[Academic Script]

Time Series Models

Subject:	Business Economics
Course:	B. A. (Hons.), 5th Semester, Undergraduate
Paper No. & Title:	Paper – 531 Elective Paper Q1 – Advanced Econometrics
Unit No. & Title:	Unit – 3 Time Series Models
Lecture No. & Title:	Lecture – 2 Time Series Models

Academic Script

1. What is Time Series?

Hello friends nice to meet you. Today we are discussing about certain models related with Time series. As you know time series has several components like trend, Seasonal, Cyclic and random effects. First of all we will study how to separate this components. Then we will talk about additive as well as multiplicative model in detail. Similarly we shall also talk about the stationary time series and then we want to know how we can detect stationarity in a given time series. There are certain test which we shall study here in detail. So we can start our lecture.

It is a series which relates the given observations pertaining to certain phenomena expressing the variation according to time.

e.g. Series of observations for production, sales, share prices, atmospheric data, rainfall, consumption, Income, Savings etc. during successive years (or months or weeks or quarters or days).

Study of such series is important in the sense that if we can examine the pattern of variation or some systematic modelling approach, then it can be made useful for further planning exercise.

Hence such a study has its own importance for government agencies as well as corporate houses, businessmen etc.

2. What are the components of a time series?

If we closely examine any given time series then we can find the existence of the following components of time series data.

(1) Trend Component

The variations or changes that occur during some longer span period is called trend component of time series. These are the

fluctuations which exhibit not in shorter period but in some longer period and they reasonably affect the given time series.
e.g. Price of commodity during 3 to 5 years period, Demand for two wheelers in some years, sales and purchases of electronic items during successive years etc. Here we find that such a variation is found not in shorter period, but for a longer period.
These changes are attributed due to apptitude, fashion, habits, living standard, financial status of people etc. Trend Component exhibits long term variation in the given time series. It is also the most prominent part in the given time series.
(2) Seasonal Component
Demand for raincoats in monsoon, demand for woolen clothes in winter, prices and demand for medicines during seasons which cause infection, fever and other causalities, number of passengers during festival days, etc. These are some examples how the variations are observed during seasons (or quarters or weeks). These are short term fluctuations in the given series.
(3) Cyclic Component
There are business cycles which exhibit changes in the observations over a certain period. These are regular changes but they are not short time changes. If we examine the series over a longer span period, cyclic fluctuations can be seen to have short time variations.
(4) Irregular or Random Component
These changes occur due to accidental circumstances. e.g. flood, famine, frost, earth quake, fire, accidents etc. Which are random events, and they can cause variation in the given time series. Generally these variations are not in our control as they occur at any time and can affect the data reasonably.

If we summarise all the above factors then given time series (for its additive model) can be expressed as composed of

(Time series) = (Trend) + (Seasonal) + (Cyclic) + (Random)

Thus (T.S.) = T+S+C+R.

3. How to separate the component of time series?

3.1 Methods to separate trend component

(1) Graphical Method

If we plot the given time series observations against time, we can examine the variational pattern of the series from the graph.

e.g. It can be linear or quadratic or curvilinear etc. This is an approximate method which can give us first hand information about general pattern of the series.

(2) Moving Average Method

This is a very simple and useful method to obtain trend component of the time series. It is based upon taking successive average obtained from the observations according to periodicity of the series. If we plot the observations against time then the average of the periods for highest (or lowest) fluctuations can determine the periodicity which can be 2 or 3 or 4 or 5 periods etc.

If the periodicity is 3 years (odd years) then we first find sum $(y_1 + y_2 + y_3)$ and divide it by 3, place it against second observation. Then take next $(y_2 + y_3 + y_4)$ and divide it by 3, and place it against third observation and so on.

If the periodicity is of 4 years (even years) we first find sum $(y_1 + y_2 + y_3 + y_4)$ place it between second and third observation, then sum two consecutive values and divide the sum by 8 and place it next and so on. We loose some observations at the top and in the bottom while doing so. We illustrate this by two examples as shown below

<u>Example 1</u> <u>odd period</u> (<u>Periodicity 3</u>)

Time	observation		<u>Sum of 3 obs.</u>	<u>Trend</u>
1	10		–	–
2	12	⟶	10+12+16=38	38/3 = 12.67
3	16	⟶	12+16+18=46	46/3 = 15.33
4	18	⟶	16+18+21=55	55/3 = 18.33
5	21	⟶	18+21+17=56	56/3 = 18.67
6	17	⟶	21+17+19=57	57/3 = 19.00
7	19	⟶	17+19+23=59	59/3 = 19.67
8	23		–	–

Example 2 Even period

(Periodicity 4)

		Sum of 4 obs.	(Sum of 2 obs.)/8	Trend
1	12		–	–
2	17		–	–
		73		
3	21		153/8	19.1
		80		
4	23		161\8	20.1
		81		
5	19		164\8	20.5
		83		
6	18		166\8	20.7
		83		
7	23		168\8	21.0
		85		
8	25		–	–

(It may be noted that moving average obtained to find trend is in a way weighted average of the successive observations.)

(3) Iterated Moving Average Method

This refers to the above note that moving average is in a way a weighted average. If we use moving average method for more than one time then it is an Iterated moving average. Here there are different methods given as per the period for weighted average. Some are known as Spencer's 15 point formula, Spencer's 21 point formula etc.

(3) Least Squares Method

Suppose that given time series has its trend given by

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \cdots + \beta_p t^p + U_t \qquad (t = 1,2,..n)$$

We can obtain β coefficients by minimizing the residual sum of squares. These are OLSE of Betas. On substitution of β coefficients, we can estimate polynomial trend given by

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2 + \cdots + \hat{\beta}_p t^p \qquad (t = 1,2,..n)$$

Example How to fit polynomial trend

$$Y_t = a + bt + ct^2 + U_t$$

We write normal equations by minimizing error sum of squares

$$\sum Y_t = na + b\sum t + c\sum t^2$$

$$\sum t\, Y_t = a\sum t + b\sum t^2 + c\sum t^3$$

$$\sum t^2\, Y_t = a\sum t^2 + b\sum t^3 + c\sum t^4$$

Hence $\hat{Y}_t = \hat{a} + \hat{b}t + \hat{c}t^2$ which is estimated trend

(4) Variate Difference Method

When we assume a polynomial equation in time of degree p, we do not know what will be the value of p. This is accomplished by Variate difference method by means of a test statistic. Thus whether the polynomial equation in t is of second or third or higher degree, this method determines it and then trend can be estimated.

(5) By fitting growth curve models

Some growth models like Gompertz, modified Gompertz, Logistic Curve etc can be fitted for trend component assuming functional relationships for trend equation.

Gompertz curve: $\log y_t = a_0 + b_0 r^t$

Logistic curve : $y_t = \frac{k}{1+b\exp(-at)}$

Here different methods are used for fitting the above curves to the data.

2. Methods of determining Seasonal Component

3.2 Methods of determining Seasonal Component

(1) Moving Average Method

Given time series can be expressed in its additive form by $y_t = T + S + R$

When we apply M.A. method to determine trend, then subtract estimated trend $\hat{T}$ from the series, so that we get $y_t' = S + R$ which consist of seasonal and random components.

Now again apply M.A. method for each year (quarter wise or season wise e.g. for 4 quarters, it is M.A. with periodicity 4).This will eliminate random component and we get seasonal component S, we can find quarter wise (or season wise) average to express respective seasonal component.

(2) Method of Seasonal Index

Suppose that given time series has observations for 5 years and for each of 12 months in an year. We express these observations by y_{ij} , $i = 1,2, \dots 12. j = 1,2,3,4,5.$

(1) First we compute yearwise monthly average $A_i = \sum_{j=1}^{5} {y_{ij}}/{5}$

(2) Next find over all mean $\bar{A} = \sum_{j=1}^{12} {A_i}/{12}$

(3) Now compute Seasonal Index $I_i = \left(\frac{A_i}{\bar{A}}\right) \times 100$

(4) We set the series which is separated from seasonal index by $x_{ij} = \left(\frac{y_{ij}}{I_i}\right) \times 100$

This contains (T+R) components.

(5) Then we can obtain Seasonal Index from $S_{ij} = y_{ij} - x_{ij}$

[that is $S = (T + S + R) - (T + R)$].

(Note: There are also other methods

Like

(a) method of link relatives,

(b) Ratio to trend method

(c) Ratio to M.A. method etc.)

3.3 How to separate Random Component?

When other components are estimated the remaining part is the random component.

Thus given time series $= y_t = T + S + R$

$$y_t' = y_t - T = S + R$$

$$y_t'' = y_t' - S = R = \text{Random component}$$

(Note: In the presence of Cyclic component also, we can separate cyclic component in a way similar to trend component in general and then step by step we can obtain random component by elimination procedure)

4. Multiplicative model for time Series

Many times we can observe that not additive but multiplicative model for time series is useful.

We expressed additive model by $y_t = T + S + C + R$

If we include all the 4 above components then multiplicative model can be written by

$T.S. = y_t = T.C.S.R$

A speciality of this model is that if a component is not present, we take it as 1.

(e.g. If S is absent, put $S = 1$, then $y_t = T.C.R$.)

Similarly if C is absent put $C = 1$, then $y_t = T.S.R$.) Such presentation of multiplicative form is useful for very long time series. Generally when the variational pattern in the series of exponential type, multiplicative model may be useful. The respective component can be obtained by division rather than subtraction in the case of additive model.

(e.g. $\frac{y_t}{S} = T \times R$, $\frac{y_t}{T} = S \times R$ etc.)

(It may be noted that our discussion up til now refers to non stationary time series models.

Now we shall discuss stationary time series and relevant results step by step in brief)

5 Stationary Time Series

In a stationary time series there are no components like trend, cyclic or seasonal factors. Such time series is in relation to stochastic processes and hence they have very wide applications. Certain statistical events which are connected with time and follow laws of probability exhibit stochastic processes.

e.g. Number of passengers travelling by air day to day, temperature recorded at some place for different days, number of absent employees in a production unit at different days etc. (These events are linked with certain probability laws).

Definition of Stationary, Stochastic Process

A stochastic process is said to be stationary if its mean and variance are constant over time and the value of the co-variance between the two time periods depends only on the distance or gap or lag between the two time periods and not the actual time at which the covariance is computed.

Let y_t be a stochastic time series with the properties

Mean $= E(Y_t) = \mu$

Variance $= E\,(Y_t - \mu)^2 = \sigma^2$

Covariance = $\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$

Where γ_k the covariance at lag k is the covariance between the values Y_t and Y_{t+k} that is between two values of Y at k periods apart.

If $k = 0\,, \gamma_k = \gamma_0 = \sigma^2$, If $k = 1, \gamma_1$ is the covariance between two adjacent values of Y

From this we understand that if a time series is stationary, its mean, variance and auto covariance (at different lags) remain the same no matter at what point we measure them. Thus a stationary time series is time invariant.

(From this we may also understand that a time series is non stationary if it is not time invariant in the sense that it will have a time varying mean or a time varying variance or both).

Purely Random or White noise process

We call a stochastic process purely random if it has zero mean, constant variance and it is serially uncorrelated. In classical linear regression model we assumed that $U_t \sim IIDN(0, \sigma^2)$, that is U_t is independently and identically distributed as a normal distribution with zero mean and constant variance, Such a process is also known as Gaussian White Noise Process.

3. Random Walk Model

It is often said that asset prices such as stock prices or exchange rates follow a random walk that is they are non stationary. (e.g. Today's stock price is equal to yesterday's stock price plus a random stock).

We have two types of random walks

(1) Random Walk without drift (that is no constant or intercept term)

(2) Random Walk with drift (that is when the constant term is present).

Random Walk Without Drift

Let U_t be a white noise error term with mean zero and variance σ^2 .Then the series y_t is said to be a random walk if

$y_t = y_{t-1} + U_t$ .... (1)

This is also called AR (1) Model.

From equation (1), $Y_1 = Y_0 + U_1$

$$Y_2 = Y_1 + U_2 = Y_0 + U_1 + U_2 \text{ \& soon}$$

Hence we get $Y_t = Y_0 + \sum U_t$ ………………….. (2)

$E(Y_t) = Y_0$ and $V(Y_t) = t\sigma^2$

Here as t increases, variance increases indefinitely. Thus RWM without drift is a non stationary stochastic process.

If we write equation (1) as

$Y_t - Y_{t-1} = U_t \Rightarrow \Delta U_t = u_t$

It can be verified easily that while Y_t is non stationary, ΔY_t is stationary.

<u>Random Walk with Drift</u>

We write $Y_t = \delta + Y_{t-1} + U_t$ …………(1)

Where δ is called <u>Drift parameter.</u>

$$\left.\begin{aligned} We\ have\ E(Y_t) &= Y_0 + t.\delta \\ \text{and}\ \ V(Y_t) &= t\sigma^2 \end{aligned}\right\} \text{………….. (2)}$$

Here , $\Delta Y_t = Y_t - Y_{t-1} = \delta + U_t$ ……………...(3)

This shows that Y_t drifts upwards or downwards depending upon δ being positive or negative. This model is also <u>AR (1) model</u>.

<u>Autocorrelation Function and Correlogram</u>

It is important to know whether the given time series is stationary or not. A simple test for stationary is based upon <u>Auto Correlation Function</u> (ACF).

ACF at lag k is defined as

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Where $\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$

$=$ Covariance at lag k and $\gamma_0 = \sigma^2$

The Autocorrelation Function defined above has the following properties

(1) $\rho_k = \rho_{-k}$

(2) $|\rho_k| \leq 1$

(3) $\rho_0 = 1$

In order that we can decide about an appropriate model for given time series, ACF will be useful. Generally we do not have value of ρ_k based upon population, hence we estimated ρ_k from the sample. If values of ρ_k are plotted against k, it is called population correlogram.

We can obtain estimates of γ_k and γ_0 based upon samples which are $\hat{\gamma}_k = \frac{\sum(y_t-\bar{y})(y_{t+k}-\bar{y})}{n}$

and $\hat{\gamma}_0 = \sum(y_t - \bar{y})^2/_n$

This gives Sample Autocorrelation Function given by

$\hat{\rho}_k = r_k = \frac{\sum(y_t-\bar{y})(y_{t+k}-\bar{y})}{\sum(y_t-\bar{y})^2}$

When we plot the values of r_k against k, we get Sample Correlogram.

Usefulness of Sample Correlogram

(1) If all the plotted values of r_k against k are very near to zero, we may conclude that the given time series is an Independent series.

(2) If value of r_1 is large and other values $r_2, r_3, \ldots$ are decreasing respectively then we may consider AR model as proper for the time series.

(3)If the plotted values of r_k do not appear to be near zero then the given time series is not stationary, but it may be of non-stationary type.

Tests of Stationarity

Now we understand that it is very important to know whether the given time series is stationary or not. There are many approaches that can be used for this purpose. We shall study some of them in brief.

(1) Q Statistic

The statistical significance of any estimated ρ_k (i.e. $\hat{\rho}_k$) can be judged by its standard error.

Bartlett has shown that if a time series is purely random (i.e. it exhibits white noise) the sample autocorrelation coefficients $\hat{\rho}_k \sim N(0,\frac{1}{n})$

From this, Q Statistic is defined as under

$Q = n\sum_{k=1}^{m} \hat{\rho}_k^2$ where $n =$ Sample size and $m =$ length of lag.

Q has χ^2 distribution with m degrees of freedom

This was given by <u>Box and Pierce.</u>

Here we can test the joint hypothesis that all ρ_k upto certain Lags are zero.

If χ^2 calculated is greater than χ^2 tabulated at chosen significance level, we can reject null hypothesis. (that means at least some of them may be non-zero.).

(2) LB Statistic

Ljung – Box (LB) statistic is a modification for Q Statistic.

$$LB = n(n+2)\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{(n-k)}$$

has χ_m^2 distribution and hence apply χ^2 test as usual. This test is more powerful than the above Q test.

(3) Unit Root Test

Unit Root Test (URT) is widely popular for its applications.

Let us write $Y_t = \rho\, Y_{t-1} + U_t$ $(-1 \leq \rho \leq 1)$ ……… (1)

Where U_t is a white noise error term?

If $\rho = 1$, equation (1) is a random walk model without drift and it is a non stationary stochastic process.

From equation (1) above, we write

$$Y_t - Y_{t-1} = (\rho - 1)Y_{t-1} + U_t$$

thus $\Delta Y_t = \delta Y_{t-1} + U_t$ ……………(2)

If we test the hypothesis $\delta = 0$ (i.e. $\rho = 1$) we have a unit root and hence given time series is non stationary.

4. Dickey Fuller Test (DF test)

Dickey and fuller have shown that under the null hypothesis $H_0\!:\delta = 0$, the estimated coefficienct of Y_{t-1} divided by its standard error is a Γ statistic. They have computed tables for DF statistic for different sample sizes (Tables are extended by Mackinnon) giving critical values at chosen significant level.

Test is carried out as under

(1)Regress equation (2) above for Y_t against Y_{t-1} and estimate coefficient of Y_{t-1} (i.e.$\hat{\delta}$)

(2) Obtain DF statistic Γ by computing $\Gamma = \hat{\delta}/(S.E\ of\, \hat{\delta})$

Generally this may be negative =--

(3) Find $|\Gamma|$ i.e. absolute value of DF statistics

(4) If the computed absolute value of DF statistic exceeds the absolute DF critical value obtained from DF tables, reject the null hypothesis that $\delta = 0$ (i.e. $\rho = 0$). Hence time series is stationary.

(5) If the value so obtained does not exceed the tabulated value then we do not reject the null hypothesis. Hence the time series is non stationary.

(Note that this test is also further revised as Augmented DF test).

<u>**COINTEGRATION**</u>

Let us define DPI = Real disposable personal income

PCE = Real personal consumption expenditure

We take logarithmic values of these variables denoted by LDPI and LPCI respectively. (Note that both the above variables have their respective time series). We consider an equation as under

$LPCE_t = \beta_1 + \beta_2 LDPI_t + U_t$ .... (1)

This is a very well known equation in macroeconomic theory. β_2 is the elasticity of real consumption expenditure with respect to real disposable personal income. We can call it as consumption elasticity.

We write from equation (1) as under

$U_t = LPCE_p - \beta_1 - \beta_2 LDPI_t$ .... (2)

Note that both $LPCE$ and $LDPI$ series have unit root and thus have a stochastic trend.

If we put U_t for unit root analysis we find that it is stationary. This interesting fact is called Cointegration. Two variables will be cointegrated if they have a long term or equilibrium relationship between them. Economic theory is often expressed in equilibrium terms. Here equation (1) above is called Cointegrating equation and the slope coefficient β_2 is called Cointgrating Parameter.

Thus cointegration means that despite being individually non stationary, a linear combination of two or more time series can be stationary. Cointegration between two (or more) time series suggests that there is a long run or equilibrium relationship between them.

There are different tests like Engel Granger (EG) and augmented Engel Granger (AEG) tests to find out if two or more series are cointegrated.

There is Error Correction Mechanism (ECM) given by Engel and Granger which is a means of reconciling the short run behavior of an economic variable into long run behavior.

(Note: There is a very wide literature in respect of time series forecasting models etc. Our study is kept limited up to this stage only)