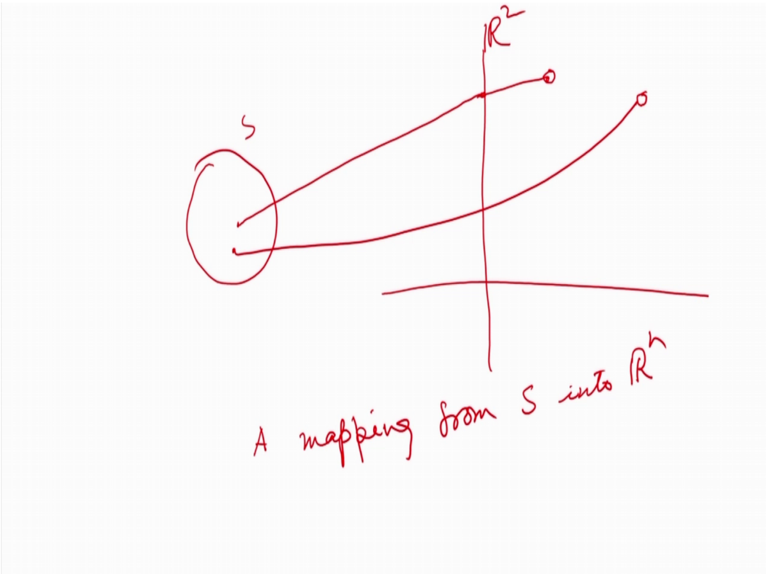
Advanced Topics in Probability and Random Processes
Prof. P. K. Bora
Department of Electrical Engineering
Indian Institute of Technology, Guwahati

Lecture – 04
Random Vectors and Random Processes

So, far we discussed about random variables, joint random variables and how to represent them. Now we will discuss about Random Vectors and Random Processes.

(Refer Slide Time: 00:41)

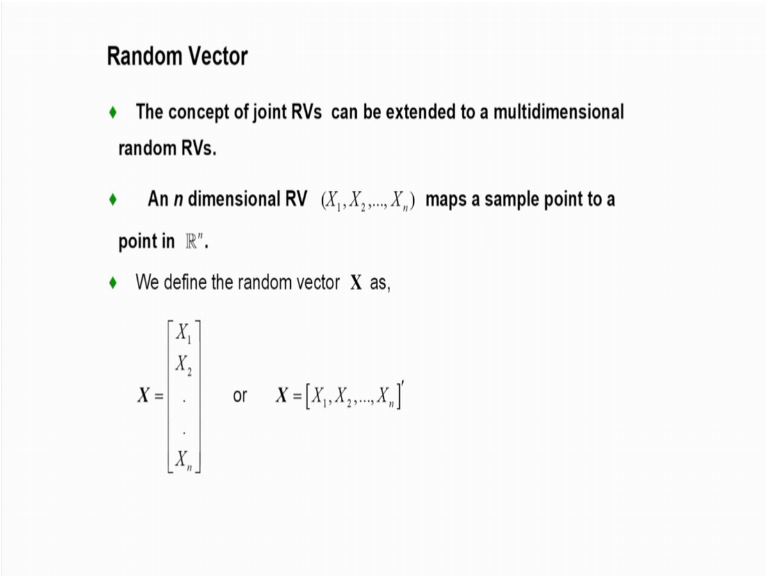


You recall that a random variable, it maps the sample space, this is suppose sample space to the real line, suppose this is the real line. So, sample space any sample point is mapped into a point on the real line. And now when; we consider joint random variable. We have R 2. So, we have R 2 and any sample point will be mapped to suppose a point in this plane. Suppose this is another sample point this will be mapped to another point in this sample real line sorry.

So another point will be mapped to the plane R 2 plane a point in the R 2 plane. So, that way we know that a random variable maps a sample point to a real line and a joint random variable maps a sample point to a point in this plane or R 2.

Now, this concept can be easily extended to suppose n random variables are there. Then we can consider that it is a mapping from the sample space today n dimensional n dimensional space R n, we will consider now a mapping from S into R n. So, this is the n dimensional random variable or we will call it a random vector.

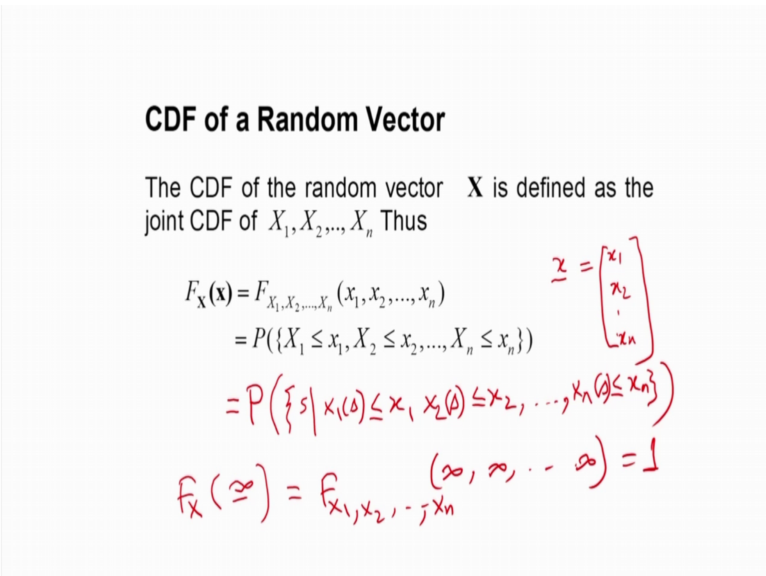
(Refer Slide Time: 02:34)



In an n dimensional random variable X 1 X 2 up to X n maps a sample point to a point in R n. So, that we have already told.

So, these n tuples X 1 X 2 up to X n can be represented by a vector X. So, we can write X this is the common notation we write a vector as a column matrix X 1 X 2 up to X n or we can write X is equal to a row matrix transpose.

(Refer Slide Time: 03:12)



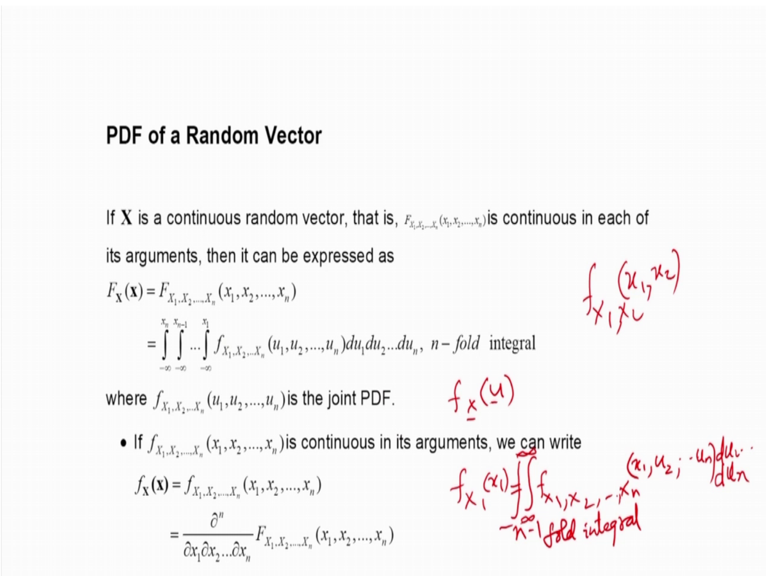
Defined a random vector X, we can find out the cumulative distribution function the CDF there of the random vector X is defined as the joint CDF of X 1 X 2 X n. So, of the random variables the CDF of the random vector x is defined as the joint CDF of the set of random variables X 1 X 2 up to X n. Thus, we can write this is the notation F x of at point x x is my point and that is equal to x 1 x 2 up to x n.

So, this at this point this joint is there is defined as that is probability that X 1 is less than equal to small x 1 X 2 is less than equal to small x 2 like this X n is less than equal to small x n. So, this probability actually means the probability that probability of the event, what is the event. As such that X 1 s is less than equal to small x 1 X 2 s is less than equal to small x 2 like that X n s is less than equal to small x n.

So, this is the event and probability of this event is known as the joint CDF of these and random variable or the CDF of the random vector X. We can extend the property of the joint CDF of two random variable to the CDF of the random vector F X of X. For example, we can say that suppose if what is the CDF of F X of suppose at point infinity. Infinity means, in this case is equal to that is F x 1 X 2 up to X n at point infinity, infinity etcetera.

So, this is equal to 1. Similarly this CDF will be a non degree decreasing function of all it is arguments. And it is right continuous of all it is arguments etcetera. So, that way n dimensional random vector can be characterized by its CDF F X of X.

(Refer Slide Time: 06:05)



Now, in the case of a random vector we can also similarly define joint probability mass function, joint probability density function etcetera. So, the PDF of a random vector because, random vector we are considering as a single variable that is why we are called calling it as a PDF if X is a continuous random vector. That is what do I mean that this joint CDF is continuous in each of its arguments. Then it can be expressed as this CDF can be expressed as the integral that is enfold integral of this non negative function f x 1 x 2 up to x n at point u 1 u 2 up to u n.

So, this function is known as the joint PDF. Just like we have defined, in the case of 2 dimensional random variable in the case of n dimensional also we can define the joint density function with the help of this integral. So, the CDF can be expressed as the integral of the joint PDF from minus infinity up to the point consider x 1 x 2 up to x n.

So, that way we can define generally define the PDF of a random vector. PDF of a random vector this quantity we can denote by and suppose f x vector of at point small u. Now suppose this joint PDF is also continuous or we continuous in it is arguments it is continuous in x 1 continuous in x 2 etcetera.

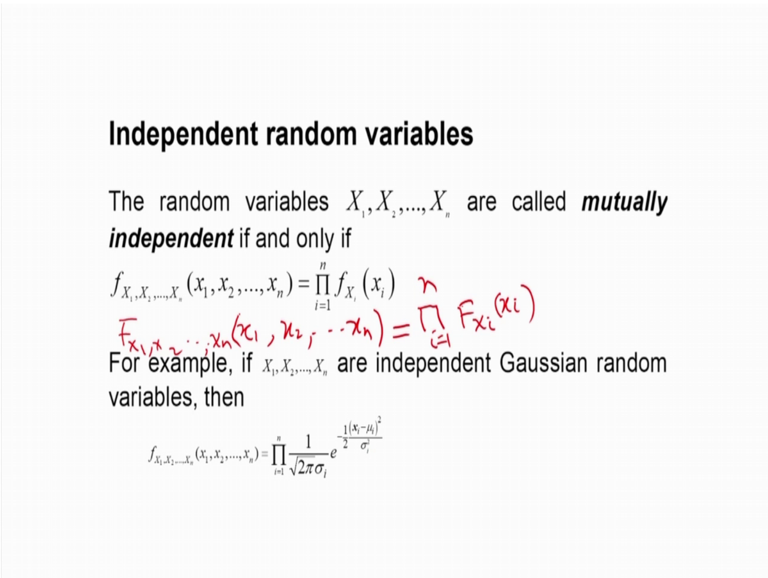
And then this PDF of the random vector f X of X can be expressed at the n th order mixed partial derivative del to the power n del x 1 del x 2 etcetera up to dell x n of the CDF. So, that way just like, in the case, of; two dimensional random variable, joint random variables, we defined the PDF at the second order mix partial derivative of 2 variables here it is the n th order first mixed partial derivative of n random n variables.

So, that we define the PDF of a random vector or joint PDF of this n random variables. Now this PDF also inherits the properties of joint PDF for example, here we can find out the marginal PDF suppose if I am I am interested to find out the marginal PDF f X1 at point x 1 how I will get it is the n full n minus 1 full integral. So, now we can write f X 1 X 2 up to X X n suppose and then at point X 1 we are interested x 1 u 1 u 2 suppose u 2 up to u n. So, we will integrate with respect to this is n full interior n minus 1 full integral so, with respect to suppose the u 2 up to the u n.

So, n n full integral if n minus 1 full integral if we perform we will get the PDF marginal PDF of X 1 at point x 1. So, this is up to infinity from minus infinity up to infinity. So, that way we can find out the any order joint PDF also suppose I want to find out the PDF of f x 1 x at point x 1 x 2 then we have to perform the n minus 2 full integral of the joint PDF.

So, that way this joint PDF completely characterized the random variables under consideration how they jointly behave or how they individually behave. Similarly, we can say about joint probability mass function etcetera. So, these are complete description of the random variables under consideration.

(Refer Slide Time: 11:35)



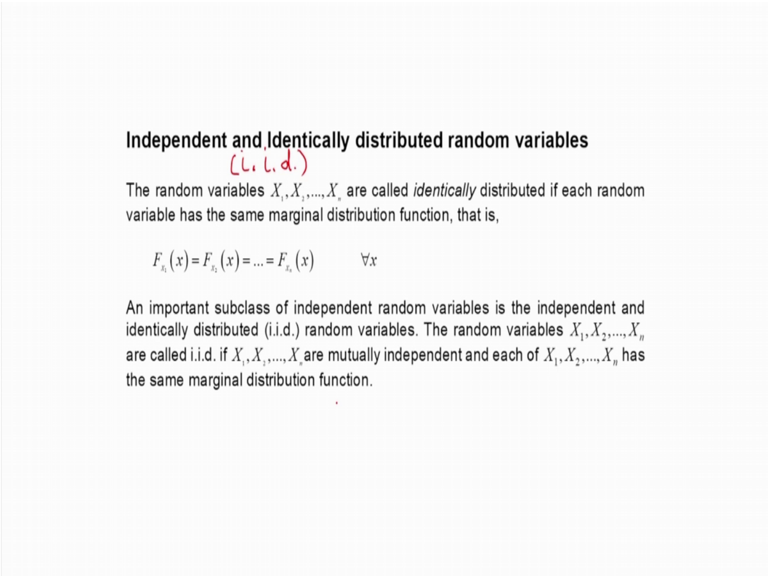
We can define independent random variables, suppose we have a collection of n random variables X 1 X 2 up to X n they are called mutually independent mutually independent if and only if the joint PDF is the product of the marginal PDF. Similarly we can write also joint main definition we can include the joint CDF F X 1 X 2 up to X n at point x 1 x 2 up to x n ok.

So, this will be the product of the marginal CDF i is equal to 1 to n F X i at point x i. Similarly, for the discrete random variables we can consider we can define the independence in terms of the probability mass function. Now this is mutually independent there is also a concept called pair wise independent X 1 X 2 etcetera.

They may be pair wise independence, but when they are mutually independent then, there will be independent for any sub collection of this random variable. Suppose X 1 and X 2 will be mutually independent x 1 x 2 and x 3 will be mutually independent like that.

But, if we consider on the pair wise independence then it does not apply mutual independence. For example, if X 1 X 2 up to X n are independent Gaussian random variables. Then we can write this joint density as the product of this individual Gaussian. So, that way it is the product of an Gaussian random variables is with respective parameter min mu i n variance sigma i square. So, we have a defined independent random variables.

(Refer Slide Time: 13:55)



One more important concept is independent and identically distributed random variables. This is abbreviated as i i d; IID random variables. The random variables collection is X1 X 2 up to X n they are called identically distributed. If each random variable has the same marginal distribution function, that is capital F x i that is the CDF of x1at point x is same as CDF of x 2 at point x like that CDF of x n at point x.

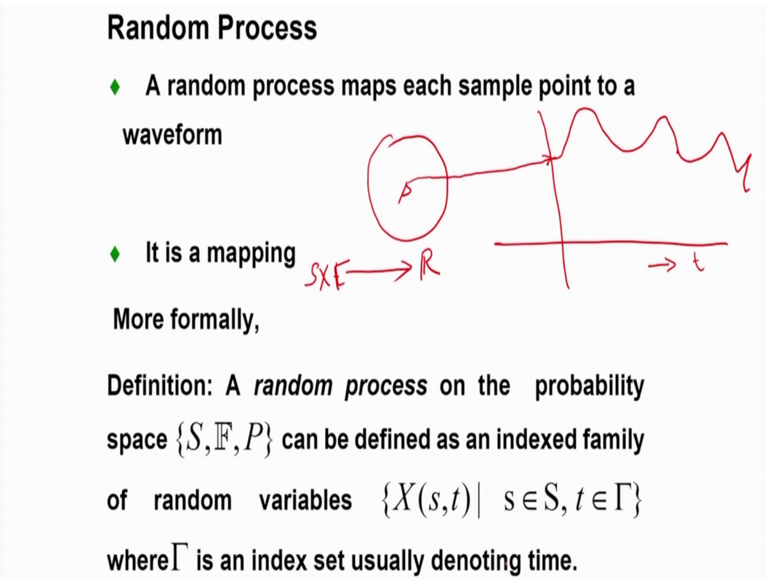
This is true for all x so in what does it mean that for all these random variables, each random variable has the same marginal CDF. Similarly we can define the identical distribution in terms of the PDF also. So, each of if these are continuous random variables then each of x i will have the x same PDF. Now an important subclass of independent random variables is the independent and identically distributed random variables.

So, these random variables are mutually independent. And each of x i s have has the same marginal distribution function or marginal PDF. So, that way we tell about IID random variables Independent and Identically Distributed random variables. We may have a number of random variables which are which has the same distribution itself which has the same distribution. And it is update random variable is the independent of other random variables in the set.

So, that way we can define IID random variables in the case of continuous random variable what the same can be extended to the script random variables. So, we can consider n discrete random variables they are independent and identically distributed. If they are joint probability mass function is the product of the individual probabilities mass function that is the independence. And each of the random variable has the same PMF at any point x if we consider each of the random variables will have the same PMF.

So, this is the definition of independent and identically distributed random variables. These will be this is a concept which is very important for our further study. So, we define the random variable joint random variable then random vector now we will define the random process.

(Refer Slide Time: 17:11)



A random process may maps each sample point to a wave form. So, suppose we have a sample space like this so, this is a simple point. Now in the case of random variable it was mapped to a point on the supposed real line. Now in the case of random process the same sample point will be mapped to a wave form means it will be maybe this type of waveform it will be mapping.

So, that way we see that it associates if I consider this time axis, then at every sample point and at every time instant t to a point on R. So, that way we can define we can say that this random process is a mapping from the sample space to the. So, we can say that a random process is a mapping from the Cartesian product as the it associates a pair s and t to a point on the real line.

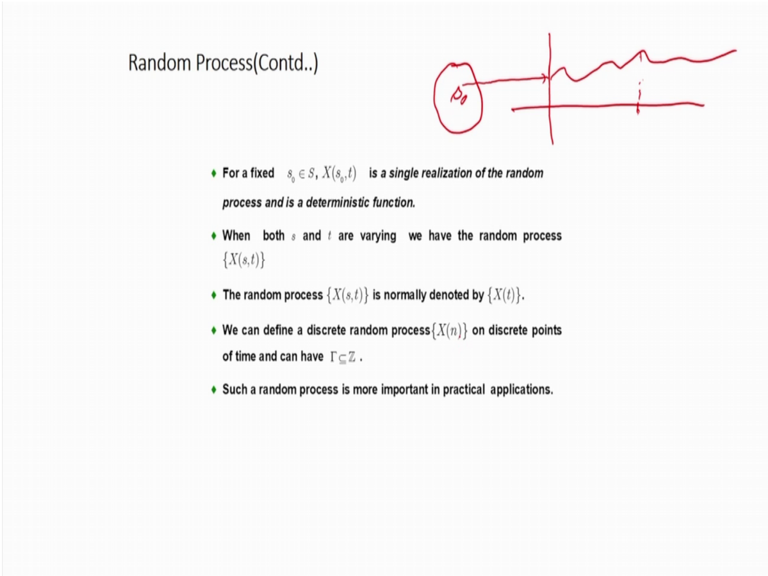
So, that it is a mapping from the Cartesian product of sample space and tau. This is the tau time generally that index set is called tau S crossed tau to the real line. So, that way we can consider the mapping in the case of random variable it is a mapping from the sample space to the real line, but in the case of random process it is a mapping from S crossed tau tau is the some index set which is usually time.

So, from S plus T to R; so we can define random process now a random process on the probability space S F P can be defined as an index family of random variable X s t such that s belongs to S t belongs to tau. Where tau is an index set of usually denote in time that is why it is a random process is a function of time.

But, tau need not be, time only it can be space for example, we can consider an image as a random process 2 dimensional random process. It is defined on suppose that sample space. And then Cartesian product of the sample space on the image plane tau need not be only time.

It can be two it can be a two dimensional plane for example, the image plane for example, an image can be considered as a random process defined on the samples, defined on the Cartesian product of sample space and the image plane. So, that way we define the random process.

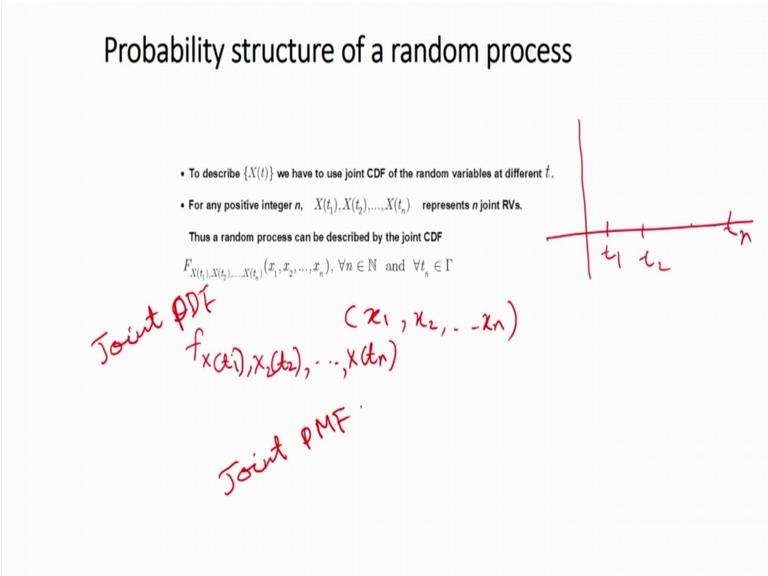
(Refer Slide Time: 20:20)



Now, a random process we see that it is a mapping from this sample space to a wave form. Now if we consider a fixed sample point. Then suppose t is varying then X s naught t is a single realization of the random process. So, we can have a single realization of the random process suppose this is some s naught. So, it is mapping to some wave from like this. So, this is a single realization of the random process which usually the observed data. So, we observe one single realization of the random process. Now when both s and t are varying we have the random process, otherwise suppose at a fixed point s naught and t naught this is we will get a single value.

Suppose if I consider this value this is a single value of the random process. Though the random process it is a function of both s and t it is normally denoted by X t only. Just like in the case of random variable we denote it by X omitting the argument s. Similarly, here also we omit the argument s. We can define a discrete random process X n on the discrete points of time we can have generally in that case tau is a subset of the integer. Such a random process is more important for practical application because normally, we get the discrete data, this discrete data can be modeled as discrete random process. So, this is a discrete time random process, this is a continuous time random process, probability structure of a random process.

(Refer Slide Time: 22:21)



Now how to describe a random process? We know that single random variable can be described by CDF joint random variable can be described by joint CDF. Similarly we can characterize a random vector by this CDF of the random vector. The same concept can be extended to random process, but in the case of random process we have a continuum of time therefore, infinite number of points in the time axis. So, characterizing the random process in terms of the joint CDF is difficult.

For example, we can consider for any positive integer n this collection we can consider X t 1 X t 2 X t n etcetera. And they will represent n n random variables. Now we can describe these points, these random variables by the joint PDF joint CDF like this. This is the joint CDF of n random variable X t 1 X t 2 up to X t n.

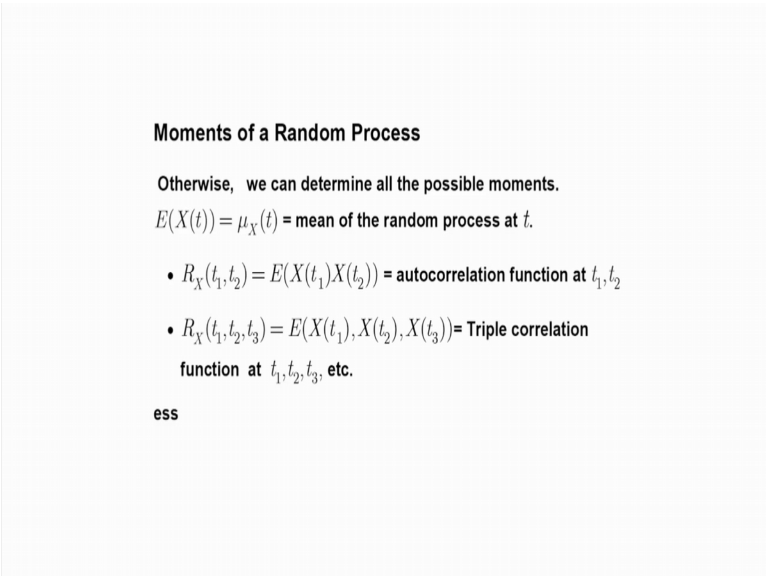
And these we have to define for a all n because any point any point supposed t 1 t 2 t 3 etcetera all possible collection we have to consider. And not only that suppose on the time axis we can take this place t 1 here t 2 here like that t n somewhere here, but we can keep t 1 t 2 etcetera anywhere on tau. So, that way also another complexity is introduced. So, thus what we see there, we can we have to define the joint CDF for all possible n and for all possible placement of the time points t n. So, that way describing a random process in terms of joint CDF is difficult.

Similarly, we can describe the random process X t by the joint PDF we can also have f X t 1 X t 2 up to X t n at point x 1 x 2 up to x n. So, we can have this joint see, there joint

PDF. So, this is also for all n belonging to the set of natural number and for all placement of the time point t n belonging to tau.

So, that way we can characterize a random process in terms of joint CDF for joint PDF. Similarly for the discrete case we can define the joint PMF also. These are the way we have to describe a random process if we are interested in the probability structure.

(Refer Slide Time: 26:12)

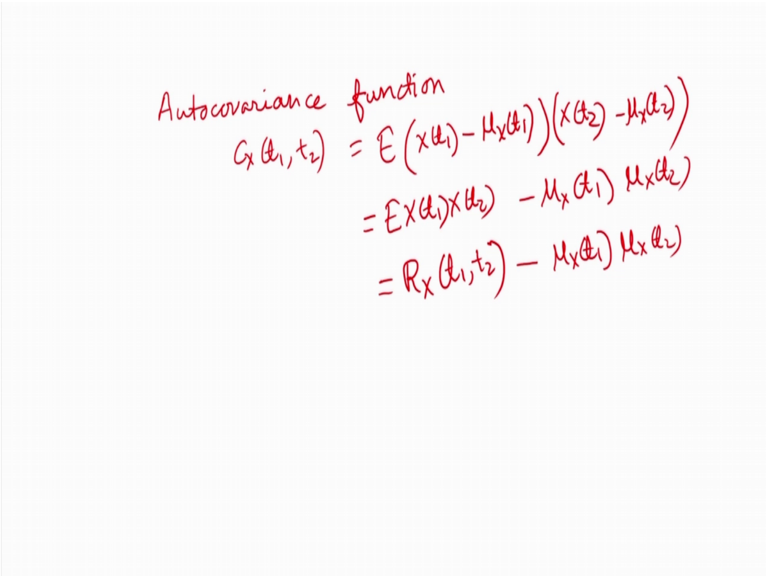


So, describing the random process in terms of joint probabilities, joint PDF, joint CDF, etcetera is difficult. Therefore, we can define a moments just like in the case of random variables, we defined moment in joint moments here also we can define some moments joint moments etcetera so that we can partially characterize the random process.

We can define the simplest moment is the mean E of X t that is the now at instant E exterior random variable. So, we can define the average value of X t that is E of X t we usually denote it by mu X of t and it is the mean of the random process at time t. So, now, since random process at different instant of time it is a different random variable. Therefore, mu X t also will be a function of time. Similarly R X of t 1 t 2 this is the autocorrelation function at point t 1 and t 2 this is defined by E of X t 1 into X t 2.

So, this is the autocorrelation function at t 1 t 2. So, that way we can define the joint moment. Since, we can define any number of random variables; we can define nth order correlation for example, triple correlation R X of t 1 t 2 up 2 t 3. R X of t 1 t 2 t 3 is define as expected value E of X t 1 into X t 2 into X t 3. So, this is the triple correlation function at time point t 1 t 2 t 3. So, that way, we can continue the definition of other order moments.

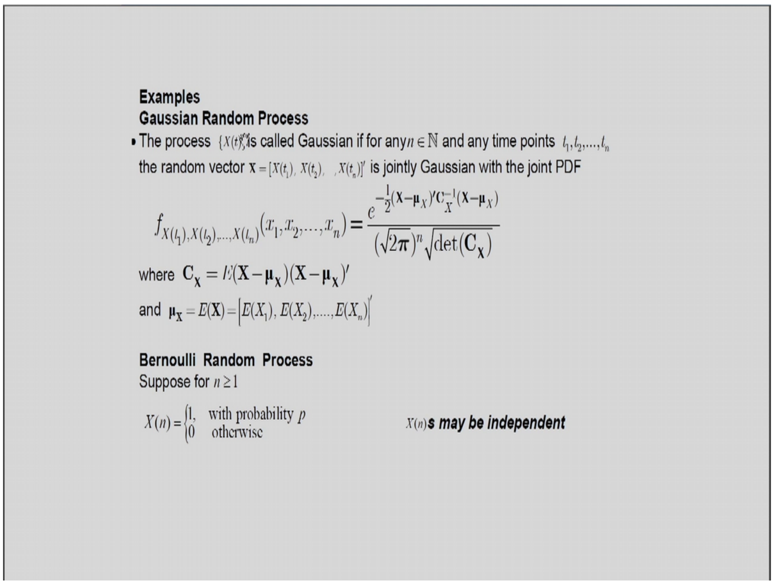
(Refer Slide Time: 28:21)



So, we can also define the auto covariance function, function that is C X of t 1 t 2 this is defined as the expected value of X t 1 minus mu X of t 1 into X t 2 minus mu X of t 2. That is equal to again we can use the property of joint moments we can define it as E of X t 1 into X t 2 minus mu X t 1 into mu X t 2. And this is nothing, but the autocorrelation function R X of t 1 t 2 minus mu X t 1 into mu X t 2.

So, that way we can define the auto correlation, auto covariance function. So, autocorrelation function, auto covariance function along with mean, they can describe the random process partially.

(Refer Slide Time: 29:55)



I will give some examples of random processes. The first example is Gaussian random process, which is one of the most important random processes. The process X t is called Gaussian. If for any n belonging to natural number N and any time points t one t 2 up to t n the random vector we define a random vector that is X t 1 X t 2 up to X t n transpose.

This is the random vectors n dimensional is jointly Gaussian. With the joint PDF the joint PDF is now given by this is the joint PDF that is equal to e to the power minus half of X minus mu X transpose C X inverse X minus mu X divided by root over 2 pi to the power n into root over determinant of C X let us define the terms. Now mu X mu ax is a vector, this is the expected value of this vector E of x. So, mu X is E of X that is equal to now component wise I will take the I will component wise I will take the I expected value that is E of X t 1 E of X t 2 up to E of X n transpose. Now C X C X is a matrix that is the covariance matrix that is E of X minus mu X X minus mu X transpose. And then we take the expectation first we have t matrix X minus mu X into X minus mu transpose.

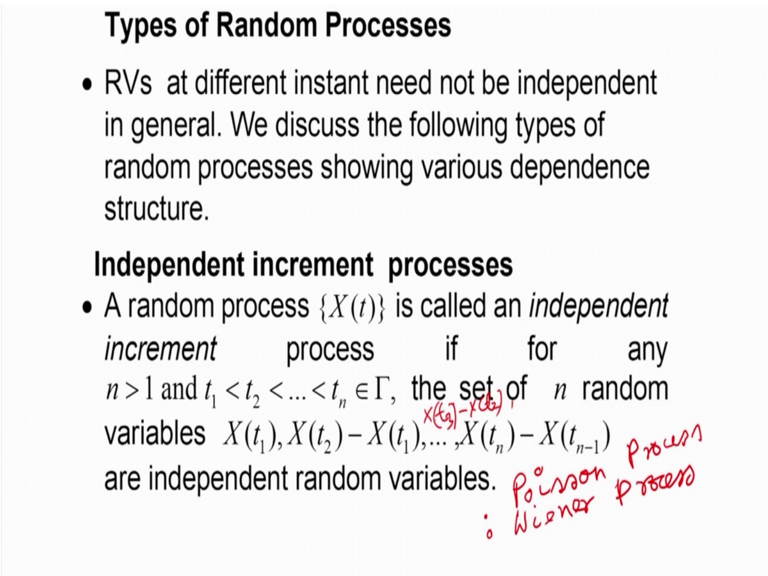
And expectation of each element in this matrix that will give us the covariance matrix so, here we determine the determinant of the covariance matrix and here inverse of the covariance matrix. So, that way we define the Gaussian random process. Next elementary example this is an elementary example of a random process what is known as the Bernoulli random process.

We know the Bernoulli random variable which can take only two values 0 and 1. Suppose X n at any instant n it takes value 1e with probability P and 0 with probability 1 minus P. Then this is a Bernoulli random variable and we may also consider that this supposed Bernoulli random variable, Bernoulli trials are independent.

If I independent indeed suppose toss a coin then if I am interested in one outcome suppose head. Then this I can take as 1 and tail as 0. So, that way we can define a Bernoulli random variable and if I repeat it toss for different time. Then we will get this sequence of Bernoulli random variable or it is called the Bernoulli random process and the in this case actions are independent.

So, we may have a discrete random process, which is where, each random variable is independent of other random variables.

(Refer Slide Time: 33:37)

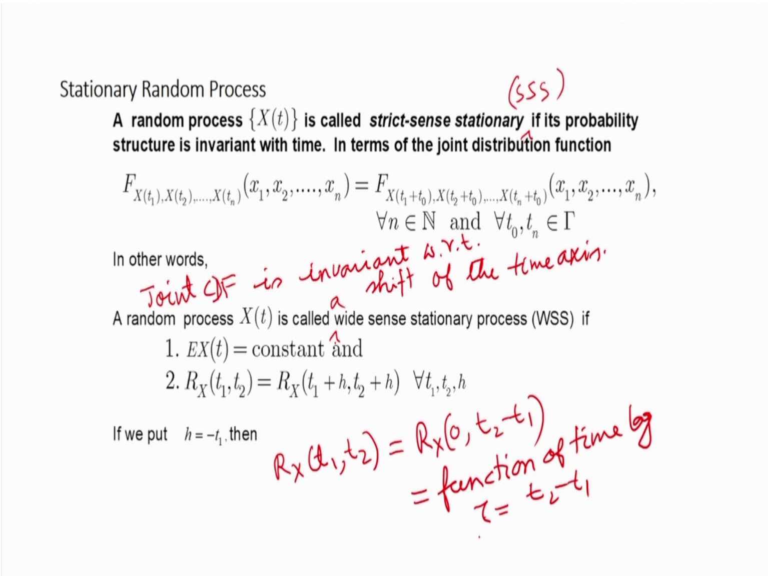


Now in the case of random process we know that at different instant we have different random variables need not be independent in general. In the case of Bernoulli process they may be independent, but they are they are generally dependent. We can discuss different type of the independent dependence we can discuss different types of dependence.

For example, independent increment random process. A random process X t is called an independent increment process if for any n suppose you consider t 1 is less than t 2 less than t 3 etcetera up to suppose t n minus 1 is less than t n. Then we can define a set of n random variable X t 1 X of t 2 minus X t 1 suppose X t 1 is a random variable X of t 2 is another random variable therefore, the increment during time t 1 to t 2 is X t 2 minus X t 1. So, this is 1 increment then next increment for example, will be X t 3 minus X t 2. So, this is another increment like that we can consider all the increments. These increments are independent random variable, then we call this process as an independent in increment process.

So, the concept of independent increment process is very important for example, we can consider the term we will consider 2 processes one is Poisson process. So, this is an example Poisson process, this is an example of independent increment random process. Similarly we can we will have another process called wiener process. So, Poisson process and wiener process are examples of independent increment processes. So, here increments are independent.

(Refer Slide Time: 36:04)



We can also consider a concept, what is known as the stationary random process. Now what is stationary? A random process X t is called strict sense stationary, this is strict sense stationary.

So, these we abbreviate as S S S, if each probability structure is invariant with time. In terms of joint CDF we can write the joint CDF at of random variable X t 1 X t 2 up to X t n at point suppose this set of points X 1 X 2 up to X n. And if we consider the joint CDF at another instant of time that is X of t 1 plus t naught X of t 2 plus t naught like that way X of t n plus t naught. Then at the same set of point X 1 X 2 X n we are considering these 2 CDF if they are equal for all n belonging to n and for all placement of t 0 t 0 and t n.

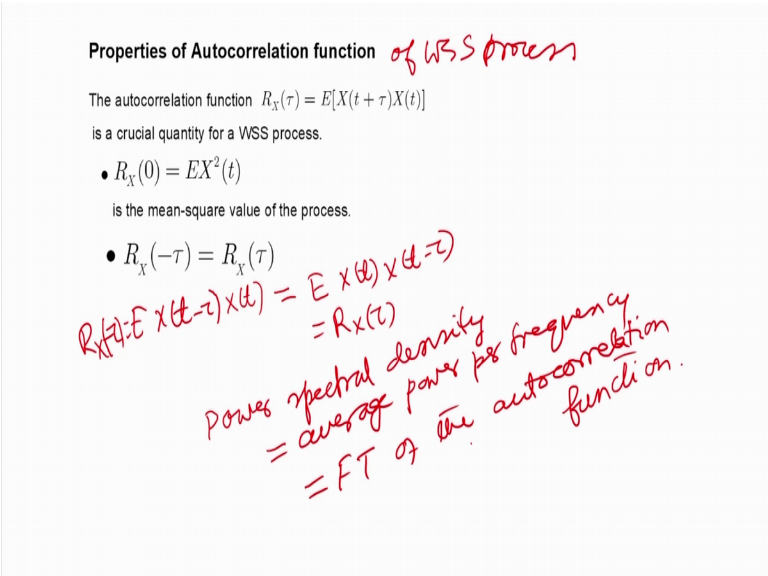
Then we call these random processes is strict sense stationary. So, what does it mean? That joint CDF for any order n is independent of this shift of the time axis. So, you consider whether at t 1 t 2 t n or t 1 plus t naught t 2 plus t naught up to t n plus t naught. So, this joint probability structure will implement invariant. So, in other words the joint CDF joint CDF is invariant with respect to the shift of the time axis.

Now, S S S process as it is seen from here it is very difficult to analyze a strict sense stationary process because, we have to consider the CDF of all possible order. So, we require a simpler process that is wide sense stationary process. In that case we do not consider the probability structure we consider only 2 moments. So, a random process X t is called wide sense stationary process a wide sense stationary process WSS if E of X t is equal to constant and R X of t 1 t 2. That is the autocorrelation at instant t 1 t 2 is same as if I shift your autocorrelation time axis by some amount is r X of t 1 plus t 2 plus is.

This is for all t 1 t 2 n is if this happens, then we say that for this random process mean is constant and autocorrelation is independent of the shift of the time axis. Then we say that this process is there WSS process wide sense stationary process. So, if I put h is equal to minus t 1 what we will have is that R X of t 1 t 2 that is equal to R X of if I put h is equal to minus t 1 this will be r X of 0 then this will be if I put minus t 1 t 2 minus t 1. So therefore, we see that in the case of WSS process this autocorrelation function is a function of function of the time lag function of the time lag t 2 minus t 1.

So, if I denote it by some single parameter tau then tau is equal to t 2 minus t 1. This autocorrelation function is a function of tau on be.

(Refer Slide Time: 40:51)



And now we can define therefore, autocorrelation function since it is a function of the lag parameter tau we can define R X of tau is equal to E of E of X of t plus tau into X t. So, it is a function because time difference will be t plus tau minus t. So, that way autocorrelation function for a WSS process autocorrelation function of a WSS portions has certain property.

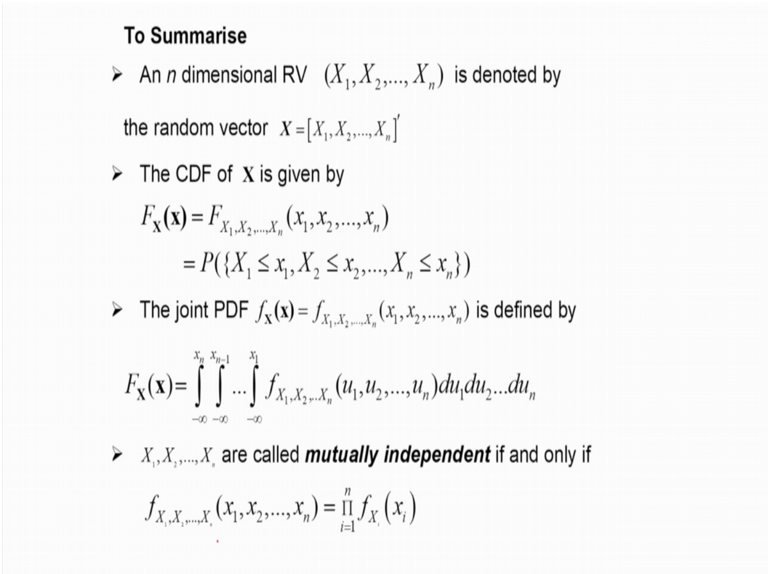
So, it is a function of like tau only we can write R X of 0 what will be R X of 0 that is equal to E of X X t into X t that is the E of X for t, this is the mean square value of the process. Similarly this autocorrelation function R X of tau suppose is an event function how? For example, R X of minus tau this is equal to E of X of t if I consider these definition t minus tau. So, R X of minus tau is a of X t minus tau into X t. And X of t minus tau and X t are both are real. So therefore, we can write that is equal to E of X t into X of t minus tau and that is equal to if I take the difference now it will be R X of tau.

So, this is R X of minus tau that is same as R X of tau. Therefore, this autocorrelation function is an even function of the lag parameter tau. Now there are other properties we will not discuss in details, but this autocorrelation function also give a frequency domain representation of the random process.

So, how to describe a WSS random process in the frequency domain you can consult any random process book, where it is shown that there is this power spectral density power spectral density. This is the average power per frequency and this is the Fourier transform of the autocorrelation function relation function.

So therefore, to analyze a WSS process in the frequency domain, we can take the help of power spectral density which is the Fourier transform of the autocorrelation function.

(Refer Slide Time: 43:55)



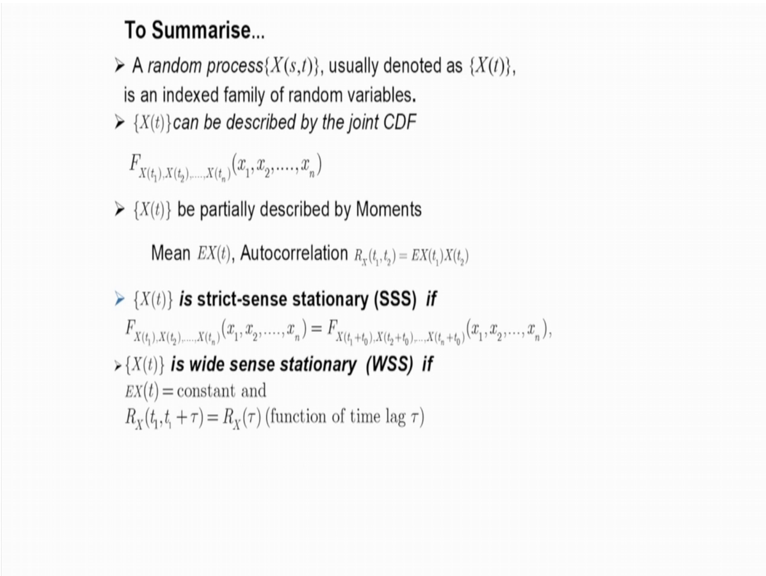
Let us summarize the lecture and n dimensional random variables. So, it is X 1 X 2 up to X n that the ordered pair is denoted by the random vector X that is we can write it as a row vector transpose. Now these n random variables can be characterized by the joint CDF that we denote by F X at point small x.

So, this is the CDF and this is actually the CDF of joint CDF of X 1 X 2 up to X n at point x 1 x 2 up to x n. So, this is defined by this probability that probability. That X 1 capital X 1 is less than equal to smaller x 1 capital X 2 is less than equal to small x 2 like that up to capital X n is less than equal to small x n.

Now this joint PDF this is notation is f X sub X. So, this joint PDF is defined through the integration this is n full integration. That is a joint CDF is the n full integration of the joint PDF with respect to it is parameters. So, that way we defined the joint CDF joint PDF and also we defined the joint P M F for a collection of random variable n random variables are there they are represented as a vector and then we got this quantity.

Now these collection of random variable X 1 X 2 up to X n are called mutually independent if and only if the joint PDF is product of the marginal PDF. So, that way we can define in terms of or the joint CDF also the joint CDF should be the product of the marginal CDFs.

(Refer Slide Time: 46:05)



Then we discussed about random process, a random process it is a function of s and t sample space it is a simple point and t is usually it is a time variable. So, it is a function of 2 arguments sample point and t. And it is usually denoted by X t X is omitted just like in the case of random variable. So, a random process is an index family of random variables.

So, what do I mean by that at a every instant t it is a random variable. X t can be described by joint CDF. So, if I consider any n instance of time t 1 t 2 up to t, n then we can find out this joint CDF, but these we have to find out for all possible n and all possible placement of t 1 t 2 t n etcetera. So, that way describing the random process in terms of joint CDF is very difficult very complex task. Therefore, we have to find out term simpler; description of the random process that way we describe the random process in terms of it is moments like mean E of X t and autocorrelation function R x t 1 t 2 this is nothing, but the expectation of X t 1 into X t 2.

So, this mean and autocorrelation are 2 important parameters of a random process. Next we define this stationarity the process X t is strict sense stationary if the joint CDF for any end is a is invariant with respect to changing the time axis. What does it mean suppose if I have the joint CDF at point t 1 t 2 t n. Then if I consider any shifted version of time t 1 plus t naught t 2 plus t naught up to t n and plus t naught then also this joint CDF will remain same.

So, this is the strict sense stationarity this we can define in terms of joint CDF joint PDF joint PM F, but this is a very difficult to analyze. Because, you have to consider it for all n and all the placement of t 1 t 2 t n so that way we consider some simpler form of stationarity we discussed about wide sense stationary process. So, X t is wide sense stationarity sorry X t is wide sense stationary WSS if E of X t that mean is constant and the autocorrelation R X t 1 t 1 plus tau.

So, this time lag is tau here t 1 t 1 plus tau and this is a function of time lag only it does not depend on t 1. So, that way if that happens then the process will be called wide sense stationary.

Thank you.