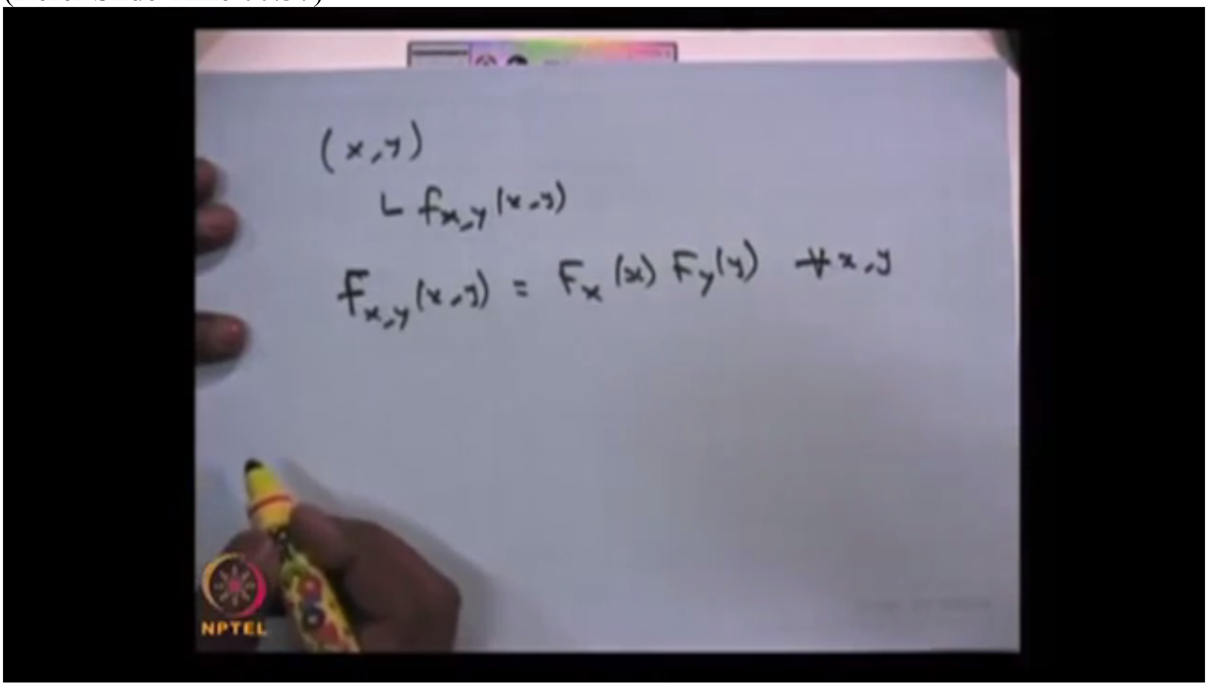
Independent Random Variables, Covariance and Correlation Coefficient and Conditional Distribution

Dr. S Dharmaraja
Department of Mathematics, IIT Delhi

Now I'm going to discuss what is the meaning of independent random variable? Suppose you have two random variables x and y and you know what is a joint probability density function or joint probability mass function based on the random variable, both are discrete or continuous, then if both the random variables are independent, then the CDF of this random variable, random vector is same as the product of CDFs of individual random variable whether it is a discrete random variable or continuous random variable and this is valid for all (x, y).

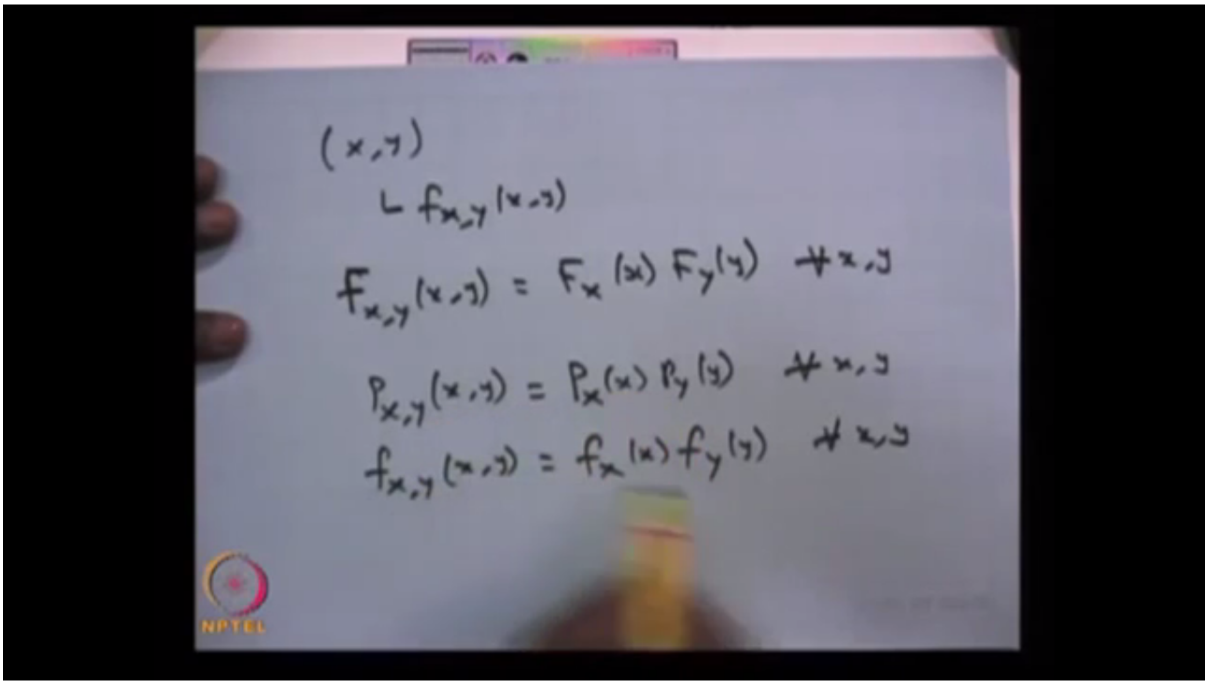
(Refer Slide Time 00:37)



That means if you have two random variables and this satisfied for all (x, y), that means the joint CDF is same as the product of CDF. This is basically a if and only if condition. If this condition is satisfied, then both the random variables are call it as a independent random variable. So this -- suppose these random variables are -- both are discrete, then you can come down from the CDF into the joint probability mass function.

The joint probability mass function, you can write it as the product of individual probability mass function for all (x, y). If both the random variables are continuous, then you have a joint probability density function, the same joint probability density function will be the product of individual probability density function.

(Refer Slide Time 01:29)



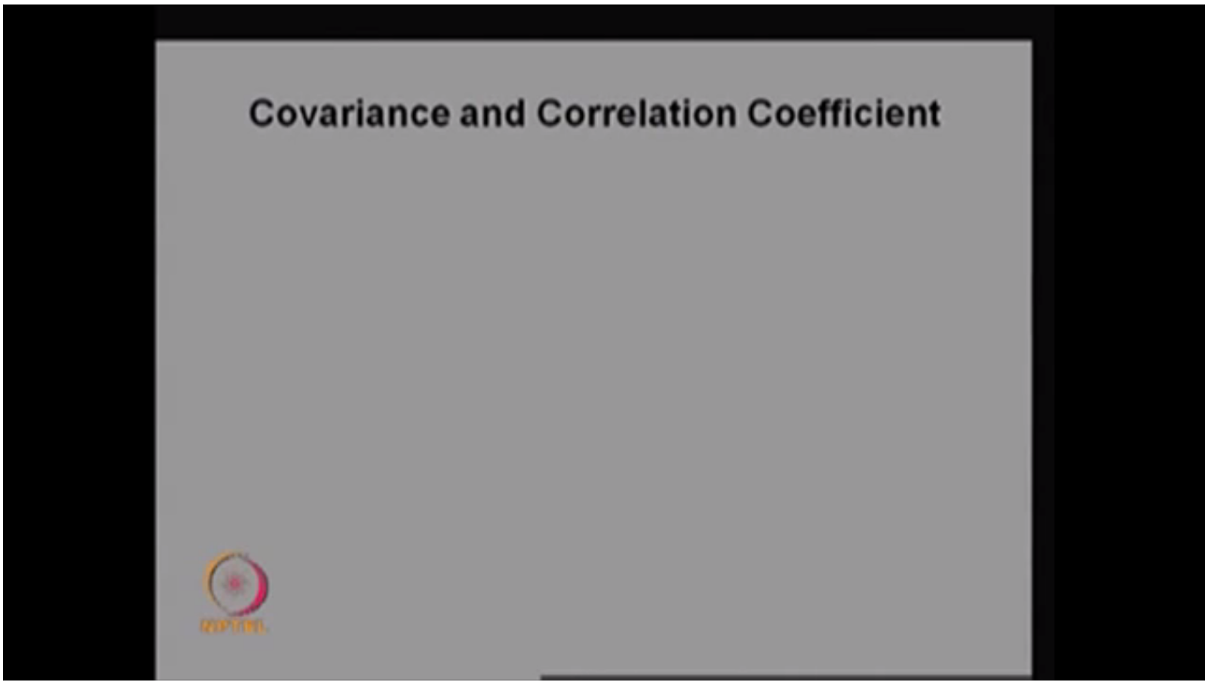
That means based on the random variables discrete or continuous, you can cross check whether this property is satisfied. So if this -- this property is satisfied, then you can conclude that random variables are independent.

Similarly, if the random variables are independent, then this property is going to be satisfied. So whether it is a discrete or continuous, you can always check in the CDF level also. If the CDF, joint CDFs and the individual CDF satisfies this property, then you can conclude that both the random variables are going to be independent random variable.

And this logic can be extended for the any n random variables. So instead of two random variables, you can go for having a n random variables, then finding out what is a joint CDF. If the joint CDF of n-dimensional random variable is going to be the product of individual random variable, then you can conclude both -- all n random variables are mutually independent random variables.

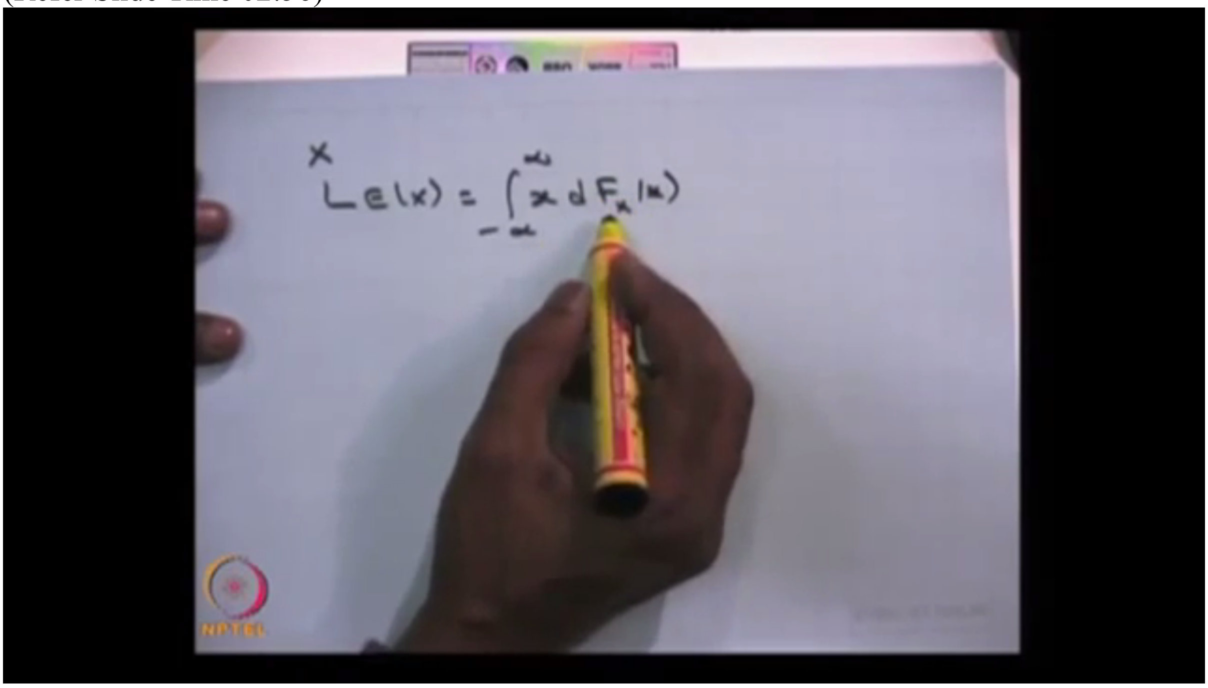
Now we are moving into the next concept.

(Refer Slide Time 02:31)



There are some moments we can find out from the random variable. The way you are computing, suppose you have a random variable x, you can able to find out the expectation of x if it exists. That means if the random variable x is there, you can always write expectation of x is from minus infinity to infinity x times d of F of x where capital F is the CDF of the random variable.

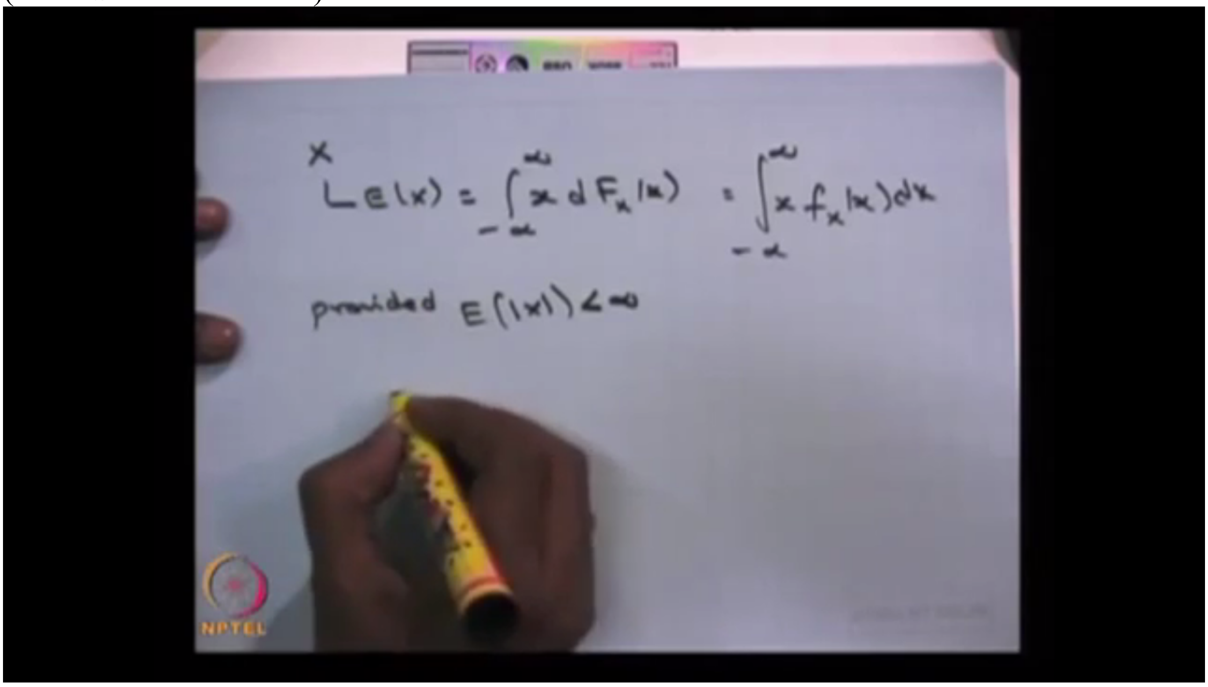
(Refer Slide Time 02:56)



So whether the random variable is a discrete or continuous or mixed type, if this integration is going to be exist, then you can able to give expectation is equal to this much. If the integration does not converges or that means the integration diverges, then you cannot go for writing expectation of x.

Suppose the random variable is a continuous random variable, then the CDF is going to be the continuous function. Therefore, this is same as -- this is same as minus infinity to infinity x times F of x dx if the random variable is a continuous random variable. In that case also, we have to cross check whether this integration is going to be see provided -- provided it says expectation of absolute x is converges.

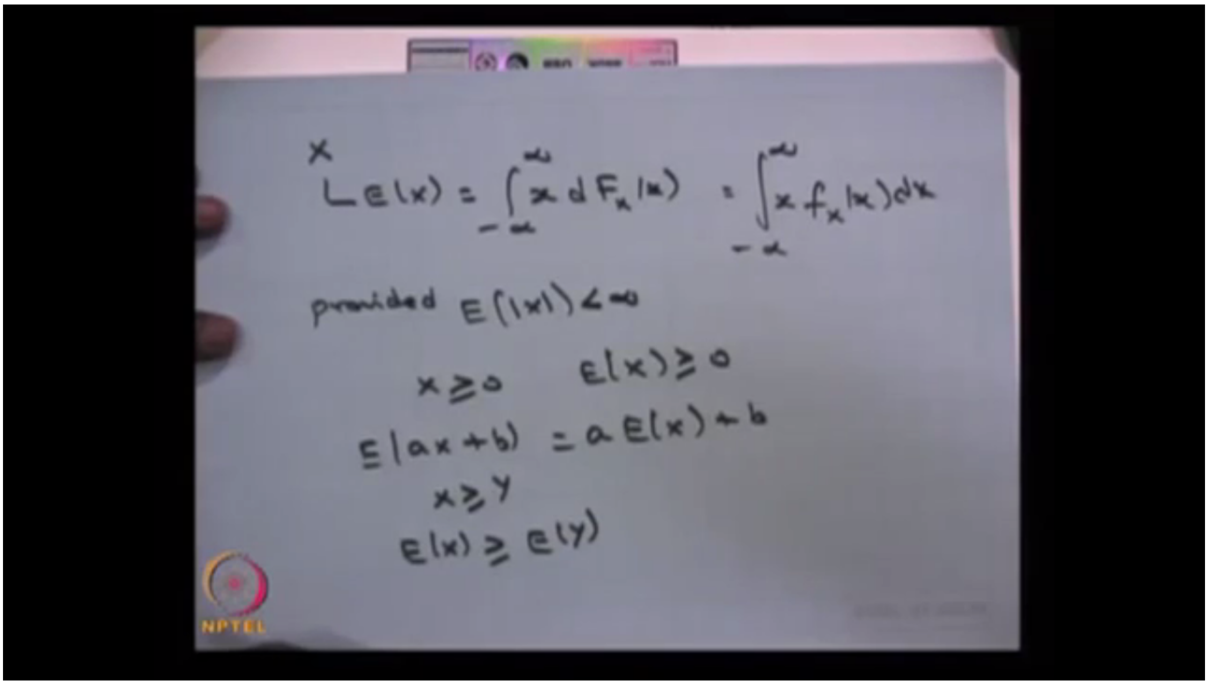
(Refer Slide Time 03:45)



This is because absolute convergence implies convergence. That means whenever you replace x by absolute of x and you find out if this provided condition is satisfied, then without absolute whatever the quantity you are going to get in the -- if it is the continuous random variable, the integration is converge -- whatever the value you are going to get, that is going to be the expectation of the random variable.

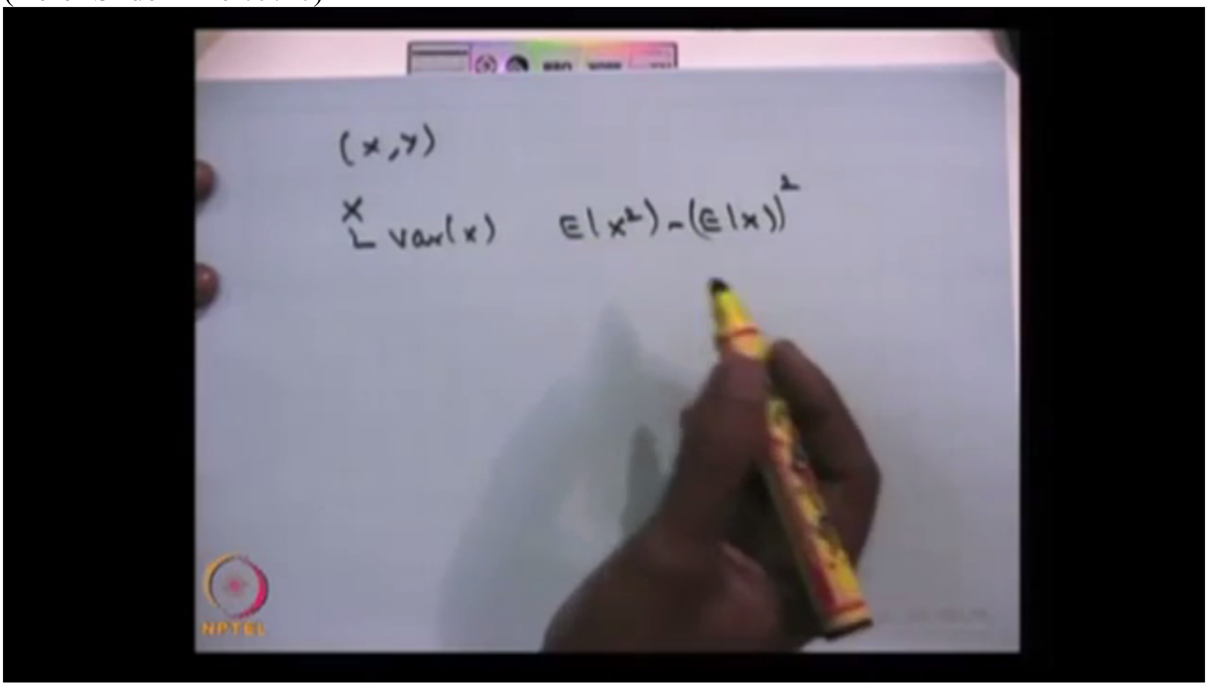
So the expectation of the random variable has -- expectation as a few property. This is going to be always a constant. This is not a random variable and the expectation of x if the random variable is greater than or equal to 0, then the expectation of x is always greater than or equal to 0, and the expectation of x has a linear property. If you have two random variables, then the expectation of x is greater than or equal to expectation of y.

(Refer Slide Time 04:41)



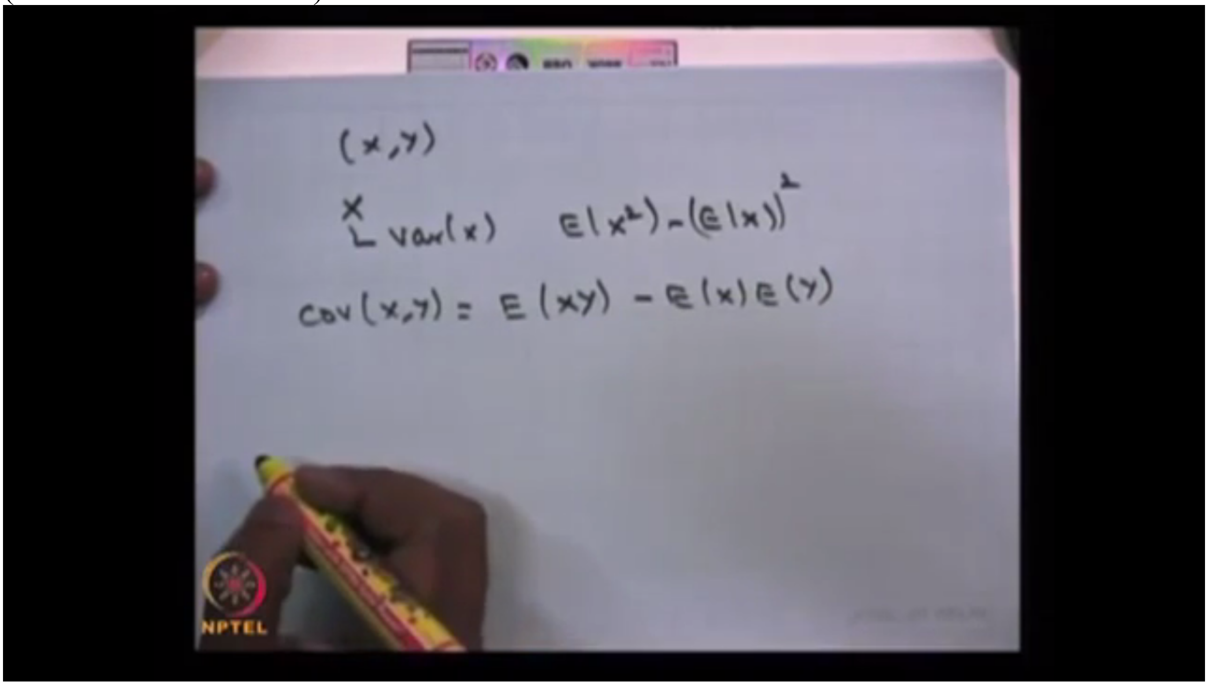
So now we are going to discuss since we have more random variables, we are going to discuss what is the -- what is other than expectation, we can go for finding out the variance of the random variable also. Variance is nothing but the second order moment. That is E of x square minus E of x whole square. So here also as long as E of x square, that means the expectation in absolute x if that is converges, then you can able to get expectation of E of x square and once you have a second order moment is exists, obviously, the all the previous order moment exists, but that does not imply the further moment exists.

(Refer Slide Time 05:26)



So now I'm going to define what is the covariance of (x, y)? So covariance of (x, y) is nothing but expectation of x into y minus expectation of x into expectation of x provided the expectation exist.

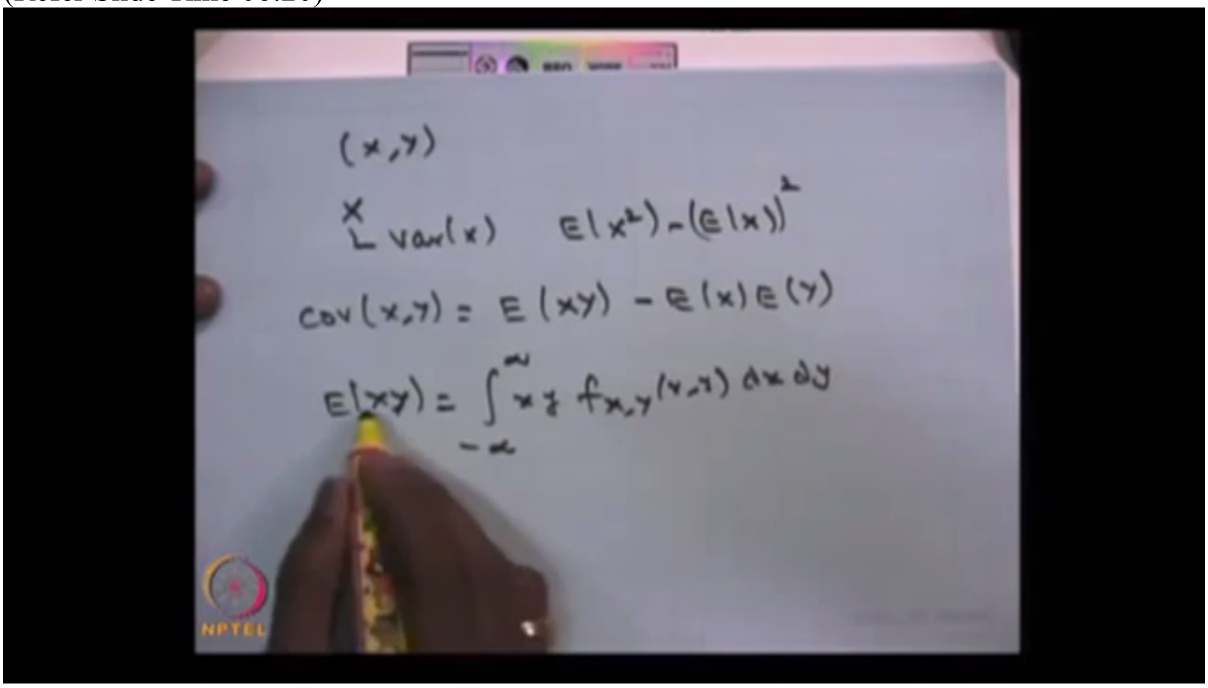
(Refer Slide Time 05:41)



So here it is expectation of x into y. That means you have to find out what is the expectation of x into y based on the random variable is a discrete or continuous, you can able to use functions of random variable method and getting the expectation. And note that even you don't know the distribution of the x into y, you can always find out the expectation of x into y.

Let me give a one situation. If both the random variables are continuous, then the expectation of x into y is going to be x into y and the joint probability density function of f of x f of y.

(Refer Slide Time 06:20)

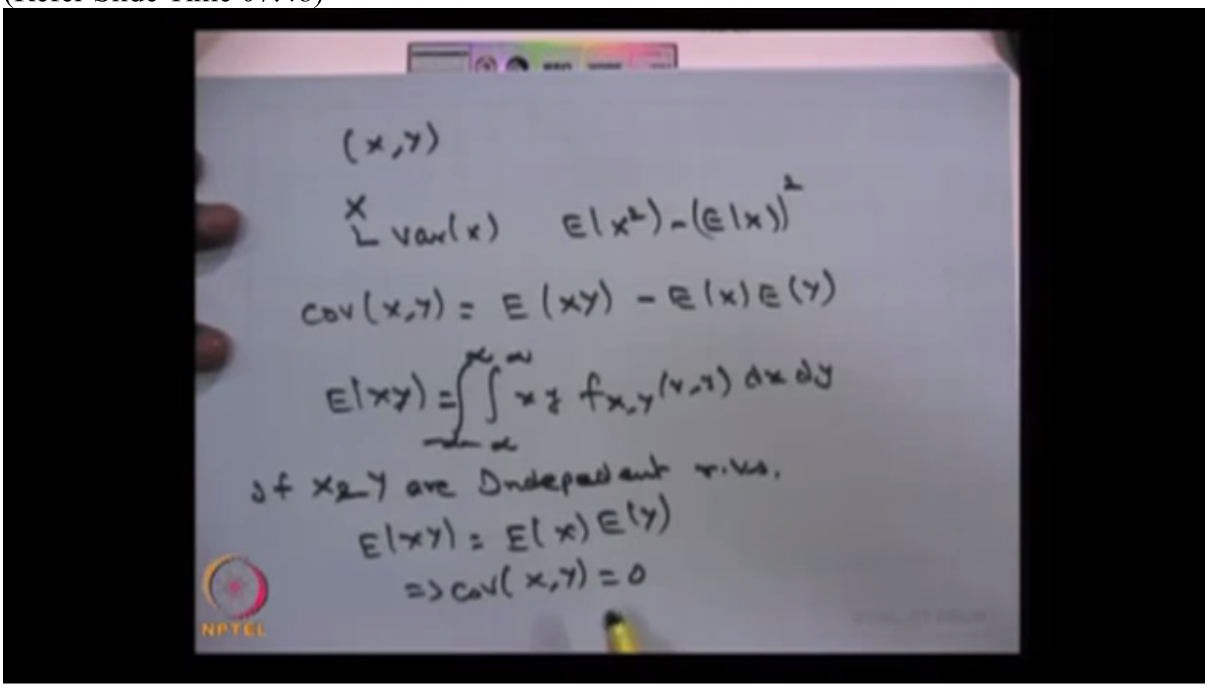


That means this is going to be the value of x, xy and what is the joint distribution of this? That means you are not finding out what is a distribution of xy, but you can -- still you can find out the expectation of xy by possible values and corresponding joint distribution you can get the expectation and here also provided the absolute sense exists, then without absolute sense it is going to be E of x into y.

Suppose x and y are independent random variable, independent -- independent random variables, then the expectation of x into y, the way I have given the situation with the both the random variables are continuous, this integration will be splitted into the two parts such a way that the f of (x, y), this is going to be, sorry, it's minus infinity to infinity, so this is going to be the product of individual probability mass function.

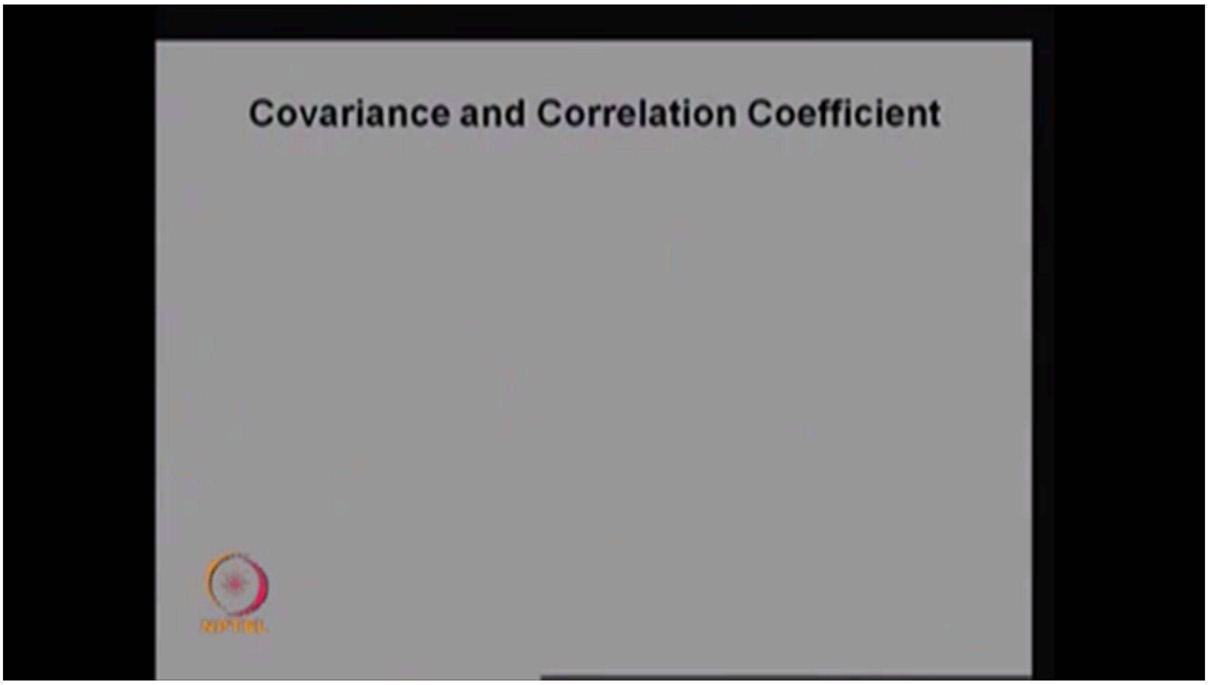
Therefore this is going to be integration will be splitted into two single integration minus infinity to infinity x times f of x and minus infinity to infinity y times f of y dy. Therefore that is nothing but the expectation of x into expectation of y. That means if two random variables are independent then implies the covariance of the two random variables is going to be zero.

(Refer Slide Time 07:48)



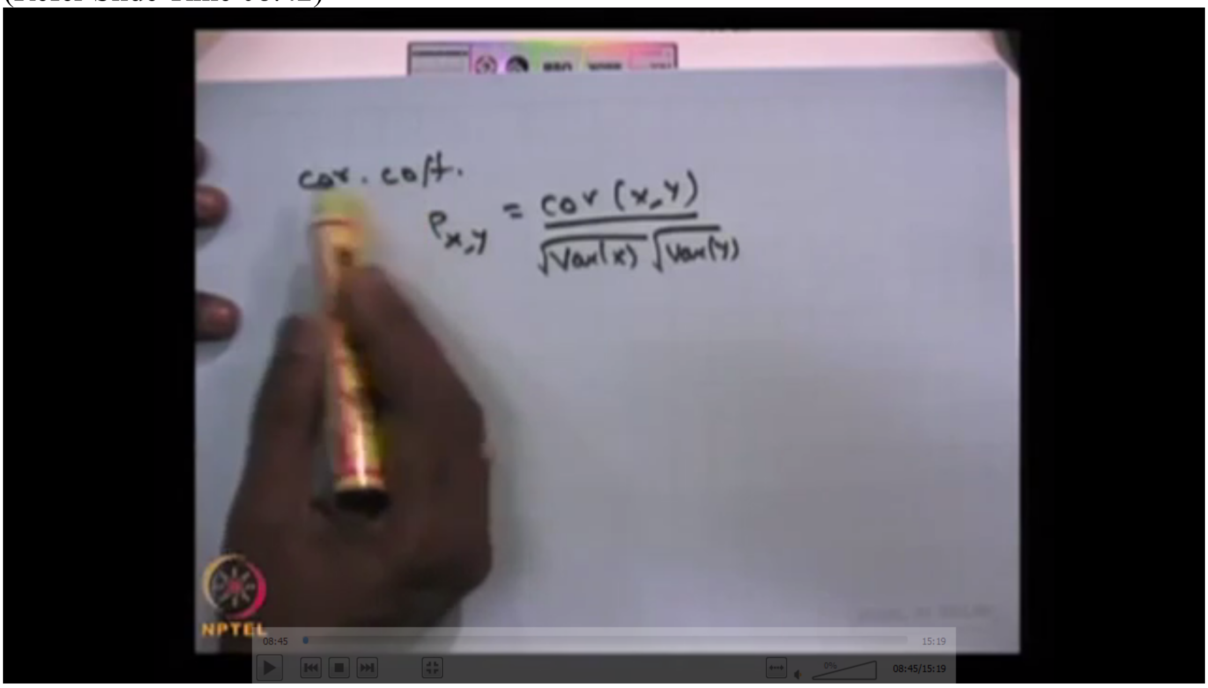
But the covariance of (x, y) equal to zero, that does not imply the random variables are independent. So this is going to be the not if and only if. If the random variables are independent, then you will come to the conclusion the covariance of (x, y) equal to zero. Not the converse.

(Refer Slide Time 08:16)



Now we are going to define the another measure that is the correlation coefficient. That is nothing but the correlation coefficient is nothing but with the letter Rho, Rho of (x, y). That means I am trying to find out what is the correlation coefficient between the random variables (x, y). That is nothing but the covariance of (x, y) divided by the square root of variance of x into square root of variance of y.

(Refer Slide Time 08:42)

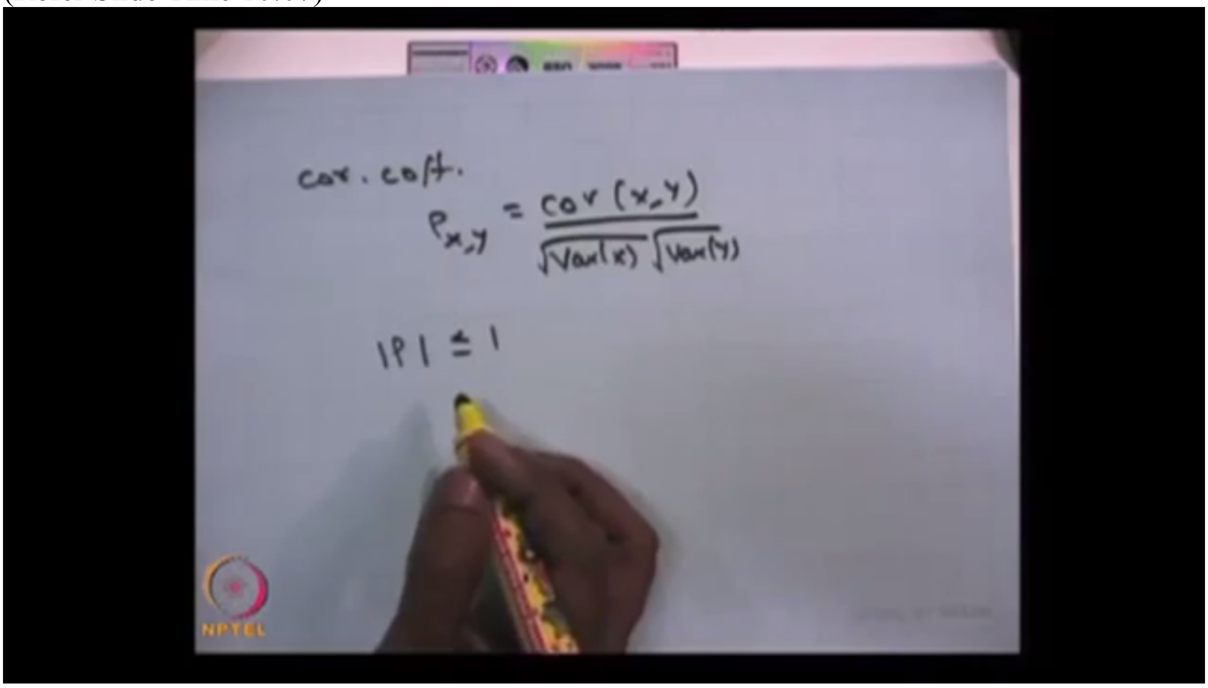


That means to have a existence of the covariance correlation coefficient, you should have -- that random variable should have at least the second order moment. So unless otherwise the second order moment does not exist, you cannot find out the correlation coefficient between these two random variables x, y because you are using the variance as well as the covariance to -- therefore, if the random variables are independent, then obviously the rho equal to zero because the numerator is going to be zero.

And since you are dividing the covariance divided by the square root of variance in x as well as y, this quantity in absolute is always lies -- is less than or equal to 1 or the rho is lies between minus 1 to 1, and the way the correlation coefficient value lies between 0 to 1, that conclude it is a positive correlated and the values lies between minus 1 to 0, it gives a negatively correlated and if the value is positive 1 or minus 1, then you can conclude the random variables x and y are linearly correlated based on the value is a positive side or the negative side, then you can conclude it is positively correlated or negatively correlated.

So other than the value minus 1 and 1, you cannot conclude what is the relation between the random variable. Only if it is 1 and minus 1, then you can conclude the random variables are correlated in the linear way.

(Refer Slide Time 10:07)

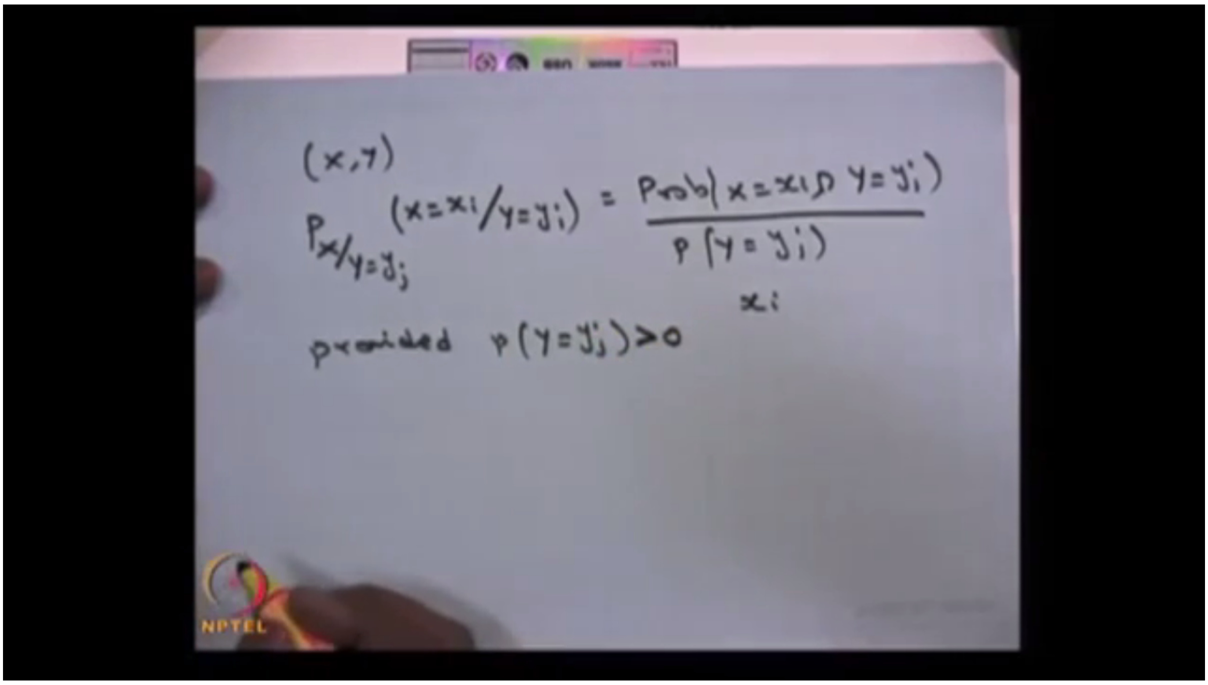


(Refer Slide Time 10:18)

Now I'm going to discuss conditional distribution because these are all the concepts are needed when you start defining some of the properties in the stochastic process. So, therefore, I am just discussing what is conditional distribution?

Suppose you have a two-dimensional random variable (x, y), you can define, suppose I make the one more assumption both are a discrete type random variable, then I can define what is the conditional distribution of the random variable x given that y takes some value y_j and here x takes the value x_i given that y takes a value y_j. That is nothing but what is the -- I can compute by finding what is the probability that x takes a value x_i intersection with the y takes a value y_j divided by what is the probability that the y takes a value y_j and here the running index is for all x_i's and this is for fixed y_j. Therefore, the provided condition provided the probability of y takes a value y_j has to be strictly greater than zero.

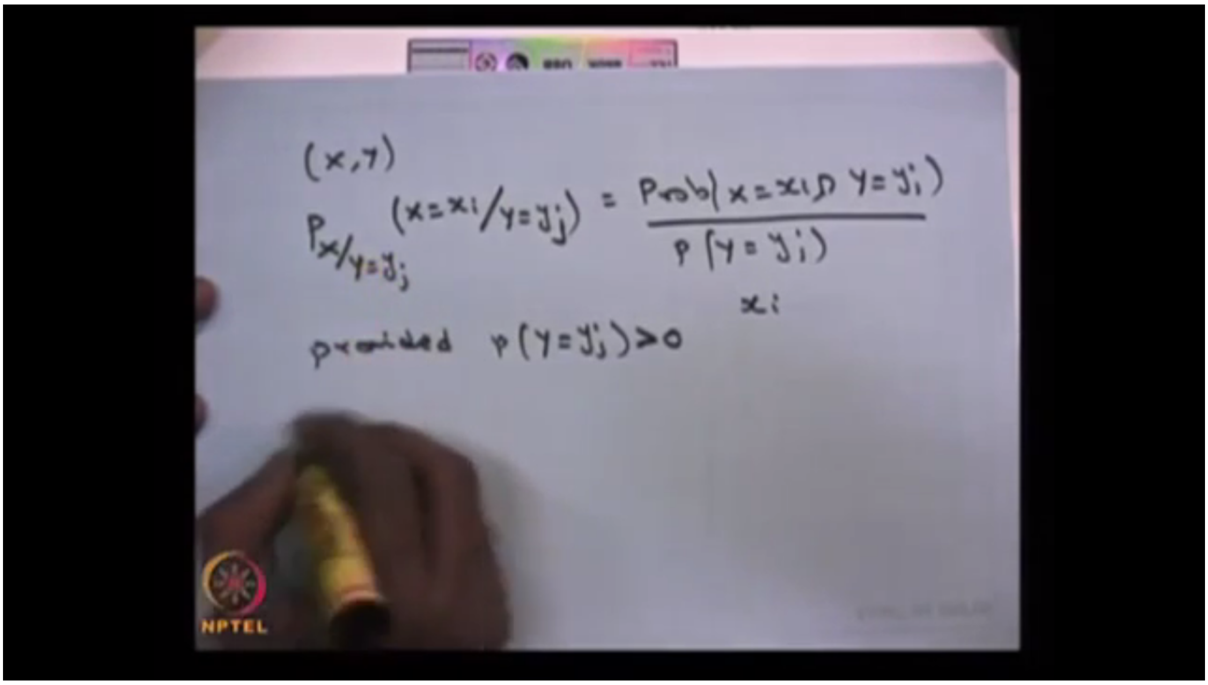
(Refer Slide Time 11:29)



That means you are making -- the way you made a conditional probability over the event, the same way we are making, this is going to be the event y is equal to y_j. So as long as the probability of the event corresponding to y is equal to y_j is strictly greater than 0 that means it is not an impossible event with the probability 0. It is the event which has the positive probability. If this happens already, then what is the probability of the random variable x takes the value x_i?

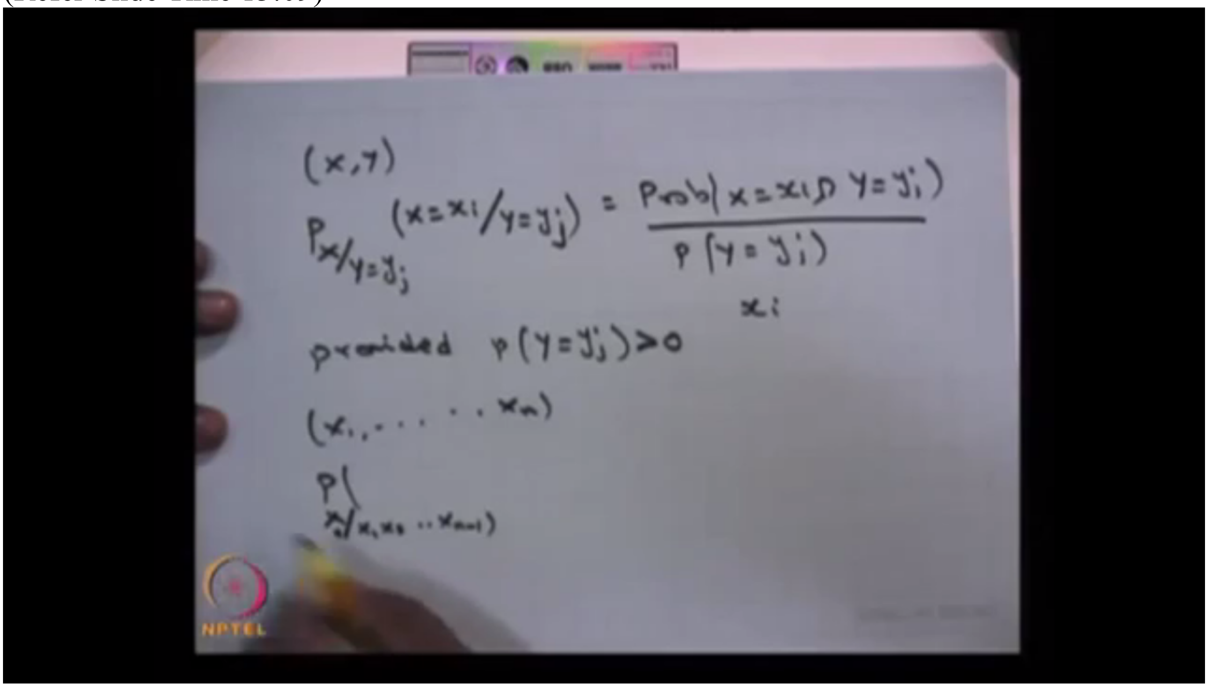
That means still our interest is to find out the distribution of x only, the random variable x with the provided or given situation that the other random variable y takes the value y_j. That means from the omega, you land up having one reduced sample space that corresponding to y is equal to y_j and from the reduced sample space, you are trying to find out what is the distribution of the random variable x for all possible values of x_i.

(Refer Slide Time 12:34)



So this we call it as a conditional distribution of x given the other random variable and this logic can be extended for more random variables. That means if you have n discrete random variable, then you can always define, suppose you have a x_1, x_2, x_n and suppose all are discrete random variable, then you can always define what is the probability distribution of x given what is the distribution of x_n given you know the distribution of x_1, x_2 till x_{n-1}.

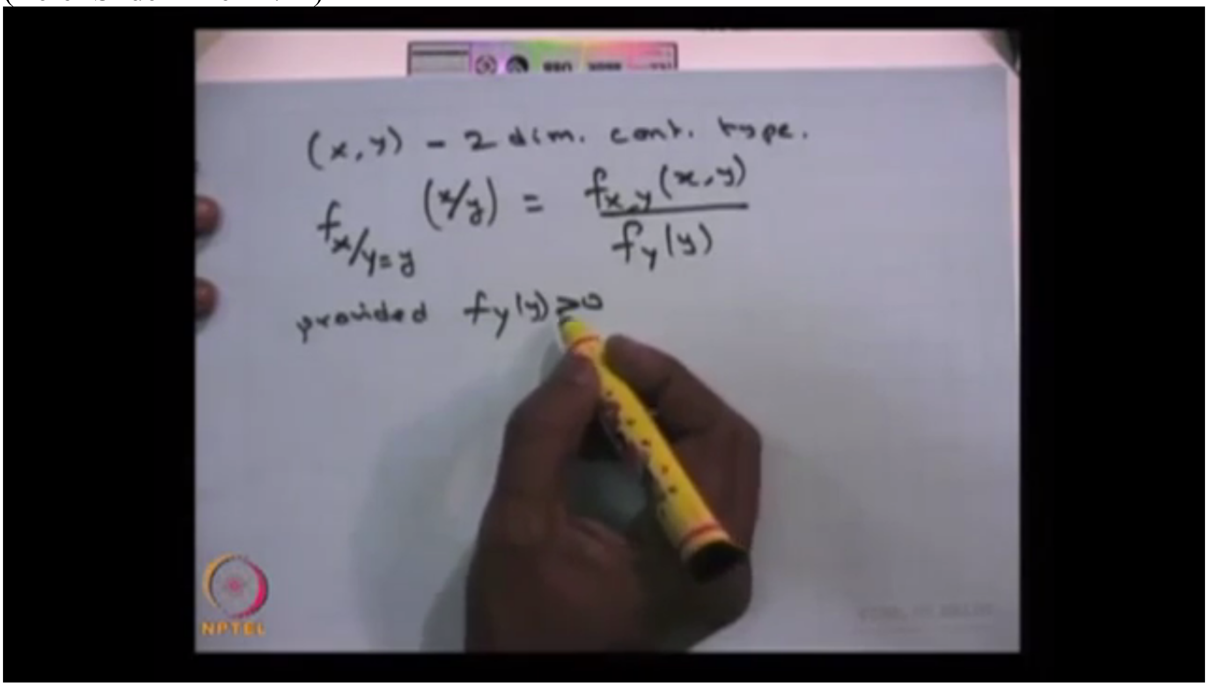
(Refer Slide Time 13:09)



That means still it is a one-dimensional random variable of x_n given that already the random variable x_1 to x_{n-1} takes some particular value. Similarly, you can go for what is a joint distribution of a few random variables given that all other random variables already taken some valuable.

Now I can go for defining, the same way, I can go for defining what is the conditional distribution of two-dimensional continuous type random variable? That means you can define what is the probability density function of a random variable x given that y takes a value y. That means that is x given y. This is nothing but what is the joint probability density function of x with y and divided by what is the marginal distribution of y, and here also the provided condition f of y is strictly greater than zero.

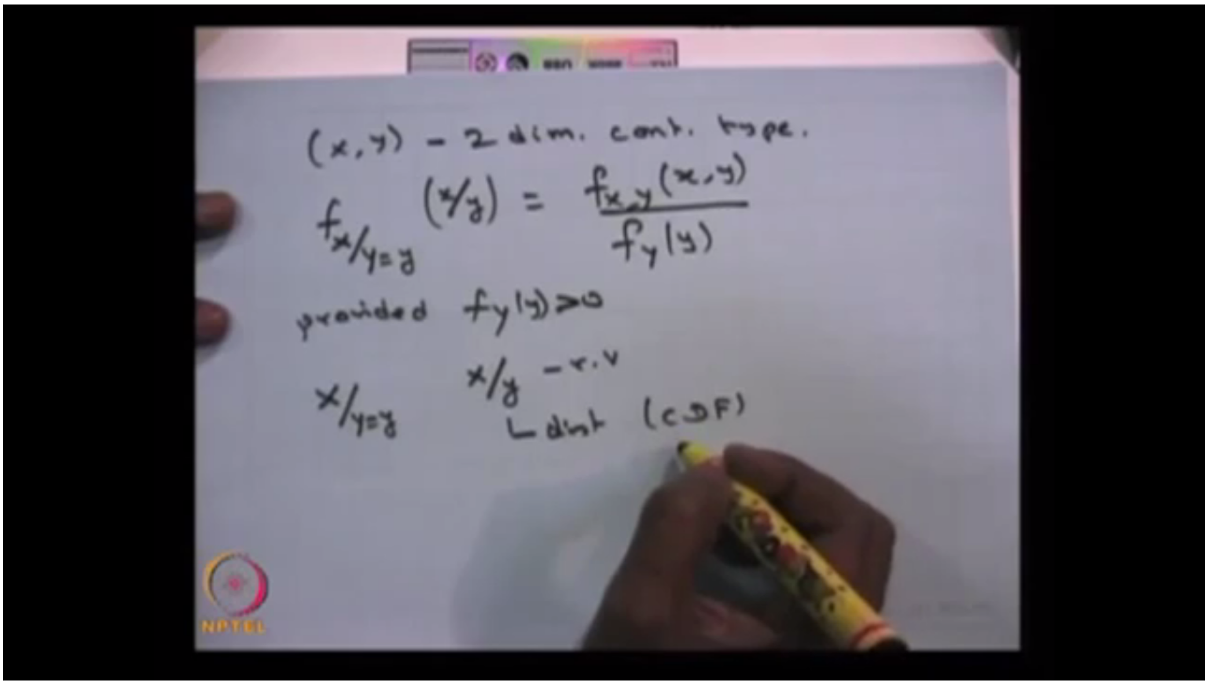
(Refer Slide Time 14:22)



That means wherever there is a density, which is greater than zero, and with that given -- given situation, you can find out the distribution of the random variable x with the given y takes the value small y. That is nothing but what is the ratio in which the joint distribution with the marginal distribution.

Once you know the conditional distribution, this is also sort of another random variable. This, that means x given y takes a value y so that I can use it as the word x small y. This is also is random variable. Therefore, you can find out what is a distribution. Therefore, this distribution is called the conditional distribution and you can find out what is the CDF of that random variable.

(Refer Slide Time 15:05)



So the way you find out the CDF of the any discrete random variable by summing what is a mass or by integrating the probability density function till that point, you will get the CDF of this conditional distribution.