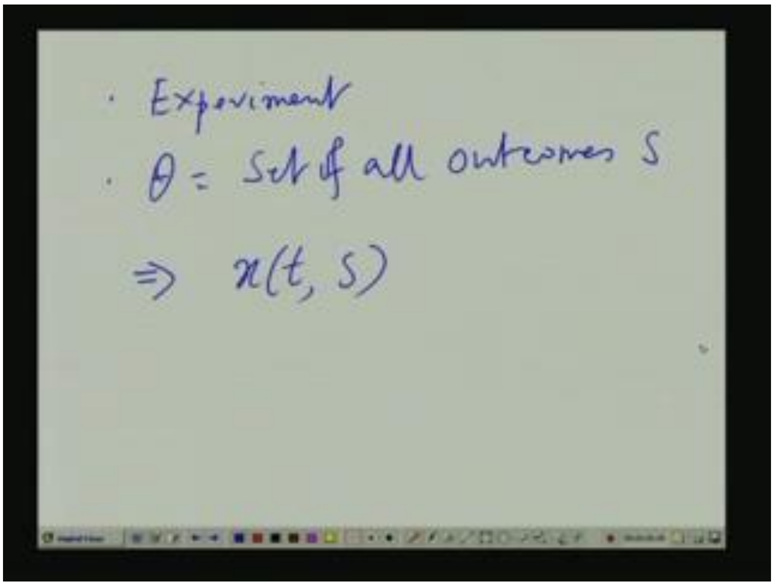
Probability and Random Variables
Prof. M. Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 29
Introduction to Stochastic Process

So, today we begin a new topic all together. You know so far we have been discussing random variables. So, one random variable, then two random variables and multiple random variables and all that and of course, their probability density functions, probability distribution functions and various relations and properties associated with them. Now, we use those concepts to explain or to interpret various natural phenomena, and that takes us to the study of a random process or which is also called stochastic process.

What is a stochastic process? Essentially you remember how we define random variables. I mean there was some experiment. You know it is not that. It could be a real experiment, but is just kind of just assume that. I mean there is an experiment and there are various outcomes coming out depending on the trial and with each e outcome, you assign a value to a variable. So, that variable then is called a random variable. Similarly, suppose we have got some experiment that is going on and every trial there is some outcome.

(Refer Slide Time: 02:15)



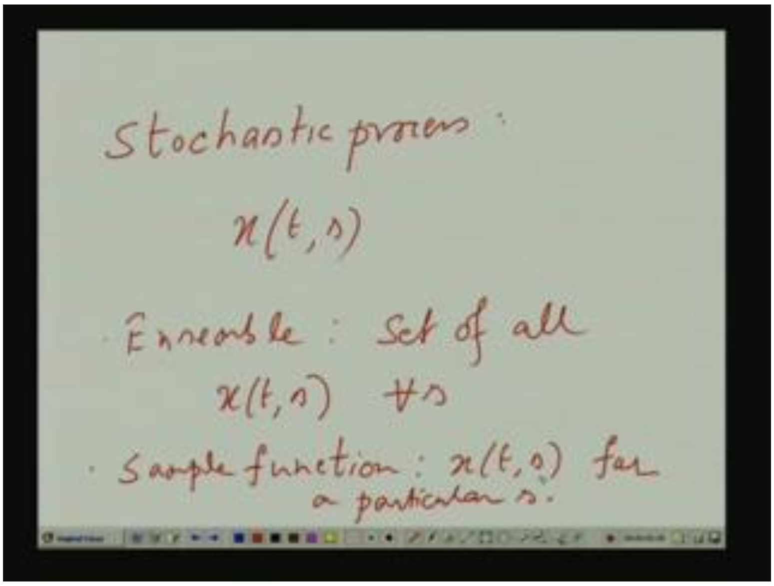
With each outcome there is first experiment that may be say theta is a set of all outcomes s. S is a typical outcome. So, wherever a particular outcome, an experimental I mean when a particular trial takes place, you get some outcome s and depending on what you get is, you assign a value to the variable x and that variable is called a random variable x, but instead of assigning a value to the random variable, suppose depending on the outcome s we generate a function x of t. A particular function that is we assign of assigning value to a variant to this outcome s, we assign a function x of t to s.

So, every time you get a new outcome, you get new function. So, that means, as experiment goes on, you get some. I mean every outcome you get some new function or some other function, another function. So, these are kind of not random variables, but random functions. The set of these random functions actually the set of all random functions are possible here. They constitute one stochastic process here. I will give examples to elaborate this. So, that means, you generate a function x as a function of some t variable t, and also it depends on the outcomes s. So, for a particular outcome, you have one particular function. For another outcome s, you have another function.

So, depending on the experimental result, you get a waveform. So, there is no fixed waveform. So, waveforms vary. For example, suppose there is a microphone and I say IIT. This is recorded and then displays it on an oscilloscope; you get some kind of waveform. So, every time I say it, you get some waveform, but using your intuition, you can easily understand that even if the same speaker goes on saying it again, every time you will not get the same waveform, because a speaker is speaking, you know I mean after all he is human being and not robot.

So, things are changing you know I mean his throat in his voice generation can. Everything is changing a little bit on each trial. As a result that waveform which represents the sound, that also varies. So, that means with each trial, with each outcome where outcome is a particular word, the way it is pronounced that is the way it is pronounced. I mean that particular wave is one outcome. So, with that a waveform is associated. Next time I say the same thing, but things differ a little bit because as I told I am a human being. I am not a robot. I get another outcome and the waveform changes. The set of all such possible waveforms that corresponds to all possible outcomes actually that together is called ensemble. I will come to that. This process of assigning random functions to outcomes is called stochastic process.

(Refer Slide Time: 05:58)

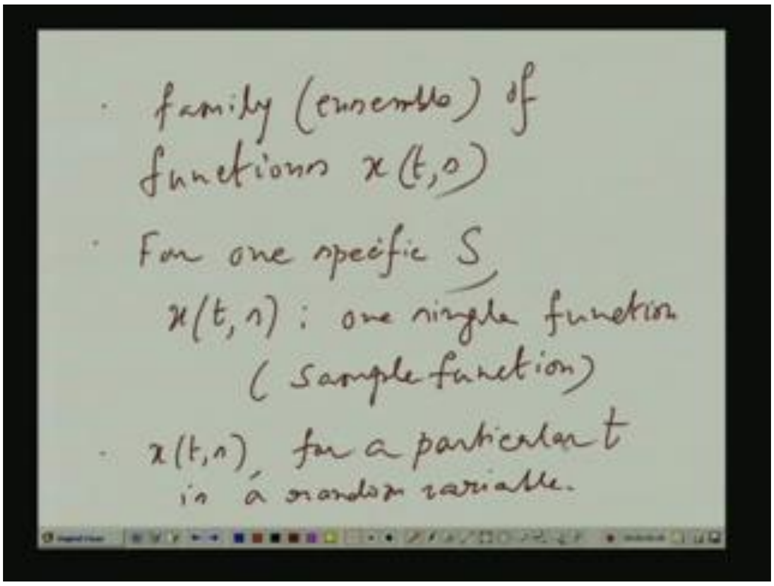


So, the family of all these functions belonging to all possible outcomes, they constitute the random process. So, I will just say xt, s. Let me make it clear. Now, if I just say xt, s is a random process that means these things and there is some experiment and every time the experiment is conducted, there a trial is made or generated. That waveform is some function xt, s, but stochastic process means the random process means every time you know perform the experiment, you get a new outcome. The waveform changes as it is not a fixed waveform in general. In symbol this is another definition for all s. That is you consider the set of all possible waveforms, all possible means depending on all outcomes. So, it is for one outcome one waveform for another outcome another outcome and another waveform. So, I am collecting all possible outcomes, and therefore correcting all the waveforms.

So, the set of all possible such waveforms for this experiment, they form one set and that set is called ensemble for this stochastic process. Stochastic process is this phenomenon say it might appear to be a little vague. Stochastic process, the definition actually is this phenomenon. What is the phenomenon? It is this phenomenon of three random functions depending on the experimental outcome. For a particular experiment, there are several outcomes possible. Every time you conduct the experiment given one outcome or another outcome, and then you get one waveform or another waveform or another waveform. This phenomenon is called the stochastic process.

A particular waveform, suppose s is a particular outcome and for that outcome, you have got a particular waveform xt, s. I will call that particular waveform that specific waveform for a particular s and sample function. So, ensemble consists of all possible sample functions. So, sample function is a particular waveform for a particular experiment outcome.

(Refer Slide Time: 09:47)

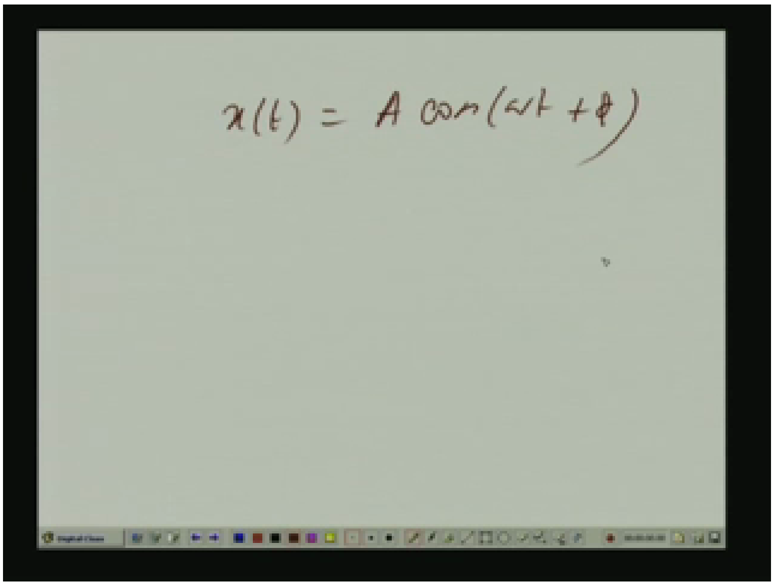


So, just to sum up we have these points. What is the stochastic process? Stochastic process is a family which is also called ensemble of functions x t, s where s stands for experimental outcome, and this familiar ensemble of all these functions actually ensemble or set out. They consider the stochastic process. Then, for one specific s, sorry one single function, obviously for particular outcome you get only one function or specific function called sample function. Then on the other hand xt, s for a particular t is a random variable. It is very important.

Why simply you are fixed where the point t. So, at that t for one particular experimental outcome is x. You have one value of this function. Next time the outcome changes; you get another value of the function. So, this then corresponds to a random variable only. It takes various values depending on the particular outcome at hand. So, these are random variable, right. If t and s both are fixed, then it is just a number. If we are fixing t also one particular time and one particular experimental outcome only, then there is nothing random about it. It is just a number. There are some processes which are called regular processes and some processes are called predictable processes. I will just touch upon this now.

Suppose we consider the Brownian motion that there is a fluid and some particles are moving and random, and they are colliding with the molecules and again changing the direction randomly. So, here if I try to describe the motion of a particular particle, obviously this random, it is every time changing its direction and also speed. It is random, but it is not given by some fixed set of parameters and all that. Also, it is not predictable like in the sense that if you know its past trajectories, you cannot say what its future trajectories would be.

(Refer Slide Time: 13:30)



So, this is actually a pure regular unpredictable process, but if I give you a signal, process like this. Xt is say A cosine omega t plus phi. I have omitted a variable s here. I mean it is quite implicit here. This is A and phi. There is amplitude and phase. Suppose they are random variables. That means, every time you conduct the experiment here, actually there is no particular experiment. Every time suppose you receive the waveform.

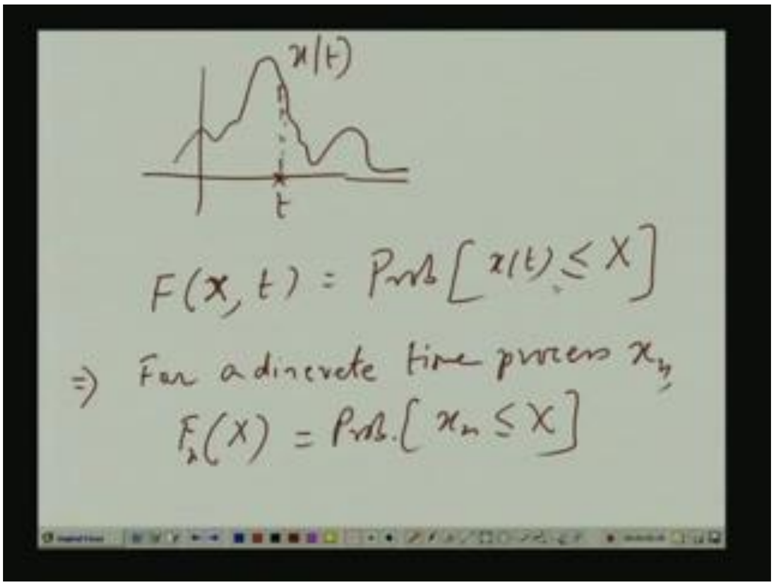
Receiving a waveform is the experiment, is the trial of the experiment. May be you get particular A or every time it generates a waveform may be by an oscillator. You get particular A and particular phi and therefore, particular xt.

So, for a particular A and phi, there is for a particular experimental outcome where for a particular experiment xt is fixed, you get one waveform. Next time A and phi changes, you get another waveform, but nature of waveform is you know there is something interesting here that the nature of waveform still remains sinusoidal. In this case, you know this waveform. I mean I will not call it regular because this random process is actually generated. It is just completely described by few parameters, just few random parameters A and phi omega is fixed.

So, if A and phi are given to you, you know what it is and then, it is also predictable for its past in the sense that suppose you know x of t, I mean for one period fully or may be x of t up to some point of times t naught, then since it is sinusoidal and therefore, periodic, it is easy to see that you can easily predict all the future values that is for all t greater than t naught also. The waveform is known to me without any error that I know because of the periodicity. So, this process is called predictable process. So, we will talk about this more in detail. You know predictable processes and regular processes all that later. Then, we carry forward our previous portion of you know a probability distribution functions and probability density functions to stochastic processes. Earlier we defined these things for random variables.

Now, obvious the question comes that when you have no particular random variable, but a function, then how will you apply the theory at the notion of probability density function and probability distribution function and all that. Here I would say that suppose I am getting waveforms like this.

(Refer Slide Time: 16:07)



You know a particular waveform and I look at a particular time t. Then, obviously the value that we find these waveform xt, I am not writing xt, s, but it is implied x of t. There is xt, s. S is for the particular experimental outcome. So, for a particular sample function i, I get at the specific t. I get this much value for the sample. Next time with another experimental outcome the waveform changes. These value also changes and likewise. So, as I said earlier that for if t is fixed, then x of t, s or simply x of t if I just ignore s for time being is a random variable. So, that random variable I can obviously define what is known as probability distribution function, probability density function.

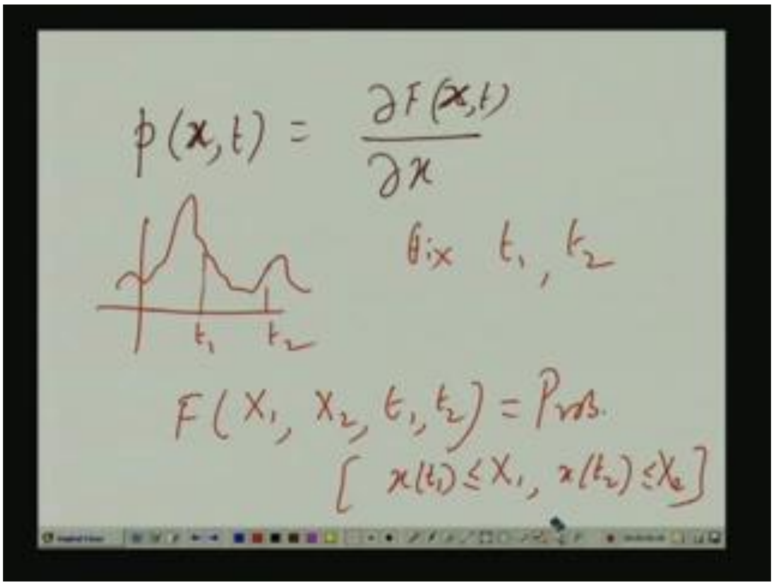
So, that means, we can define simply like this capital X and for a particular t x is nothing but probability of this random sample xt taking value less than equal to capital X. So, x of t is a random variable as long as t is fixed and then, probability of its taking values less than equal to capital X is indeed the probability distribution function. So, it is the function of the capital X you have chosen, but you have also put t as another variable because obviously this depends on the t you have chosen. So, it is very much same as before, but only thing is I mean it is depended on two factors. Now, one is of course the capital X that you choose. Another is of course the time variable because this probability distribution function is now varying about if you call t as time is varying about t, varying about time and various points of time.

Of course, I forgot to mention that you know the waveforms that you are seeing here, these are actually continuous random variables or continuous random process and not continuous random variable continuous random process. In the sense that the variable t is chosen from the entire real axis which is continuous, I mean which corresponds to mean continuous t. You see t is taken from the entire real axis and obviously, t is a continuous variable and x of t is called a continuous random process. On the other hand, if t is taken to be integers on this axis say t equal to 0 or t equal to 1 or t equal to 2 and like that, but only those integers specific integers. Then, it is a discrete time random process also called random sequence or in our examples, we are considering continuous time random processes, but you know these same notions can be carried forward to discrete type random process also very easily.

For instance, if you want to find out this probability distribution function for a random sequence or for a discrete random process, then we will not have t because t is not taking any value. We will have rather n because t can be an integer n. T is integer. So, t can be 0 1 2 dot dot dot to n. So, for a particular integer point n, we will have F within bracket x, n or maybe you can change your style. You can put n in the subscript Fn within bracket x is nothing but probability of xn, because then I will not be denoted at xt. I will be calling at xn, so probability of xn for a fixed n probability of xn less than equal to capital X.

So, I write it actually. This shows that you know this concept can be generalized to or extended due to the discrete end phase. So, for a discrete time process say xn, it is an integer. So, I can have at n equal to 0, some value x zero n equal to 1 and x 1 like that. So, depending on a particular experimental outcome, we get one sequence. Next time the outcome changes, we get another sequence. I say sequence and not function. Earlier I had function. Now, it is just the sequence, because it is discrete time, but in a same manner. We can define its probability density probability distribution function, that is F capital X of course and you can put n here or maybe you can put a subscript n. Either way it so dependent on n. This is nothing but probability of that particular sample xn in a sample xn less than equal to capital X. It is just analogous to the continuous case. So, if you have got probability distribution function, obviously you can have the probability density function also.

(Refer Slide Time: 20:50)



So, p again I am back to continuous time random process. So, p xt is nothing but derivative of since I am deriving, I can put small x again back here. Not a specific capital X, but again there is general variable x and differentiated with respect to n. In fact, I would rather for correctness, I would put partial derivative because use a function of two variables, these are probability density function. The meaning is same as before that is if you want to find out the net probability of the random variable x lying within a string between x and x plus dx at the same time t of course, then that probability will be this density function times dx same notion, but every where they are dependents on t because I have targeted the random variable located at time point t.

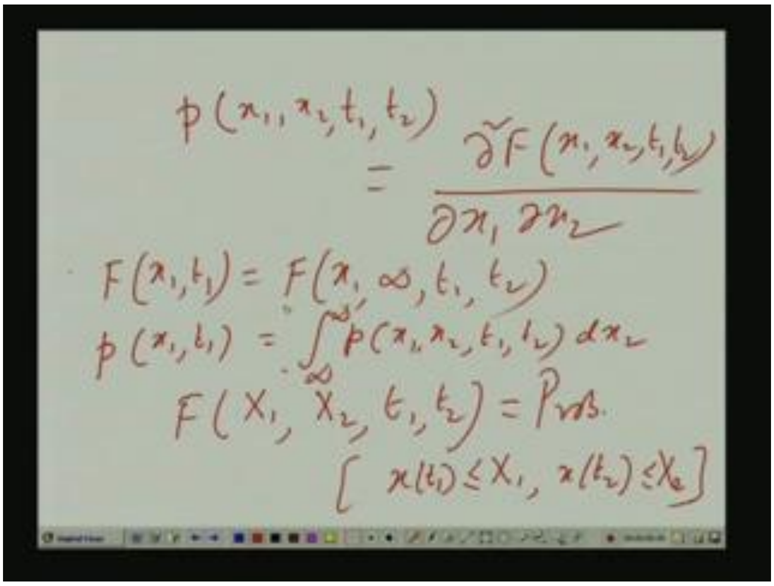
That is the particular random variable actually, random process, any continuous value random process, any continuous value random process you can view also as that uncountable infinity count. Uncountable infinity of random I mean random variables. Then, let me explain actually what is countable and what is uncountable. Of course, if you got a finite set of numbers, you know 1, 2, 3 or may be 10, 15, 17, 21 and all that, then obviously it is countable. You can count. Then, again you can make the set infinity, but still it is countable in the sense if it is say 0, 1, 2, 3, 4, 5 dot dot dot is countable, but that will be called countable infinite.

First one was countable finite, then next one is countabl infinite. So, in the case of discrete type time random process that was countable infinite. That was corresponding to countable infinite set of random variables. At every time point, there is a random variable standing, but whenever it comes to this continuous value, it is not countable. It is continuous. You know it may be two points. Again there is I mean have got infinite points. There is a case of continuous random variables, right. So, a continuous value random variable random process is nothing but a non-countable infinity of random variables, and then if you have these things, we can again you know I mean just use the previous notions in the same way.

For instance, so far we have defined the probability distribution of just one random variable x of t located at the particular t and corresponding density also defined. Now, you can make it two random variables, that is we can define joint distribution and joint density and like you know I mean you can find out, we can consider say a particular sample function. Say t 1 and t 2 and fix t 1 t 2 fix t 1 t 2. So, this x of t 1 and x of t 2, there are two random variables because depending on a particular waveform or a particular outcome if there is take various some values, next time the waveform changes, they take another value.

So, they are random and they are jointly random they are varying jointly hand in hand.
So, that means, we can define is an analogous way. They are joint distribution and joint density functions like that, but then joint distribution function will depend on two things. X 1 x 2, but also the two time points, t 1 t 2 because the definition as you know understand will be probability of two things joint probability, that is xt xt1 falling below x 1 xt 2 falling below x 2.

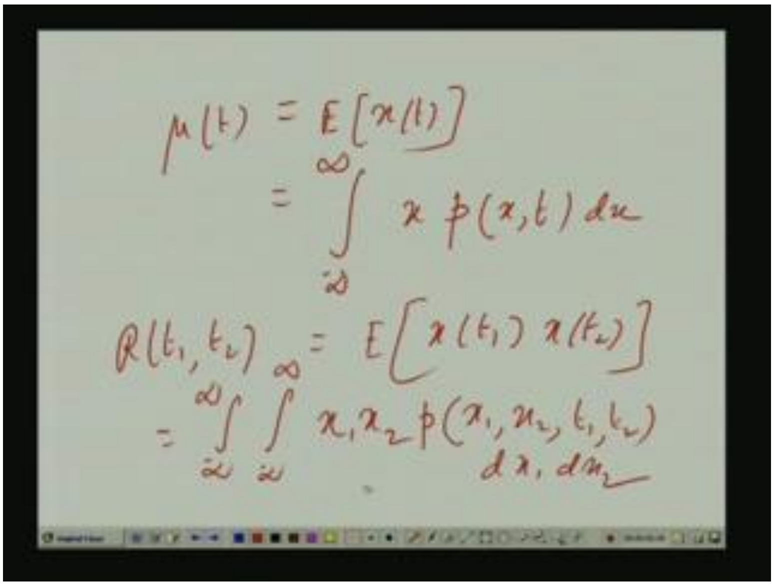
(Refer Slide Time: 25:53)



Corresponding joint density function will be now again I bring back small x 1 x 2 because I want to define the probability density, not with as specific x 1 x 2, but the general variables small x 1 small x 2. These are nothing but del square F double differentiation del square F x 1 x 2 t 1 t 2. Remember t 1 t 2, they are not changing. They are fixed, but they are just present everywhere. Then, again the same manner you can say or I am just showing that all our previous notions of probability density distribution apply here in this manner. If you just view x t 1 and x t 2 be to the random variables, you know varying jointly.

So, you can easily see I can write the sum of the further some other identity also. There is F x 1 t 1 you can always write as F x 1. The joint distribution x 2 taking and this infinity t 1 and any t 2, because this means this is another variable, this is the joint distribution. This is a single distribution is a in the case of joint distribution. If the other variable x 2 and x 2 lying below infinity that covers all possibilities of x 2, so obviously you get just a function of single random variable. I mean I get a single random variable probability distribution function and not joint. This you have done earlier.

(Refer Slide Time: 28:32)

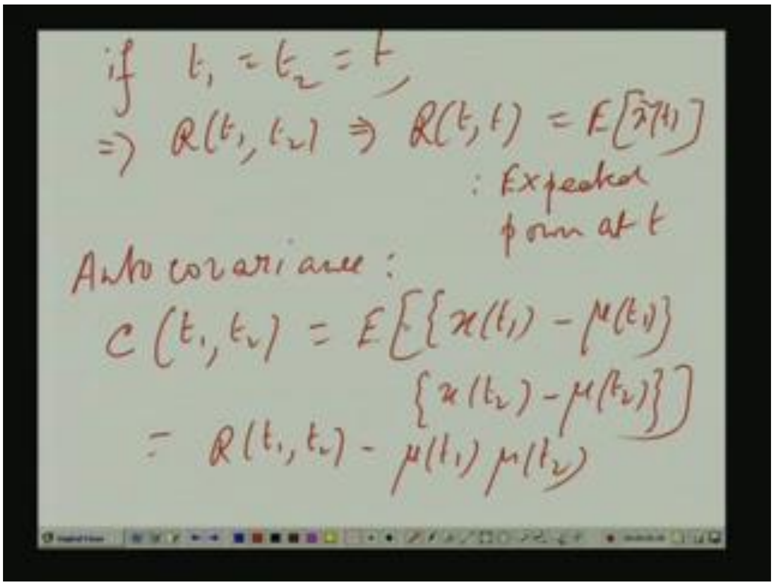


Similarly, you have taken any joint distribution joint density function, but integrate it with respect x 2 from minus infinity to infinity. Obviously, x 2 will disappear. You get the single density function and not joint obviously. So, therefore, in the similar method I can define mu t as the mean of the random variable xt that is nothing but E of xt for a particular t which is nothing but for a particular t, the probability density multiplied by the value x integrate with respect to x. Similarly, we can define, of course I am assuming real valued case, but we have to generalize these two complex cases that we will do little later.

Similarly, you can define autocorrelation r t 1 t 2. This is strictly for the real case. For imaginary case, you know have to make bring some complex conjugation somewhere that we will do a little later. Just we see it is nothing but we have got two random variables x t 1 x t 2 and is just up expectation of the product. In the real value case, it is simply the product and no complex conjugation anywhere and expected value which is nothing but the double integral minus infinity to infinity. Suppose it takes x 1, the other one takes x 2. This product is multiplied by the joint density, where t 1 t 2 will continue and dx 1 dx 2. You can also see that this is a function of two arguments.

So, you can plot it in a three-dimensional plot and if you take the straight line t 1 equal to t 2 say equal to t 1 equal to t 2 if you put as if we call it t. So, for any t 1 equal t 1 equal to t 2, you have got e of x t 1. Again x t 1 which is nothing but the expected power of the random variable at a point t 1, that is if t 1 equal to t 2 equal to t.

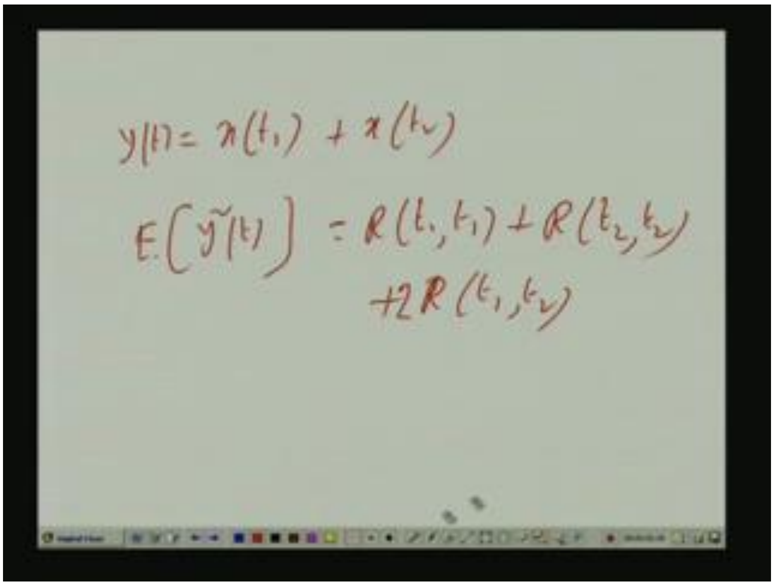
(Refer Slide Time: 30:38)



Then, R t 1 t 2 leads to simply R t, t which is E x square t which is nothing but expected power. X square t is the instantaneous power that is for the particular waveform the power at time point t. Then, if average it, what the entire ensemble that is taken for all the sample functions belonging to the ensemble and average out, then there is an expected power. So, it is expected power at t. In a similar manner you can define cross auto correlation auto co-variants. C is nothing but xt 1. This is random variable minus its mean times. Again I am doing for the real value case. The complex case, we will generalize, we will bring some conjugation and all that, and if you expand it we have done it earlier with the stat x and y. It is very easy to see. This becomes nothing but R t 1 t 2 minus the product of the two means.

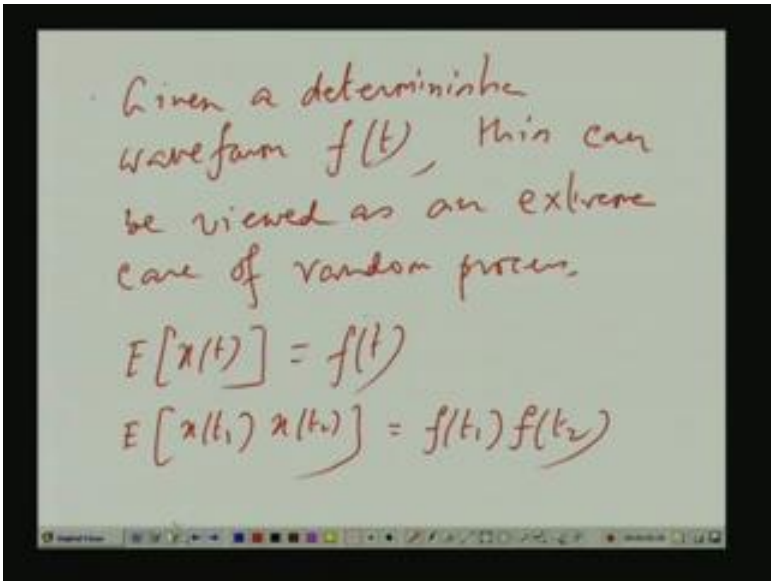
You can answer this question as to if R t, t is what is important to us. There is a power. Why do we consider a more general function R t 1, t 2 was you always consider this function R of t 1, t 2 as if R is a function of two variables. Why it is? So, why not just R t, t. Answer is you know suppose I am passing a random process through a linear system and I want to find out the power in the output.

(Refer Slide Time: 33:16)



Further, suppose the system just takes one sample, another sample and as it is a linear system, if just take two samples at a gap of t 1 minus t 2 as they are given in the output. Suppose this is the output. If we have to find out this, see if you just square this up, you will get x square t 1 plus x 2 square t 1 plus twice x t 1 x t 2. Then, put expectation operator, apply the expectation operator on each of the three terms. Obviously what you get is R t 1, t 1 and then, R t 2, t 2, but also twice R t 1, t 2. So, this is where we need it. Some further examples if you are given a deterministic waveform F of t within that you viewed as an extreme case of random process.

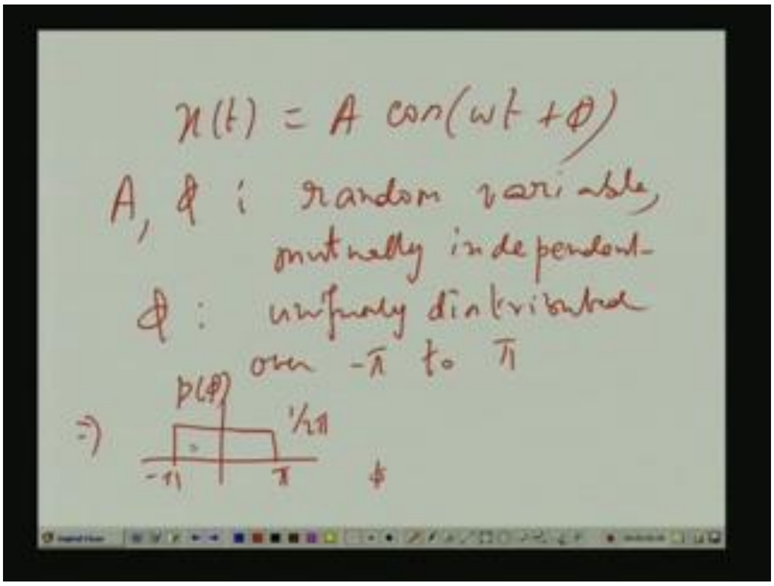
(Refer Slide Time: 34:40)



This can be viewed as an extreme case of random process. It means that suppose we are conducting experiments, we conduct every experiment lists the same waveform, that is with every outcome is same waveform executed. That means for a particular time position t x of t is nothing but the value Ft. So, x of t always takes the value F of t with the probability 1. This is the meaning of this because every time the experiment gives rise to the same waveform, say x of t takes the value f of t, a particular value with probability 1. This is a meaning of this. That means, actually mathematically E of x of t which is nothing but E of f of t, but f of t x of t always takes the value ft.

If it is not required to write this xt always take the value ft. So, expected value of xt also is ft. Then, it is f of t 1 for a particular experimental outcome f of t 1 f of t 2, but for any outcome also, it is f t 1, f t 2. So, if you take the expected value, this will take the xt 1 will take value ft 1 with probability 1 or this product will take the value ft 1 into ft 2 with probability 1. So, if you multiply by the probability density and integrate, you will get only ft 1 into ft 2. This is the special case and extreme case of random process, where the experimental outcomes yield the same function only always or if put in a different way, given the fixed function deterministic function, we call it deterministic function. We can all still view it as a special case of random process, where all experiments or all trials lead to the same function.

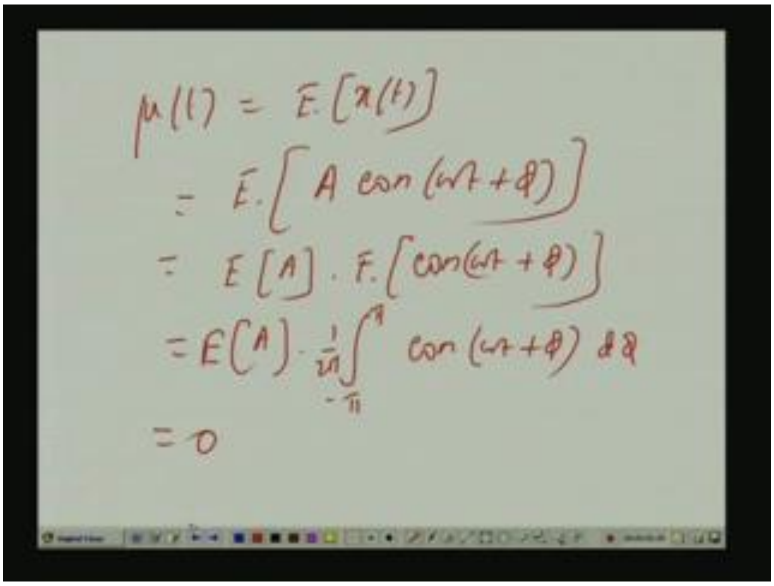
(Refer Slide Time: 37:44)



On the other hand, consider this that suppose you are given I am coming to the same example A cosine omega t plus phi A and phi random variables, given they are mutually independent. That is statically independent and phi uniformly distributed over minus pi to pi with a p of phi. This is phi axis just it is like this constant from pi to minus pi. Obviously area has to be 1. So, height is 1 by 2 pi. This as I told you, this x of t is a random process because every time you generate it or you make the experiment and observe it or generate it or whatever, you get because of some new A and new phi, you get a new waveform. They may appear to be same because they are sinusoidal, but actually they are different because the starting phase is different, amplitude is different.

So, it is a random process, but I told you it is very special kind of random process. It is a predictable random process that is if the entire past is known or at least one full period is known, interrupt of future can be predicted without any error, because it is periodic. Now, in this case we can just find out you know mean and auto correlation and all that. In particular say auto correlation what will be the mean. Firstly what is the mean for a particular t?

(Refer Slide Time: 40:03)

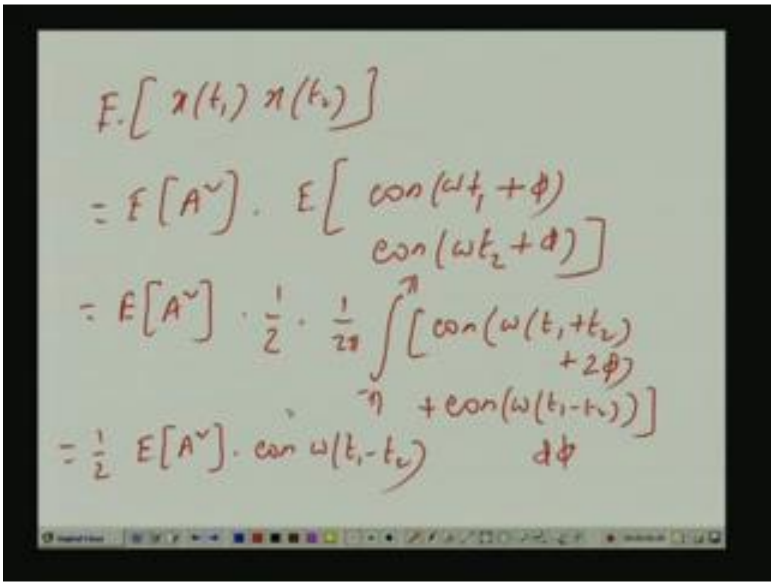


What is mu? It is t E of xt for a particular t, that is E of A cosine omega t plus phi. So, actually you have to multiply this by the probability density joint density of A and phi and integrate joint double integral. What is the entire range of phi and entire range of A, but these are we had given the fact that A and phi, they are statistically independent. So, that joint density between A and phi is nothing but a product of the probability density of A and probability density of phi. So, that means this is nothing but EA because the probability joint density will be then broken into two parts. Probability density of A into A integral with respect to A that goes on one side and probability density of phi times cosine omega t plus phi and one integral from minus phi to phi because phi is minus phi to phi that is another one. These two gave two separate expected values.

I repeat again how can you calculate the expectation? First remove this E and all that, multiply this thing by the joint density of A and phi and double integral with respect to A and phi, but the joint density is nothing but product of individual densities of A and individual density of phi because they are statically independent. So, double integral can be separated into product of two integrals. One of them is just with respect to A, A multiplied by its own density. Another one is cosine omega t plus phi multiplied by the probability density of phi. That obviously gives rise to this which I do not know what is E of A is. It is some value because we are not told about the probability density of A. So, we just keep it as E of A, but here we multiply cosine omega t plus phi by the probability density. That is uniform.

So, value is 1 by 2 pi over from minus pi to pi. That was the probability density uniform from minus pi to pi at within that range probability density of phi is taking the value constant value 1 by 2 pi which goes outside the integral, and if this integral, this is sinusoidal integral, right. It is varying with phi and you are integrating it over one full period of phi from minus pi to pi, obviously its sinusoidal integral will be 0. You all know any sinusoidal function when integrated over its period whether is sin or cos whatever, then it leads to 0. So, this will give rise to 0. So, mean is 0. This gives an idea about how to carry out these jobs actually.

(Refer Slide Time: 43:10)

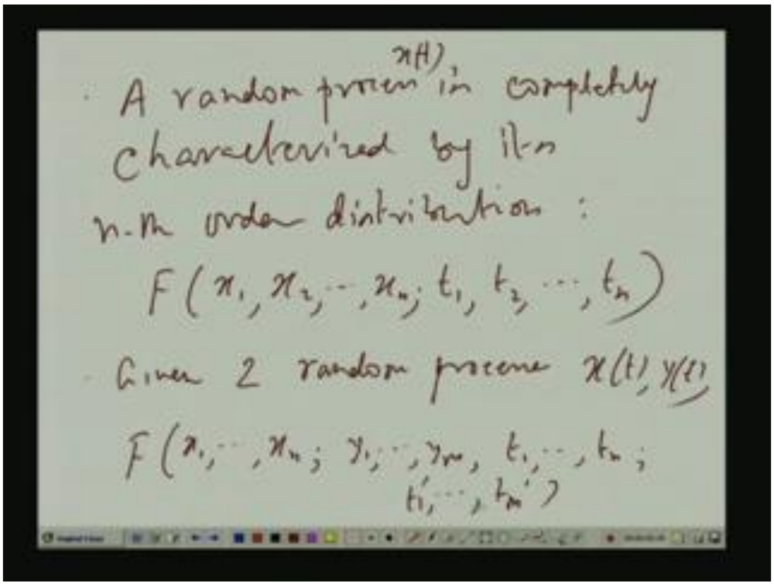


How about auto correlation? Take two time points t 1 and t 2 are multiplied. Obviously, there will be A square term that will give rise to a. So, separate E of A square because of reason I have already explained, because A and phi, they are statistically independent. So, joint density will be broken into will be written as product of two individual densities. One with respect to A that will come with A square integrals. That will gives rise to E of A square. Another term will be expected value of two terms cosine omega t 1, sorry then cosine omega t 2 again same phi.

So, E of A square and you break it as two things half and this is expectation you know is nothing but multiplying this by probability density of phi and integrating probability density of phi is 1 by 2 pi within the range minus pi to pi and another is 0. So, it is nothing but 1 by 2 pi minus pi to pi. Why this half? It is because cos A cos B is nothing but half cosine A plus B plus cosine A minus B. So, this will give rise to two terms cosine omega t 1 plus t 2 plus 2 phi plus cosine. Just omega phi cancels. Now, it is t 1 minus t 2 d phi. You see the first term cosine omega t 1 plus t 2 plus 2 phi, you are integrating with respective phi. So, this function actually is a function. This is now reaction function of phi.

So, within the period minus phi to phi, this function completes two full periods because we have 2 phi here, and obviously integral of two full periods is again 0. So, first time gives rise to 0. Second time has no phi left. So, this term can be brought outside the integral. Integral will give rise to 2 pi 2 and pi cancels. So, simply get half E A square cosine omega t 1 minus t 2. So, if t 1 and t 2 are same, so equal to t and cosine, this value becomes 1 because it is 0. Obviously, the power in that case will be half of half times E of mod A square E of A square. It can be easily generalized to the complex case anyway.

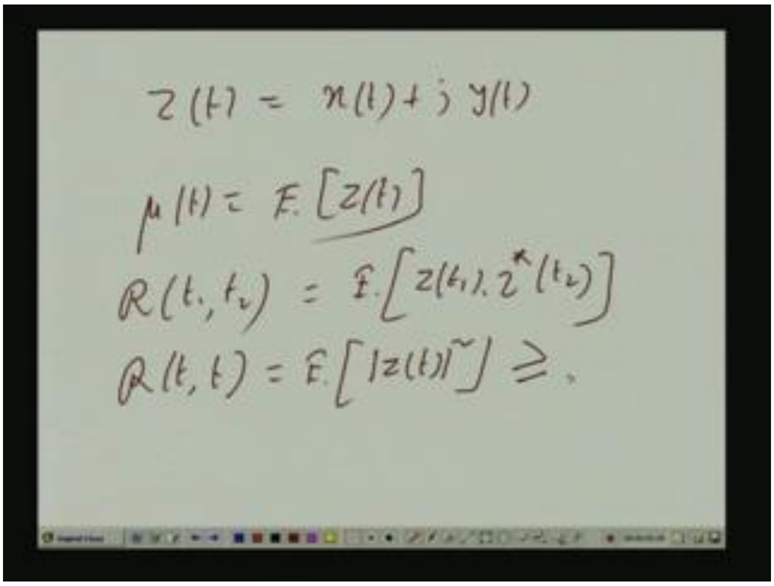
(Refer Slide Time: 46:48)



Remember random process is completely characterized say xt characterized by its nth order, where t 1 t 2 tn, they are arbitrary points and again n is taken arbitrary. So, this is true for all the n and all possible choice of t 1, t 2, tn. Then, this is completely characterized. Similarly, if you are given two random processes x and y, they are completely characterized by the joint distribution of this, that is F of this is joint distribution x 1 to xn.

Similarly y 1 dot dot dot ym t 1 to tn. They corresponds to x 1 to xn and some other points, t 1 prime dot dot dot tm prime, they corresponds to these. For arbitrary choice of t 1 to tn, for all possible choices of t 1 to tn and t 1 prime to tn prime and then, for all in a name, but you know this is more theoretical in practice. We are more bothered by the movements mean and auto position or auto co-variants. We have already seen it. We have just given the more generalized definition which is applicable to the case of complex value processes.

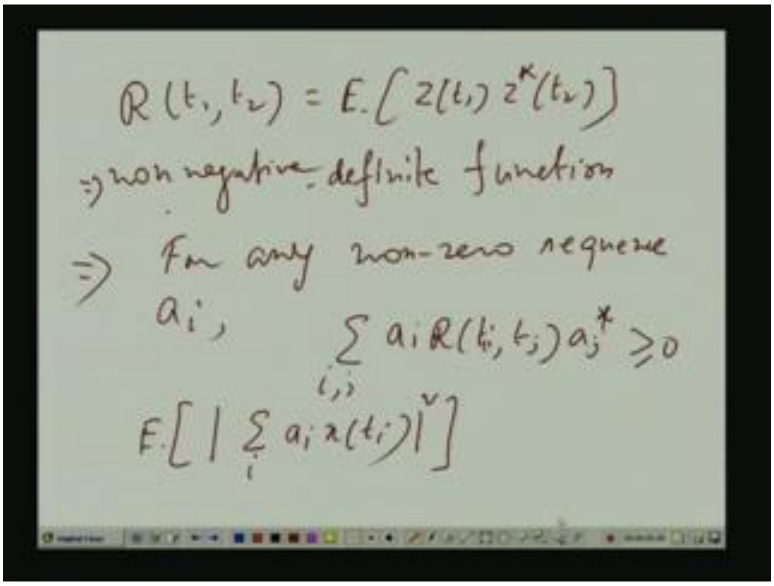
(Refer Slide Time: 49:31)



Complex value process means Z t. Actually at every point of time, it takes complex value, some complex value. So, you can also view it as some kind of summation, two waveforms. One is a real waveform; another waveform corresponds to the imaginary part. For this kind of processes mu say mu t remains as before E of say zt, but how about R t 1 t 2. Here you take zt 1 into z star t 2.

Why you start from? It is because we all know that when t 1, t 2 are same, this should give rise to power expected power. Suppose t 1 and equal to t 2 equal to t. So, we have got Rt, t and that should give rise to power expected power at the point t. The power is real if I do not put the conjugate here. I simply get zt whole square which is again a complex number. Expected value of that also is a complex number, but if I put star here, there are t. We have got just mod zt square expected value which is real.

(Refer Slide Time: 51:31)

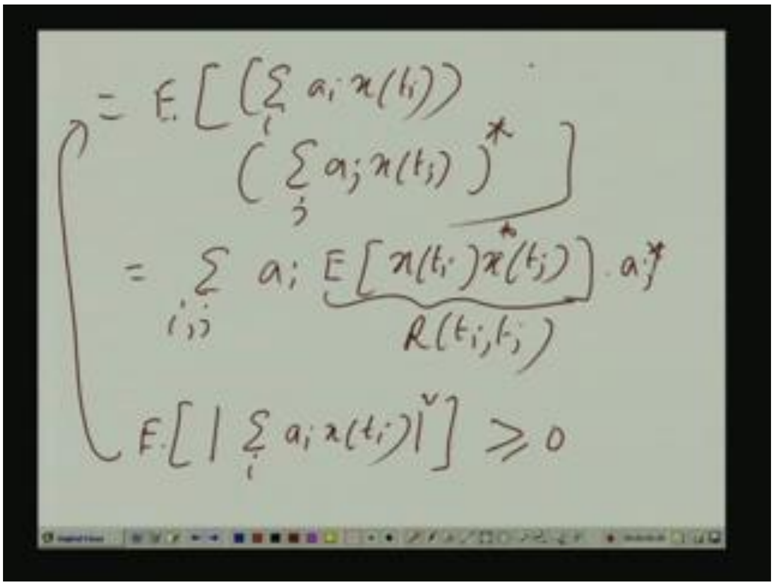


That means in general suppose Rt 1, t 2 say for a complex process is given z, then this is called a positive definite sequence. After all in the positive definite function, sorry is not sequence. In the area of discrete time process, it could have been a sequence because I would have at summing that n and m which taking real values. Here t 1 and t 2. So, is a function of two continuous valid random variables t 1 and t 2. So, you call it positive definite function. This implies for any non-zero sequence ai, where i can be 0, i can be one whatever you have this is always satisfied, that is summation over i, j ai Rti, tj aj star. This number is real and it is greater than equal to 0.

In fact, there it is called non-negative definite function, if it is strictly greater than 0. I change and call it non-negative. What does it mean? It means that you have given the function R of t 1, t 2. You just follow your sequence. They are picking up arbitrary i and j i can be 10. Just the sum and from the sequence actually in the form of i, and if the sequence ai you pick up any non-zero sequence ai, a 0, a 1, a 2, a 3 like that where you only guarantee one thing that not all s, not all the values of this sequence is 0. Then, for any choice of inj if you carry out this summation ai or ti, tj, aj star, this will be real and this is greater than equal to 0. This is obvious you know because this thing is nothing but mod.

If you carry out this some choice of i, you see this is more general. I can be 1, then i can be 27, then i can be 31. It is not that i is 0, then 1 or 2 is not. I is only integer. You are picking up and the discrete points are known to you. What is t 1? What is t 2? What is t 3? So, it is a very general thing. If you take this square, this is nothing but I will just take a minute and wind up today.

(Refer Slide Time: 55:19)



This is nothing but E of, but this we know this number cannot be, this is a real number. It is mod square at expected value. So, it can never be negative. This is always greater than equal to 0, but this thing is equal to E and then, aj xtj star and over j. This is over i and this is nothing but summation. May be you can put double summation or just write i over j ai aj are not random. So, they remain as it is xti xtj expect xti into x star tj expected value of that into aj star and this is nothing but R of ti tj. So, this proves it.

Conversely, it can also be proved that giving the positive definite function, you can find out a random process whose auto correlation function is given by this positive derivative function. This we might show later. That is all for today. Thank you very much.

Preview of next lecture. So, in the last class, we have given introduction to what is called stochastic process that is earlier we used to consider random variables, where we have an experiment and depending on the outcome of the experiment, a variable takes a value. So, its outcomes are random. The variable takes values in a random way and we call it random variable. So, we studied single random variable, we studied double random variable and functions of random variables and all that, there probability densities and probability distribution functions.

In the case of process, it was extended to function that is depending on a particular experiment. You got just a function. Function could be continuous, first function could be discrete, but function discrete in time. A function emerges and that function is random because depending on I mean depending on the particular outcome at hand, you get just one function, right. So, that set of all functions is called the process. This phenomenon is called the process. The set of all such possible random functions is called the ensemble. You can just stick to a particular time point, maybe t or t 1 say on the time x is. So, another point, every time you have the experimental outcome, you get one value because a function takes a particular value next time the value changes.

So, the value that this function takes you may call the function x of t, that x of t 1 for a fixed t 1. It will be the random variable, and that is true for all such t 1 of the time axis. So, basically random process is nothing but an infinite collection of random variables which take different values depending on the experimental outcome at hand. So, if basically is a random functions, then in the last class we also discussed the meaning of you know joint probability density joint probability distribution functions, given a stochastic process or random process, and then correlation co-variants and related things also I defined. Today we will be considering a particular class of random process called stationary process.