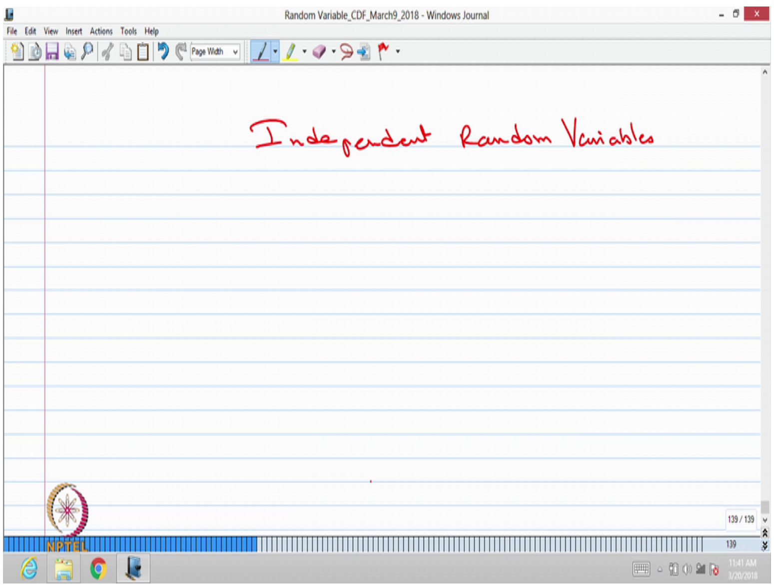
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture - 27

In this lecture we are going to discuss the independent random variables. When we discuss many random variables, sometimes some random variable they may have a relations within themselves or sometimes may not. So, it is important to study whether these random variables are having some dependency or not. So, this dependency can be studied by using a nice mathematical way through the CDF that is a joint CDF. If the joint CDF of two dimensional random variable or n dimensional random variable satisfy some conditions then we can conclude those random variables are independent.

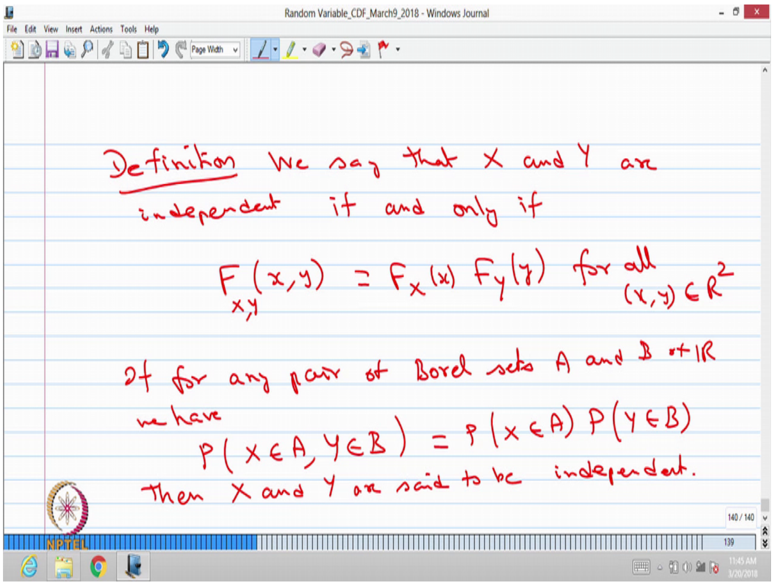
Then the next question comes why do you need to study the independent of random variables. If the random variables are independent, then some of the prediction or some of the sum of finding the probabilities of those random variables will be easy when those random variables are independent. So, let me start with the concept called independent random variables.

(Refer Slide Time: 01:21)



Let me start with the definition of independent random variables, then few more properties when this random variables are independent at the end I will give one or 2 examples for the conceptual understanding the definition.

(Refer Slide Time: 01:54)



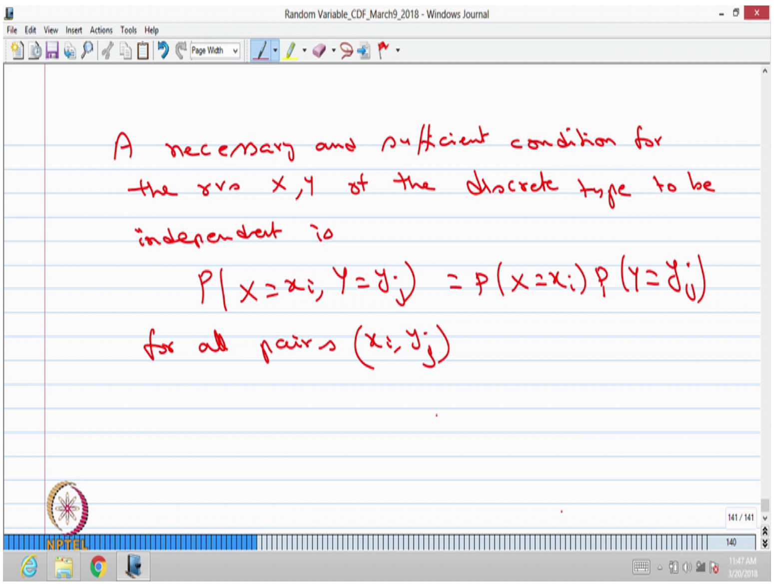
Let me start with the two dimensional random variable, then the same concept can be extended to the n dimensional random variable. So, that way it is easy to explain the concept. We say that the random variable X and the random variable Y or independent if and only if the CDF of two dimensional random variable that is same as the CDF of one dimensional random variable with the product. Whenever I write the suffix; that means, the CDF is corresponding to that random variables. So, F suffix x comma y; that means, it is the CDF of two dimensional random variable F suffix x.

That means x is the CDF of the random variable x, F suffix capital Y; that means, CDF of the random variable y. If the product of CDFs of the random variable x and y that is same as the CDF of two dimensional random variable x comma y for all x comma y in R 2. Then we can conclude this is a independent random variables they say if and only if condition; that means, if two random variables are independent this condition will be satisfied, if this condition is satisfied then we can conclude both the random variables are independent. It is a immaterial of the random variables are of the discrete type or continuous type or mixed type because the CDF is always exist whatever be the type of random variable. Here we are saying the two random variables are independent if and only if the condition on the CDFs.

That means, if I have a for any pair of Boral sets. Suppose I keep the Boral set A and another Boral set B of real line then we have if two random variables are independent, then the probability of x belonging to the Boral set A and y belonging to the Boral set B that is same as the probability of x belonging to the Boral set A multiplied by the probability of y belonging to Boral set B. That means, if two random variables are independent if and only if conditions is satisfied.

Whenever two random variables are independent, then we can always get the probability of x belonging to Boral set and y belonging to another Boral set that is same as X belonging to the one Boral set multiplied by probability of Y belonging to the other Boral set. This is if this condition is satisfied then X and Y or set b independent whereas, the first condition is if and only if condition we are not saying the random variable is a discrete type or continuous type. Now, I am going to make a condition on whether the random variables are of the discrete type or continuous type and how this if and only if condition changes that is a necessary.

(Refer Slide Time: 06:27)

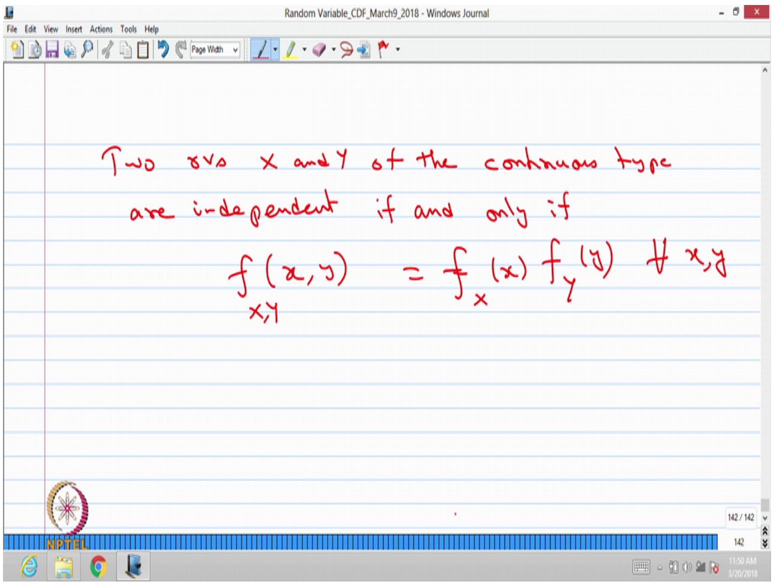


And sufficient condition for that random variables x and y of the discrete type to be independent is the joint probability mass function of x and y.

That is same as the product of the probability mass function of x with the probability mass function of y this is for all pairs xi comma yj. That means, this is a if and only if condition if both the random variables are of the discrete type, then the c condition independent condition of CDF can be replaced by the independent condition on the joint probability mass function or joint probability mass function is same as product of probability mass functions of x and y. This is also if and only if condition; that means, if two random variables are of the discrete type or independent then this condition will be satisfied we call this condition as a independent condition.

Similarly, if this condition is satisfied for all pairs, then both the random variables are of the discrete type and they are independent.

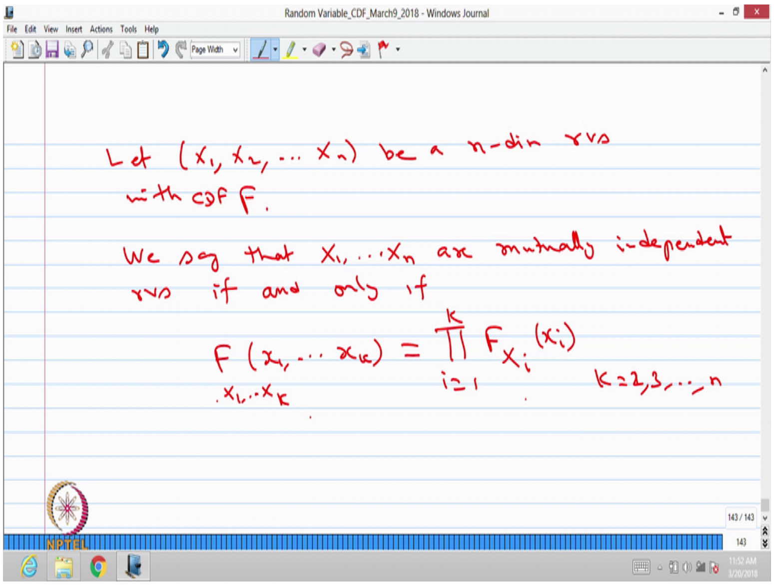
(Refer Slide Time: 08:40)



The similar results is for the continuous type also. two random variables x and y of the continuous type or independent, if and only if and only if you have to replace the condition of the CDF by the probability density function. So, the joint probability density function when I write suffix x comma y; that means, it is a joint probability density function of x comma y, that is same as a the probability density function of x multiplied by probability density function of y this is for all x comma y. So, if this condition is satisfied then two random variables are of the continuous type or independent.

If two random variables are of the continuous type or independent, then this condition will be satisfied for all x comma y. Therefore, the condition for independent random variables either in the level of a CDF or if it is a discrete random variable in the form of probability mass function if the random variables are of the continuous type, then it is a probability density function for. So, all are all 3 are going to be if and only if condition not the only one side it is in the both side. So, even though we have explained through the 1 d two dimensional random variables this can be extended to the multi dimensional random variable also.

(Refer Slide Time: 10:49)



That means, let x 1 comma x 2 comma so, on be a n dimensional random variables with the CDF capital F is a function of n variables. We say that the random variable x 1 x 2 xn are mutually independent are mutually independent random variables, if and only if you take any fewer random variables CDF k, for K is equal to 2 3 and so, on till n, that is same as product of i is equal to 1 to k the CDF of those random variables CDF. This means if you take any two random variables the CDF of those two random variable is same as a product of a CDF of only those two random variables.

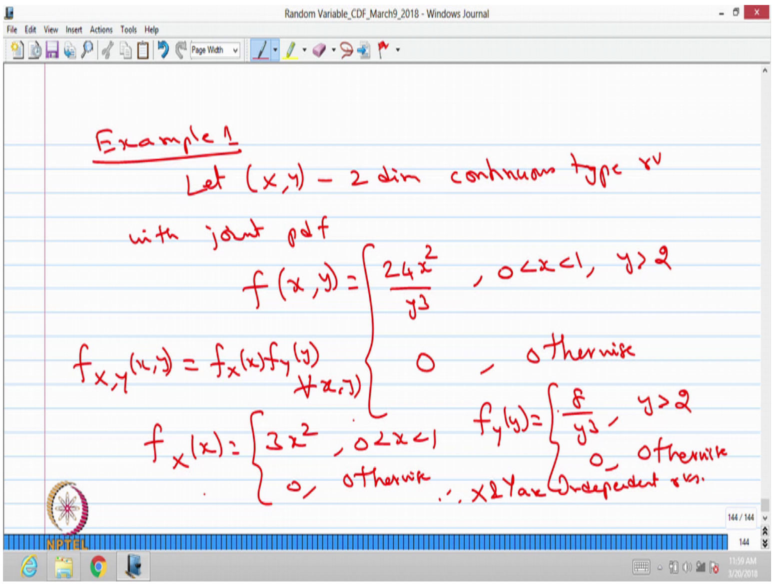
If I take any 5 random variables, when n is greater than 5 then CDF of 5 random variables is same as product of those 5 random variables CDF, then we conclude they are mutually independent. It is same as the mutually independent it is same as the mutually independent events. If you have a n events and once you say that they are pairwise dependent. That means, any 2 ra events satisfies the independent concept or independent condition then they are called pairwise. If it means mutually if they are mutually independent; that means, whatever be the collection of events you take the independent condition is satisfied then we conclude their mutually independent events.

The same thing here if you have n random variables whatever be the number of random variable you take it from those n random variables, that satisfies independent condition then it start from any 2 till all the random variable then we conclude they are mutually independent. Whenever we say more than two random variables are independent; that means, by default they are mutually independent whenever we say more than two random variables are independent; that means, by default they are mutually independent.

Many times we would not write again and again mutually independent word when we discuss more than two random variables, but whenever we have more than two random variables when we use the word independent random variable. That means, they are mutually dependent; that means, the independent condition is satisfied for all the forms of collection of random variables, satisfying the independent condition.

Now, we will move into some problem of explaining how the independent random variable is playing a role.

(Refer Slide Time: 14:41)



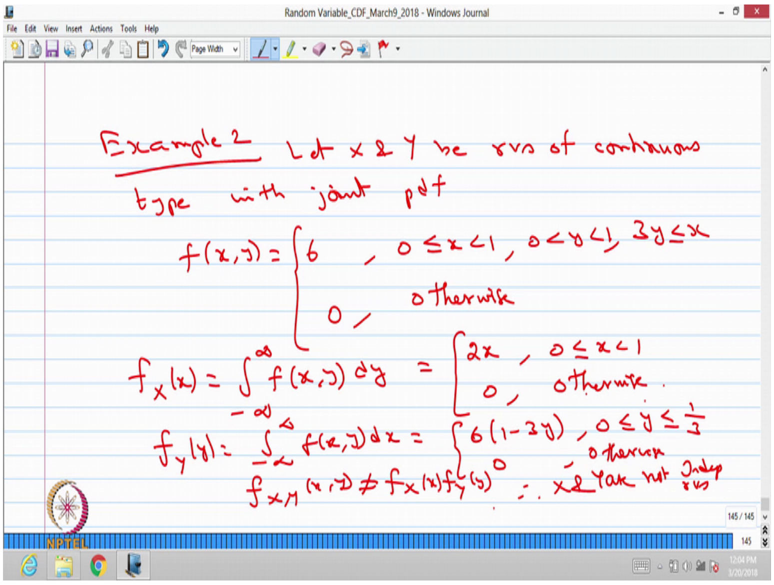
The first example that is same as the example which we have considered earlier, that is let x comma y be a continuous type random variable two dimensional continuous type random variable with joint probability density function is of the form f of x comma y, that is after we find the value of k we got a 24 x square divided by y cube, when x is lies between 0 to 1 and y is greater than 2 otherwise it is 0. So, this problem just now we have discussed when two random variables are of the continuous type for the 2 we found the k, that k value was 24 and we found the probability of x between some interval.

Here, we will verify whether these two random variables are independent or not. So, this is the joint probability density function already in the same example we got the probability density function of x, that is 3 times x square when x is lies between 0 to 2 1 otherwise 0. Similarly we got the probability density function of y that is 8 divided by y cube when y is greater than 2 0 otherwise.

Easily it can be verified in this example the joint probability density function is same as product of probability function of random variables x comma y. Because it is 3 times x square the other one is 8 divided by y cube, not only that the 3 x square is a range between 0 to 1 and 8 divided by y cube the range is great y is greater than 2, which is same as if you make a product that is 24 x square divided by y cube and the range of x is 0 to one and the range of y is 2 to infinity that is same as the joint probability density function of a 24 x square divided by y cube when x lies between 0 to 1 and y is greater than 2.

So, their interval matches and the value matches 0 otherwise matches therefore, for all x comma y the joint probability density function is immersed the product of density functions of x and y therefore, these two random variables independent. Therefore, x and y are independent random variables. Each one is of the continuous type, we can check it from the CDF also that is if and only if condition for the CDF. But, since the joint probability density function is given you can find the probability density function of x and y, then you can verify the independent conditions on probability density function that is satisfied. Therefore, both the random variables are independent.

(Refer Slide Time: 18:40)



We will go for one more example, example 2 again we have a continuous type let x and y be a random variables of continuous type with joint probability density function that is f of x comma y, which takes a value 6 when x is lies between 0 to 1 and y is lies between 0 to 1 as well as 3 y is less than x less than or equal to x. So, the joint probability density function is greater than 0, that is 6 when x is lies between 0 to 1, y is lies between 0 to 1 and 3 y is less than or equal to x 0 otherwise.

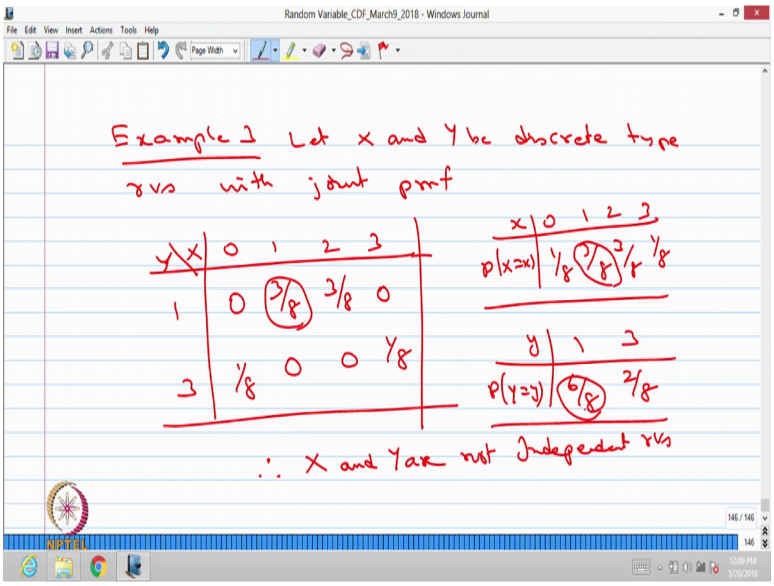
Before we proceed the problem we can always verify whether this is correct joint probability density function; that means, if you integrate if you integrate double integration over x and y this has to be 1. So, one can verify this is going to be double integration is one. Therefore, this is the correct joint probability density function. From these one can find the marginal distribution of x by integrating a the joint probability density function with respect to y. If you do the little simplification you can get the answer that is 2 times x when x is lies between 0 to one otherwise it is 0.

You can verify this result also whether this is a correct probability density function of x by integrating 0 to 1 to x you will get the value 1 and it is greater than or equal to 0 therefore, this is the probability density function of x. Similarly you can compute the probability density function of y by integrating the joint probability density function with respect to x. Here also am skipping the integration part one can get the answer that is a 6 times 1 minus 3 y, when y is lies between 0 to one third 0 otherwise; that means, within this interval 0 to one third the probability density function is greater than 0, 6 times 1 minus 3 y otherwise 0.

By seeing the probability density function of x probability density function of y this product is not going to be the joint; that means, the f of x comma y is not equal to the product of probability density function of x and y. For all x comma y if this condition is satisfied equal to then you can conclude they are independent.

But since by seeing this you can say it is a 2 times x 6 times one minus 3 y whereas, the nine probability density function is 6 obviously you can say they are not equal. Therefore, x and y are not independent random variables. So, I have given the first example which in which they are independent random variable by finding the marginal distribution whereas, in this example by finding the marginal distribution of x and y, we are concluding condition is not satisfied independent condition is not satisfied therefore they are not independent random variables.

(Refer Slide Time: 23:27)



We will go to the one more example that is example 3, because already we discussed 2 problems of the continuous type we will see one problem of the discrete type also, that is let x and y be a discrete type.

Let x and y be discrete type random variables with joint probability mass function is given by. If you recall this is same as the problem which we have discussed in the last class, the possible values of x is 0 1 2 3 and the possible values of y is 1 and 3 where x denotes a number of heads obtained when we tossing a unbiased coin 3 times. And y is the difference in absolute of a number of heads and number of tails of time therefore, the possible values of y is 1 and 3 and the possible values of x is 0 1 2 and 3, and in that problem we have got the joint probability mass function that is 0 3 by 8, 3 by 8, 0 then 1 by 8, 0 0 and 1 by 8 and in that problem we have got the marginal distribution of x and y also if you recall.

So, for possible values of x that is 0 1 2 and 3 and the probability of x takes a value of x that is going to be 1 by 8, 3 by 8 again 3 by 8 and 1 by 8. So, this is a probability mass function of x and similarly probability mass function of y that is 1 and 3. So, 1 it takes a value 6 by 8, and 3 takes a value 2 by 8. So, this is the probability mass function of y. Now you can verify whether this two random variables are independent suppose x takes a value 1 y takes a value one that probabilities 3 by 8.

Here x takes a value 1 and y takes a value 1 that is this much. So, if you make product of 3 by 8 into 6 by 8 that is not equal to 3 by 8. Even at one pair it is not satisfied then you cannot conclude it is independent random variable. If all the pairs the joint probability mass function is same as product of probability mass functions of x and y, then only you can conclude their independent. Since any one pair does not satisfy then you can immediately conclude both the random variables are not independent.

This will be a set of obvious because the random variable y is defined as difference in absolute with the number of heads and number of tails whereas, the random variable x is defined number of heads; that means, the y itself is a function of x; that means, y is at dependent on x. Therefore, there is a dependency between the random variable y and x in the definition itself. Therefore, from there itself you can conclude they are not a independent random variable, but that we have concluded from the distributions also.