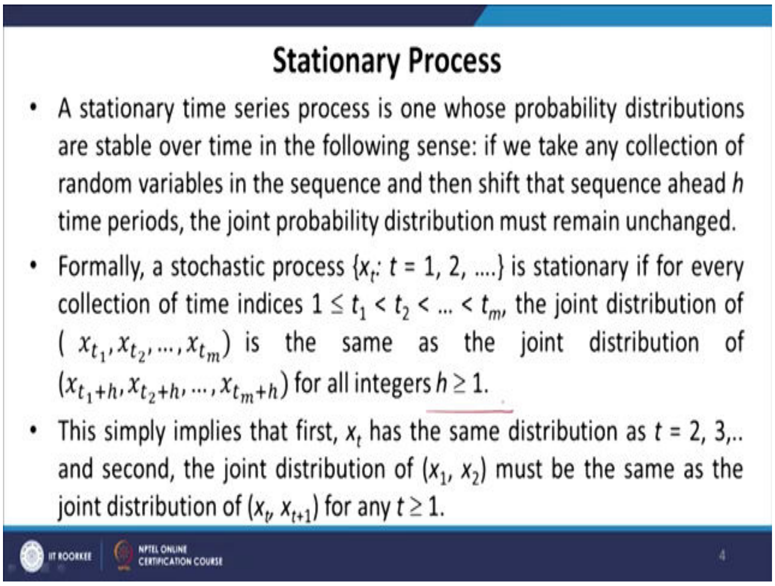
Econometric Modelling
Professor Sujata Kar
Department of Management Studies
Indian Institute of Technology Roorkee
Lecture 38
Stationarity & Unit Root Testing - I

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)

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.

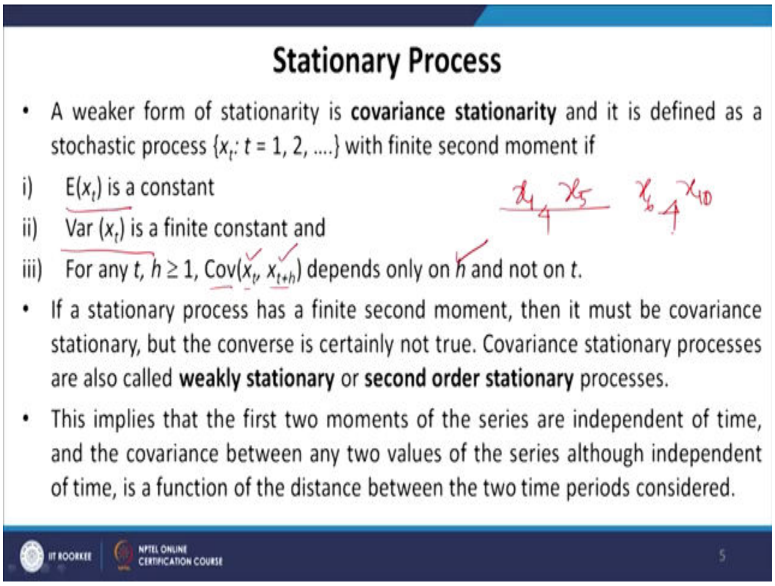
(Refer Slide Time: 0:58)



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)



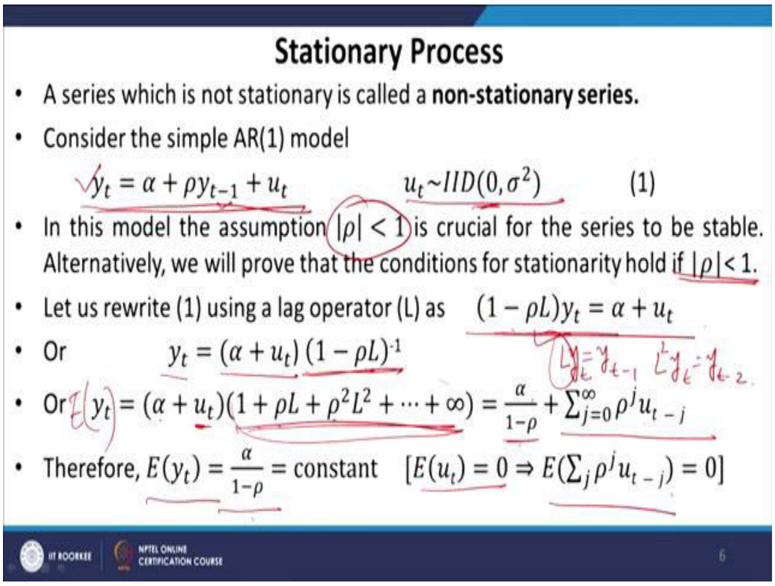
Often, we deal with something which is actually a weaker form of stationarity and that is denoted by covariance stationarity. This stationarity was introduced also while we were discussing the univariate time series modeling, that is AR, MA, and ARMA processes.

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 betweenx_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)

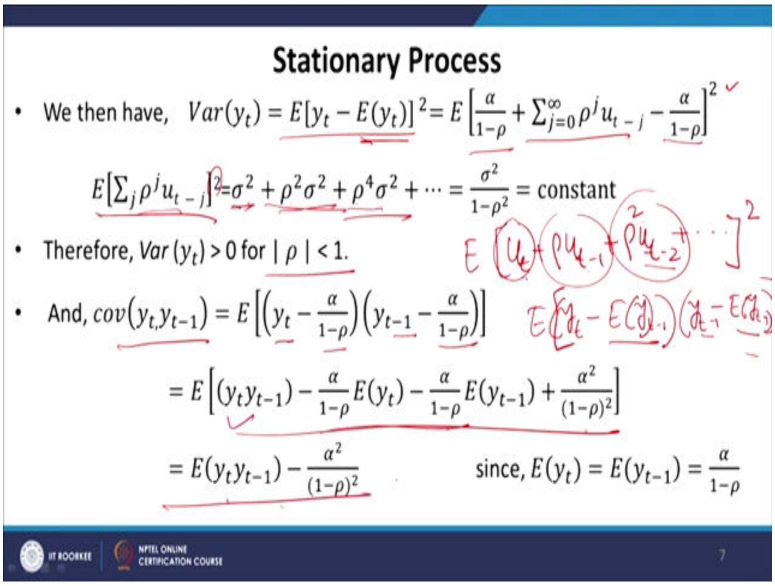


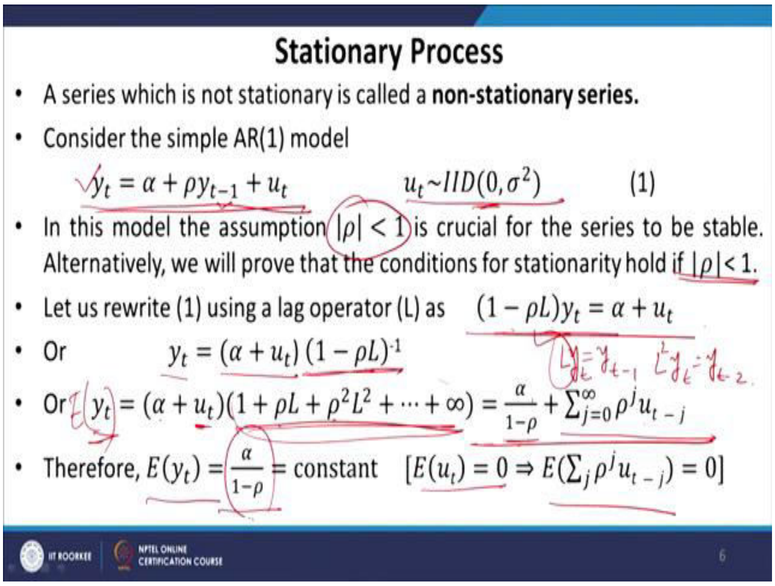
A series that is not stationary is called a non-stationary series. Now, we consider a simple AR(1) model, this is how we used to write an AR(1) model, where we are considering 1lag length of the AR term, that is an autoregressive process, the endogenous variable. We assume that u_t, the error term is 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)





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 $\gamma(\tau)$ $for\ \tau = 0 \ldots t$ Multiplying both sides of (1) by y_{t-1} and taking expectations, we obtain

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

- 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 $\gamma(2) = \frac{\alpha^2}{1-\rho} + \rho\gamma(1) \qquad (3)$

IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 8

Stationary Process

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

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

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

IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 6

Stationary Process

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

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

- Therefore, $Var(y_t) > 0$ for $|\rho| < 1$.
- And, $cov(y_t, y_{t-1}) = E\left[\left(y_t - \frac{\alpha}{1-\rho}\right)\left(y_{t-1} - \frac{\alpha}{1-\rho}\right)\right]$

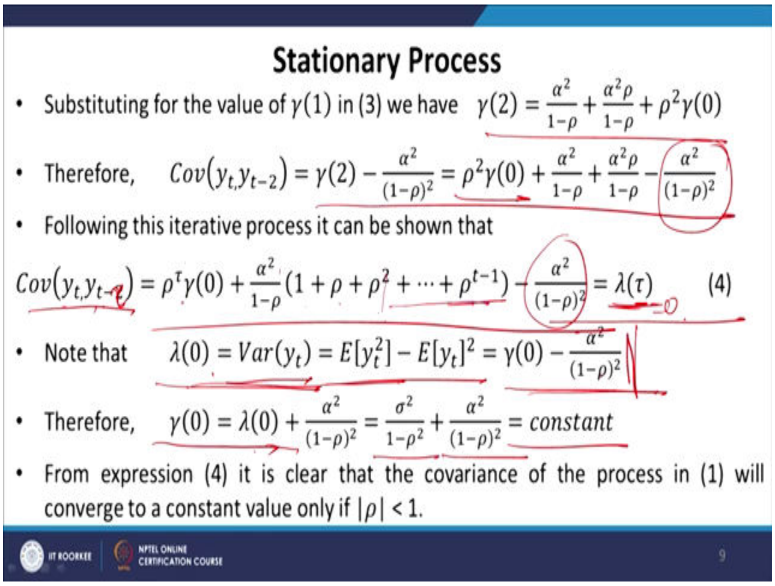
$$= E\left[(y_t y_{t-1}) - \frac{\alpha}{1-\rho} E(y_t) - \frac{\alpha}{1-\rho} E(y_{t-1}) + \frac{\alpha^2}{(1-\rho)^2}\right]$$

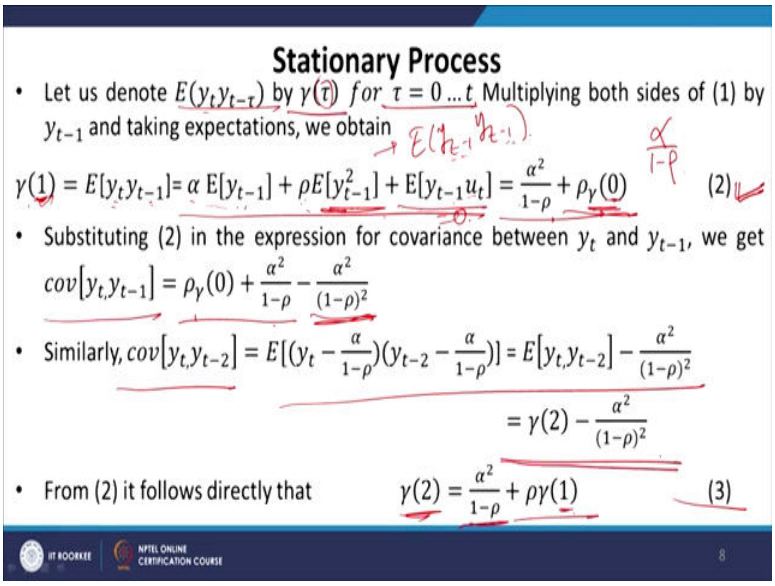
$$= E(y_t y_{t-1}) - \frac{\alpha^2}{(1-\rho)^2} \qquad \text{since, } E(y_t) = E(y_{t-1}) = \frac{\alpha}{1-\rho}$$

IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 7

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

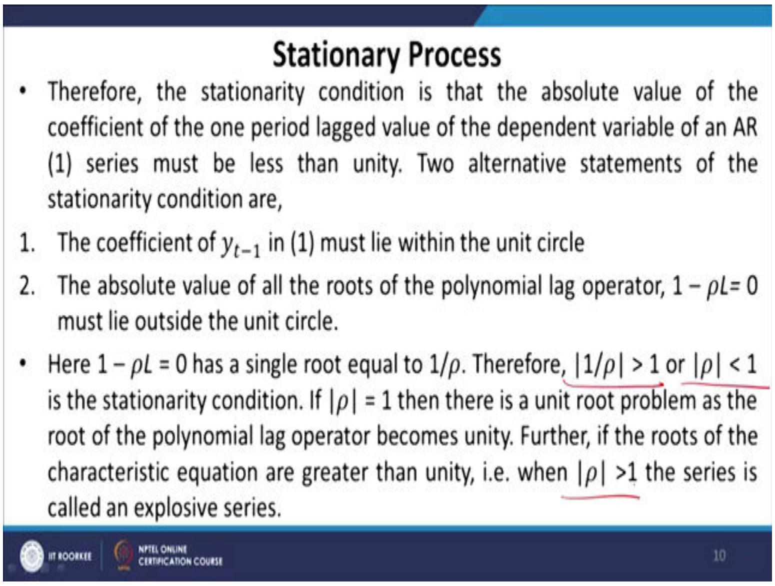
(Refer Slide Time: 14:45)





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

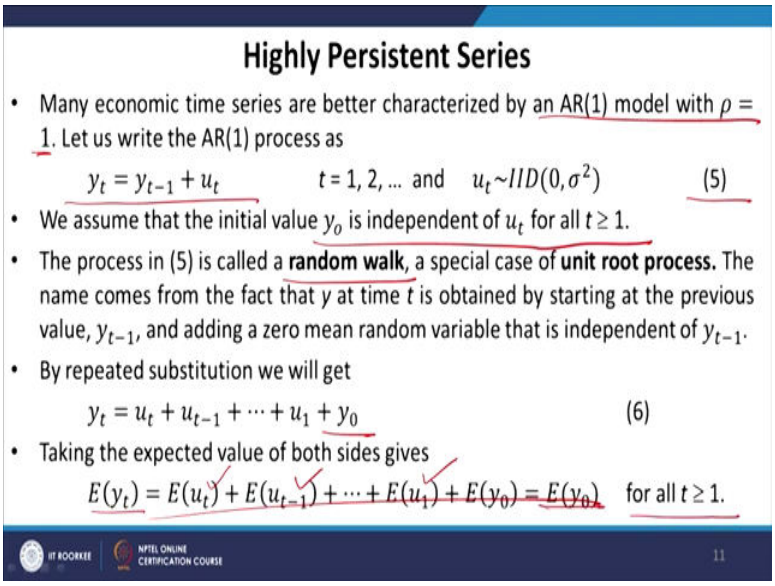
(Refer Slide Time: 17:31)



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)

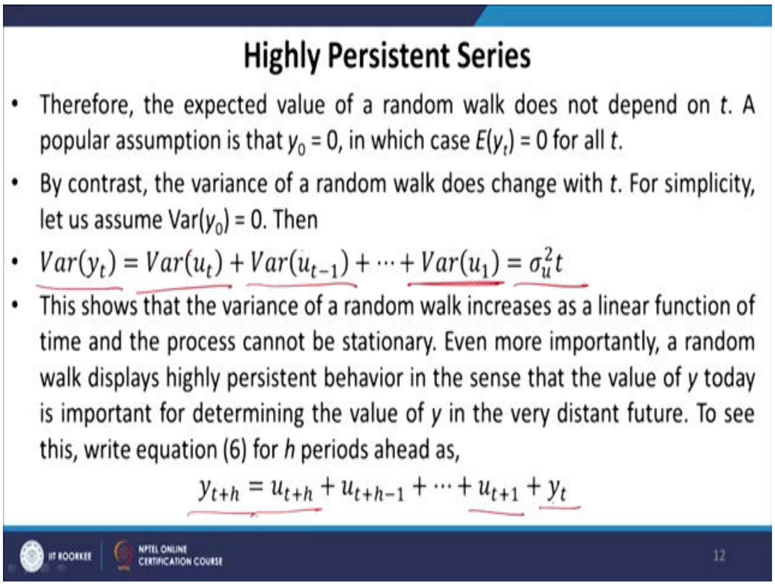


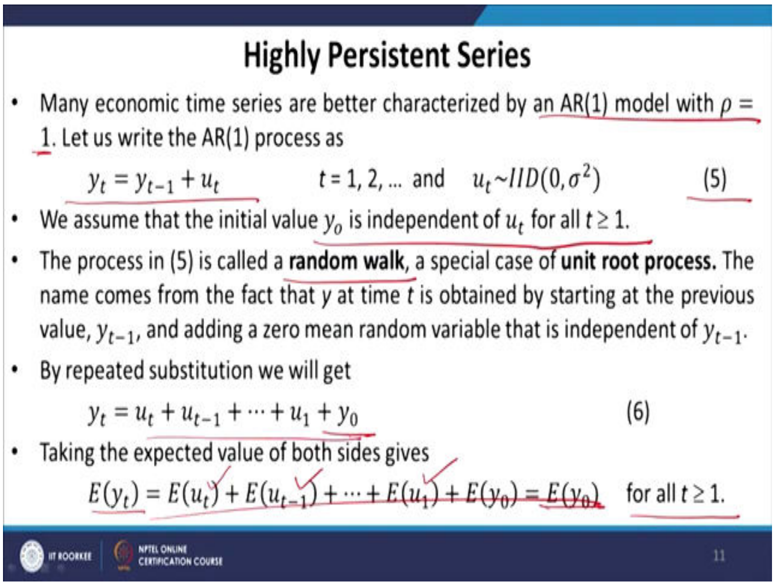
So, now we talk about the properties of having a non-stationary process. So, non-stationary processes are generally highly persistent. So, now I discuss why do we call it a highly persistent series. So, many economic time series are better characterized by an AR(1) model (refer slide time: 19:00- 19:49).

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)



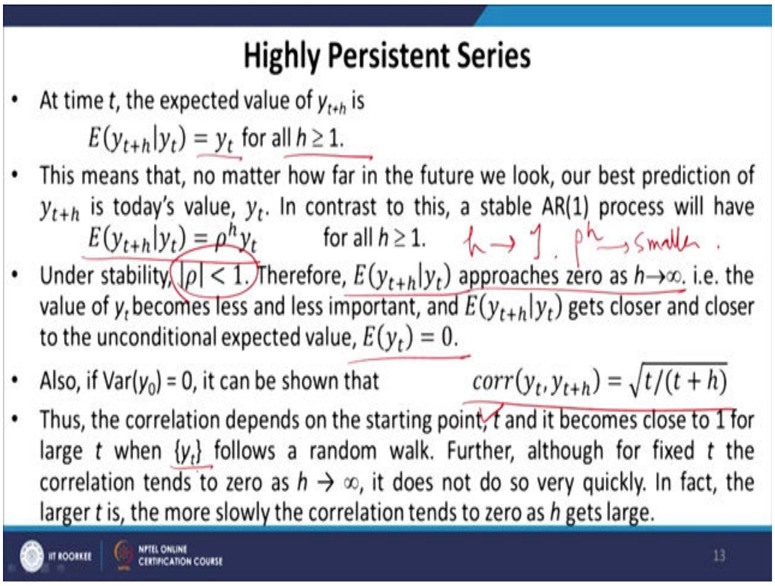


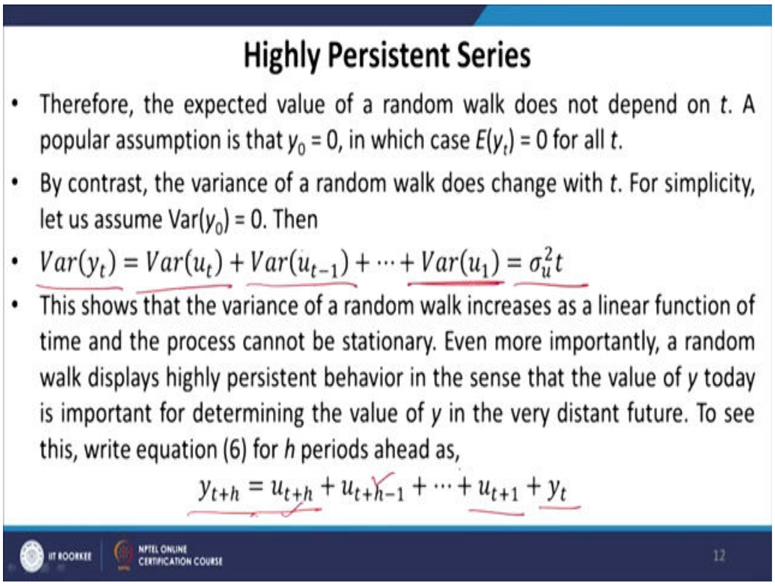
Therefore, the expected value of a random walk does not depend on t. A popular assumption is that (refer slide time: 21:03- 21:47).

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)

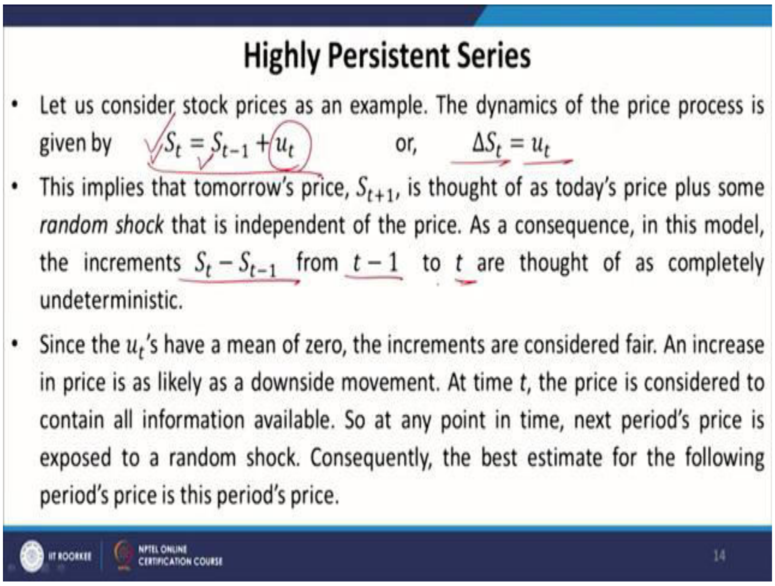




At time t, (refer slide time: 22:30- 24:11).

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)

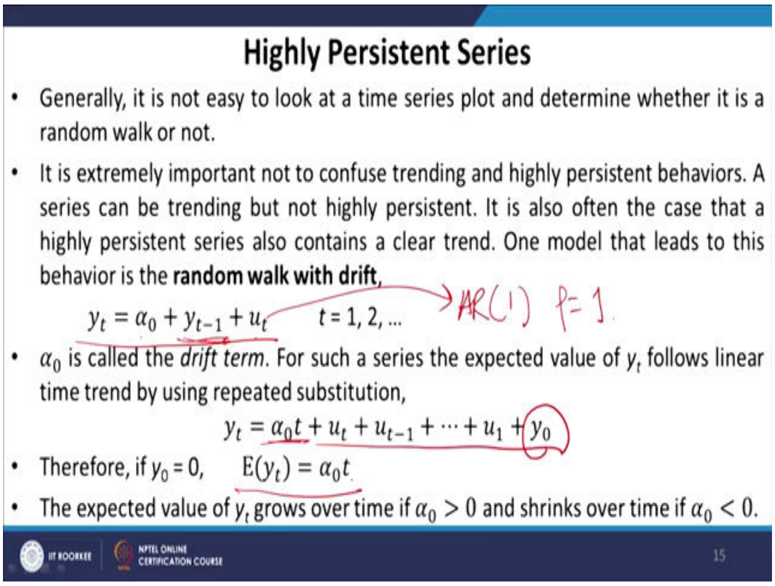


Now, let us consider stock prices as an example, the dynamics of the price process are given by (refer slide time: 24:36). So, S_t refers to stock price, so today's stock price is yesterday's stock price plus some random component, or the change in today's stock price is actually equal to today's random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

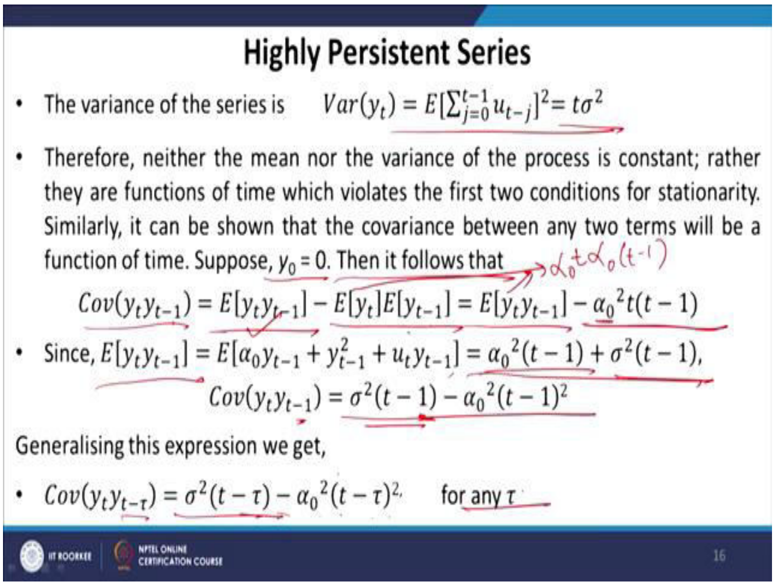


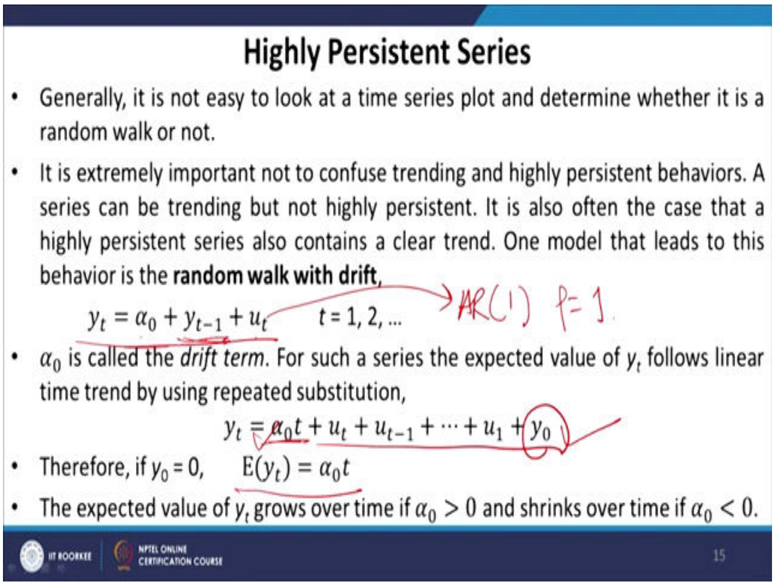
Generally, it is not easy to look at a time series plot and determine whether it is a random walk, or not. So, there is no shortcut or straightforward way of looking determining it just by looking at the series. It is extremely important not to confuse trending and highly persistent behaviors. A series 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 ρ 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)



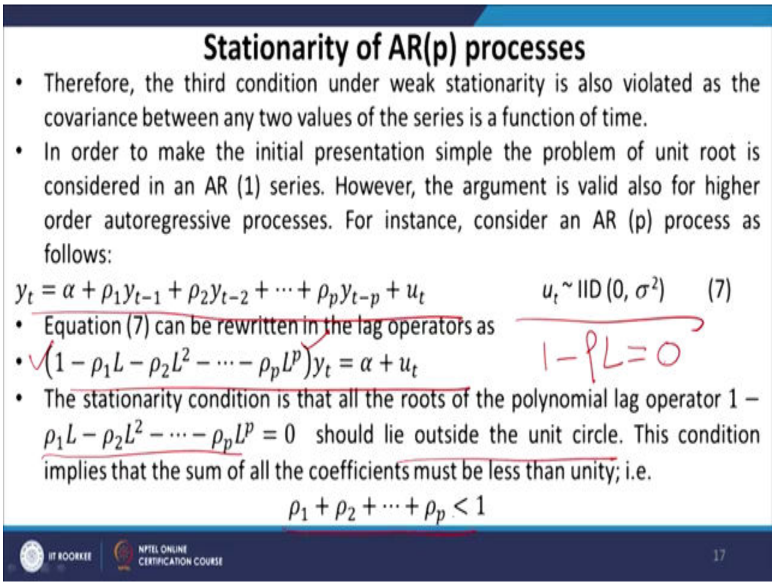


The variance of the series is (refer slide time: 28:00- 28:39)

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).

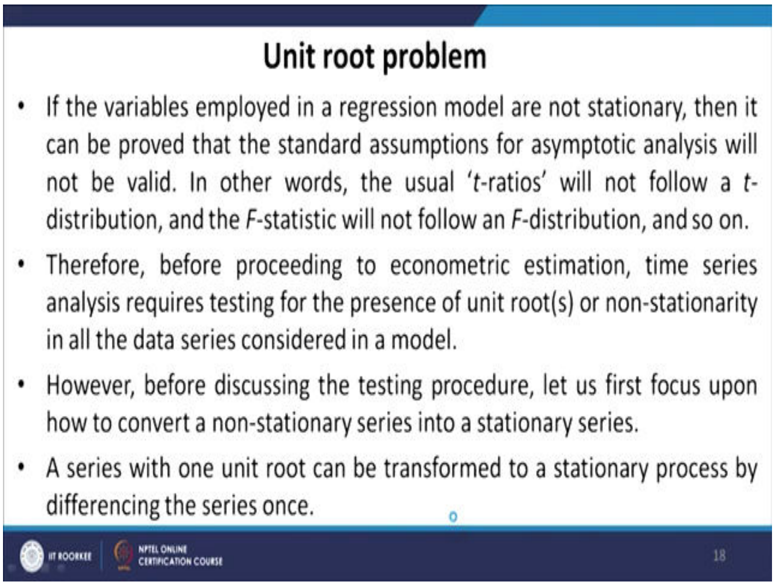
(Refer Slide Time: 30:30)



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).

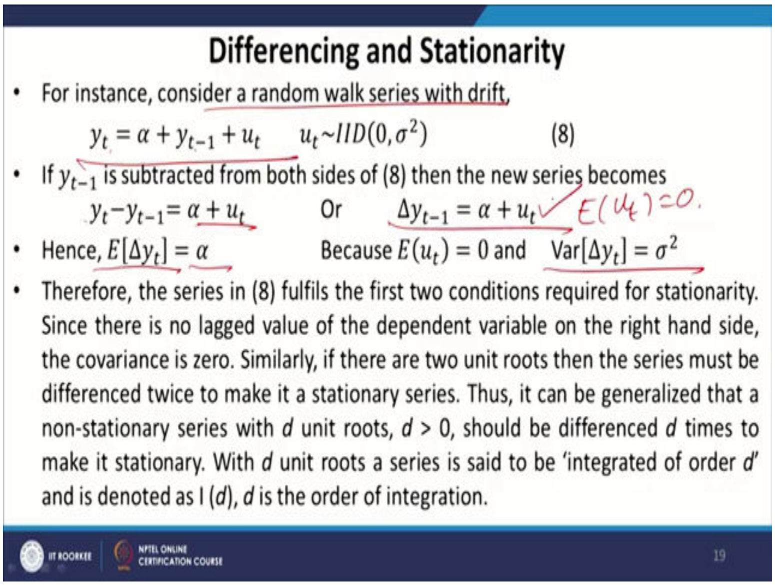
(Refer Slide Time: 31:48)



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.

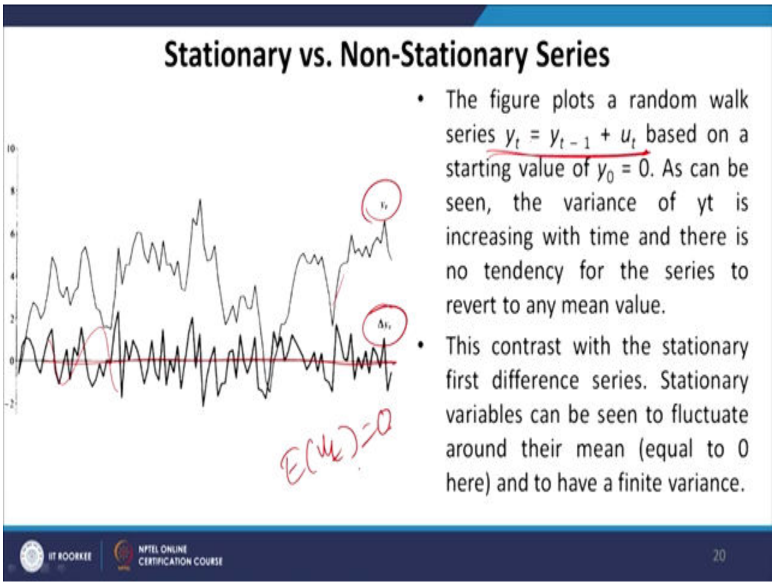
However, before discussing the testing procedure let us first focus on how to convert a non-stationary series into a stationary series. A series with one unit can be transformed into a stationary process by differencing the series once.

(Refer Slide Time: 32:37)



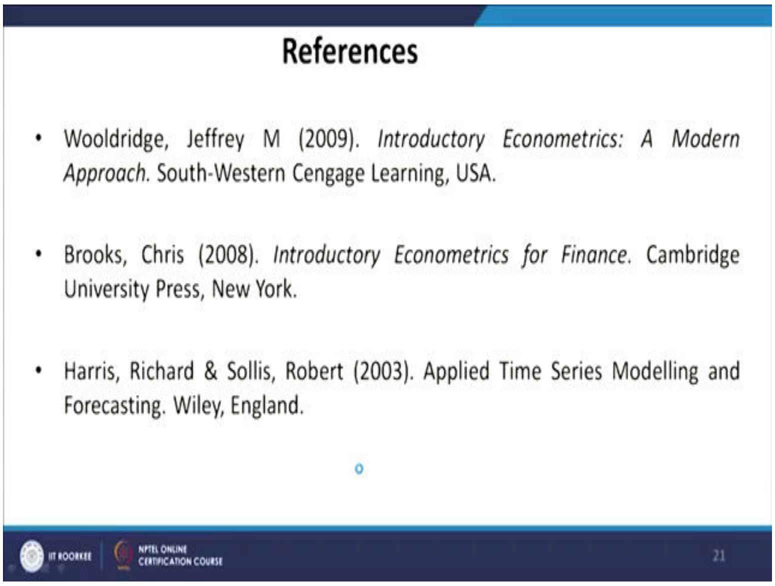
So, for instance, we consider a random walk series with drift. (Refer slide time: 32:42-33:12).Therefore, the series in 8 fulfills the first two conditions required for stationarity. Since there is no lag value of the dependent variable on the right-hand side, the covariance is 0. Similarly, if there 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)



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.