

**Introduction to Probability Theory and Stochastic Processes**  
**Prof. S. Dharmaraja**  
**Department of Mathematics**  
**Indian Institute of Technology, Delhi**

**Lecture – 57**

Now we are moving into the 4th definition that is, Autocorrelation, Autocorrelation function.

(Refer Slide Time: 00:08)

4. Auto correlation function

$$R(s,t) = \frac{E[x(t)x(s)] - E[x(t)]E[x(s)]}{\sqrt{\text{Var}(x(t))} \sqrt{\text{Var}(x(s))}}$$

Assume  $R(s,t)$  depends only on  $|t-s|$   
we have

$$R(\tau) = \frac{E[(x(t)-\mu)(x(t+\tau)-\mu)]}{\sigma^2}$$

$\mu(t) = E[x(t)] = \mu$  ;  $\text{Var}(x(t)) = \sigma^2$

The way we have defined the covariance function, now we are defining the autocorrelation function. It is defined with the notation  $R$  of  $s$  comma  $t$  that is, nothing but or we can write it in the terms of expectation of  $x$  of  $t$   $x$  of  $s$  minus expectation of  $x$  of  $t$  into expectation of  $x$  of  $s$  divided by the square root of variance of  $x$  of  $t$  and the square root of variance of  $x$  of  $s$ .

So, the numerator can be written covariance of  $x$  of  $t$  comma  $s$  divided by square root of variance of  $x$  of  $t$  square root of variance of  $x$  of  $s$ . So, this is going to be used in with the notation  $R$  of  $s$  comma  $t$ , and this is going to be autocorrelation function for the random variable  $x$  of  $t$  and  $x$  of  $s$ . So, it is basically describes the correlation between values of process at the different time points  $s$  and  $t$ .

Sometimes we assume the, we assume  $R$  of  $s$  comma  $t$  depends; only on absolute of  $t$  minus  $s$ . In the later case when you are discussing the stationary process it is going to be

depends only on the interval length not the actual time therefore, the  $R$  of  $s$  comma  $t$  is going to be depend only on the length of the  $t$  minus  $s$  in absolute not the actual  $s$  and  $t$ .

Therefore, by assuming  $R$  of  $s$  comma  $t$  is going to be a only depends on  $t$  minus  $s$ , we can have  $R$  of instead of a 2 variables I can use the one variable as the  $R$  of  $\tau$ ; that is nothing but, the expectation of  $x$  of  $t$  minus  $\mu$  multiplied by  $x$  of  $t$  plus  $\tau$  minus  $\mu$  the expectation of that product divided by  $\sigma$  square.

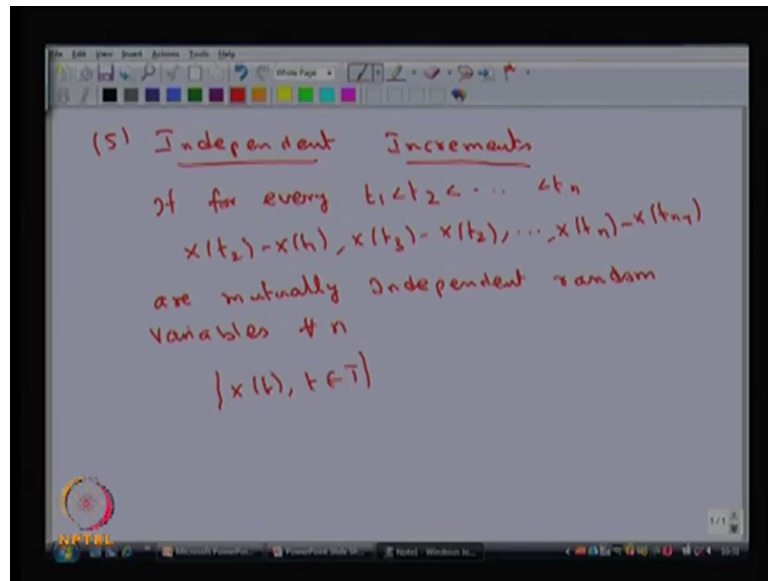
So, here I have made the one more assumption, the  $m$  of  $t$  that is nothing but the expectation of  $x$  of  $t$  that is going to be  $\mu$ . And variance of  $x$  of  $t$  is going to be  $\sigma$  square with that assumption only the  $R$  of  $\tau$  is going to be expectation of this product divided by  $\sigma$  square, where the variance of  $x$  of  $t$  is going to be not a function of  $t$  it is a constant that is  $\sigma$  square. And similarly the mean function expectation of  $x$  of  $t$  is going to be  $\mu$  that is also independent of  $t$ .

Therefore, I can simplify this  $R$  of  $s$  comma  $t$  the product expectation minus individual expectation that can be simplified as expectation of this product and ah. So, basically this is evaluated at  $x$  of  $t$  and  $x$  of  $t$  plus  $\tau$  and that difference is going to be  $\tau$ . And this is also going to be a even function; that means, it has  $R$  of  $\tau$  is same as a  $R$  of a minus  $\tau$ . And this autocorrelation function is used in time series analysis as well as a signal processing.

In the signal processing, we assume that the signal the corresponding a time series satisfying the stationary property therefore, the stationarity property implies the autocorrelation function is going to be depends only on the length of the interval not the actual time. Therefore, these  $R$  of a  $\tau$  will be used in the signal processing as well as in general time series analysis also.

The fifth definition so be we are covering the different definitions which we want.

(Refer Slide Time: 05:10)



The fifth definition, first we started with the main function, second we started with the second order stochastic process, then third we start third we have given the covariance function and the 4th we have given the autocorrelation function.

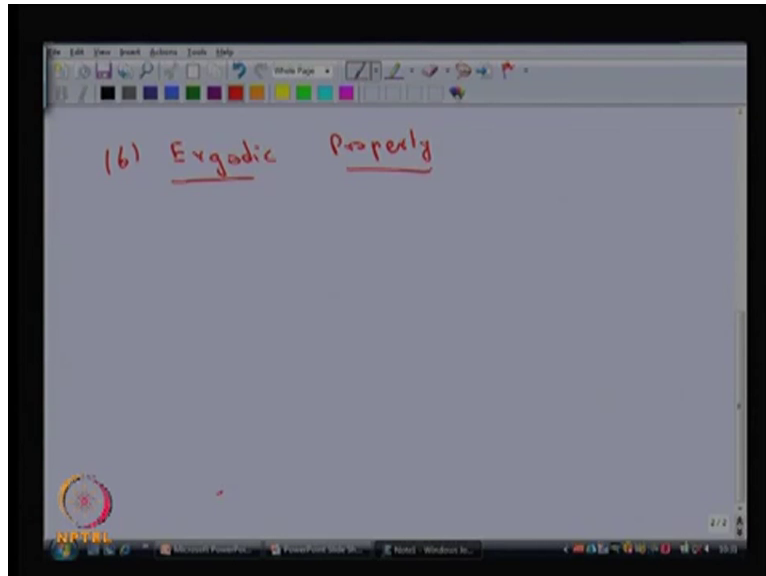
Now, we are giving the fifth definition that is, a independent increments. If for every  $t_1 < t_2 < \dots < t_n$ , the random variables  $x(t_2) - x(t_1)$ ,  $x(t_3) - x(t_2)$ , ...,  $x(t_n) - x(t_{n-1})$  are mutually independent random variables for all  $n$  then we say; the corresponding stochastic process is having independent increment property.

So, whenever you take a few  $t_1, t_2, \dots, t_n$  and the increments that is  $x(t_2) - x(t_1)$  like that till  $x(t_n) - x(t_{n-1})$ . So, these are all going to be the increment and each one is a random variable therefore, the increment is also going to be a random variable. And you have a  $n$  such random variables and suppose these  $n$  random variables are mutually independent random variable for all  $n$ . So, this is mixed for one  $n$ , like that if you go for all  $n$  if this property is satisfied then we can conclude the corresponding stochastic process having the property of independent increments.

So, the independent increment that does not imply, some other properties, but here what we are saying is the increment satisfies the mutually independent property; that means, if you find out the CDF of the joint CDF of this random variable that is same as the product

of the individual CD of it that is property satisfied by all the  $n$  then you can conclude that stochastic process has the independent increment.

(Refer Slide Time: 08:30)

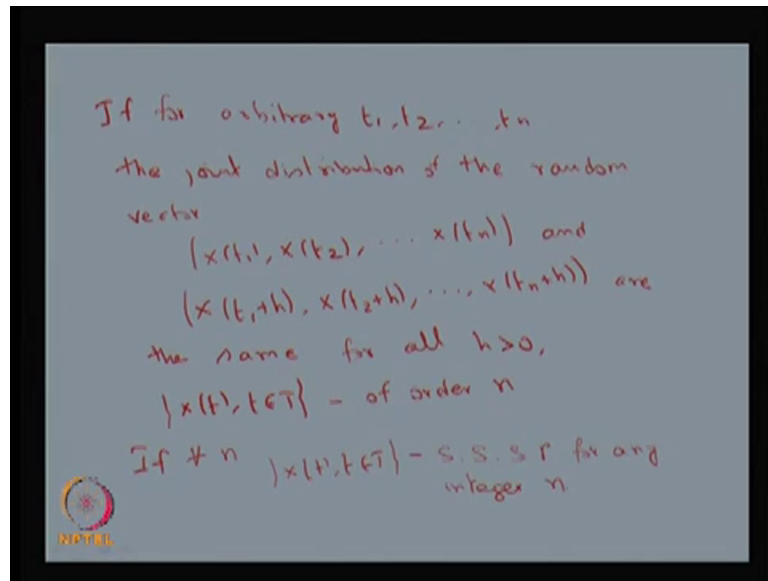


The next property or the next to definition is a Ergodic Property. What is the meaning of Ergodic property ? It says the time average of a function along a realization or sample exists, almost everywhere and is related to the space average. What it means? Whenever the system or the stochastic process is Ergodic the time average is the same for all almost initial points that is, the process evolved for a longer time forgets it is initial state.

So, statistical sampling can be performed at one instant across a group of identical processes or sampled over time on a single process, with no change in the measured result. We will discuss the Ergodic property for the Markov process in detail later, but this Ergodic property is going to be very important, when we when you study the Markov cross property or when you study the stationarity property therefore, these Ergodic properties always goes along with the stationarity property or goes along with the Markov, Markov property therefore, the stochastic property is going to behave in a different way, and that we are going to discuss later.

the most important stationary process that is a Strict sense stationary process.

(Refer Slide Time: 10:15)



First let me start with the strict sense stationary process of order  $n$ , then I will define the strict sense stationary process for all order  $n$  or the exist strict stationary process itself. If for arbitrary  $t_1$  comma  $t_2$  and so on  $t_n$ , the joint distribution of the random vector that is  $x$  of  $t_1$  comma  $x$  of  $t_2$  and so on,  $x$  of  $t_n$ .

And the another random vector that is  $x$  of  $t_1$  plus  $h$  comma  $x$  of  $t_2$  plus  $h$  and so on  $x$  of  $t_n$  plus  $h$  are the same; for all  $h$  which is greater than 0 then we say, the stochastic process is a strict sense stationary of order  $n$  because here we restricted with the  $n$  random variable.

So, we take  $n$  random variable taken at the points  $t_1$   $t_2$  and  $t_n$  and a find out the joint distribution of  $x$  of  $t_1$   $x$  of  $t_2$  and  $x$  of  $t_n$ . So, you can find out what is a joint distribution of this  $n$  random variable. Also you find the joint distribution of  $n$  random variable shifted by  $h$ ; that means, earlier the random variable  $x$  of  $t_1$ , now you have a random variable  $x$  of  $t_1$  plus  $h$  with the same shift  $h$  you do it with the  $t_2$  therefore, the random variable  $x$  of  $t_2$  plus  $h$ .

Similarly, the  $n$ th random variable is  $x$  of  $t_n$  earlier, now you have a random variable  $x$  of  $t_n$  plus  $h$ . So, you have a another random vector with  $n$  random variables and find out the joint distribution of that. If the joint distribution of this first  $n$  random variable as well as the joint distribution of the shifted by  $h$  that random variable. If both the distributions are same; that means, they are identically distributed the joint distributions are going to

be identical, then you can conclude this stochastic process is a strict sense stochastic process of order  $n$  because you use the  $n$  random variable.

If this is going to be satisfied, the other property is going to be satisfied for all  $n$ ; then you can conclude the stochastic process is going to be a strict sense stationary process for any integer  $n$  this is going to be a strict sense a stationary process for any integer  $n$ . So, we start to cross checking the joint distribution of  $n$  random variable. So, if it is satisfying by only with the maximum sum integer, then it is going to be a strict sense stationary process of order that  $n$  if it is going to be satisfied for all  $n$  then for any integer  $n$ , then it is going to be call it as a just strict sense stationary process.