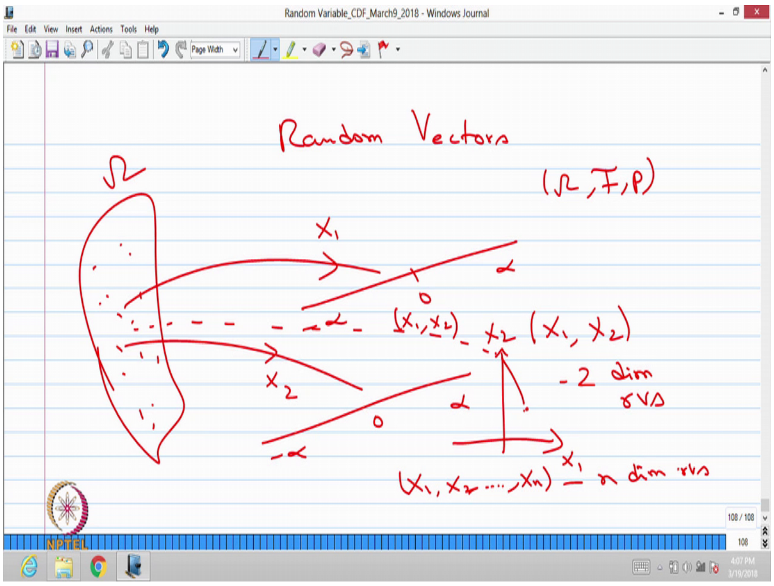
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Module - 05
Higher Dimensional Distributions
Lecture - 24

We are moving into the 5th module; let me recall first module we discussed the basics of probability. And the second module we discussed the random variable and the third module we discussed the movements and inequalities and the fourth module we discussed the standard distributions both the discrete and continuous type.

(Refer Slide Time: 00:53)



Now, we move in to the 5th module that is random vectors in most of the random experiment you may have a omega in which you may able to create many random variables simultaneously. And also sometimes the interest will be finding the some measures together or given one set of measures takes this value what could be the distribution of other random variable? Or sometimes you may interest to find the distribution of more than one random variable at the same time. So, till now we have discussed only one random variable and then later we have discussed a function of a random variable; that means, again that is another one random variable. There is a possibility you may need to study more than one random variables at the same time or simultaneously or together, in that case you need a random vector.

When we use the word random vector; that means, it is n dimensional random variable; that means, we put a n random variables together that we call it as a n dimensional random variables or random vectors of size n. That means, each one is random variable and we are studying n random variables together or jointly. Many real world problem you may need to study the distributions of a more than one random variable jointly therefore, we need a random vector. That means, whatever we have understood the concept of a random variable and their distributions and movements; the same concept has to be extended from one dimensional to multi dimensional.

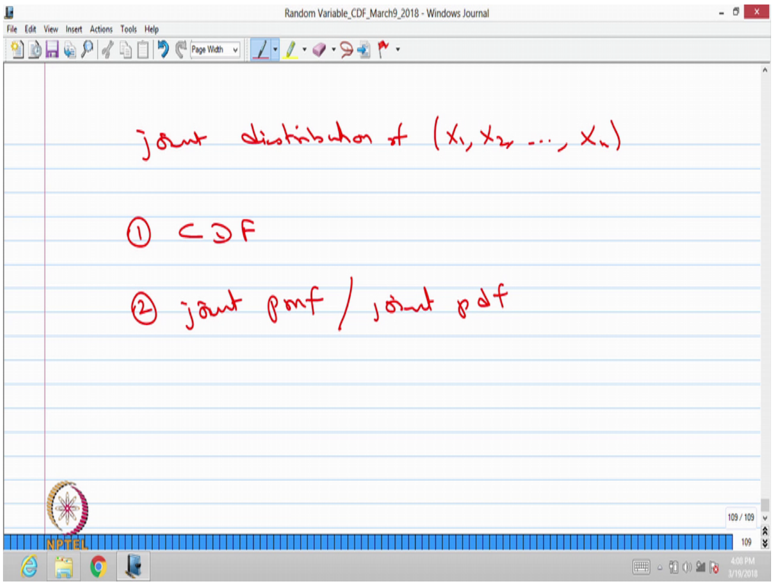
So, whatever the calculus we have used calculus means finding the integration or finding the derivative and so on; whatever we have done the calculus part with the one dimensional variable or one variable, now we have n dimensional random variable. So, the corresponding calculus of several variables has to be used. Let me explain it how the random vector coming into the picture. The omega consist of many samples, it could be finite or countable infinite or uncountable many. The random variable one random variable is defined as a X 1.

Another random variable defined from the omega that is X 2 like that there are many more random variables defined in the omega that is a real valid function from omega to r satisfying the condition that is X inverse of minus infinity to small x semi closed interval belonging to F then only the real value function is a random variable. That means, you have a probability space omega F P, omega is a collection of all possible outcomes and F is the sigma filled on omega and P is the set function satisfying the 3 conditions the Kolmogorov accemetrix condition. Therefore, P is the probability; so, this is the probability space.

In this probability space X 1 is defined the same probability space X 2 is defined like that many more random variables are defined in the same omega. Now we are going to create a random vector by making a few random variables together or jointly; that means, if you go for X 1 comma X 2 together then this is called 2 dimensional random variables or it is a random vector of size 2 X 1 and X 2; that means, the same omega X 1 comma X 2 is defined such that the possible values is going to be in the 2 dimensional plane.

This is X 1 and this is X 2; earlier the X 1 is defined from omega to r now the X 1 comma X 2 is defined from omega to r cross r. So, again each one X 1 is random variable X 2 is a random variable together we call it as a random vector of size 2 or 2 dimensional random variable. The same way one can define n dimensional random variable X 1 X 2 and so on till X n; each one is a random variable therefore, this is a n dimensional random variable in short r vs random variables.

(Refer Slide Time: 06:45)



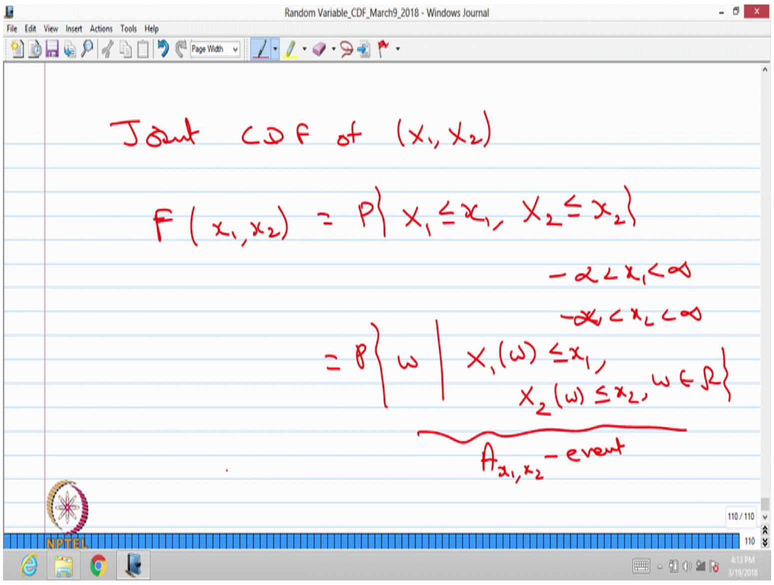
That means, we can go for finding what is the joint distribution of n dimensional random variable X 1 comma X 2 so on X n.

Earlier we used a word distribution of X 1, now we are studying the distribution of n random variable together or jointly. Therefore, it is called a joint distribution of n dimensional random variables. So first we can discuss; what is the CDF Cumulative Distribution Function; if it is a discrete type random variable one can discuss; what is a joint probability mass function; if each random variables of the continuous type then one can study what is a joint probability density function of n dimensional random variable.

Initially, we study all the random variables of that discrete type or all the random variables are of the continuous type. Later we will see one random variable of the discrete type and the other one is of the continuous type and so on that we will do it in the later.

So, our interest is to find the joint distribution of n dimensional random variables in which each one is a random variable; that means, that satisfies the definition of random variable. Therefore, this is a n dimensional random variable first let us go for how to give the definition of joint CDF joint.

(Refer Slide Time: 08:31)



CDF of instead of going for the n dimensional random variable first we will restrict to the 2 dimensional random variables. Once we discuss the 2 dimensional random variables CDF, then it is easy to visualize or easy to understand for more than 2 dimensional random variable ok. It is X 1 comma X 2 the joint CDF can be written as the 2 functions x 1 comma x 2 is nothing but the probability of capital X 1 is less than or equal to small x 1, capital X 2 is less than or equal to small x 2.

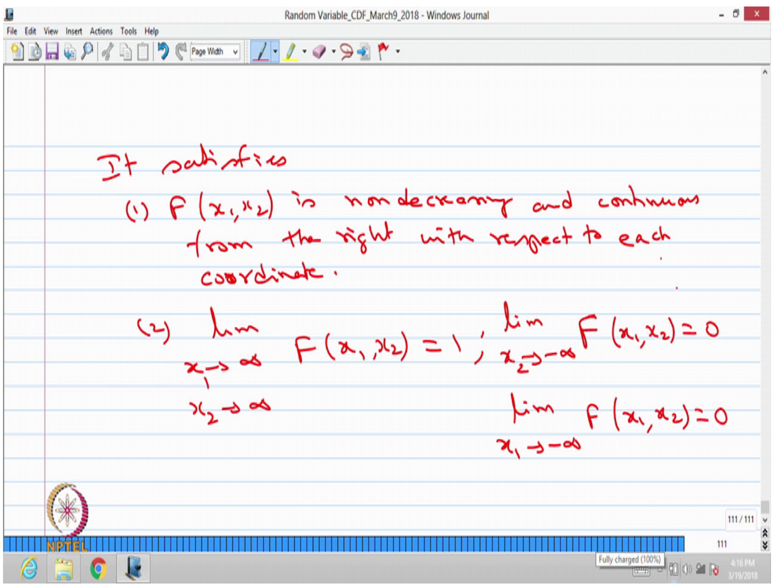
Both x 1 and x 2 lies between minus infinity to infinity. Similar way we have defined the CDF of 1 dimensional random variable, now we are defining the CDF of 2 dimensional random variable with the 2 variables small x 1 comma small x 2 that is nothing but probability of capital X 1 is less than or equal to small x 1 and capital X 2 is less than or equal to small x 2.

That is nothing but P of collection of w such that under the operation X 1 the w has to be less than or equal to x 1; that means, this value can lie between minus infinity to small x 1. Also the same w under the operation X 2, it gives the value small x 2. And the w's belonging to omega R; that means, we are collecting a few possible outcomes satisfying the condition under the operation X 1 it should give the value less than or equal to x 1 under the operation of X 2 it gives a values less than or equal to small x 2.

So, we are collecting those possible outcomes then finding the probability and that probability of these possible outcomes satisfying this condition that is going to be the CDF at the point X 1 comma X 2 where small x 1 can lie between minus infinity to infinity small x 2 also lies between minus infinity to plus infinity. The collection of w such that this condition we can label this as the A suffix x 1 comma x 2 this is a event because whenever you collect a few possible outcomes; that is nothing but the event.

So, the event A suffix x 1 comma x 2 that is a event; that means, the probability of event by using a Kolmogorov accemetric definition probability of any event or P of any event always greater or equal to 0 and P of omega is equal to 1 and P of union of Ai's that is equal to summation of P of Ai's as long as Ai's are mutually disjoint events. With the same logic the P of any event getting from the different values of small x 1 and x 2; this is also going to satisfy the Kolmogorov accemetric definition. Therefore, one can make what are all the conditions or what are all the properties is going to be satisfied by this capital F.

(Refer Slide Time: 12:43)

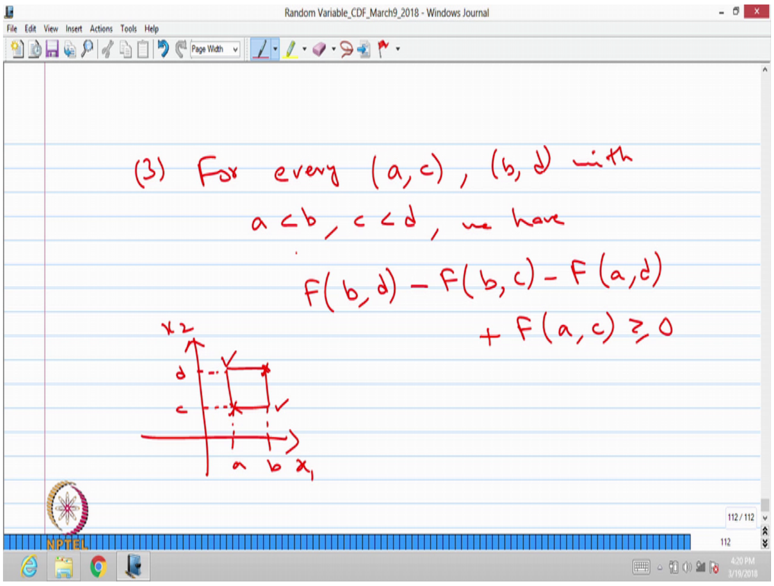


So, let me write down; it satisfies, it satisfies the first condition the capital F of x 1 comma x 2 that is non decreasing and continuous from the right.

That means it is a right continuous with respect to with respect to each coordinate, each coordinate that is x 1 x 2 it is a non-decreasing as well as a continuous from right with respect to each coordinate that is a first point. Second point the limit of x 1 tends to plus infinity x 2 tends to plus infinity; the CDF x 1 comma x 2 limit x 1 tends to plus infinity x 2 tends to plus infinity, F of x 1 comma x 2 is always 1.

And the limit x 2 tends to minus infinity of F of x 1 comma x 2 that is equal to 0 and the limit x 1 tends to minus infinity F of x 1 comma x 2 that is also 0. When both becomes a positive infinity, it becomes 1 when either one of them is minus infinity and the limit extends x 2 tends to minus infinity or limit x 1 tends to minus infinity that function of x 1 comma x 2 that is going to be 0.

(Refer Slide Time: 15:11)

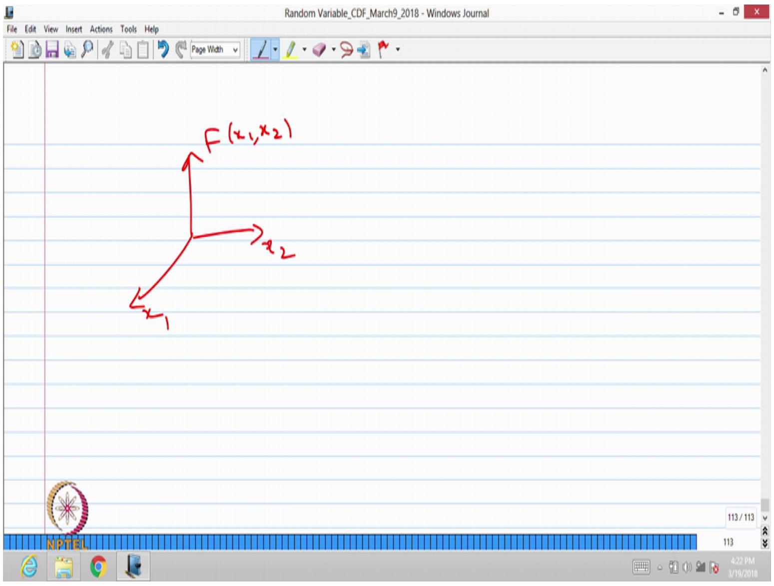


The third point for every a comma c and b comma d with a is lesser than b and c is lesser than d; we have I can just draw the diagram first then I can go for it that is easy. So, I can take as small box in which this is going to be a and this is going to be b and this is going to be c and this is going to be d correct. I can go for after drawing the diagram. Now I can go for F of b comma d minus F of b comma c minus F of a comma d plus F of a comma c which is greater or equal to 0.

So, the cross is going to have a positive symbol and a tick mark has a negative symbol; that means, F evaluated at b comma d minus F evaluated at a comma d and b comma c with the minus sign then plus F evaluated at a comma c that value has to be greater or equal to 0. Whenever you have a 2 dimensional random variable whose CDF always satisfies a these 3 conditions. The third condition is a very important condition in the sense even you may have a real valued function with a 2 variables satisfying these 2 conditions may not be a CDF of 2 dimensional random variable unless otherwise it satisfies a third condition.

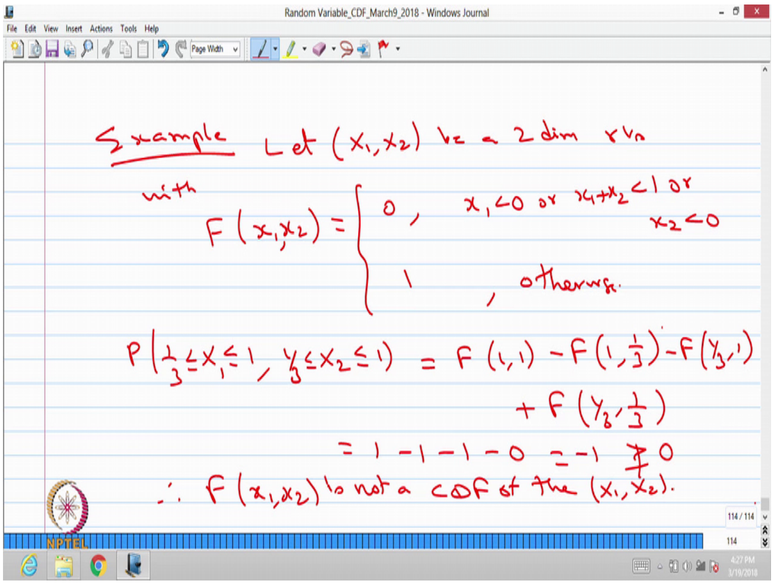
But if you have a 2 dimensional random variable, you will have a always unique CDF with the 2 variables that satisfies all these 3 conditions. CDF of 2 dimensional random variables can be represented in the form of graphical.

(Refer Slide Time: 18:01)



So, one can visualize the CDF in the form of x axis x 1, x 2 and this is going to be F of x 1 comma x 2. So, this is possible only for 2 dimensional x 1 is one random variable, x 2 is another random variable. So, z axis that is a CDF of x 1 comma x 2 whereas, you cannot visualize you cannot make a graphic representation of more than 2 dimensional random variable.

(Refer Slide Time: 18:44)



Let us go for one simple example in which we can conclude whether this is going to be CDF or not. As example let X 1 X 2 be a 2 dimensional random variables with CDF x 1 comma x 2 that is either 0 when x 1 is less than 0 or x 1 plus x 2 is lesser than 1 or x 2 is lesser than 0. It takes a value 1 otherwise verify whether the capital F is going to be the CDF of 2 dimensional random variable; that means, this is real valued function with a 2 variables.

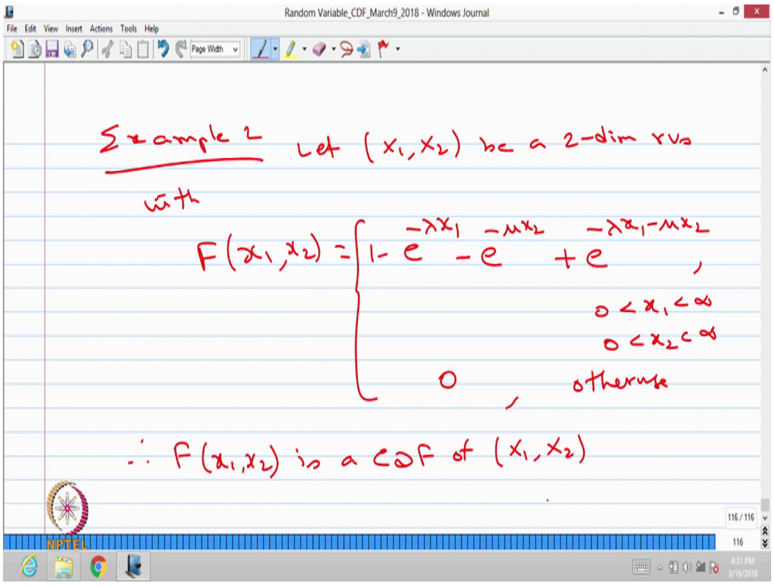
Whether this satisfies the 3 properties which we have given; if all these 3 properties are satisfied then you can conclude this will be the CDF of 2 dimensional random variable. You can easily verify the first 2 properties the function is a non decreasing as well as continuous from the right. Similarly, you can easily verify the limit of x 1 tends to plus infinity, x 2 tends to plus infinity that value is 1. Either x 1 is minus infinity or x 2 is minus infinity the limit is going to be 0 that also can be easily verified.

We have to verify the third condition that is the for any 2 points; the CDF of the difference had a 2 points with a positive sign, 2 points with the negative sign has to be greater or equal to 0. For example, you do it with the probability of X lies between X 1 lies between one third to 1 comma X 2 is lies between one third to 1.

If you compute this probability that is nothing but F at the point 1 comma 1 minus F at the point 1 comma one third; I am using the same property minus F one third comma 1 plus F one third comma one third. If you substitute the value F of 1 comma 1 that is 1, F of 1 comma one third that is again 1, F of one third comma 1 again 1, F of one third comma one third that is 0. And this value is going to be minus 1 which is not greater or equal to 0. So, for any arbitary points this third property has to be satisfied.

So, since the third property is not satisfied; you can conclude a this F is not a CDF of the random vector X 1 comma X 2 or the 2 dimensional random variables X 1 comma X 2.

(Refer Slide Time: 22:58)

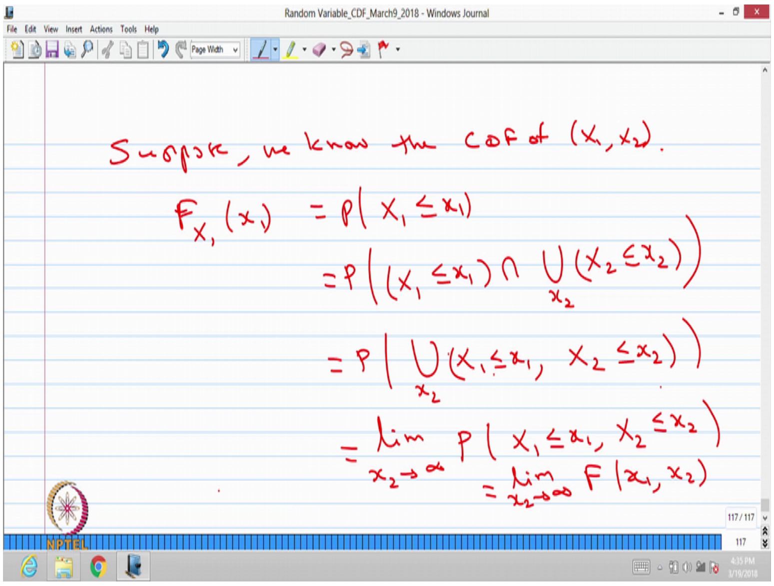


We will give some example in which that is going to be CDF. So, example 2 let X 1 comma X 2 be a 2 dimensional random variables with the function F of x 1 comma x 2 is 1 minus e power minus lambda x 1 minus e power minus mu times x 2 plus e power minus lambda x 1 minus mu x 2.

So, this is going to be the value when both x 1 and x 2 lies between 0 to infinity; otherwise it is 0. We can verify whether this is going to be the CDF of 2 dimensional random variables. By seeing the function you can easily say when x 1 and x 2 is the positive infinity, it becomes 1; either x 1 or x 2 is going to be minus infinity that is going to be 0. And it is a non decreasing function and continuous from the right therefore, the properties 1 and 2 are easily satisfied.

For different values one can able to verify the third property also will be satisfied therefore, hope one can conclude this is going to be the CDF of 2 dimensional random variables. So, I am not giving the proof of the third property. But that can be verified we can conclude this is going to be the CDF of 2 dimensional random variables.

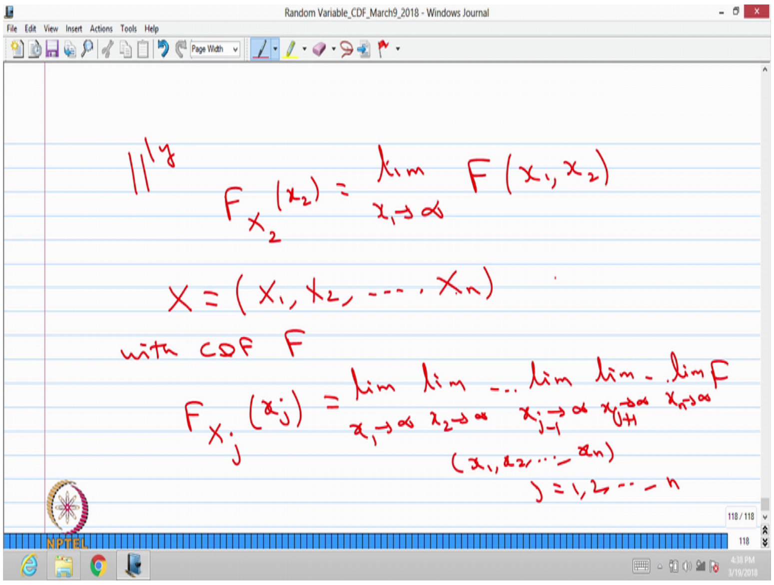
(Refer Slide Time: 25:08)



Suppose we know the CDF of X 1 comma X 2; one can able to find the CDF of any one random variable, that is one can find the CDF of random variable X 1 as a function of small x 1 that is nothing but probability of X 1 less than or equal to small x 1. That is same as probability of X 1 less than or equal to small x 1 which intersect all the union of X 2 less than or equal to small x 2; for all possible values of x 2. That is same as the probability of union of all possible values of small x 2, X 1 is less than or equal to x 1 and X 2 is less than or equal to small x 2.

That is same as the P of union is nothing but limit X 2 tends to infinity P of X 1 less than or equal to small x 1 capital X 2 less than or equal to small x 2. That is same as limit X 2 tends to infinity capital F of x 1 with x 2. That means, if you know the CDF of x 1 comma x 2 by taking limit of the other variable tends to plus infinity that will give the CDF of the other random variable.

(Refer Slide Time: 27:41)



Similarly, the CDF of the random variable X 2 has a function of small x 2 that is nothing but limit x 1 tends to plus infinity of CDF evaluated at the points x 1 comma x 2; so, this is valid for 2 dimensional random variable. So, this (Refer Time: 28:11) concept can be extended to n dimensional random variable. That means, suppose I denote capital X as a n dimensional random vector; instead of again and again writing n dimensional random variable, I can write capital X with the CDF capital F; that means, F also has a n variables. Then we can find the CDF of any one random variable suppose I want j th random variable.

Suppose, I want to find out the CDF of j-th random variable that is nothing but limit x 1 tends to infinity limit x 2 tends to infinity I assume that j is in between 1 to n. Therefore, limit x j minus 1 tends to infinity, limit x j plus 1 tends to infinity and so on, limit x n tends to infinity of capital F; which has a elements just I will write down in the next line x 1 comma x 2 x n.

So, this is the way one can get the CDF of 1 dimensional random variable from the CDF of n dimensional random variable, where j can be 1 2 and so on till n. In this exercise, I have not said whether the random variable is of the discrete type or the continuous type or mixed type. So, based on the each random variable or of the discrete continuous or mixed; one can discuss the joint probability mass function or joint probability density function and so on. So, at present we do not know; what is a type of each random variable.

Therefore, we stopped it at the CDF of the n dimensional random variable. Once we know the type of each random variable, then one can go for studying the probability mass function jointly, probability density function jointly that is going to be the second lecture.