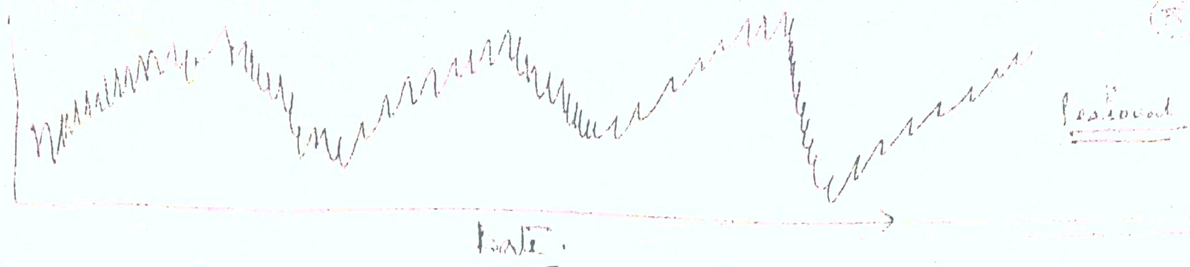
more or less at a constant level. For example an upward trend maybe observed in data on population while a downward trend in money value and a series of barometric readings of a particular place may remain more or less at the same level.

(ii) Seasonal Variations: It is a periodic movement (i.e. a movement that repeats itself at regular time intervals) of a series where the period is not longer than one year. It is found in most most series of economics statistics for which weekly, monthly or quaterly values are obtainable. For example, monthly expenditures of a family, quaterly sale in a departmental store, no. of books issued from a library on different days of a week, etc. These variations in economic time series is possibly due to two broad factors

(a) climatic changes of various seasons (for example the sale of ice creams, demand of electric fans goes up in summer)

(b) habits & customs followed by people at different times (for example sale of certain consumer goods increases during festivals)

(iii) Cyclical Fluctuations: It is an oscillatory movement in a time series with period of oscillation more than a year. One complete period is known as a cycle. These variations though more or less regular are not necessarily periodic. The cyclical fluctuations in a time series are usually attributed to a business cycle comprising 4 successive phases —— prosperity (boom), recession, ~~deg~~ depression & recovery. The swing from boom to



recovery & back again to boom is found to vary in time span. Prices, wages, volume of production, etc are affected by business cycles.

(iv) Irregular ~~Variation~~ Fluctuations: Apart from the regular variations, all the series contain another factor known as random / irregular which are not accounted for by trend, seasonal and cyclical variations. These variations are purely random or completely unaccountable or are caused by unpredictable events like floods, wars, earthquakes, etc which are beyond the control of human hand.

□ Non-Stationary Time Series

A. Measurement of Trend

Ans. 2

Trend can be measured by eliminating the remaining components from the time series using the following methods.

(i) Method of free-hand curve fitting
(ii) Method of moving averages
(iii) Method of mathematical curves

(i) Method of free-hand curve fitting: The simplest method of describing a trend is by inspection. Here we first draw the line diagram for the data and then we draw a free hand smooth curve which fits the data best. This method is used both for linear & non-linear trend. This method is quite subjective and there is no mathematical law behind it.

(ii) Method of moving averages:

In this method a series of arithmatic means each of k successive obs. of the given data is computed and these means are referred to as moving averages of period k. To begin with we take the first k values then, exclude the first and include the (k+1)th value & so on. We repeat this process until we reach at the last set of k values. Each mean is placed against the mid-point of the time interval it covers. If k is odd the moving averages correspond to tabulated times for which the time series is given. On the other hand, when k is even each moving average falls midway between two tabulated time values. Hence a subsequent two item moving averages is computed to make the resultant moving averages values correspond to the given times. (This is known as centering). This moving averages are the trend values for the corresponding items.

Example: 1. Determine trend in 3 yearly moving average

Year	Profit (in ₹1000)	3-yearly moving total	3-yearly moving average (Trend values)
1991	[illegible]	—	—
1992	90	272	90.66
1993	94	280	93.33
1994	96	288	96
1995	98	287	95.66
1996	96	—	—

2. Determine trend in 4 yearly moving average

Year	Profit (in ₹ 1000)	4 yearly moving total	2 item moving total	4 yearly moving average
1991	85			
1992	90	365		
1993	97	378	743	92.875
1994	93	384	762	95.25
1995	98			
1996	96			

28.08.2017

Note:- This method is very simple and flexible. It does not involve complicated calculations. Further, the inclusion of a few additional observations to the given series simply results in the computation of some more averages. The previous calculations remain unaffected. This method possess the merit of objectivity, since the period of the moving average can be more or less objectively determined.

However, this method is satisfactory when the trend is more or less linear, but unsatisfactory for non-linear trend. Moving average method cannot be used for forecasting or predicting future trend, which is the main objective of trend analysis, because of the absence of any specific mathematical equation. Another drawback of this method is that some trend values at each end of series cannot be estimated.

(iii) Method of Mathematical Curve :

The method of moving average cannot be used for forecasting purposes, as there does not exist any mathematically expressed trend equation. Thus an attempt is made to fit the observed time series with a simple mathematical curve. In this method, a suitable trend equation from the graphical presentation of the given time series is first selected, and then the constants associated with the equation are estimated on the basis of the given data.

(a) Straight Line trend (Linear trend)

$$T_t = a + bt \quad , t \neq 0$$

(b) Second Degree Polynomial (Parabolic trend)

$$T_t = a + bt + ct^2 \quad , c \neq 0$$

(c) Exponential Curve

$$T_t = a \cdot b^t \quad , a > 0$$

Linear Trend:

Y_t : observed value of the time series at time t

$T_t = a + bt$ (estimated from the graphical representation)

"Method of Least Squares"

$$S = \sum_t (Y_t - T_t)^2 \quad \text{will be min. w.r.t } a \,\&\, b$$

$$S = \sum_t (Y_t - a - bt)^2$$

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

$$\Rightarrow \sum_t Y_t = na + b\sum_t t \qquad \text{——— (i)}$$

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

$$\Rightarrow \sum_t t.Y_t = a\sum_t t + b\sum_t t^2 \qquad \text{——— (ii)}$$

Second degree Polynomial (Parabolic Trend)

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

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

$$\left.\begin{aligned} \sum Y_t &= na + b\sum t \\ \sum t Y_t &= a\sum t + b\sum t^2 \\ \sum t^2 Y_t &= a\sum t^2 + b\sum t^2 + c\sum t^3 \end{aligned}\right\}$$

Let us suppose that the trend is linear. Let the trend equation be,

$$T_t = a + bt \quad , \quad b \neq 0$$

Here, the constants 'a' & 'b' are estimated by the method of least squares. Suppose we are given the values y_t for 'n' equidistant values of t. We minimise

$$S = \Sigma (y_t - T_t)^2$$

$$= \Sigma (Y_t - a - bt)^2 \quad \text{w.r.t. 'a' \& 'b'}$$

The normal equations are :-

$$\Sigma Y_t = na + b \Sigma t \quad \text{——— (i)}$$
$$\Sigma t Y_t = a \Sigma t + b \Sigma t^2 \quad \text{——— (ii)}$$

If $\hat{a}$ & $\hat{b}$ are the estimates of 'a' & 'b' resp., then the trend equation becomes,

$$T_t = \hat{a} + \hat{b} t$$

Note: When Y_t increases or decreases by equal absolute amounts, a straight line trend is used.

Second degree Polynomial

Let $T_t = a + bt + ct^2 \quad , \quad c \neq 0$

The normal equations are :-

$$\Sigma Y_t = na + b \Sigma t \quad \text{——— (i)}$$
$$\Sigma t Y_t = a \Sigma t + b \Sigma t^2 \quad \text{——— (ii)}$$
$$\Sigma t^2 Y_t = a \Sigma t + b \Sigma t^2 + c \Sigma t^3 \quad \text{——— (iii)}$$

Let, $\hat{a}, \hat{b}, \hat{c}$ be the estimates of a, b & c resp.

Scanned with CamScanner

Then the trend equation becomes $T_t = \hat{a} + \hat{b}t + \hat{c}t^2$

For computational purpose, we will consider the mid-point of the entire time-span as origin and as a result, sum of 't' values can be reduced to zero.

Eg. 1. Fit a straight line (linear trend) to the following figures:-

Year	1980	1981	1982	1983	1984	1985
Production (in 1000 tonnes)	75	83	109	129	134	148

Also find the production in the year 1988 & the trend values for these years.

Year	Production	t	t^2	tY_t
1980	75	-5	25	-375
1981	83	-3	9	-249
1982	109	-1	1	109
1983	129	1	1	129
1984	134	3	9	402
1985	148	5	25	740
			70	538

$t = \dfrac{(\text{Year} - \text{Midpoint})}{\text{Unit}}$, Unit $= \frac{1}{2}$, even no. of years

$= 1$, odd " "

$\Sigma Y_t = 6a \quad \therefore 678 = 6a \quad \Rightarrow a = 113$

$\Sigma tY_t = a\Sigma t + b\Sigma t^2 = b\Sigma t^2$

$\Rightarrow 538 = 70b \quad \Rightarrow b = 7.69$

Hence, the linear trend equation is —

$T_t = 113 + 7.69t$, with origin at the midpoint of 1982 to 1983

and half year as unit of t.

For the year 1988, $t=11$ and hence,

$T_{1988} = 113 + 7.69(11)$
$= 197.6$ [in 1000 tonnes]

▶ Exponential Curve

Ans-14.

Here, the trend equation will be given by
$Y_t = a \cdot b^t$, where $a, b > 0$

Therefore, $\log Y_t = \log a + t \log b$
i.e. $\log Y_t$ is a linear function of t.

Note that $\dfrac{Y_t}{Y_{t-1}} = b$, i.e. the exponential curve indicates a constant ratio of change

If $0 < b < 1$, Y_t values gradually decay but if $b > 1$, Y_t values gradually increases

Fitting of Exponential Curve

Let the trend equation be $T_t = a \cdot b^t$

$\Rightarrow \log T_t = \log a + t \log b$

$\Rightarrow Y_t' = A + Bt$.

$$\left[\begin{aligned} Y_t' &= \log T_t \\ A &= \log a \\ B &= \log b \end{aligned} \right]$$

The normal equations are

$$\sum Y_t' = nA + B\sum t$$
$$\sum t Y_t' = A\sum t + B\sum t^2$$

Let, $\hat{A}$ and $\hat{B}$ be the estimates from (i) and (ii). Therefore,

$$\hat{A} = \log \hat{a} \quad \text{and} \quad \hat{B} = \log \hat{b}$$
$$\Rightarrow \hat{a} = e^{\hat{A}} \quad \text{and} \quad \hat{b} = e^{\hat{B}}$$

Hence, the fitted trend equation is given by

$$T_t = \hat{a} \cdot \hat{b}^t$$

Modified Exponential Curve

Ans. 19.

Over long periods of time, time series are not likely to have either a constant amount of change or a constant ratio of change. It is more likely that an increasing series (decreasing series) will show an increasing (or a decreasing) amount of change, but at a decreasing ratio of change. It is also possible that an increasing series may show a decline in the amount of increase. To encounter this problem, ~~we~~ ~~shall~~, we shall study the modified exponential curve, which is given as:-

$$Y_t = k + ab^t$$

Note that $\Delta Y_t = Y_{t+1} - Y_t = a\left(b^{t+1} - b^t\right)$

$$\frac{\Delta Y_t}{\Delta Y_{t-1}} = \frac{a\left(b^{t+1} - b^t\right)}{a\left(b^t - b^{t-1}\right)} = \frac{b^t (b-1)}{b^{t-1}(b-1)} = b.$$

06.09.17

Fitting of Modified Exponential Curve

Ans. 19.ii.

The curve has 3 constants, k, a & b
Thus 3 equations are required for fitting. Here, we will use the method of partial sums. Here, we first divide the observed series, Y_t into 3 equal sections. Let the no. of terms in each of the 3 sections be n, i.e. the first one from 1 to n, (n+1) to 2n for the next section and (2n+1) to 3n for the last section. The subtotals are represented as S_1, S_2 & S_3 respectively. Therefore,

$$S_1 = \sum_{t=1}^{n} k + a.b^t$$

$$= nk + a \sum b^t$$

$$= nk + a[b + b^2 + b^3 + \dots b^n]$$

$$= nk + ab[1 + b + b^2 + \dots + b^{n-1}]$$

$$= nk + \frac{ab.b^{n-1}}{b-1}$$

$$= nk + \frac{ab.b^{n-1}}{b-1}.$$

Similarly, $$S_2 = \sum_{t=n+1}^{2n} k + ab^t$$

$$= nk + a \sum b^t$$

$$= nk + a[b^{n+1} + b^{n+2} + \dots + b^{2n}]$$

$$= nk + ab^{n+1}[1 + b + \dots b^{n-1}]$$

$$= nk + \frac{ab^{n+1}(b^n - 1)}{b-1}$$

$$S_3 = \sum_{t=2n+1}^{3n} k + ab^t$$

$$= nk + a\sum_{t=2n+1}^{3n} b^t$$

$$= nk + a[b^{2n+1} + b^{2n+2} + \dots + b^{3n}]$$

$$= nk + ab^{2n+1}\left(\frac{b^n - 1}{b-1}\right) = nk + ab^{2n+1}[1 + b + \dots b^{n-1}]$$

$$\therefore \quad s_2 - s_1 = nk + ab^{n+1}\left[\frac{b^n - 1}{b-1}\right] - nk - ab\left[\frac{b^n-1}{b-1}\right]$$

$$= ab^{n+1}\left[\frac{b^n-1}{b-1}\right] - ab\left[\frac{b^n-1}{b-1}\right]$$

$$= \left[\frac{b^n-1}{b-1}\right]\left[ab^{n+1} - ab\right]$$

$$= \left[\frac{b^n-1}{b-1}\right]\left[ab(b^n-1)\right]$$

$$= \frac{ab(b^n-1)^2}{(b-1)}.$$

$$s_3 - s_2 = nk + ab^{2n+1}\left[\frac{b^n-1}{b-1}\right] - nk - ab^{n+1}\left[\frac{b^n-1}{b-1}\right]$$

$$= \left[\frac{b^n-1}{b-1}\right]\left[ab^{2n+1} - ab^{n+1}\right]$$

$$= \left[\frac{b^n-1}{b-1}\right]\left[ab^{n+1}(b^n-1)\right]$$

$$= \frac{(b^n-1)^2}{b-1}\, ab^{n+1}$$

$\dfrac{s_3 - s_2}{s_2 - s_1} = b^n \longrightarrow$ From here, we get the value of b.

$$\therefore \quad b = \left(\frac{s_3 - s_2}{s_2 - s_1}\right)^{1/n}$$

Now, $s_2 - s_1 = ab\left(\dfrac{b^n-1}{b-1}\right)(b^n - 1)$.

Putting the value of b, we now find the value of $\hat{a}$.

$$s_1 = nk + ab\left(\frac{b^n-1}{b-1}\right)$$

Putting the value of a & b, we find the value of k

Note : If the no. of year is not a multiple of 3, slightly overlaping intervals has to be taken. For ex. if there are 10 year, first 4 years group 1, (4-7)th group 2, (7-10)th group 3, Group 4 & 7 are getting repeated.

Gompertz curve

Ans. 2(i)

Here the trend equation will be given by $T_t = k(a^b)^t$ — (1).

This curve is useful to find the trend in which the growth

[redacted]

07.09.17.

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

Let, $T_t = a + bt$ is a yearly trend equation with unit of t = 1 year & origin 2016

Module 2

Measurement of Seasonal Variation

There are 3 methods of measurement of seasonal variation.

(i) Averages of unadjusted data / Method of monthly (or quarterly) averages

*(ii) Ratio to moving average.

* (iii) Ratio to trend.

(i) Averages of unadjusted data

When the data do not contain trend or cyclical movements to any appreciable extend then to find out seasonal variation we will use this method. Here we have to eliminate the irregular variations only & that is done by averaging the monthly / quarterly / weekly values over different time points.

Let y_{ij} be the observation of the jth month of ith year. Compute the averages $\bar{y}_{ij} \Rightarrow$ $\bar{y}_j = \dfrac{\sum_{i=1}^{n} y_{ij}}{\text{no. of years}}$; $j = 1(1)12$

This $\bar{y}_j$ measures the seasonal component for the jth month, $j = 1(1)12$. To express $\bar{y}_j$ as seasonal indices they are to be shown as %s of a grand mean. Thus seasonal index for the jth month $= \bar{y}_j / \bar{y} \times 100$

Hence the total of seasonal indices for monthly data will be 1200 or 400. We have used multiplicative model. For additive model, the grand mean is to be substracted from $\bar{y}_j$. To obtain the total seasonal indices equal to 0.

11.09.2017

(ii) Ratio to Moving Average Method

Let, y_{ij} be the observation for the jth month in the ith year. Assuming multiplicative model y_{ij} is composed of ~~T_t into~~ $T_t \cdot S_t \cdot C_t \cdot I_t$. So from monthly data we compute moving averages say $y_{ij}(s)$. This $y_{ij}(s)$ gives a rough estimate in the combined effects of T_t & C_t, then the ratio $\frac{Y_{ij}}{y_{ij}}$ gives ~~an~~ an estimate of the seasonal variation with a part of irregular variation.

* (Then ~~Define~~ $r_{ij}' = \frac{Y_{ij}}{y_{ij}} \times 100$ shows some variation from year to year. The values of r_{ij}' for each month are averaged to remove irregular fluctuation i.e. we compute $r_j = \frac{\sum r_{ij}'}{\text{no. of years}}$

If $\sum_{j=1}^{12} r_j \neq 1200$, the seasonal indices of monthly variations are then obtained by adjusting the r_j's to add upto 1200. The adjustment factor is $\lambda = \frac{1200}{\sum r_j}$ and then the reqd. seasonal indices are given by $S_j = \lambda \cdot r_j$; $j = 1(1)12$)

(iii) Ratio to trend

This method is recommended when cyclical variation is known to be absent or when it is not so pronounced even if present.

Let Y_{ij} be the obs. for the jth month of the ith year. Further, let T_{ij} be the trend value obtained from the trend equation on yearly data.

Assuming multiplicative model we compute the ratios to trend $r_{ij} = \frac{Y_{ij}}{T_{ij}}$

*

Problem :-

1. Kolkata has shown no appreciable change in the annual rainfall over the years. Discuss how you will find the seasonal variation in rainfall in Kolkata given the last 5 years monthly data.

2. The daily flow of traffic has observed for 865 days of a year on a particular road in the city. The data revealed that

   (i) the flow varied over the different days of a week and

   (ii) the flow had a gradual increase over the weeks

   Describe a method to analyse the data.

Residual Series

21.09.17.

The elimination of ~~two~~ trend from time series data is known as the detrending of the time series. The elimination of seasonal variation is known as the deseasonalisation of the time series. When trend & seasonal variation have been removed from the data we are left with the series which will present fluctuations of a more or less regular kind known as the residual series.

Kendall's Test of Randomness

Ans. 8

Given an ordered series of obs. $u_1, u_2, \ldots, u_n$

in that order, let us count the no. of pairs for which $u_j > u_i$ for all $j > i$. Let this be p, there [redacted]

12.10.17

Module 4

Probabilistic / Stochastic Time Series

A time series maybe defined as a collection of random variables which are ordered in time and defined at a set of time points. Here a time series is looked upon a real valued function $x_t(\omega)$ defined on a parametric space Ω and the index set T.

□ Stationary Time Series: A time series $\{X_t : t \in T\}$ is said to be strictly stationary if the joint distribution of $X_{t_1}, X_{t_2}, \ldots, X_{t_m}$ is the same as the joint distribution of $X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_m+h}$.

$$f(X_{t_1}, X_{t_2}, \ldots, X_{t_m}) = f(X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_m+h})$$

In particular, if $m=1$ strict stationarity implies the dist. of X_t is same for all t. & hence $E(X_t) = \mu$ (say) & $V(X_t) = \sigma^2$ (say)

i.e. mean & variance are constant.

□ Autocovariance function: The covariance between X_t & X_{t+h} seperated by h intervals of time is called the autocovariance function at lag h & is defined as

$$\gamma(h) = Cov(X_t, X_{t+h})$$

$$= E[(X_t - \mu)(X_{t+h} - \mu)]$$

~~Simill~~ Similarly the autocorrelation function at lag h is the correlation between X_t & X_{t+h} & is defined as

$$\rho(h) = Cov(X_t, X_{t+h}) / \sqrt{V(X_t) V(X_{t+h})}$$

$$= \frac{\gamma(h)}{\sqrt{\gamma(0)\gamma(0)}}$$

$$= \frac{\gamma(h)}{\gamma(0)}$$

$\because \gamma(0) = Cov(X_t, X_t) = Var(X_t)$

$\rho(0) = 1.$

Properties of autocorrelation:-

(i) $\rho(h) = \rho(-h)$ i.e. $\rho(h)$ is an even fn. of the lag h

Pf. $\gamma(h) = cov(X_t, X_{t+h})$ $\quad cov(X,Y) = cov(Y,X)$

$= cov(X_{t-h}, X_t)$

$= cov(X_t, X_{t-h})$

$= cov(X_t, X_{t+(-h)})$

$= \gamma(-h)$

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\gamma(-h)}{\gamma(0)} = \rho(-h)$$

Hence proved.

(ii) $|\rho(h)| \leq 1$

Pf. We know, $V(X_t \pm X_{t+h}) \geq 0$

$\Rightarrow var(X_t) + Var(X_{t+h}) \pm 2\,cov(X_t, X_{t+h}) \geq 0$

$\Rightarrow 2\gamma(0) \pm 2\gamma(h) \geq 0$

$\Rightarrow 2\gamma(0)\left[1 \pm \frac{\gamma(h)}{\gamma(0)}\right] \geq 0$

$\therefore 1 \pm \frac{\gamma(h)}{\gamma(0)} \geq 0 \quad [\because 2\gamma(0) \geq 0]$

$\Rightarrow 1 \pm \rho(h) \geq 0$

$\Rightarrow \rho(h) \geq -1$ & $\rho(h) \leq 1$

$\Rightarrow -1 \leq \rho(h) \leq 1 \Rightarrow |\rho(h)| \leq 1$

Hence proved.

Weakly Stationarity

16.10.17.

Ans. 5(i) A time series is said to be weakly stationary of order r if the moments of the process ~~upto~~ upto order r depends only on time differences.

Particular case: A time series is said to be second order stationary (i.e. weakly stationary of order 2) if its mean is constant & the autocovariance function depends only on the lag between the time differences,

i.e. $E(X_t) = \mu \quad \forall t$

& $\operatorname{cov}(X_t, X_{t+h}) = \gamma(h)$

Here put $h=0$, we get

$\operatorname{cov}(X_t, X_t) = \gamma(0)$

$\Rightarrow \operatorname{Var}(X_t) = \gamma(0)$ = a constant

Correlogram

The autocorrelation coefficient $\rho(h)$ as a fn. of lag h is known as the autocorrelation fn. of the process. The graph obtained by plotting $\rho(h)$ as ordinate against the lag h as abscissa is called the correlogram.

From a correlogram we can have an idea about the nature of internal dependence of the time

Some useful time series.

(i) A purely random process: A discrete time series $\{\epsilon_t\}$ is called a purely random process if it consists of uncorrelated random variables ϵ_t with mean 0 & constant variance.

i.e. $E(\epsilon_t) = 0$ & $Var(\epsilon_t) = \sigma_\epsilon^2$ (a constant ind. of t)

as ϵ_t's are uncorrelated,

$$\therefore \ Cov(\epsilon_t, \epsilon_{t+h}) = \begin{cases} 0, & h \neq 0 \\ \sigma_\epsilon^2, & h = 0 \end{cases}$$

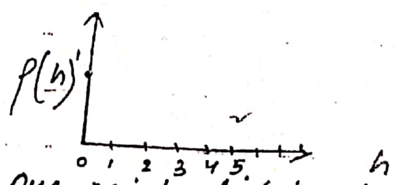
$$\therefore \ \gamma(h) = \begin{cases} 0, & h \neq 0 \\ \sigma_\epsilon^2, & h = 0 \end{cases}$$

Thus the mean is constant & the autocovariate fn. does not depend on time. Hence the process is second order stationary. The autocorrelation fn. $\rho(h)$,

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 0, & h \neq 0 \\ 1, & h = 0 \end{cases}$$

$\cancel{\rho(h) = \{}$

Such a process is also known as <u>White Noise</u>.



One point distribution
(degenerate ")

(ii) Random Walk : suppose that ϵ_t is a purely random process with mean 0 & variance σ_ϵ^2. A process / time series $\{X_t\}$ is said to be a random walk if $X_t = X_{t-1} + \epsilon_t$ with $X_0 = 0$

Ans. 11

$\therefore \quad X_1 = \epsilon_1$

$X_2 = X_1 + \epsilon_2 = \epsilon_1 + \epsilon_2$

.of t)

$\therefore$ In general

$$\boxed{X_t = \sum_{i=1}^{t} \epsilon_i}$$

$$E(X_t) = E\left[\sum_{i=1}^{n} \epsilon_i\right]$$

$$= \sum_i E(\epsilon_i) = 0.$$

$$Var(X_t) = Var\left(\sum_{i=1}^{t} \epsilon_i\right)$$

$$= \sum_{i=1}^{t} var(\epsilon_i) \quad [\text{as } \epsilon_i\text{'s are uncorrelated}]$$

te fn.
econd

$$= t\,\sigma_\epsilon^2$$

i.e. $Var(X_t)$ depends on the time t i.e. it is not a constant. Hence the process is non-stationary. However the first differences of X_t, is given by,

$$\Delta X_t = X_t - X_{t-1} = \epsilon_t$$

h
on
)

ΔX_t is a purely random process & hence stationary.

Recap:

- $\{X_t\}$ is weakly stationary if $E(X_t) = 0$

  $\gamma(h) = Cov(X_t, X_{t+h})$ = depends only on h & not on t

  auto ~~correlation~~ covariance fn.

  or $Var(X_t) = \gamma(0) = \sigma$ constant.

- $\{X_t\}$ is purely random process (white noise)

  if $E(X_t) = 0$ & $Var(X_t) = \sigma^2$

  & X_t's are ind.

  $\rho_h \to y$ axis } diagram

$Cov(aX+bY, cX+dY)$
$= ac\,V(X) + ad\,Cov(X,Y)$
$+ bd\,V(Y) + bc\,Cov(X,Y)$

$\beta_h \epsilon_t$ $\qquad$ $\beta_h \epsilon_t$

$Cov(X_t, X_{t+h})$

$$= Cov\left(\beta_0\epsilon_t + \beta_1\epsilon_{t-1} + \cdots + \beta_q\epsilon_{t-q},\ \beta_0\epsilon_{t+h} + \beta_1\epsilon_{t+h-1} + \cdots + \beta_q\epsilon_{t+h-q}\right)$$

$$= \beta_0\beta_h\, var(\epsilon_t) + \beta_1\beta_{h+1}\, var(\epsilon_{t-1}) + \cdots + \beta_{q-h}\beta_q\, var(\epsilon_{t-q})$$

$$= \begin{cases} \beta_0\beta_h\sigma^2 + \beta_1\beta_{h+1}\sigma^2 + \cdots + \beta_{q-h}\beta_q\sigma^2 \;\cancel{(\epsilon_{t+h-q})} & ; h \leq q \\ 0 & ; o.w \end{cases}$$

These terms will have all the ϵ_t's common other terms vanishes as $Cov(\epsilon_t, \epsilon_{t+h}) = 0$.

$\therefore$ $E(X_t) = 0$ & $Cov(X_t, X_{t+h})$ depends only on h & not on t

Hence $MA(q)$ process is weakly stationary.

The autocorrelation fn is given by

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

$$= \begin{cases} 1 & ; h = 0 \\ \sum_{i=0}^{q-h}\beta_i\beta_{h+i} \Big/ \sum_{i=0}^{q}\beta_i^2 & ; h = 1, 2, \ldots, q \\ 0 & h > q \end{cases}$$

$\gamma(h) = \sigma^2\sum_{i=0}^{q-h}\beta_i\beta_{h+i}$

$\gamma(0) = var(X_t) = \sigma^2\sum_{i=0}^{q}\beta_i^2$

Also, $\rho(-h) = \rho(h)$

Thus we see a autocorrelation fn of an $MA(q)$ process is zero beyond the lag q i.e. the order of the process. In other words $\rho(h)$ has a cut off at lag $q+1$

30.10.17.

Special case :

1. MA(1) : It is given by,

$$X_t = \beta_0 \epsilon_t + \beta_1 \epsilon_{t-1} \; ; \text{ with } \beta_0 = 1$$

(i) $E(X_t) = 0$

(ii) $Var(X_t) = \sigma^2 \sum_{i=0}^{1} \beta_i^2 = \sigma^2 (\beta_0^2 + \beta_1^2) = \sigma^2 (1 + \beta_1^2)$

(iii) $cov(X_t, X_{t+h}) = Cov(X_t, X_{t+1}) = \sigma^2 \sum_{i=0}^{0} \beta_i \beta_{h+i}$

$$= \sigma^2 \beta_1 = \cancel{\gamma(h)}\, \gamma(1)$$

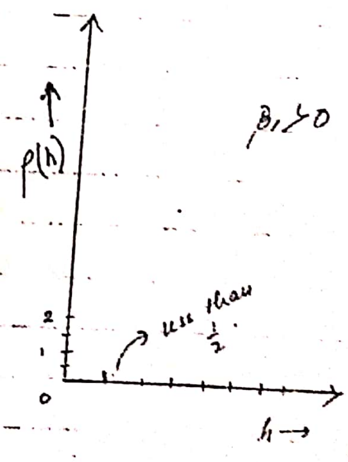
(iv) $$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 1 & ; h = 0 \\ \frac{\beta_1}{1+\beta_1^2} & ; h = 1 \\ 0 & ; h \geq 2 \end{cases}$$

$$\rho(1) = \frac{\beta_1}{1+\beta_1^2}$$

$$\Rightarrow \rho(1) + \beta_1^2 \rho(1) - \beta_1 = 0$$

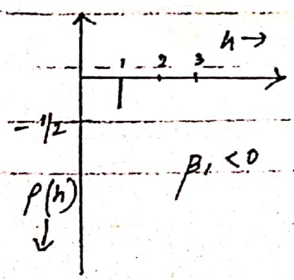
$$\Rightarrow \beta_1^2 \rho(1) - \beta_1 + \rho(1) = 0$$

$$\therefore \beta_1 = \frac{+1 \pm \sqrt{1 - 4\rho^2(1)}}{2\rho(1)}.$$



$\because$ β_1 is real

$\therefore 1 - 4\rho^2(1) \geq 0$



$$\Rightarrow \rho^2(1) \leq \frac{1}{4}$$

$$\Rightarrow |\rho(1)| \leq \frac{1}{2}. \quad \underline{OR} \quad -\frac{1}{2} \leq \rho(1) \leq \frac{1}{2}.$$

Note: Moving Average Process cannot be identified uniquely from a given autocorrelation function.

Ex. Consider a MA(1) process

$$X_t = \beta_0 \epsilon_t + \frac{1}{\beta_1} \epsilon_{t-1} \; ; \text{ where } \beta_0 = 1$$

$$\rho(h) = \frac{1/\beta_1}{1 + 1/\beta_1^2} = \frac{1/\beta_1}{1 + \beta_1^2/\beta_1^2} = \frac{\beta_1}{1+\beta_1^2}$$

which is the autocorrelation fn. of the usual MA(1) process

Thus we cannot identify a MA process uniquely from a given autocorrelation fn.

2. MA (2) $\longrightarrow$ Assignment.

Ex. Find the autocorrelation fn. of a MA(2) process given by $X_t = \epsilon_t + 0.7 \epsilon_{t-1} + 0.2 \epsilon_{t-2}$ where ϵ_t is a purely random process with mean 0 & ~~varior~~ variance σ^2. Also draw the correlogram.

Ans. 29

$$E(X_t) = 0$$

$$Var(X_t) = \sigma^2 + (0.7)^2 \sigma^2 + (0.2)^2 \sigma^2 \quad [\text{as } Cov(\epsilon_t ; \epsilon_{t+h}) = 0 \; \forall h \neq 0]$$

$$= \sigma^2 [1 + 0.49 + 0.4] = \sigma^2 (0.53 + 1)$$

$$= (1.53) \sigma^2 = \gamma(0)$$

$$\text{cov}(X_t, X_{t+1}) = \text{cov}(\epsilon_t + 0.7\epsilon_{t-1} - 0.2\epsilon_{t-2},\ \epsilon_{t+1} + 0.7\epsilon_t - 0.2\epsilon_{t-1})$$

$$= 0.7\,\text{Var}(\epsilon_t) - 0.14\,\text{Var}(\epsilon_{t-1})$$

$$= 0.7\sigma^2 - 0.14\sigma^2$$

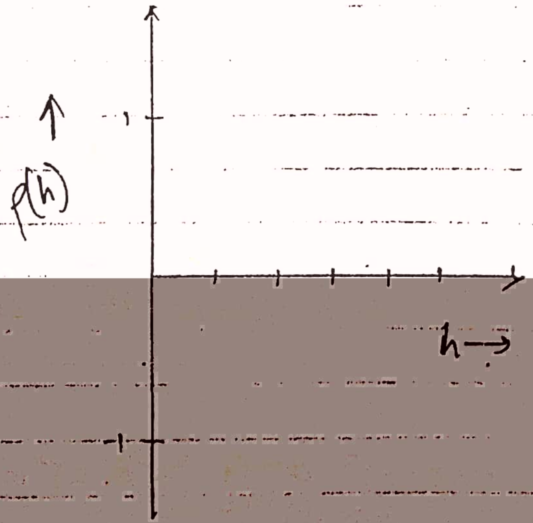
$$= (0.56)\sigma^2 = \gamma(1)$$

$$\text{cov}(X_t, X_{t+2}) = \text{cov}(\epsilon_t + 0.7\epsilon_{t-1} - 0.2\epsilon_{t-2},\ \epsilon_{t+2} + 0.7\epsilon_{t+1} - 0.2\epsilon_t)$$

$$= \cancel{-0.2\epsilon_t}$$

$$= -0.2\sigma^2 = \gamma(2)$$

$$\therefore \rho(h) = \begin{cases} 1 & ; h = 0 \\ \frac{0.56}{1.53} & ; h = 1 \\ \frac{-0.2}{1.53} & ; h = 2 \\ 0 & ; h \geq 3 \end{cases}$$



06.11.2017.

$\epsilon_t - 0.2\,\epsilon_{t-1})$

Simple Moving Average Process

Ans 17, Ans. 20. In moving average queue (MAQ) → MA(q) process $\{X_t\}$; $X_t = \beta_0 \epsilon_t + \beta_1 \epsilon_{t-1} + \cdots + \beta_q \epsilon_{t-q}$ with $\beta_0 = 1$. If the weight β_i's are equal and $\sum_{i=0}^{q} \beta_i = 1$ i.e. if $\beta_i = \frac{1}{q+1}$, $i = 0, 1, 2, \ldots, q$

$1 \mp 0.2\epsilon_t)$

then the process reduces to $X_t = \frac{1}{q+1}\left[\epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_{t-q}\right]$

it is known as simple moving average process of order q.

$$E(X_t) = 0$$

$$Var(X_t) = \frac{1}{(q+1)^2}\sum_{i=0}^{q}\sigma^2 = \frac{(q+1)\sigma^2}{(q+1)^2} = \frac{\sigma^2}{q+1}$$

$$Cov(X_t, X_{t+h}) = \sigma^2 \sum_{i=0}^{q-h} \beta_i \beta_{h+i} \; ; \; h \le q$$

$$= \sigma^2 \sum_{i=0}^{q-h} \frac{1}{(q+1)^2}$$

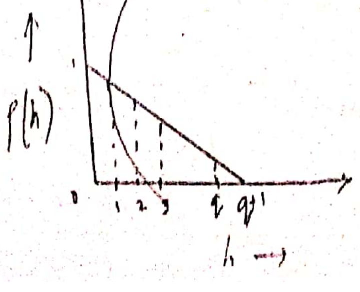
$$= \sigma^2 \frac{q-h+1}{(q+1)^2} = \gamma(h)$$

not a fn of 't'

$$\gamma(h) = \begin{cases} \frac{\sigma^2(q-h+1)}{(q+1)^2}, & h \le q \\ 0, & h > q \end{cases}$$

As $E(X_t) = 0$ and $\gamma(h)$ does not depend on time 't' hence the process is weakly stationary.

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 1, & h = 0 \\ \frac{q-h+1}{q+1}; & h \le q \\ 0; & h > q \end{cases}$$



13·11·2017

Autoregressive Process

Ans 2(i) A time series X_t is said to be an autoregressive process of order p ($AR(p)$) if it can be given as

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \epsilon_t$$

where ϵ_t is the white noise with

$$E(\epsilon_t) = 0$$

$V(\epsilon_t) = \sigma^2$ for all t where $\alpha_1, \alpha_2, \dots, \alpha_p$ are model parameters.

cases:

(i) $AR(1)$ = Markov Process

(ii) $AR(2)$ = Yule Process

AR(1) : Markov Process

Ans. 2.7: Ans. 2(ii): $X_t = \alpha X_{t-1} + \epsilon_t$; ϵ_t = white noise

$$E(\epsilon_t) = 0$$

$$V(\epsilon_t) = \sigma^2$$

$$|\alpha| < 1$$

$$X_t = \alpha X_{t-1} + \epsilon_t$$

$$X_t = \alpha(\alpha X_{t-2} + \epsilon_{t-1}) + \epsilon_t$$

$$= \alpha^2 X_{t-2} + \alpha \epsilon_{t-1} + \epsilon_t$$

$$= \epsilon_t + \alpha \epsilon_{t-1} + \alpha^2 (\alpha X_{t-3} + \epsilon_{t-2})$$

$$= \epsilon_t + \alpha \epsilon_{t-1} + \alpha^2 \epsilon_{t-2} + \alpha^3 X_{t-3}$$

$$\vdots$$

$$= \epsilon_t + \alpha \epsilon_{t-1} + \alpha^2 \epsilon_{t-2} + \alpha^3 \epsilon_{t-3} + \dots)$$

$= MA(\infty)$ Process

$-1/2$

$\rho(h)$ ↓

Scanned with CamScanner

$MA(q) = X_t = \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \quad \beta_q \epsilon_{t-q}$

comparing with MA(q) process we get,

$\beta_0 = 1$
$\beta_1 = \alpha$
$\beta_2 = \alpha^2$
$\beta_3 = \alpha^3$
$\vdots$

$\beta_i = \alpha^i \;;\; i = 0, 1, 2, \ldots, \infty \; [\beta_0 = \alpha^0 = 1]$

for $AR(1) = MA(\infty)$

$E(X_t) = 0$

$V(X_t) = \sigma^2 \sum_{i=0}^{\infty} \alpha^{2i} \quad [\text{as } \beta_i = \alpha^i]$

$= \sigma^2 \{1 + \alpha^2 + \alpha^4 + \ldots\}$

$= \sigma^2 \cdot \frac{1}{1-\alpha^2} = \frac{\sigma^2}{1-\alpha^2} \quad \because |\alpha| < 1$

$Cov(X_t, X_{t+h}) = \sigma^2 \sum_{i=0}^{\infty} \alpha^{i} \alpha^{h+i}$

$= \sigma^2 \alpha^h \sum_{i=0}^{\infty} \alpha^{2i}$

$= \sigma^2 \alpha^h (1 + \alpha^2 + \alpha^4 + \ldots + \alpha^\infty)$

$= \sigma^2 \alpha^h \frac{1}{1-\alpha^2}$

$= \frac{\alpha^h \sigma^2}{1-\alpha^2} = \gamma(h)$ does not depend on 't'

Now as $E(X_t) = 0$

$Cov(X_t, X_{t-h})$ does not depend on 't'

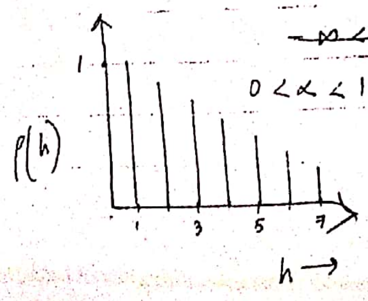
So $AR(1)$ is weakly stationary, provided $|\alpha| < 1$

Autocorrelation fn. $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 1 & ; h = 0 \\ \alpha^h & ; h > 0 \end{cases}$

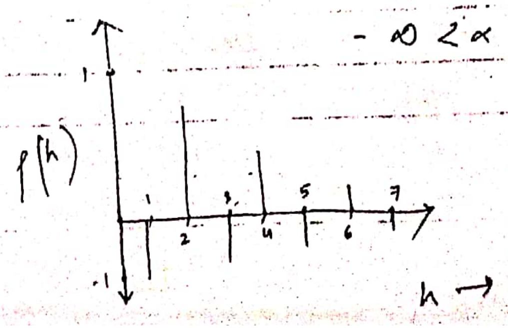
$\therefore \rho(h) = \alpha^h \;;\; h > 0$

we know $\rho(-h) = \alpha^h$;

$\therefore \rho(h) = \alpha^{|h|} \;;\; \forall h \;;\; \alpha < 1$



$-\infty < \alpha < 0$



AR(2) : Yule Process

Then auxillary equ. will be given by

$$r^2 - \alpha_1 r - \alpha_2 = 0$$

$$\Rightarrow r = \frac{\alpha_1 \pm \sqrt{\alpha_1^2 + 4\alpha_2}}{2}$$

 te

e

0

$= \sigma^2$

(case 1) Roots are real & unequal

i.e. $\alpha_1^2 + 4\alpha_2 > 0$

Let, ~~γ_1 & γ_2~~ be the 2 roots be r_1 & r_2

$\therefore$ General soln. ~~$\rho_k = A_1\gamma_1^k + A_2\gamma_2^k$~~ $\rho_k = A_1 r_1^k + A_2 r_2^k$

A_1 & A_2 are constants to be estimated.

Putting $k=0$

~~$\rho_0 = A_1\gamma_1^0 + A_2\gamma_2^0$~~ $\rho(0) = A_1 r_1^0 + A_2 r_2^0$

$= A_1 + A_2 = 1$ ——(i)

Putting $k=1$

~~$\rho_1 = A_1\gamma_1^1 + A_2\gamma_2^1$~~ ——(ii) $\rho(1) = A_1 r_1^1 + A_2 r_2^1$ —(ii)

$= A_1 r_1 + A_2 r_2$

$t\ X_{t-k}$

$(t-2)+\cdots$

Putting $k=1$ in Yule Walker equ, we get

$\rho_1 = \alpha_1\rho_0 + \alpha_2\rho_1$

$\Rightarrow \rho_1 = \alpha_1 + \alpha_2\rho_1$

$\Rightarrow \rho_1 = \frac{\alpha_1}{1-\alpha_2} = \frac{\gamma_1+\gamma_2}{1+\gamma_1\gamma_2} \quad \frac{r_1+r_2}{1+r_1r_2}$

$- X_{t-1})$

$X_{t-2})$

$-k)$

from equ. (ii)

$\rho_1 = A_1 r_1 + (1-A_2) r_2 = A_1 r_1 + (1-A_1) r_2$

$A_1 = \frac{\rho_1 - r_2}{r_1 - r_2} = \frac{r_1+r_2}{1+r_1r_2} - r_2 \Big/ r_1 - r_2$

$$= \frac{r_1(1-r_2^2)}{(1+r_1r_2)(r_1-r_2)}$$

$$A_2 = 1 - A_1 = \frac{-r_2(1-r_2^2)}{(r_1-r_2)(1+r_1r_2)}$$

in $\rho(k)$ of

$$\therefore \rho_k = \hat{A}_1 r_1^k + \hat{A}_2 r_2^k$$

on.

ρ_{k+1}

ρ_{k+2} & so on

$2r = \alpha_1$

18 · 11 · 2017

Case 2 Roots are real & equal

i.e. $\alpha_1^2 + 4\alpha_2 = 0$

Let the common root be 'r'

General solu: $\rho_k = (c_1 + c_2 k) r^k$

Putting $k = 0$

$\rho_0 = c_1$

$\Rightarrow 1 = c_1$

Putting $k = 1$

$\rho_1 = (c_1 + c_2) r$

$= (1 + c_2) r$

Putting $k = 1$ in Yule Walker equ,

$\rho_1 = \alpha_1 \rho_0 + \alpha_2 \rho_{-1} = \alpha_1 + \alpha_2 \rho_1$

$= \cancel{\rho_1 (1 - \alpha_2)}$

$\rho_1 = \dfrac{\alpha_1}{1 - \alpha_2}$ ✓

NO

$\therefore \dfrac{\alpha_1}{1 - \alpha_2} = (1 + c_2) r$

The characteristic equ. was $r^2 - \alpha_1 r - \alpha_2 = 0$

$\Rightarrow c_2 = \dfrac{\alpha_1}{r(1 - \alpha_2)} - 1$

$2r = \alpha_1$

$\Rightarrow c_2 = \dfrac{\alpha_1}{r(1 + r^2)} - 1 = \dfrac{2r}{r(1+r^2)} - 1$

$= \dfrac{1 - r^2}{1 + r^2}$

$\therefore \rho_k = \left(1 + \dfrac{1 - r^2}{1 + r^2} \cdot k\right) r^k$

Case 3. Roots are complex conjugate

i.e. $\alpha_1^2 + 4\alpha_2 < 0$

Complex roots r_1 & r_2

i.e. $r_1 = p(\cos\theta + i\sin\theta)$

$r_2 = p(\cos\theta - i\sin\theta)$

Scanned with CamScanner

$$\rho_k = A_1 r_1^{\,k} + A_2 r_2^{\,k}$$
$$= A_1 \{ p^k (\cos k\theta + i \sin k\theta) \} + A_2 \{ p^k (\cos k\theta - i \sin \theta k) \}$$

(Applying de-Moivre's Theorem.)

$$= p^k \{ \underbrace{(A_1 + A_2)}_{A^*} \cos k\theta + \underbrace{i(A_1 - A_2)}_{B^*} \sin k\theta \}$$

$$= p^k \{ A^* \cos k\theta + B^* \sin k\theta \}$$

we need to find A^* & B^*

Putting $k = 0$

$$\rho_0 = A^* = A_1 + A_2$$
$$A^* = 1$$

Putting $k = 1 \Rightarrow \rho_1 = p \{ A^* \cos\theta + B^* \sin\theta \}$ —— (i)

Again $k = -1 \Rightarrow \rho_{-1} = \frac{1}{p} \{ A^* \cos(-\theta) + B^* \sin(-\theta) \}$

$$= \frac{1}{p} \{ A^* \cos\theta - B^* \sin\theta \} \quad \text{—— (ii)}$$

Now $\rho_1 = p \{ A^* \cos\theta + B^* \sin\theta \} = p \{ \cos\theta + B^* \sin\theta \}$ & $\rho_{-1} = \frac{1}{p} \{ \cos\theta - B^* \sin\theta \}$

Putting $\rho_1 = \rho_{-1}$

eqn was
$K_2 = 0$
$= -K_2$
$= K_1$
$-1 = \frac{2}{1+\gamma^2} - 1$

$$p \{ \cancel{A^*} \cos\theta + B^* \sin\theta \} = \frac{1}{p} \{ \cancel{A^*} \cos\theta - B^* \sin\theta \}$$

$$p (\cos\theta + B^* \sin\theta) = \frac{1}{p} (\cos\theta - B^* \sin\theta)$$

$$p \cos\theta + p B^* \sin\theta = \frac{1}{p} \cos\theta - \frac{B^*}{p} \sin\theta$$

$$B^* \left(p \sin\theta + \frac{\sin\theta}{p} \right) = \frac{1}{p} \cos\theta - p\cos\theta$$

$$\Rightarrow B^* = \frac{\cos\theta \left(\frac{1}{p} - p \right)}{\sin\theta \left(1 + \frac{1}{p} \right)} = \cot\theta \left(\frac{1 - p^2}{1 + p^2} \right)$$

$$\therefore \rho_k = p^k \left\{ \cos k\theta + \frac{1 - p^2}{1 + p^2} \cot\theta \sin k\theta \right\}$$

Problems :-

1\. Consider the following AR(2) processes

(a) $X_t = \frac{1}{12} X_{t-1} + \frac{1}{12} X_{t-2} + \epsilon_t$

(b) $X_t = X_{t-1} - \frac{1}{2} X_{t-2} + \epsilon_t$

Show that autocorrelation fn. are resp. given as

(a) $\rho_k = \frac{45}{77} \cdot \frac{1}{3}^{|k|} + \frac{32}{77} \left(\frac{-1}{4}\right)^{|k|}$; $k = 0, \pm 1, \pm 2, \ldots$

(b) $\rho_k = \left(\frac{1}{\sqrt{2}}\right)^k \left\{ \cos k \frac{\pi}{4} + \frac{1}{3} \sin k \frac{\pi}{3} \right\}$

20.11.2017

Ans 6. Estimating the parameters of the autoregressive process

(i) AR(1)

$X_t = \alpha X_{t-1} + \epsilon_t$

The Yule Walker equation for the process is

$\rho_k = \alpha \, \rho(k-1)$

$= \alpha (\alpha \, \rho_{k-2})$

$= \alpha^2 \rho_{k-2}$

$= \alpha^2 (\alpha \, \rho_{k-3})$

$= \alpha^3 \rho_{k-3}$

$\vdots$

$\therefore \rho(k) = \alpha^k \rho_0 = \alpha^k$

$\rho_1 = \alpha$

$\therefore \hat{\alpha} = r_1$

The sample auto correlation of lag 1.

(ii) AR(2)

Ans. 24. ii. $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \epsilon_t$

The Yule Walker eqns. for the process is

$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2}$

If $k=1$ $\therefore \rho_1 = \alpha_1 \rho_0 + \alpha_2 \rho_{(-1)}$

$\Rightarrow \rho_1 = \alpha_1 + \alpha_2 \rho_1$ —— (1)

$\Rightarrow \alpha_1 = \rho_1 (1-\alpha_2)$

$\cancel{\rho_1 - \alpha_2 \rho_1 = \alpha_1}$

$\cancel{\rho_1 (1-\alpha_2) = \alpha_1}$

Putting $k=2$

$\rho_2 = \alpha_1 \rho_1 + \alpha_2 \rho_0$

$= \alpha_1 \rho_1 + \alpha_2$ —— (2)

$\rho_2 = \alpha_1(\alpha_1 + \alpha_2 \rho_1) + \alpha_2$

$= \alpha_1^2 + \alpha_1 \alpha_2 \rho_1 + \alpha_2$

$= \rho_1 (1-\alpha_2) \rho_1 + \alpha_2$

$\therefore \rho_1$ & ρ_2

$\rho_2 = \rho_1^2 (1-\alpha_2) + \alpha_2$

$\rho_2 = \rho_1^2 - \rho_1^2 \alpha_2 + \alpha_2$

$\rho_1^2 \alpha_2 - \alpha_2 = \rho_1^2 - \rho_2$

$\alpha_2 (\rho_1^2 - 1) = \rho_1^2 - \rho_2$

$\therefore \alpha_2 = \dfrac{\rho_1^2 - \rho_2}{\rho_1^2 - 1} = \dfrac{\rho_2 - \rho_1^2}{1-\rho_1^2}$

$\cancel{\left\{ \rho_1 \left(1 - \dfrac{\rho_1^2 - \rho_2}{\rho_1^2 - 1}\right) = \rho_1 \left(\dfrac{\rho_1^2 - 1 - \rho_1^2 + \rho_2}{\rho_1^2 - 1}\right) \right.}$

$\cancel{\alpha_1 = \dfrac{\rho_1 \rho_2 - \rho_1}{\rho_1^2 - 1} = \dfrac{\rho_1 - \rho_1 \rho_2}{1-\rho_1^2}}$

$$\therefore \hat{\alpha}_1 = \frac{r_1(1-r_2)}{1-r_1^2}$$

$$\hat{\alpha}_2 = \frac{r_2 - r_1^2}{1-r_1^2}$$

Stationary conditions for AR(2)

Ans. 21.(iii)

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \epsilon_t.$$

define an operator L as:

$$L X_t = X_{t-1}$$

$$\therefore X_t = \alpha_1 L X_t + \alpha_2 L^2 X_t + \epsilon_t$$

$$(1 - \alpha_1 L - \alpha_2 L^2) X_t = \epsilon_t$$

Let $(1-\alpha_1 L - \alpha_2 L^2) = (1-\lambda_1 L)(1-\lambda_2 L)$

$$\therefore X_t = (1-\alpha_1 L - \alpha_2 L^2)^{-1} \epsilon_t$$

$$= \{(1-\lambda_1 L)(1-\lambda_2 L)\}^{-1} \epsilon_t$$

$$\lambda_1 + \lambda_2 = \alpha_1$$

$$\lambda_1 \lambda_2 = -\alpha_2$$

$$\lambda_1, \lambda_2 = \frac{\alpha_1 \pm \sqrt{\alpha_1^2 + 4\alpha_2}}{2}$$

Now, $|\lambda_i| < 1 \quad \forall i = 1, 2$

$$\Rightarrow -1 < \frac{\alpha_1 \pm \sqrt{\alpha_1^2 + 4\alpha_2}}{2} < 1$$

$$\Rightarrow -2 - \alpha_1 < \pm\sqrt{\alpha_1^2 + 4\alpha_2} \leq 2 - \alpha_1$$

$$\therefore \sqrt{\alpha_1^2 + 4\alpha_2} \leq (2 - \alpha_1) \quad \text{— (i)}$$

$$\& \quad -\sqrt{\alpha_1^2 + 4\alpha_2} > (-2 - \alpha_1) \quad \text{— (ii)}$$

$$\Rightarrow \sqrt{\alpha_1^2 + 4\alpha_2} < (2 + \alpha_1) \quad \text{— (iii)}$$

Now (i) $\Rightarrow \alpha_1^2 + 4\alpha_2 \leq (2-\alpha_1)^2$

$$\alpha_1^2 + 4\alpha_2 \leq 4 - 4\alpha_1 + \alpha_1^2$$

$$\alpha_1 + \alpha_2 < 1 \quad \text{— (a) first condition}$$

$= 1/2$

$\rho(h)$

from (ii) we have

$$(\alpha_1^2 + 4\alpha_2) < 4 + \alpha_1^2 + 4\alpha_1$$

$$(\alpha_2 - \alpha_1) < 1 \quad \text{— (b)} \quad \text{2nd condition}$$

Generally, a time series has broadly two parts or components. One which is determined by a systematic part or deterministic, rule, is called systematic part & the other, which remains.

where ϵ_t is a white noise, i.e. the random component (or random disturbance) with

$$E(\epsilon_t) = 0$$
$$V(\epsilon_t) = \sigma^2 \quad \text{(unknown)}$$

$$\therefore \Delta^k X_t = \Delta^k \epsilon_t \quad (\because \underbrace{\Delta^k(a_0 + a_1 t + \dots + a_{k-1} t^{k-1})}_{=0})$$

$$= (E-1)^k \epsilon_t$$
$$= \left[E^k - \binom{k}{1} E^{k-1} + \binom{k}{2} E^{k-2} - \dots + (-1)^k\right]\epsilon_t$$

$$= \epsilon_{t+k} - \binom{k}{1}\epsilon_{t+k-1} + \binom{k}{2}\epsilon_{t+k-2} - \dots + (-1)^k \epsilon_t$$

$$\therefore E(\Delta^k X_t) = 0 \quad [\because E(\epsilon_t) = 0, \forall t]$$

$$\therefore V(\Delta^k X_t) = E(\Delta^k X_t)^2$$
$$= E\left(\sum_{i=0}^{k} (-1)^i \binom{k}{i} \epsilon_{t+k-i}\right)^2$$

$$= E\left[\sum_{i=0}^{k} \binom{k}{i}^2 \epsilon^2_{t+k-i} + \sum_{\substack{i=0 \\ i\neq j}}^{k}\sum_{j=0}^{k} (-1)^{i+j}\binom{k}{i}\binom{k}{j}\epsilon_{t+k-i}\,\epsilon_{t+k-j}\right]$$

$$= \cancel{\sum_{i=0}^{k} \binom{k}{i}^2 \sigma^2}$$

$$= \sigma^2 \sum_{i=0}^{k} \binom{k}{i}^2 + \sum_{\substack{i=0 \\ i\neq j}}^{k}\sum_{j=0}^{k}$$

part becomes oscilatory polynomial of degree fails

$$= \sigma^2 \sum_{i=0}^{k} \binom{k}{i}^2 \quad [\because \epsilon_t\text{'s are iid the significance } cov(\epsilon_t, \epsilon_{t'}) = 0 \ \forall t \neq t']$$

Now, $\binom{k}{0}^2 + \binom{k}{1}^2 + \dots + \binom{k}{k}^2 = \binom{2k}{k}$ (check)

$$\therefore V(\Delta^k X_t) = \mu_2'(\Delta^k X_t) = \sigma^2 \binom{2k}{k}$$

h)

$$\therefore \sigma^2 = \frac{\mu_2'(\Delta^k X_t)}{\binom{2k}{k}}$$

Now using MME, we can write

$$\hat{\sigma}^2 = \frac{m_2'(\Delta^k X_t)}{\binom{2k}{k}} \; ; \text{ provided } k \text{ is obtained}$$

where m_2' denotes 2nd order sample raw moment, i.e.

$$m_2'(\Delta^k X_t) = \frac{1}{N-k} \sum_{i=1}^{N-k} (\Delta^k X_{ti})^2$$

where N = no. of sample obs. on time series

By trial we can fix the choice of k in this

Forecasting :-

Ans. 10. (i) Forecasting the future values of an obs. time series is

1 an important problem in many cases / areas, including

$$\hat{x}(N,1) = \alpha x_N + \alpha(1-\alpha)x_{N-1} + \alpha(1-\alpha)^2 x_{N-2} + \cdots \longrightarrow (**)$$

infinite no. of past obs. but in practice infinite there will be a finite no.

Strictly speaking, equation (**) implies an infinite no. of past observations, but in practice there will only be a finite no. So equation (**) is rewritten in the form:

$$\hat{x}(N,1) = \alpha x_N + (1-\alpha)\{\alpha x_{N-1} + \alpha(1-\alpha)x_{N-2} + \cdots\}$$

$$= \alpha x_N + (1-\alpha)\hat{x}(N,1) \quad \text{———} \quad (***)$$

if $\hat{x}(1,1) = x_1$ ⇓ *** can be used recursively to compute forecast *** reduces amt. of arith. involved as it can be updated using only obs.

If we set $\hat{x}(1,1) = x_1$, then equ. (***) can be used recursively to compute forecasts. Equation (***) also reduces the amount of arithmetic involved since forecasts can easily be updated using only the latest obs. & the previous forecast. The procedure defined by equ (***) is called exponential smoothing.

The value of the smoothing constant α depends on the properties of the given time series. Value between 0.1 & 0.3 are commonly used & produ... us a forecast which depends on a large no. of

[redacted]

Repeat this procedure for other values of α between 0 & 1, say in steps of 0.1 & select the value which minimizes $\sum_{i=2}^{N} e_i^2$

The Holt-Winters Forecasting Procedure

$\hat{x}(t,k) = (L_t + k T_t) I_{t-12+k}$, for $k = 1(1)12$

In order to implement the method, the user must

(a) provide starting values for L_t, T_t & I_t at the beginning of the series

(b) estimate values for α, γ, δ by minimizing $\sum e_t^2$ over a suitable fitting period for which data are available

(c) decide whether or not, to normalize the seasonal indices at regular intervals.

Question & Answer:

Q. Show that 2 models A: $X_t = Z_t + \theta Z_{t-1}$ & B: $X_t = Z_t + \theta^{-1} Z_{t-1}$ have same ACF.

Sol

Model A:

$X_t = Z_t + \theta Z_{t-1}$

here $\beta_0 = 1$ & $\beta_1 = \theta$.

(i) $E(X_t) = 0$

(ii) $V(X_t) = \sigma^2 \sum_{i=0}^{1} \beta_i^2$

$= \sigma^2(\beta_0^2 + \beta_1^2)$

$= \sigma^2(1+\theta^2) = \gamma(0)$

(iii) $cov(X_t, X_{t+h})$

$= cov(X_t, X_{t+h})$

$= \sigma^2 \sum_{i=0}^{1-h} \beta_i \beta_{h+i} = \sigma^2[\beta_h + \beta_1\beta_{h+1}]$

$= \sigma^2 \beta_1 = \sigma^2 \theta = \gamma(1)$

(v) $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 1 & , h=0 \\ \frac{\theta}{1+\theta^2} & , h=1 \\ 0 & , h \geq 2 \end{cases}$

Model B:

$X_t = Z_t + \frac{1}{\theta} Z_{t-1}$

here $\beta_0 = 1$ & $\beta_1 = \theta^{-1}$

(i) $E(X_t) = 0$

(ii) $V(X_t) = \sigma^2 \sum_{i=0}^{1} \beta_i^2 = \sigma^2(1+\beta_1^2)$

$= \sigma^2(1+\frac{1}{\theta^2})$

$= \sigma^2\left(\frac{\theta^2+1}{\theta^2}\right)$

(iii) $Cov(X_t, X_{t+h})$

$= Cov(X_t, X_{t+h})$

$= \sigma^2 \sum_{i=0}^{1-h} \beta_i \beta_{h+i}$

$= \sigma^2 \beta_1 = \frac{\sigma^2}{\theta} = \gamma(1)$

(iv) $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 1 & , h=0 \\ \frac{\theta}{1+\theta^2} & , h=1 \\ 0 & , h \geq 2 \end{cases}$

$\therefore$ Both the models have same ACF (proved).

Q. Kolkata has shown no appreciable change in the total annual rainfall over the years. Discuss how you will find the seasonal variation in rainfall in Kolkata given last five years monthly data.

Since the given data do not contain trend or cyclical movement to any appreciable ~~extend~~ extent, we use the method of averages of adjusted data.

Here we eliminate the irregular variations only by averaging the monthly values over different time points.

Let, y_{ij} = obs. of the jth month of ith year

$$\bar{y}_j = \frac{\sum_i y_{ij}}{\text{No. of years}} = \frac{\sum_i y_{ij}}{5} \quad ; \; j = 1, 2, 3, \ldots, 12$$

This $\bar{y}_j$ measures the seasonal component for the jth month, $j = 1(1)12$. To express $\bar{y}_j$ as seasonal indices, they are shown as %s of the grand mean $\bar{Y}$.

[redacted]

Given, $x_t = (-1)^t e_t$

Now, $E(x_t) = (-1)^t E(e_t)$
$= (-1)^t . 0$
$= 0$

$Var(x_t) = [(-1)^t]^2 \, var(e_t)$
$= Var(e_t)$
$= \sigma^2$

$Cov(x_t, x_{t+h}) = 0 \; , \forall t$ [$\because$ e_t are mutually uncorrelated]

Hence, mean & variance of the time series are constant values.

$\therefore$ The time series is stationary.

Q. AR v/s MA

AR	MA
(i) A time series X_t is said to be an ~~autoregg~~ autoregressive process of order p i.e. AR(p) if it can be given as $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} + \epsilon_t$ where ϵ_t is white noise with $E(\epsilon_t) = 0$, $Var(\epsilon_t) = \sigma^2 \; \forall t$ where $\alpha_1, \alpha_2, \cdots, \alpha_p$ are model parameters	(i) In MA(q) process $\{x_t\} \ni x_t = \beta_0 \epsilon_t + \beta_1 \epsilon_{t-1} + \cdots + \beta_q \epsilon_{t-q}$ with $\beta_0 = 1$. If the weight β_i's are equal $\frac{1}{q} \sum_{i=0}^{q} \beta_i = 1$ i.e. if $\beta_i = \frac{1}{q+1}, i = 0(1)q$ then the process reduces to $x_t = \frac{1}{q+1}[\epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_{t-q}]$ is known as simple moving average process of order q. $E(x_t) = 0$, $Var(x_t) = \frac{1}{q+1} \sum_{i=0}^{q} \sigma^2 = \frac{(q+1)\sigma^2}{(q+1)^2} = \frac{\sigma^2}{q+1}$
(ii) Depends on errors (residuals) of the previous forecasts we made to make current forecasts.	(ii) Depends on the lagged values of the data we are modeling to make forecast

Q. Define Moving Average process of order 2. Find its autocorrelation function.

MA(2) process is given by,

$X_t = \beta_0 \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2}$ where ϵ_t is a purely random process with

$E(\epsilon_t)=0$ & $Var(\epsilon_t) = \sigma^2$

& $\beta_0 = 1$.

(i) $E(X_t)=0$

(ii) $Var(X_t) = \sigma^2 \sum_{i=0}^{2} \beta_i^2 = \sigma^2(\beta_0^2+\beta_1^2+\beta_2^2)$

$= \sigma^2(1+\beta_1^2+\beta_2^2) = \gamma(0)$

(iii) $Cov(X_t, X_{t+1}) = Cov(\epsilon_t, \epsilon_{t+1})$ $= \sigma^2 \sum_{i=0}^{q-h} \beta_i \beta_{h+i}$ $= \sigma^2(\beta_0\beta_1 + \beta_1\beta_2)$ $= \gamma(1)$

$\cancel{= \beta_1\sigma^2 + \beta_2\sigma^2}$

$\cancel{= \sigma^2(\beta_1+\beta_2) = \gamma(2)}$

$cov(X_t, X_{t+2}) = cov(\epsilon_t, \epsilon_{t+2})$

$= \sigma^2 \sum_{i=0}^{q-h} \beta_i \beta_{h+i} = \cancel{\sigma^2(\beta_0\beta_1+\beta_1^2+\beta_2^2)}$

$= \sigma^2 \beta_2 \beta_0 = \gamma(2)$

(iv) $\rho(h) = \begin{cases} 1 & , h=0 \\ \dfrac{\beta_1+\beta_1\beta_2}{1+\beta_1^2+\beta_2^2} & , h=1 \\ \dfrac{\beta_2}{1+\beta_1^2+\beta_2^2} & , h=2 \\ 0 & , h \geq 3 \end{cases}$

Q. Consider an AR(2) process given by $X_t = X_{t-1} - \frac{1}{2} X_{t-2} + e_t$
Find its autocorrelation fn. & draw correlogram:

AR(2) process is given by:

$$X_t = X_{t-1} - \frac{1}{2} X_{t-2} + e_t$$

~~the auxilliary equs. will be given by~~
$\cancel{r^2 - \alpha_1 r - \alpha_2 = 0}$

Premultiplying by X_{t-k} we get

$\Rightarrow X_{t-k} \cdot X_t = E(X_{t-k} \cdot X_{t-1}) - \frac{1}{2} E(X_{t-k} \cdot X_{t-2}) + \cdots + E(\epsilon_t \cdot X_{t-k})$

$\Rightarrow \{E(X_{t-k} \cdot X_t) - E(X_{t-k}) E(X_t)\} = \{E(X_{t-k} \cdot X_{t-1}) - E(X_{t-k}) E(X_{t-1})\}$
$- \frac{1}{2}\{E(X_{t-k} X_{t-2}) - E(X_{t-k}) E(X_{t-2})\} + E(\epsilon_t) E(X_{t-k})$

as $E(X_t) = 0 \quad \forall t$ & ϵ_t is white noise

$\Rightarrow \gamma(k) = \gamma(k-1) - \frac{1}{2} \gamma(k-2)$

$\cancel{\Rightarrow \gamma(k) = \gamma(k-1) - \frac{1}{2} \gamma(k-2)}$

dividing by $V(X_t)$
$\Rightarrow \gamma(0)$ we get

$$\rho(k) = \rho(k-1) - \frac{1}{2} \rho(k-2) \quad \text{—— } (*)$$

This is a homogeneous diff. eqn. in $\rho(k)$ of order 2

Eqn. is known as Yule Walker Eqn.

Now define an operator E as $E(\rho_k) = \rho_{k+1}$
$E^2(\rho_k) = \rho_{k+2}$ & so on

$\therefore$ (*) becomes

$$E^2(\rho_{k-2}) - E(\rho_{k-2}) + \frac{1}{2}(\rho_{k-2}) = 0$$

$$\left(E^2 - E + \frac{1}{2}\right)(\rho_{k-2}) = 0$$

The auxilliary eqn. will be given by

$$r^2 - r + \frac{1}{2} = 0$$

$$r = \frac{+1 \pm \sqrt{1-2}}{2}$$

$\because 1-2 < 0$ the root are complex conjugate complex roots = r_1 & r_2

let, $r_1 = p(\cos\theta + i\sin\theta)$

$r_2 = p(\cos\theta - i\sin\theta)$

Ans. 28. $X_t = X_{t-1} - 0.5X_t + e_t$

Here $\alpha_1 = 1$ & $\alpha_2 = 0.5$

The auxiliary eqn will be given by $r^2 - \alpha_1 r - \alpha_2 = 0$

$r^2 - r + 0.5 = 0$

Now, here $\alpha_1^2 + 4\alpha_2 = 1 + 4(-0.5) = 1 - 2 = -1 < 0$

$\therefore$ The roots are complex conjugate

Let, the complex roots be r_1 & r_2

Let, $r_1 = p(\cos\theta + i\sin\theta)$

$r_2 = p(\cos\theta - i\sin\theta)$

Now, $r^2 - r + 0.5 = 0$

$$\therefore r = \frac{1 \pm \sqrt{1 - 4(0.5)}}{2} = \frac{1 \pm \sqrt{-2}}{2} = \frac{1 \pm \sqrt{2}i}{2}$$

i.e. $r_1 = \frac{1+\sqrt{2}i}{2}$ & $r_2 = \frac{1-\sqrt{2}i}{2}$

Now, $r_1 + r_2 = \frac{1+\sqrt{2}i + 1 - \sqrt{2}i}{2} = 1$

$\therefore 1 = 2p\cos\theta$ —(i)

& $r_1 - r_2 = \frac{\cancel{1} + \sqrt{2}i \cancel{-1} + \sqrt{2}i}{2} = \frac{2i}{2} = i$

$\therefore i = 2pi\sin\theta$

$1 = 2p\sin\theta$ —(ii)

$\therefore \frac{2p\sin\theta}{2p\cos\theta} = 1 \Rightarrow \tan\theta = 1 \Rightarrow \theta = \tan^{-1}(1) = \pi/4$

Hence, $1 = 2p\sin(\pi/4) \Rightarrow \frac{1}{2} = p\sin(\pi/4) \Rightarrow \frac{1}{2} = p\left(\frac{1}{\sqrt{2}}\right)$

$\Rightarrow p = \frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$.

Now, auto correlation fn. is given by

$$\rho_k = p^k \left\{ \cos k\theta + \left(\frac{1-p^2}{1+p^2}\right) \cot\theta \sin k\theta \right\}$$

$$= \left(\frac{1}{\sqrt{2}}\right)^k \left\{ \cos k\frac{\pi}{4} + \frac{1}{3} \cot\theta \sin k\theta \right\}$$

$$= \frac{1}{\sqrt{2}}^k \left\{ \cos \frac{k\pi}{4} + \frac{1}{3} \cot\frac{\pi}{4} \sin \frac{k\pi}{4} \right\}$$

$$= \left(\frac{1}{\sqrt{2}}\right)^k \left\{ \cos \frac{k\pi}{4} + \frac{1}{3} \sin \frac{k\pi}{4} \right\}$$

$$\Rightarrow \gamma(k) = \frac{1}{12}\gamma(k-1) + \frac{1}{12}\gamma(k-2)$$

dividing by $V(X_t)$

$$\Rightarrow \rho(k) = \frac{1}{12}\rho(k-1) + \frac{1}{12}\rho(k-2) \quad \text{---} \ (*)$$

This is a homogeneous diff. eqn. of order 2 in $\rho(K)$
Eqn. is known as Yule Walker Eqn.
We define operator E as $E(\rho_k) = \rho(k+1)$
$E^2(\rho_k) = \rho(k+2)$ & so on.

(*) becomes

$$E^2[\rho(k-2)] - \frac{1}{12}E[\rho(k-2)] - \frac{1}{12}[\rho(k-2)] = 0$$

The auxilliary eqns. are

$$r^2 - \frac{r}{12} - \frac{1}{12} = 0.$$

$$r = \frac{\frac{1}{12} \pm \sqrt{\frac{1}{144} + \frac{1}{3}}}{2}$$

$\because \frac{1}{144} + \frac{1}{3} > 0$

The roots are real & unequal.

$$r = \frac{\frac{1}{12} \pm \sqrt{\frac{49}{144}}}{2}$$

$$= \frac{\frac{1}{12} \pm \frac{7}{12}}{2}$$

$$r_1 = \frac{\frac{\cancel{8}\,4}{12}}{\cancel{2}\,1} = \frac{1}{3} \qquad r_2 = \frac{-\frac{1}{2}}{2} = -\frac{1}{4}.$$

Now, $\hat{A}_1 = \dfrac{r_1(1-r_2^2)}{(1+r_1r_2)(r_1-r_2)} = \dfrac{\frac{1}{3}\left(1-\frac{1}{16}\right)}{\left(1-\frac{1}{12}\right)\left(\frac{1}{3}+\frac{1}{4}\right)} = \dfrac{\frac{15}{3\times16}}{\frac{11}{12}\times\frac{7}{12}} = \dfrac{32}{77}$

$\ne \cancel{\frac{17\times 3}{11\times 7}} \cancel{\frac{51}{77}}$

$\hat{A}_2 = 1 - A_1 = 1 - \frac{32}{77} = \frac{77-32}{77} = \frac{45}{77}.$

$\therefore \rho(k) = \cancel{45}\ \hat{A}_1 r_1^{|k|} + \hat{A}_2 r_2^{|k|} \ ; k = 0, 1, 2, \ldots$

$= \frac{45}{77}\left(\frac{1}{3}\right)^{|k|} + \frac{32}{77}\left(-\frac{1}{4}\right)^{|k|}$

Hence proved.

8. Let $X_t = Y\cos\theta t + Z\sin\theta t$, where Y & Z are two uncorrelated random variables each with mean 0 & variance unity & $\theta \in [-\pi, \pi]$. Is the time weakly stationary? Give reasons for your answer.

$X_t = Y\cos\theta t + Z\sin\theta t \ ; \ E(Y) = 0 \ ; \ Var(Y) = 1$

$E(Z) = 0 \ ; \ Var(Z) = 1$

$\rho_{Y,Z} = 0 \Rightarrow Cov(Y, Z) = 0.$

Now, $E(X_t) = \cos\theta t E(Y) + \sin\theta t E(Z) = 0$

$V(X_t) = \cos^2\theta t\, Var(Y) + \sin^2\theta t\, Var(Z)$

$= \cos^2\theta t + \sin^2\theta t = 1.$

$Cov(X_t; X_{t+h}) = Cov\left(Y\cos\theta t + Z\sin\theta t,\ Y\cos\theta(t+h) + Z\sin\theta(t+h)\right)$

$= Cov\left(Y\cos\theta t + Z\sin\theta t,\ Y\cos\theta t\cos\theta h - Y\sin\theta t\sin\theta h + Z\sin\theta t\cos\theta h + Z\cos\theta t\sin\theta h\right)$

$= \cos^2\theta t\cos\theta h\, Var(Y) - \sin\theta t\cancel{\cos\theta t}\sin\theta h\, Var(Y)$
$+ \sin^2\theta t\cos\theta h\, Var(Z) + \sin\theta t\cos\cancel{\theta t}\sin\theta h\, Var(Z)$

$= 1.\cos\theta h$ = ind. of t

hence weakly stationary

Stationary Data — a time series variable exhibiting no significant upward or downward trend over time.

Non Stationary Data — — — — exhibiting a significant

Seasonal Data — — exhibiting a repeating patterns at regular intervals over time.

Aims

- to describe & summarise time series data
- fit models & make forecasts.

Regression: y on x → linear
time Data are not independent.
Series often change with time, so bigger isn't always better.

$$E\left[\frac{1}{n-1}\sum(x_i-\bar{x})^2\right] = \sigma^2$$

$$= \lambda$$

$$E \quad V(\bar{x})$$

$$\frac{1}{n-1}\sum(x_i-\bar{x})^2$$

— 1/2