

**Econometric Modelling**  
**Professor Sujata Kar**  
**Department of Management Studies**  
**Indian Institute of Technology Roorkee**  
**Lecture 38**  
**Stationarity & Unit Root Testing - I**

Hello everyone this is module 38 of the course on econometric modeling. We are at the last leg of the course.

(Refer Slide Time: 0:30)



<b>Part 1: Introduction to Econometrics</b> Module 1: An Overview Module 2: Formulation of Econometric Modelling Module 3 & 4: Review of Basic Concepts Module 5: Types of Data	<b>Part 5: Univariate Time Series Modeling</b> Module 25, 26, 27: Problem of Serial Correlation Module 28: AR, MA & ARMA Processes Module 29: Modelling Seasonal Variations
<b>Part 2: Overview of Classical Linear Regression Model</b> Module 6 & 7: Simple Regression Module 8: Assumption of Classical Linear Regression Module 9: Properties of OLS Estimators Module 10: Hypothesis Testing	<b>Part 6: Models with Binary Dependent and Independent Variables</b> Module 30: Spline Function & Categorical Variables Module 31: Linear Probability Models Module 32: Probit and Logit Models Module 33: Tobit & Multinomial Logit Models
<b>Part 3: Multiple Regression Analysis &amp; Diagnostic Tests</b> Module 11 & 12: Multiple Regression Module 13 & 14: Problems of Multicollinearity Module 15 & 16: Omitted Variables & Parameter Stability Module 17 & 18: Problem of Heteroscedasticity	<b>Part 7: Multivariate Models</b> Module 34: Panel Data Methods Module 35 & 36: Simultaneous Equations System Module 37: Introduction to VARs
<b>Part 4: Statistical Inference</b> Module 19: t-test Module 20 & 21: Wald test Module 22 & 23: F-test Module 24: Chow test	<b>Part 8: Modelling Long Run Relationships</b> Module 38 & 39: Stationarity & Unit Root Testing Module 40: Basics of Cointegration

So, this is the last part, modelling long-run relationships and in that, the first two modules are on stationarity and unit root testing. So, the first module of these two that is module 38 is on primarily the concept of stationarity and how we can deal with or remove stationarity. In the module 39, I will be discussing how to test for stationarity.

(Refer Slide Time: 0:58)

### Stationary Process

- A stationary time series process is one whose probability distributions are stable over time in the following sense: if we take any collection of random variables in the sequence and then shift that sequence ahead  $h$  time periods, the joint probability distribution must remain unchanged.
- Formally, a stochastic process  $\{x_t; t = 1, 2, \dots\}$  is stationary if for every collection of time indices  $1 \leq t_1 < t_2 < \dots < t_m$ , the joint distribution of  $(x_{t_1}, x_{t_2}, \dots, x_{t_m})$  is the same as the joint distribution of  $(x_{t_1+h}, x_{t_2+h}, \dots, x_{t_m+h})$  for all integers  $h \geq 1$ .
- This simply implies that first,  $x_t$  has the same distribution as  $t = 2, 3, \dots$  and second, the joint distribution of  $(x_1, x_2)$  must be the same as the joint distribution of  $(x_t, x_{t+1})$  for any  $t \geq 1$ .

 IIT ROORKEE  NPTEL ONLINE CERTIFICATION COURSE 4



So, a stationary time series process is one whose probability distributions are stable over time in the following sense, that if we take any collection of random variables in the sequence and then shift that sequence ahead  $h$  time periods, the joint probability distribution must remain unchanged.

Formally, a stochastic process (refer slide time: 1:24- 2:25). This definition of stationarity refers to strict stationarity.

(Refer Slide Time: 2:32)

### Stationary Process

- A weaker form of stationarity is **covariance stationarity** and it is defined as a stochastic process  $\{x_t; t = 1, 2, \dots\}$  with finite second moment if
  - i)  $E(x_t)$  is a constant
  - ii)  $\text{Var}(x_t)$  is a finite constant and
  - iii) For any  $t, h \geq 1$ ,  $\text{Cov}(x_t, x_{t+h})$  depends only on  $h$  and not on  $t$ .
- If a stationary process has a finite second moment, then it must be covariance stationary, but the converse is certainly not true. Covariance stationary processes are also called **weakly stationary** or **second order stationary** processes.
- This implies that the first two moments of the series are independent of time, and the covariance between any two values of the series although independent of time, is a function of the distance between the two time periods considered.

 IIT ROORKEE  NPTEL ONLINE CERTIFICATION COURSE 5

Often, we deal with something which is actually a weaker form of stationarity and that is denoted by covariance stationarity. This stationarity was introduced also while we were discussing the univariate time series modeling, that is AR, MA, and ARMA processes.

So, covariance stationarity is defined as a stochastic (refer slide time: 2:55) that is we are considering the stochastic process in the two time periods, then the covariance between the two time periods should actually depend on the distance between the two time periods, that is  $h$ , and not on the time periods themselves.

So, this implies that, if I am considering time period 1 and time period 5. So, there is a distance of 4, the covariance between them will be equal to the covariance between  $x_6$  and  $x_{10}$ , where again we have a difference of 4. If a stationary process has a finite second moment, then it must be covariance stationary, but the converse is certainly not true.



Covariance stationary processes are also called weakly stationary or second-order stationary processes. This implies that the first two moments of the series are independent of time and the covariance between any two values of the series although independent of time is a function of the distance between the two time periods considered.

(Refer Slide Time: 4:33)

### Stationary Process

- A series which is not stationary is called a **non-stationary series**.
- Consider the simple AR(1) model
 
$$y_t = \alpha + \rho y_{t-1} + u_t \quad u_t \sim IID(0, \sigma^2) \quad (1)$$
- In this model the assumption  $|\rho| < 1$  is crucial for the series to be stable. Alternatively, we will prove that the conditions for stationarity hold if  $|\rho| < 1$ .
- Let us rewrite (1) using a lag operator (L) as  $(1 - \rho L)y_t = \alpha + u_t$
- Or  $y_t = (\alpha + u_t)(1 - \rho L)^{-1}$
- Or  $y_t = (\alpha + u_t)(1 + \rho L + \rho^2 L^2 + \dots + \infty) = \frac{\alpha}{1 - \rho} + \sum_{j=0}^{\infty} \rho^j u_{t-j}$ 

$L y_t = y_{t-1}$     $L^2 y_t = y_{t-2}$
- Therefore,  $E(y_t) = \frac{\alpha}{1 - \rho} = \text{constant}$     $[E(u_t) = 0 \Rightarrow E(\sum_j \rho^j u_{t-j}) = 0]$



6

A series that is not stationary is called a non-stationary series. Now, we consider a simple AR(1) model, this is how we used to write an AR(1) model, where we are considering 1 lag length of the AR term, that is an autoregressive process, the endogenous variable. We assume that  $u_t$ , the error term is identically and independently distributed with 0, mean and constant variance at sigma square.

So, in this model the assumption that,  $|\rho| < 1$ , is crucial for the series to be stable. Alternatively, we will prove that the conditions for stationarity hold, if  $|\rho| < 1$ . So, we have mentioned so far the conditions that are required for stationarity, that a constant mean, a



constant variance, that is they are not dependent on time and covariance which is basically a function of the distance between the two time periods, and again not a function of time.

So, we can prove that for values of (refer slide time: 5:40- 8:07).

(Refer Slide Time: 8:07)



### Stationary Process

- We then have,  $Var(y_t) = E[y_t - E(y_t)]^2 = E\left[\frac{\alpha}{1-\rho} + \sum_{j=0}^{\infty} \rho^j u_{t-j} - \frac{\alpha}{1-\rho}\right]^2$   
 $E\left[\sum_{j=0}^{\infty} \rho^j u_{t-j}\right]^2 = \sigma^2 + \rho^2 \sigma^2 + \rho^4 \sigma^2 + \dots = \frac{\sigma^2}{1-\rho^2} = \text{constant}$
- Therefore,  $Var(y_t) > 0$  for  $|\rho| < 1$ .
- And,  $cov(y_t, y_{t-1}) = E\left[\left(y_t - \frac{\alpha}{1-\rho}\right)\left(y_{t-1} - \frac{\alpha}{1-\rho}\right)\right]$   
 $= E\left[\left(y_t y_{t-1} - \frac{\alpha}{1-\rho} E(y_t) - \frac{\alpha}{1-\rho} E(y_{t-1}) + \frac{\alpha^2}{(1-\rho)^2}\right)\right]$   
 $= E(y_t y_{t-1}) - \frac{\alpha^2}{(1-\rho)^2}$  since,  $E(y_t) = E(y_{t-1}) = \frac{\alpha}{1-\rho}$



7

### Stationary Process

- A series which is not stationary is called a **non-stationary series**.
- Consider the simple AR(1) model  
 $y_t = \alpha + \rho y_{t-1} + u_t$       $u_t \sim IID(0, \sigma^2)$      (1)
- In this model the assumption  $|\rho| < 1$  is crucial for the series to be stable. Alternatively, we will prove that the conditions for stationarity hold if  $|\rho| < 1$ .
- Let us rewrite (1) using a lag operator (L) as  $(1 - \rho L)y_t = \alpha + u_t$
- Or  $y_t = (\alpha + u_t)(1 - \rho L)^{-1}$
- Or  $y_t = (\alpha + u_t)(1 + \rho L + \rho^2 L^2 + \dots + \infty) = \frac{\alpha}{1-\rho} + \sum_{j=0}^{\infty} \rho^j u_{t-j}$
- Therefore,  $E(y_t) = \frac{\alpha}{1-\rho} = \text{constant}$       $[E(u_t) = 0 \Rightarrow E(\sum_j \rho^j u_{t-j}) = 0]$



6

Now, we consider (refer slide time: 8:08- 11:55).

(Refer Slide Time: 11:54)

### Stationary Process

- Let us denote  $E(y_t y_{t-\tau})$  by  $\gamma(\tau)$  for  $\tau = 0 \dots t$ . Multiplying both sides of (1) by  $y_{t-1}$  and taking expectations, we obtain  $E(y_t y_{t-1})$

$$\gamma(1) = E[y_t y_{t-1}] = \alpha E[y_{t-1}] + \rho E[y_{t-1}^2] + E[y_{t-1} u_t] = \frac{\alpha^2}{1-\rho} + \rho \gamma(0) \quad (2)$$

- Substituting (2) in the expression for covariance between  $y_t$  and  $y_{t-1}$ , we get

$$\text{cov}[y_t, y_{t-1}] = \rho \gamma(0) + \frac{\alpha^2}{1-\rho} - \frac{\alpha^2}{(1-\rho)^2}$$

- Similarly,  $\text{cov}[y_t, y_{t-2}] = E[(y_t - \frac{\alpha}{1-\rho})(y_{t-2} - \frac{\alpha}{1-\rho})] = E[y_t y_{t-2}] - \frac{\alpha^2}{(1-\rho)^2}$   
 $= \gamma(2) - \frac{\alpha^2}{(1-\rho)^2}$

- From (2) it follows directly that  $\gamma(2) = \frac{\alpha^2}{1-\rho} + \rho \gamma(1)$  (3)

### Stationary Process

- A series which is not stationary is called a **non-stationary series**.
- Consider the simple AR(1) model

$$y_t = \alpha + \rho y_{t-1} + u_t \quad u_t \sim IID(0, \sigma^2) \quad (1)$$

- In this model the assumption  $|\rho| < 1$  is crucial for the series to be stable. Alternatively, we will prove that the conditions for stationarity hold if  $|\rho| < 1$ .

- Let us rewrite (1) using a lag operator (L) as  $(1 - \rho L)y_t = \alpha + u_t$

- Or  $y_t = (\alpha + u_t)(1 - \rho L)^{-1}$

- Or  $y_t = (\alpha + u_t)(1 + \rho L + \rho^2 L^2 + \dots + \infty) = \frac{\alpha}{1-\rho} + \sum_{j=0}^{\infty} \rho^j u_{t-j}$

- Therefore,  $E(y_t) = \frac{\alpha}{1-\rho} = \text{constant}$  [ $E(u_t) = 0 \Rightarrow E(\sum_j \rho^j u_{t-j}) = 0$ ]

### Stationary Process

- We then have,  $\text{Var}(y_t) = E[y_t - E(y_t)]^2 = E\left[\frac{\alpha}{1-\rho} + \sum_{j=0}^{\infty} \rho^j u_{t-j} - \frac{\alpha}{1-\rho}\right]^2$

$$E\left[\sum_j \rho^j u_{t-j}\right]^2 = \sigma^2 + \rho^2 \sigma^2 + \rho^4 \sigma^2 + \dots = \frac{\sigma^2}{1-\rho^2} = \text{constant}$$

- Therefore,  $\text{Var}(y_t) > 0$  for  $|\rho| < 1$ .

- And,  $\text{cov}(y_t, y_{t-1}) = E\left[\left(y_t - \frac{\alpha}{1-\rho}\right)\left(y_{t-1} - \frac{\alpha}{1-\rho}\right)\right]$

$$= E\left[\left(y_t y_{t-1} - \frac{\alpha}{1-\rho} E(y_t) - \frac{\alpha}{1-\rho} E(y_{t-1}) + \frac{\alpha^2}{(1-\rho)^2}\right)\right]$$

$$= E(y_t y_{t-1}) - \frac{\alpha^2}{(1-\rho)^2} \quad \text{since, } E(y_t) = E(y_{t-1}) = \frac{\alpha}{1-\rho}$$

Let us denote (refer slide time: 11:56- 14:45)



(Refer Slide Time: 14:45)

### Stationary Process

- Substituting for the value of  $\gamma(1)$  in (3) we have  $\gamma(2) = \frac{\alpha^2}{1-\rho} + \frac{\alpha^2\rho}{1-\rho} + \rho^2\gamma(0)$
- Therefore,  $Cov(y_t, y_{t-2}) = \gamma(2) - \frac{\alpha^2}{(1-\rho)^2} = \rho^2\gamma(0) + \frac{\alpha^2}{1-\rho} + \frac{\alpha^2\rho}{1-\rho} - \frac{\alpha^2}{(1-\rho)^2}$
- Following this iterative process it can be shown that

$$Cov(y_t, y_{t-\tau}) = \rho^\tau\gamma(0) + \frac{\alpha^2}{1-\rho}(1 + \rho + \rho^2 + \dots + \rho^{t-1}) - \frac{\alpha^2}{(1-\rho)^2} = \lambda(\tau) \quad (4)$$

- Note that  $\lambda(0) = Var(y_t) = E[y_t^2] - E[y_t]^2 = \gamma(0) - \frac{\alpha^2}{(1-\rho)^2}$
- Therefore,  $\gamma(0) = \lambda(0) + \frac{\alpha^2}{(1-\rho)^2} = \frac{\sigma^2}{1-\rho^2} + \frac{\alpha^2}{(1-\rho)^2} = \text{constant}$
- From expression (4) it is clear that the covariance of the process in (1) will converge to a constant value only if  $|\rho| < 1$ .





9

### Stationary Process

- Let us denote  $E(y_t, y_{t-\tau})$  by  $\gamma(\tau)$  for  $\tau = 0 \dots t$ . Multiplying both sides of (1) by  $y_{t-1}$  and taking expectations, we obtain  $E(y_{t-1}y_t) = \alpha E(y_{t-1}) + \rho E(y_{t-1}^2) + E(y_{t-1}u_t)$

$$\gamma(1) = E[y_t y_{t-1}] = \alpha E[y_{t-1}] + \rho E[y_{t-1}^2] + E[y_{t-1}u_t] = \frac{\alpha^2}{1-\rho} + \rho\gamma(0) \quad (2)$$

- Substituting (2) in the expression for covariance between  $y_t$  and  $y_{t-1}$ , we get  $cov[y_t, y_{t-1}] = \rho\gamma(0) + \frac{\alpha^2}{1-\rho} - \frac{\alpha^2}{(1-\rho)^2}$
- Similarly,  $cov[y_t, y_{t-2}] = E[(y_t - \frac{\alpha}{1-\rho})(y_{t-2} - \frac{\alpha}{1-\rho})] = E[y_t, y_{t-2}] - \frac{\alpha^2}{(1-\rho)^2} = \gamma(2) - \frac{\alpha^2}{(1-\rho)^2}$
- From (2) it follows directly that  $\gamma(2) = \frac{\alpha^2}{1-\rho} + \rho\gamma(1) \quad (3)$



8

So, substituting the value of (refer slide time: 14:46- 17:30).

(Refer Slide Time: 17:31)

## Stationary Process

- Therefore, the stationarity condition is that the absolute value of the coefficient of the one period lagged value of the dependent variable of an AR (1) series must be less than unity. Two alternative statements of the stationarity condition are,
  1. The coefficient of  $y_{t-1}$  in (1) must lie within the unit circle
  2. The absolute value of all the roots of the polynomial lag operator,  $1 - \rho L = 0$  must lie outside the unit circle.
- Here  $1 - \rho L = 0$  has a single root equal to  $1/\rho$ . Therefore,  $|1/\rho| > 1$  or  $|\rho| < 1$  is the stationarity condition. If  $|\rho| = 1$  then there is a unit root problem as the root of the polynomial lag operator becomes unity. Further, if the roots of the characteristic equation are greater than unity, i.e. when  $|\rho| > 1$  the series is called an explosive series.



Therefore, the stationarity condition is that the absolute value of the coefficient of the one period lagged value of the dependent variable of an AR(1), series must be less than unity. Two alternative statements of the stationarity conditions are, (refer slide time: 17:49- 18:25)

Further, if the roots of the characteristic equation are greater than unity, that is when  $|\rho| > 1$ , the series can be called an explosive series, or the series is actually an explosive series.

(Refer Slide Time: 18:40)

### Highly Persistent Series

- Many economic time series are better characterized by an AR(1) model with  $\rho = 1$ . Let us write the AR(1) process as
$$y_t = y_{t-1} + u_t \quad t = 1, 2, \dots \text{ and } u_t \sim IID(0, \sigma^2) \quad (5)$$
- We assume that the initial value  $y_0$  is independent of  $u_t$  for all  $t \geq 1$ .
- The process in (5) is called a random walk, a special case of unit root process. The name comes from the fact that  $y$  at time  $t$  is obtained by starting at the previous value,  $y_{t-1}$ , and adding a zero mean random variable that is independent of  $y_{t-1}$ .
- By repeated substitution we will get
$$y_t = u_t + u_{t-1} + \dots + u_1 + y_0 \quad (6)$$
- Taking the expected value of both sides gives
$$E(y_t) = E(u_t) + E(u_{t-1}) + \dots + E(u_1) + E(y_0) = E(y_0) \quad \text{for all } t \geq 1.$$

  11

So, now we talk about the properties of having a non-stationary process. So, non-stationary processes are generally highly persistent. So, now I discuss why do we call it a highly persistent series. So, many economic time series are better characterized by an AR(1) model (refer slide time: 19:00- 19:49).

So, basically, the entire process, the changes, whatever we are observing in  $y_t$ , that is generated by the random component. So, it is absolutely random, if we remove that randomness my current observation would be exactly equal to my previous observation, the entire movements in  $y_t$  are caused by the random term that is the disturbance term, and that is why it is called a random walk process.

By repeated substitution in equation 5, we get, (refer slide time: 20:23- 20:57).

(Refer Slide Time: 20:57)



## Highly Persistent Series

- Therefore, the expected value of a random walk does not depend on  $t$ . A popular assumption is that  $y_0 = 0$ , in which case  $E(y_t) = 0$  for all  $t$ .
- By contrast, the variance of a random walk does change with  $t$ . For simplicity, let us assume  $\text{Var}(y_0) = 0$ . Then
- $\text{Var}(y_t) = \text{Var}(u_t) + \text{Var}(u_{t-1}) + \dots + \text{Var}(u_1) = \sigma_u^2 t$
- This shows that the variance of a random walk increases as a linear function of time and the process cannot be stationary. Even more importantly, a random walk displays highly persistent behavior in the sense that the value of  $y$  today is important for determining the value of  $y$  in the very distant future. To see this, write equation (6) for  $h$  periods ahead as,

$$y_{t+h} = u_{t+h} + u_{t+h-1} + \dots + u_{t+1} + y_t$$



## Highly Persistent Series

- Many economic time series are better characterized by an AR(1) model with  $\rho = 1$ . Let us write the AR(1) process as

$$y_t = y_{t-1} + u_t \quad t = 1, 2, \dots \text{ and } u_t \sim \text{IID}(0, \sigma^2) \quad (5)$$

- We assume that the initial value  $y_0$  is independent of  $u_t$  for all  $t \geq 1$ .
- The process in (5) is called a random walk, a special case of unit root process. The name comes from the fact that  $y$  at time  $t$  is obtained by starting at the previous value,  $y_{t-1}$ , and adding a zero mean random variable that is independent of  $y_{t-1}$ .
- By repeated substitution we will get

$$y_t = u_t + u_{t-1} + \dots + u_1 + y_0 \quad (6)$$

- Taking the expected value of both sides gives

$$E(y_t) = E(u_t) + E(u_{t-1}) + \dots + E(u_1) + E(y_0) = E(y_0) \quad \text{for all } t \geq 1.$$



Therefore, the expected value of a random walk does not depend on  $t$ . A popular assumption is that (refer slide time: 21:03- 21:47).

This shows that the variance of a random work increases as a linear function of time and the process cannot be stationary. Even more importantly a random walk displays highly persistent behavior in the sense that, the value of  $y$  today is important for determining the value of  $y$  in the very distant future.

To see this write equation 6 (refer slide time: 22:11).

(Refer Slide Time: 22:28)

## Highly Persistent Series

- At time  $t$ , the expected value of  $y_{t+h}$  is  

$$E(y_{t+h}|y_t) = y_t \text{ for all } h \geq 1.$$
- This means that, no matter how far in the future we look, our best prediction of  $y_{t+h}$  is today's value,  $y_t$ . In contrast to this, a stable AR(1) process will have  

$$E(y_{t+h}|y_t) = \rho^h y_t \text{ for all } h \geq 1. \quad h \rightarrow \uparrow, \rho^h \rightarrow \text{smaller.}$$
- Under stability,  $|\rho| < 1$ . Therefore,  $E(y_{t+h}|y_t)$  approaches zero as  $h \rightarrow \infty$ , i.e. the value of  $y_t$  becomes less and less important, and  $E(y_{t+h}|y_t)$  gets closer and closer to the unconditional expected value,  $E(y_t) = 0$ .
- Also, if  $\text{Var}(y_0) = 0$ , it can be shown that  $\text{corr}(y_t, y_{t+h}) = \sqrt{t/(t+h)}$
- Thus, the correlation depends on the starting point,  $t$  and it becomes close to 1 for large  $t$  when  $\{y_t\}$  follows a random walk. Further, although for fixed  $t$  the correlation tends to zero as  $h \rightarrow \infty$ , it does not do so very quickly. In fact, the larger  $t$  is, the more slowly the correlation tends to zero as  $h$  gets large.



## Highly Persistent Series

- Therefore, the expected value of a random walk does not depend on  $t$ . A popular assumption is that  $y_0 = 0$ , in which case  $E(y_t) = 0$  for all  $t$ .
- By contrast, the variance of a random walk does change with  $t$ . For simplicity, let us assume  $\text{Var}(y_0) = 0$ . Then
- $$\text{Var}(y_t) = \text{Var}(u_t) + \text{Var}(u_{t-1}) + \dots + \text{Var}(u_1) = \sigma_u^2 t$$
- This shows that the variance of a random walk increases as a linear function of time and the process cannot be stationary. Even more importantly, a random walk displays highly persistent behavior in the sense that the value of  $y$  today is important for determining the value of  $y$  in the very distant future. To see this, write equation (6) for  $h$  periods ahead as,

$$y_{t+h} = u_{t+h} + u_{t+h-1} + \dots + u_{t+1} + y_t$$



At time  $t$ , (refer slide time: 22:30- 24:11).

Further, although for fixed  $t$  the correlation tends to 0 as  $h$  tends to infinity it does not do so very quickly. In fact, the larger  $t$  is, the more slowly the correlation tends to 0 as  $h$  gets larger. So, that is why we call this series you know highly persistent series.

(Refer Slide Time: 24:30)

## Highly Persistent Series

- Let us consider stock prices as an example. The dynamics of the price process is given by  $S_t = S_{t-1} + u_t$  or,  $\Delta S_t = u_t$
- This implies that tomorrow's price,  $S_{t+1}$ , is thought of as today's price plus some *random shock* that is independent of the price. As a consequence, in this model, the increments  $S_t - S_{t-1}$  from  $t-1$  to  $t$  are thought of as completely undeterministic.
- Since the  $u_t$ 's have a mean of zero, the increments are considered fair. An increase in price is as likely as a downside movement. At time  $t$ , the price is considered to contain all information available. So at any point in time, next period's price is exposed to a random shock. Consequently, the best estimate for the following period's price is this period's price.

Now, let us consider stock prices as an example, the dynamics of the price process are given by (refer slide time: 24:36). So,  $S_t$  refers to stock price, so today's stock price is yesterday's stock price plus some random component, or the change in today's stock price is actually equal to today's random component.

So, it is not at all explained by anything other than that, this implies that tomorrow's price that is (refer slide time: 25:00- 25:22).

Since the  $u_t$ 's have a mean of 0, the increments are considered fair. An increase in a price is likely a downside movement. At time  $t$ , the price is considered to contain all information available. So, at any point in time, the next period's price is exposed to a random shock. Consequently, the best estimate for the following period's price is this period's price or today's price.

(Refer Slide Time: 25:48)

## Highly Persistent Series

- Generally, it is not easy to look at a time series plot and determine whether it is a random walk or not.
- It is extremely important not to confuse trending and highly persistent behaviors. A series can be trending but not highly persistent. It is also often the case that a highly persistent series also contains a clear trend. One model that leads to this behavior is the **random walk with drift**,

$$y_t = \alpha_0 + y_{t-1} + u_t \quad t = 1, 2, \dots \rightarrow \text{AR}(1) \quad \rho = 1.$$

- $\alpha_0$  is called the *drift term*. For such a series the expected value of  $y_t$  follows linear time trend by using repeated substitution,

$$y_t = \alpha_0 t + u_t + u_{t-1} + \dots + u_1 + y_0$$

- Therefore, if  $y_0 = 0$ ,  $E(y_t) = \alpha_0 t$ .
- The expected value of  $y_t$  grows over time if  $\alpha_0 > 0$  and shrinks over time if  $\alpha_0 < 0$ .



Generally, it is not easy to look at a time series plot and determine whether it is a random walk, or not. So, there is no shortcut or straightforward way of looking determining it just by looking at the series. It is extremely important not to confuse trending and highly persistent behaviors. A series can be trending, but not highly persistent. It is also often the case that a highly persistent series also contains a clear trend.

So, the trend the way we have understood it so far or discussed previously that it is a deterministic trend. And it is sometimes possible that a highly persistent series, that a random walk series also contains a clear cut trend. One model that leads to this behavior is the random walk with drift. So, this is a series with a random walk with drift, you can see that this is very similar to the AR(1) model, that was considered in the beginning, except for the fact that here  $\rho$  has taken the value 1.

So, (refer slide time: 26:50- 27:50). But, overall the fact is that the expected value is actually a function of time and it is not constant the way we want covariance stationary series to be.

(Refer Slide Time: 27:58)

## Highly Persistent Series

- The variance of the series is  $Var(y_t) = E[\sum_{j=0}^{t-1} u_{t-j}]^2 = t\sigma^2$
- Therefore, neither the mean nor the variance of the process is constant; rather they are functions of time which violates the first two conditions for stationarity. Similarly, it can be shown that the covariance between any two terms will be a function of time. Suppose,  $y_0 = 0$ . Then it follows that

$$Cov(y_t, y_{t-1}) = E[y_t y_{t-1}] - E[y_t]E[y_{t-1}] = E[y_t y_{t-1}] - \alpha_0^2 t(t-1)$$

- Since,  $E[y_t y_{t-1}] = E[\alpha_0 y_{t-1} + y_{t-1}^2 + u_t y_{t-1}] = \alpha_0^2 (t-1) + \sigma^2 (t-1)$ ,  
 $Cov(y_t, y_{t-1}) = \sigma^2 (t-1) - \alpha_0^2 (t-1)^2$

Generalising this expression we get,

- $Cov(y_t, y_{t-\tau}) = \sigma^2 (t-\tau) - \alpha_0^2 (t-\tau)^2$  for any  $\tau$



## Highly Persistent Series

- Generally, it is not easy to look at a time series plot and determine whether it is a random walk or not.
- It is extremely important not to confuse trending and highly persistent behaviors. A series can be trending but not highly persistent. It is also often the case that a highly persistent series also contains a clear trend. One model that leads to this behavior is the **random walk with drift**,

$$y_t = \alpha_0 + y_{t-1} + u_t \quad t = 1, 2, \dots \rightarrow AR(1) \quad \rho = 1$$

- $\alpha_0$  is called the *drift term*. For such a series the expected value of  $y_t$  follows linear time trend by using repeated substitution,

$$y_t = \alpha_0 t + u_t + u_{t-1} + \dots + u_1 + y_0$$

- Therefore, if  $y_0 = 0$ ,  $E(y_t) = \alpha_0 t$
- The expected value of  $y_t$  grows over time if  $\alpha_0 > 0$  and shrinks over time if  $\alpha_0 < 0$ .



The variance of the series is (refer slide time: 28:00- 28:39)

Therefore, neither the mean nor the variance of the process is constant rather they are functions of time that violates the first two conditions for stationarity.

Similarly, it can be shown that the covariance between any two terms will be a function of time. Suppose, (refer slide time: 28:53- 30:28).

(Refer Slide Time: 30:30)



### Stationarity of AR(p) processes

- Therefore, the third condition under weak stationarity is also violated as the covariance between any two values of the series is a function of time.
- In order to make the initial presentation simple the problem of unit root is considered in an AR (1) series. However, the argument is valid also for higher order autoregressive processes. For instance, consider an AR (p) process as follows:

$$y_t = \alpha + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + \rho_p y_{t-p} + u_t \quad u_t \sim \text{IID}(0, \sigma^2) \quad (7)$$

- Equation (7) can be rewritten in the lag operators as
- $(1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_p L^p) y_t = \alpha + u_t$
- The stationarity condition is that all the roots of the polynomial lag operator  $1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_p L^p = 0$  should lie outside the unit circle. This condition implies that the sum of all the coefficients must be less than unity; i.e.

$$\rho_1 + \rho_2 + \dots + \rho_p < 1$$



17



Therefore, the third condition under weak stationarity is also violated as the covariance between any two values of the series is a function of time. In order to make the initial presentation simple, the problem of unit root is considered in an AR(1) series, however, the argument is valid also for higher-order autoregressive processes.

For instance, consider an (refer slide time: 30:50- 31:47 ).

(Refer Slide Time: 31:48)

### Unit root problem

- If the variables employed in a regression model are not stationary, then it can be proved that the standard assumptions for asymptotic analysis will not be valid. In other words, the usual 't-ratios' will not follow a t-distribution, and the F-statistic will not follow an F-distribution, and so on.
- Therefore, before proceeding to econometric estimation, time series analysis requires testing for the presence of unit root(s) or non-stationarity in all the data series considered in a model.
- However, before discussing the testing procedure, let us first focus upon how to convert a non-stationary series into a stationary series.
- A series with one unit root can be transformed to a stationary process by differencing the series once.



18

If the variable employed in a regression model are not stationary, then it can be proved that the standard assumptions for asymptotic analysis will not be valid. In other words, the usual 't-ratios' will not follow a t-distribution and the F-statistic will not follow an F-distribution, and so on. Therefore, before proceeding to econometric estimation time series analysis



requires testing for the presence of unit roots or non-stationarity in all the data series considered in a model.

However, before discussing the testing procedure let us first focus on how to convert a non-stationary series into a stationary series. A series with one unit can be transformed into a stationary process by differencing the series once.

(Refer Slide Time: 32:37)

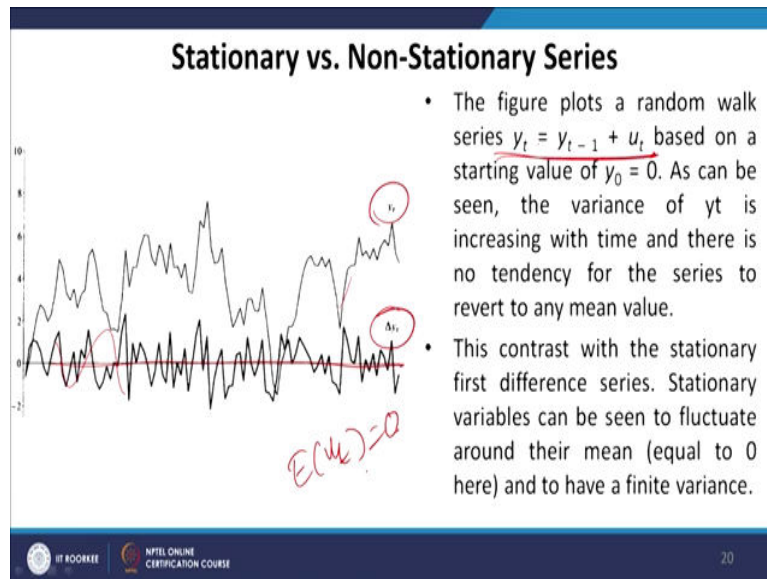
### Differencing and Stationarity

- For instance, consider a random walk series with drift,
 
$$y_t = \alpha + y_{t-1} + u_t \quad u_t \sim IID(0, \sigma^2) \quad (8)$$
- If  $y_{t-1}$  is subtracted from both sides of (8) then the new series becomes
 
$$y_t - y_{t-1} = \alpha + u_t \quad \text{Or} \quad \Delta y_{t-1} = \alpha + u_t \quad E(u_t) = 0.$$
- Hence,  $E[\Delta y_t] = \alpha$  Because  $E(u_t) = 0$  and  $\text{Var}[\Delta y_t] = \sigma^2$
- Therefore, the series in (8) fulfills the first two conditions required for stationarity. Since there is no lagged value of the dependent variable on the right hand side, the covariance is zero. Similarly, if there are two unit roots then the series must be differenced twice to make it a stationary series. Thus, it can be generalized that a non-stationary series with  $d$  unit roots,  $d > 0$ , should be differenced  $d$  times to make it stationary. With  $d$  unit roots a series is said to be 'integrated of order  $d$ ' and is denoted as  $I(d)$ ,  $d$  is the order of integration.



19

So, for instance, we consider a random walk series with drift. (Refer slide time: 32:42-33:12). Therefore, the series in 8 fulfills the first two conditions required for stationarity. Since there is no lag value of the dependent variable on the right-hand side, the covariance is 0. Similarly, if there are two unit-roots then the series must be differenced twice to make it a stationary series. Thus, it can be generalized that a non-stationary series with  $d$  unit-roots,  $d > 0$  should be differenced  $d$  times to make it stationary. With  $d$  unit-roots a series is said to be integrated of order  $d$  and is denoted by  $I(d)$ ,  $d$  is the order of integration.

(Refer Slide Time: 33:55)



Now, I conclude this module with this graph which basically plots stationary series against a non-stationary series. So, first of all, this (refer slide time: 34:00- 34:55). So, you do not observe the series to deviate much over a period of time from its mean value the deviation remains constant. So, that is primarily about the stationary series, what kind of problem it poses and how we can take care of stationarity.

(Refer Slide Time: 35:18)

### References

- Wooldridge, Jeffrey M (2009). *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning, USA.
- Brooks, Chris (2008). *Introductory Econometrics for Finance*. Cambridge University Press, New York.
- Harris, Richard & Sollis, Robert (2003). *Applied Time Series Modelling and Forecasting*. Wiley, England.

IT ROORKEE    NPTEL ONLINE CERTIFICATION COURSE    21

These are the references that I have consulted. So, in the next module, I will continue with the testing of stationarity and once we test for the presence of stationarity, then we have already discussed how we can correct for the presence of stationarity. An alternative is basically going for co-integration or co-integrated models that will be taken up in the last module. Thank you.