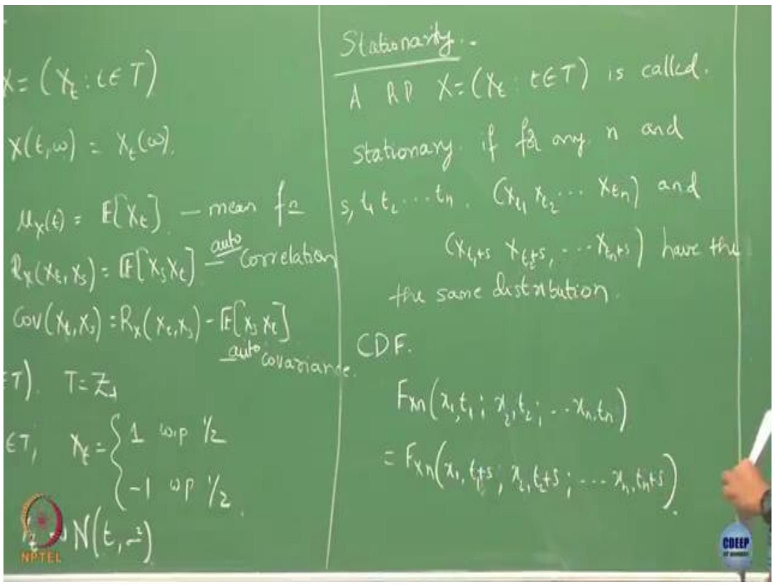
Introduction to Stochastic Process
Professor Manjesh Hanawai
Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay
Lecture 17
Properties of Random Process

(Refer Slide Time: 00:40)



So today we are going to study some more properties of Random Process, so just. So index could be time right, so if index is time then the way we are interpreting random variable x and then we said my random process x can get out of T omega, what is this? This is x of t omega, so if you are going to treat this random process x as I said, it can be thought of function of two arguments. One argument is the index and other is the sample point.

The index could be time itself, in which case it is going to talk about that Tth random variable and the value it takes at sample point omega. The random process we also defined Mu x of t this is and so this is for all. So recall that this T could be just integers in which case it is unaccountably sorry it is finite sorry infinite but still countable or it could be an interval itself in which it is uncountable, okay.

So and we distinguish these two cases, discrete random process and continuous random process based on whether my number of elements, the number of indices indeed is countable or uncountable. And then we have this question of, (one more question) yeah.

Student: If this T an interval,

Professor: Again there are two discreteness that has coming into picture here, right. One is the in terms of the indices whether indices are taken discrete value or they are taking continuous values. Based on that we are classifying our random process as either discrete or random process. But each of the random variable in this process itself they could themselves be discrete or continuous further. So for example, so let us take an example, let me just first complete this, this is covariance is.

So let me take x where my T is equal to Z that is all positive Z plus let us say all 1, 2, 3 up to infinity and my for T belonging to T let us say x of t is equal to 1 with probability half and let us say minus 1 with probability half. So this random process here, here by our definition it is discrete random process and further each of the random variables here themselves are discrete, okay.

But now we can say that my X t is, let us say Gaussian distributed with some mean t which depends on the index let us say and also variance, some variance. Here for each index my distribution is what, is taking continuous value outcome. So here even though it is a discrete discrete random process but each of my random variable is continuous here. Because I know already my Gaussian random variable is continuous, right?

So we said that this is a mean function, this is what? Correlation and this is called covariance. So yesterday in my class I think I made one small error inside that one these two random variables are the same we said it is going to be called as further auto covariance and auto-correlation but that is not the correct, the thing is we are talking about one random process here right, if you are just talking about one random process then in general we are going to call instead of correlation we will also call it as auto-correlation and auto-covariance.

So if we are talking about only one random process and they are defined for this particular random process. So this is auto-correlation, the auto is coming here because we are talking about the same random process, in that random process we are looking at a correlation between two random variable. So that is we can sometimes it is also called auto-correlation. So we will see another notion called cross-correlation and cross-covariance.

When we have two different random processes, so that we will come to it later. Now for this random process I want to define one important notion called stationary, a random process. So what I was saying is take any random process T and you take any n random variables, so take n, n is telling okay, it is telling what is the number of random variables you are interested and then take any t1, t2, t n this is going to tell you the indices which are those random variable interest you are interested.

For any such n and for such any vector if you are going to look at the joint distribution of these n random variables, taken a t1, t2 up to t n and look at the joint distribution of the shifted versions that is t1 plus s, t2 plus s and t n plus s, if they have the same distribution then we are going to call a random process stationary. So what it means in a way is like what we are doing is, this set of random variables joint distribution, now we are looking at another set of random variables delayed by time S or shifted by time S.2

Then if this again have the same distribution then we are going to call our random process stationary. That means if my random process is shift invariant whatever let us say I am looking at let us say I have a process over at the each random variable correspond to one particular day, if you are going to look at the joint distribution of let us say day 2, day 5 and day 7 this join distribution of these 3 and then you look at the joint distribution day 3, day 6 and day 8 that is everybody just shifted by the joint distribution of first 3 will have the same joint distribution of the last 3.

In that case we are going to call it as stationary, okay. So what is this basically means is in terms of CDF functions, this is going to be same as, so if you just shift the time indices for all them by the same amount then the joint distribution does not change. Okay, let us now understand what is the meaning of this stationarity? What is the thing that this stationarity implies.

(Refer Slide Time: 09:47)

Okay now let us take us second order, so what we mean by a second order random process? So we say that its second moment it is going to be finite for all possible indices, right? If you are going take a second order random process, (okay)now let us apply this definition to n equals to 1, so remember this, this is for any n it should happen, okay.

So let us take this n equals to 1 and then apply this definition then what it say is, f of n equals to x of let us say x1 let us say x1 and then x some time t1 this should be same as this n is equals to 1 here, this should be same as f of x of n and then x1 that is t1 plus s and this should be true for all S, right.

Because this should be also true for all S, what is this saying is you take the random variable at time t1 or you take any shift at any other index, so if I have a t1 choosing these different-different S I should be able to get all other possible indices also, right? So what is this means then basically, the CDF of all the random variables are the same, right? If this stationarity definition has to hold for every possible n and for every possible m, so this has to hold.

So for every random variable has the same CDF, is this clear? Why this should be true? So if every random variable has the same CDF then what does that mean? In terms of the mean will they have the same mean all of the x ts? So then Mu of a process x t will be like something like Mu x only right, it does not depend on t. what was t here, the indices but what we now said is, the CDF is going to be the same at any index we are going to look at, that is the meaning of this condition.

Because of that the mean at any of the random variable at all the indices is going to be the same, okay. So mean is constant, so remember that the mean of a general random process need not be constant, it is a function of indices, given index. But if its stationary process it does not depend on which index you are looking at, it is going to be the same for everybody. Then, now let us look at the covariance.

Okay, before that, now let us apply this the same thing for n equals to two case, what that means is now f of x2 x1 t1, x2 t2 is equal to f of x2 x1 t1 plus s and x2 t2 plus s right. Now the claim is if you are going to look at any two pair of random variables their joint distribution is going to be the same for any pair of random variable, is that true from this condition? Shifted by the same amount, shifted by the same amount, so pick random variable at t1 and t2, look at the CDF, now what we are saying this.

You shift both t1 and t2 by the same amount S then the joint distribution remains the same, right? So that is what this condition told us, so because of that if you take the two random variables and shift by them same amount the distribution is going to the same, why? Because of this what does it say about my correlation, are they going to be the same? If I am going to shift the random variables by the same amount then the expected value should also be the same right?

Why is that? Because these two have the same joint CDF by definition, the CDFs are going to be the same then the expectation is also going to be the same, right? So then what it means? If I am going to look at these random variable joint distribution and the shifted version of them it does not depend on what is the amount of the shift but it makes still look at what is the two pair of random variables you would be looking at.

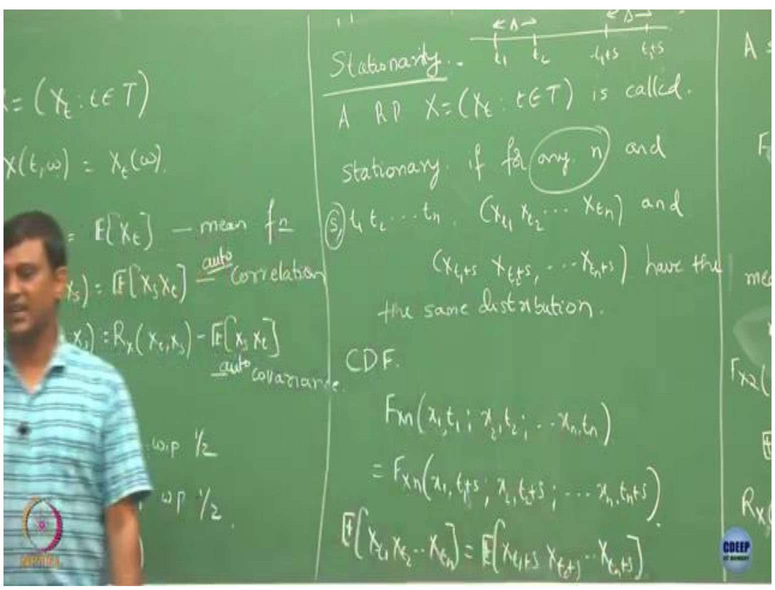
Suppose you change t1, t2 to some other value that correlation may be different but if you what it is juts saying is you are looking at this t1, t2 and other two random variable which are just shifted versions of these 2 then a correlation is going to remain same. So suppose you say this is t1 and t2 you shift both of them by the same amount okay, so whatever this delta time difference is there, here also the same delta time difference is there, right.

It is just like both of them have been shifted by some amount S, but what we are just saying is this shift does not matter then what is (mat) then what should be the only thing that should be governing this covariance? The only thing that should be governing this covariance is this length of this interval, right. So add x1 at t1 and t2 should be simply I should be able to write t1 and t2.Or if let us say t2 is going to be, we will just follow this combination.

It only depends on the length of that interval rather than by what amount these random variables are shifted. So because of that here also the shift is the difference in this interval in this time t1 plus s and t2 pus s is again going to be t2 minus t1 and that is also the same here. So this correlation, auto-correlation function is only functions of the length of the interval of these of the at a point where this indices are taken.

Okay, fine, so if you go on I just did it for n is equal to 1 and n equals to 2 right, you could go on doing this for any number of n, right. You can go and do it for n equals to 3.

(Refer Slide Time: 18:33)



That means if we take 3 random variables and shift all of them by a same amount then the it should be the case that x of t1, x of t2 up to x of t n should be same as expectation of x of t1 plus s, x of t2 plus s up tp x of t n plus s. So stationarity is a very-very strong property, right?

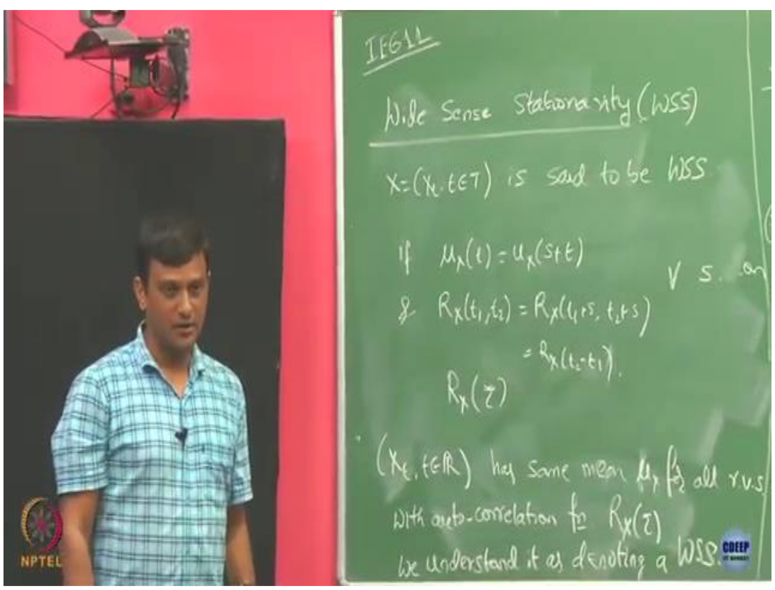
It makes the process kind of shift invariant and the again the joint statistics of any order is going to be independent of the shift. So we are going to say the means of the values in only single random variable as the first order statistics .Like for example mean we have already define, so when we are going to look at 2 random variables, let us say x t1 and x t2, t1, t2 are two random time indices we are going to call this as a second order statistics.

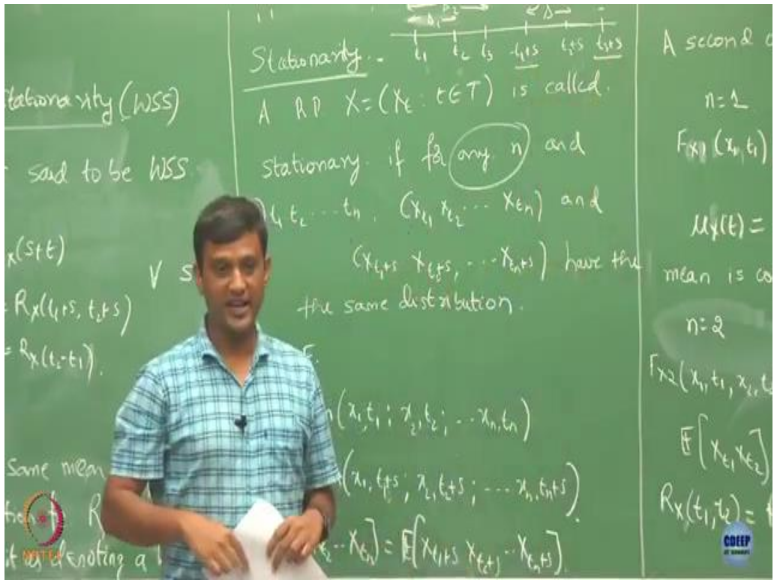
So here, it is basically saying this is the m power statistics, right, m power statistics here is shift invariant. So stationarity is basically saying that my m power statistics are shift invariant and it should be true for n equals to 1, 2 all the way up to infinity. So this is the much stronger property, so what.

Student: Sir, what is the significance of the second order process first, why did you mention it specifically here?

Professor: The second order random process, okay, so what is the second order statistics we said that expectation of t is going to be finite for all t, right. Here you could as well choose x t1 equals to t2 in this case you are following that case and we need it to be finite, okay. So as you see the stationarity is a much-much stronger property which requires all other statistics to be kind of shift invariant, instead of that one can look at a slightly weaker version of the stationarity called as Wide Sense Stationarity.

(Refer Slide Time: 21:03)





What wide stand stationarity asks is only shift invariance in the first and second order statistics, it do not care about higher order statistics, okay. That is it wants that this means to be shift invariants, Mu at index t is going to be the same as Mu at index x plus t that basically mean that my Mu x is constant right, this is going to be the same for all possible indices and then it says that the correlation is again shift invariant. So that means as we already discussed this only depends on the length of this interval.

So as long as you take any 2 random variables that has same, that has separated by the same amount then the Cos relation is going to be the same, okay. So basically wide sense stationarity is only restricting this condition to be (equal) hold only for n is equal to 1 and n equals to 2 whereas stationary wanted it to hold for all n, all possible values of n, okay, fine. So because of this property that my upper correlation here only depends on the interval or the two difference of the indices rather than the actual values of the indices itself.

Affirmative is given in terms of first single random variable, after that you can write it as function which takes only one argument. So here it is given, it is taking two argument, right, but two argument is essentially translating to only difference in these two argument that means I can as well think of it is in function of a single random variable, sorry single argument, okay. So if I define a process, if I say that I have a random process x with mean constant as Mu and say that its auto-correlation function is r of x.

Now here tau is just a single variable then it should be by default will take it as, what? White sense stationary random processes, okay. So let me just make this more clear. So we will just say sometimes (2) 1 if you say that x of t, t equals to R has same mean Ux for all random variables with auto-correlation. So then we already take it as like we understand it, as denoting a white sense stationary process, okay.

So note that even for the stationary case my R function here auto-correlation function is again only depends on the single random variable right, sorry single argument here because it just depends on but it also for that I also need to unsure that for n greater than 2 also all this condition holds but if I just take only path of the mean and upper correlation function and further I say that the mean is a constant or the same for all random variables and that its auto-correlation function is just a function of a single variable then we understand that this is already indicating a white sense stationary, this is just a convention, okay.

Student: (())(26:15)

Professor: No, it does not matter which t1 and t2 you are talking about right, so for reference as I said here you have t1 and t2 here and you have t1 plus t2 S here if you look at the correlation of these two random variables they are going to be the same, it just depend on this interval, yes it is going to change. So suppose now let us say you have another t3 here and now you are interested in correlation between, so this was let us say call delta and let us call this as delta 2.

We are going to look at the upper correlation between t1 and t3, it is going to be function of delta 2 now, right? So and in similarly so here also if now let see t3 s and the auto-correlation of t1 and some t3 as it is going to be the same as R of delta 2 because what matters is only their separation, where they are with the same separation where they are is does not matter.

Student: (())(27:35)

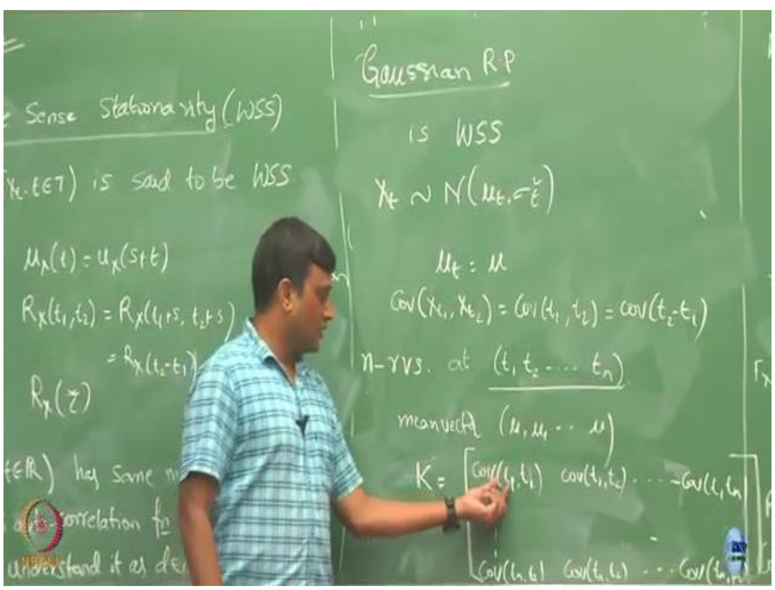
Professor: No, it is not, so okay let me see, if you want to calculate auto-correlation between two random variables, what you are going to tell me is, their indices, right? All I need to know is the difference in the indices then I already have auto-correlation value for your two random variables. For example as I said if you want to compute auto-correlation function of at index 5 and 10 okay, it needed all I need to tell you that is the auto-correlation function value at argument 5.

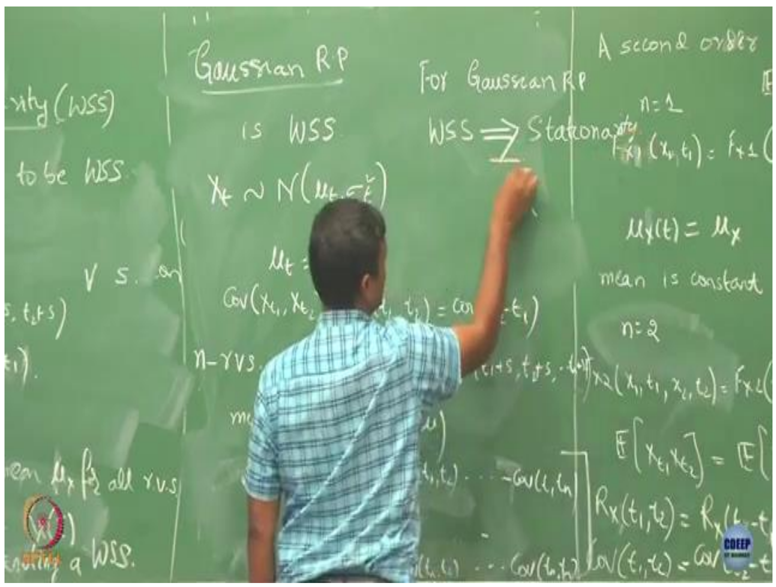
So and if you have another set of random variable, to pair a finite number 1 at time index 4 and another at time index 10, what is the value of the auto-correlation function I need to give to which argument? 6, right? That is the only thing matters and I am going to give you this value for all possible values okay. So this is for all possible values of indices, so I just need to that is why this is function of single random variable.

One more thing I want to mention about this is, in the last class, so suppose if I say the process is stationary does it means already means it is white sense stationary? It is true, right? Because white sense stationary needs only small area, the requirement it only needs this CDF to be invariant shift only for n equals 1 and 2 but whereas stationarity needs the all them.

Now is the upper direction is also true? Is white sense stationary imply stationarity? Not necessarily, right? Because stationary is a much stronger (condi) but as now we will see that if I have a Gaussian random process even white sense stationarity imply stationarity, okay let us see why is that?

(Refer Slide Time: 30:04)





So do you recall what I mean by a Gaussian random process? So Gaussian random process is if a random process is such that it all its random variables are jointly Gaussian then we are going to call that as a Gaussian random process, right.

Okay now let us see, suppose let say my I have a Gaussian close random process which is white sense stationary. Okay let us say it is white sense stationary, okay. Now I know that if it is white sense stationary already, so what are the other things now? If it is a Gaussian random process, it will necessary that each of the random variable at any of the these indices is again Gaussian distributed, right.

Now and now I further assuming it is white sense stationary that means each of this random Gaussian random variable should have a same mean right, so we know that x of t for all is Gaussian distributed, okay and now we know that all this Mu t has to be the same Mu for some Mu that is nothing that is coming from your white sense stationary, right because this Mu t is independent of Mu and now further.

So another thing here you look at this correlation, what about the covariance? Its covariance is again a function of the difference in t2 minus t1 only, right. So I can also write covariance to be t2 minus t1, so I just verify that. Now what about the covariance between two random variable x t1 and x t2? We have already said that this is nothing but covariance of 2 by t2 by definition and this is going to be covariance of t2 minus t1.

What is that? So by white sense stationary is a definition we already got the means are going to be the same for all this Gaussian random variable and this covariance is going to be just depends on the difference in the length of their interval. Now let us look at any n set, n random variables. Now I am interested in, so to now I want to go from n equals to 1, n equals to 2 to any n and let us say n random variables at t1, t2 and t n.

Okay, and this t 1 to t n random variable now I want to look at their distribution okay and the distribution is what? We already know what is this distribution right. The distribution was defined in terms of the mean vector and the covariance matrix. So what is the mean vector here? The mean vector is simply, mean vector is simply Mu, Mu, Mu right everywhere because each of these random variable has the same Mu.

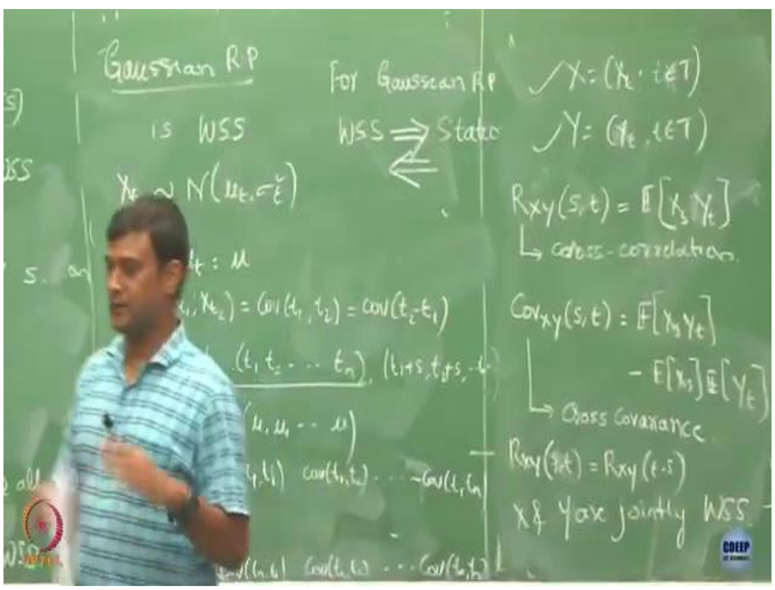
Now what is the covariance matrix for this? The covariance matrix for this is covariance for t1, t2 covariance of t1 (sorry) t1, t2 and covariance of t1 and t2 and then covariance of t n covariance of, right, the joint distribution of this n random variable depends on the mean vector and the covariance matrix, right? Now you should look at each of these term in this covariance matrix, this covariance each of this term only (different) depends on the difference of these two terms, right, nothing else, okay.

So now if I am going to and what we know? We know how to define the join distribution of the this set of random variables, right, in terms of the Mu and k we also did it in the last class. Now suppose if I shift each of these point indices by some amount W upper is not going to change for the random (vector) variable, so this time induces, right because it is going to be the same respective. Now what about the covariance matrix? Does it depend on the shift S?

No, right, because the difference if you just take t1 plus s and t2 plus s the difference is only t2 minus t1 now, the s has no role in this case. So as you see that the joint distribution is now independent of the shift we are going to look at right and that is what and this is true for any n. So then what does this mean? The our requirement of stationarity is satisfied? Right, the PDF is going to be the same, the PDF is in this case is anywhere n that means also like CDF can (argu) that CFD is also invariant, right also the shifts then my stationarity property satisfied.

So for a Gaussian and a process wide sense stationarity implies stationarity and anyway this other ways always true by default, right but if you have a Gaussian random process, the weaker notion of white sense stationarity already implies the stationarity, okay. Okay, so just one last thing about this is, so far we have been talking about one random process right mean so it may happen that you have to deal with multiple random process.

(Refer Slide Time: 38:10)



So let us say, this is my one random process and this is my another random process where all the random variables are defined on the same probability space. Right, so when we say the random process, we mean that all these x ts are defined on the same probability space. So now let take another random variable, another random process where each of the slight is are defined on the still the same probability space.

Now when we have such things may be you may still want to define what is a correlation, similarly and covariance in this case, right? So in this case we are going to define R x y at point s of t to be expectation of the first random variable, computed at time index S from the x process and the second random variable is the coming from the second random process at time index t and this is called cross correlation.

And similarly you can define cross covariance okay this is and one last thing is x t is only (x y) then we call x and y are jointly wide sense stationary, okay. So this is just like extension of notion from one random process to multiple random process when we have to deal with. So it may so happen that for example let us say you are dealing with the stock exchanges at every day, the may be Bombay Stock Exchange how it is varying, everyday and so you are going to model this as a one random process.

And may be you feel that whether as implication on the Bombay Stock Exchange, so then you may want to model another process which is like a weather and each of these days as another random process, okay. So then you may want to see like how they are correlated, covariance between them? So each of this process could be corresponding to some aspects of error, something that you want to model like a outcome of an experiment which has, which is like a many-many indices.

So that one experiment and if another experiment similarly which has many-many indices if you want to like understand how one has influence over the other or what is their like correlation or covariance you want to look for all such properties.