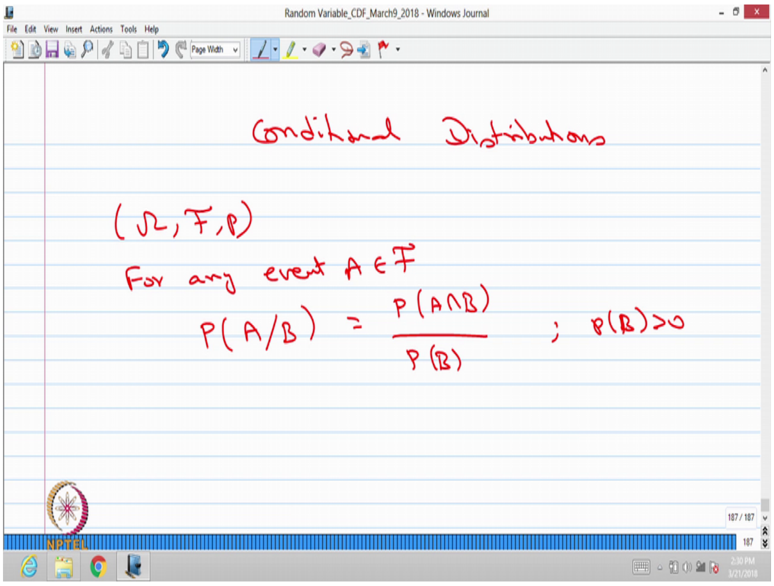
Introduction to Probability Theory and Stochastic Processes

Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture – 33

In the last class in this module that is a Functions of Several Random Variables. In the lecture 1 we discussed the distributions of functions of several random variables and in the second lecture we have discussed order statistics; that means, we have created n dimensional random variable, which is set of order statistics. And we discussed distribution of a order statistics and we have discussed some simple examples also.

(Refer Slide Time: 00:44)

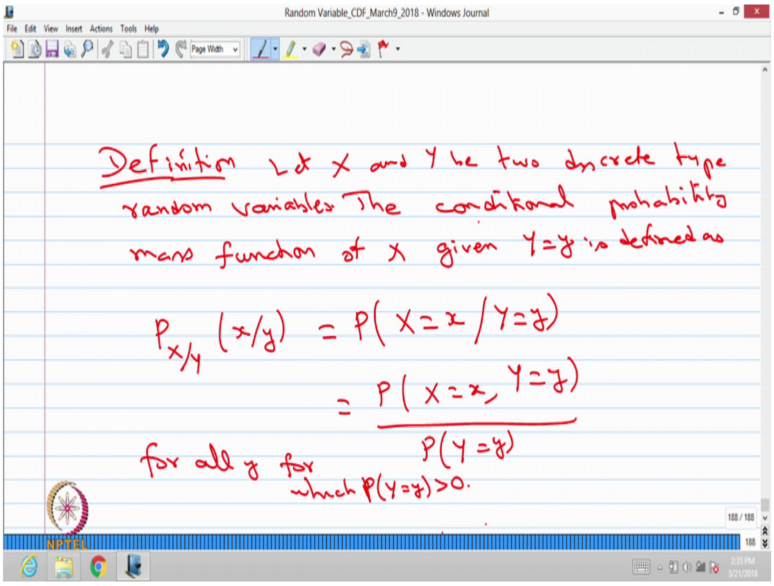


In this lecture we are going to discuss conditional distribution followed by random sum. So, these two topics which we are going to discuss in the in this lecture let us start with the conditional distribution.

Conditional distributions we have already studied conditional of a events, let me recall if you have a probability space omega F P for any event A belonging to F the probability of a given B, B is also event that is same as the probability of A intersection B divided by probability of B provided probability of B is greater than 0. This is for a conditional events conditional probability of event A given event B that is say probability of A intersection B divided by probability of B, provided probability of B is greater than 0.

In the same concept we are going to introduce for the random variables whatever we have discussed in the module 1 for the events, we are going to discuss same thing for random variables. Therefore, conceptually it is same only thing is now we are going to discuss through the random variables. Let me start with the discrete type random variables first the definition of conditional distribution.

(Refer Slide Time: 03:29)



Then I give one example then I move into continuous type random variable and the conditional distribution of continuous type random variables, then one more example let start with the definition.

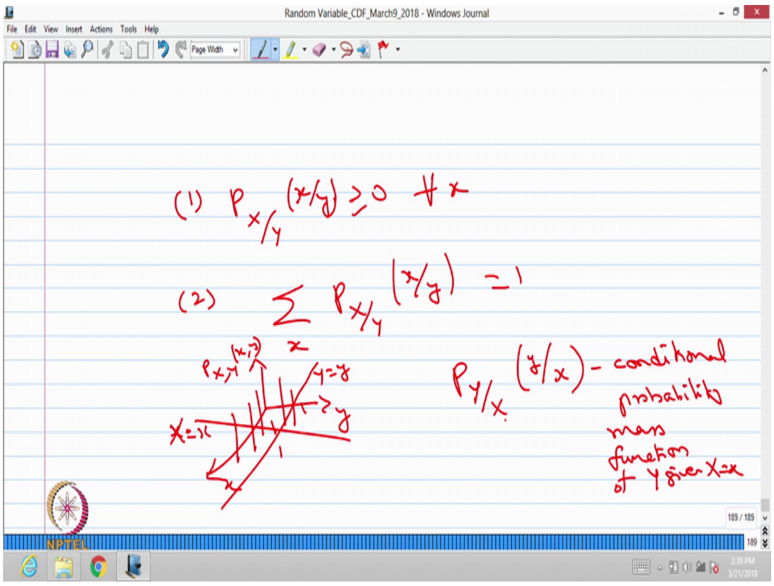
Let X and Y be two discrete type random variables, the conditional probability mass function of the random variable X given the other random variable Y takes the value small y, that is defined as P X slash Y; that means, it is a conditional distribution of X given the other random variable takes a value capital Y is equal to small y. As a function of x and y, but we write it as a x slash y it is not a x divided by y x slash y; that means, you have to treat small y as a constant and the function is function of X the conditional probability mass function of X given Y takes a value small y here you gave to treat X as a variable and Y as a constant.

That is same as the probability of X takes a value small x given that Y takes a value small y that is same as the conditional probability means the probability of X takes a value small x and Y takes a value small y divided by probability of Y takes a value small

y. This is for all y for which the probability of Y is equal to small y has to be strictly greater than 0.

So, the variable is x you have to treat a small y as a constant and this is defined whenever the probability of Y takes a value 0 is equal to small y has to be greater than 0. We call this as the conditional probability mass function of X given the other random variable takes a value Y is equal to small y. Since I use the word probability mass function one can easily verify this satisfies the properties of probability mass function that is a this probability, the conditional probability mass function satisfies p of X given small y that is always going to be greater than or equal to 0 for all x; x given y that probability value is always going to be greater than or equal to 0 for all x.

(Refer Slide Time: 07:03)



The second condition since it is a conditional probability mass function of x, the summation over x, the conditional probability mass function for different values of x, that summation is going to be 1; that means, this conditional probability mass function also satisfies the these two conditions.

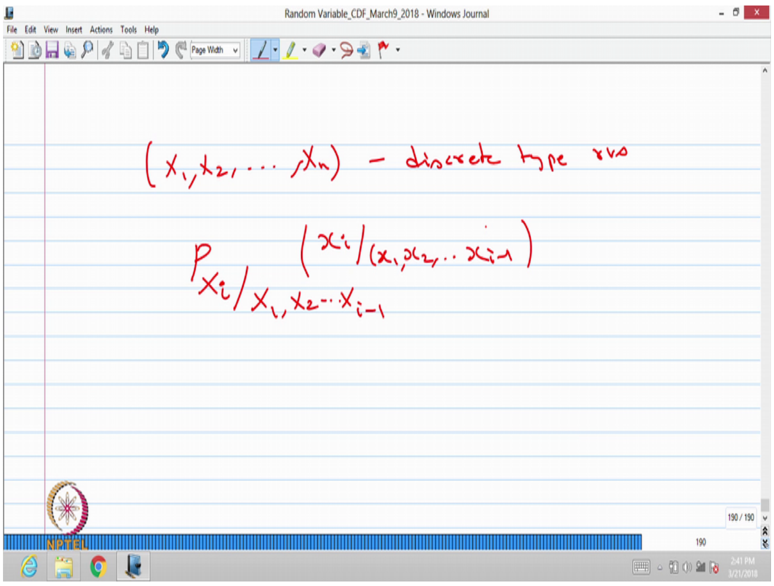
Since we are defining this random variable as x given y takes a value small y therefore, we should use the word conditional probability mass function. One can visualize if it is 2 dimensional random variable with x y; the joint probability mass function for a different values of x comma y. So, this is a different heights are nothing, but the joint probability mass function for the different values of x comma y.

The conditional distribution can be visualized suppose we make a capital Y takes a value small y; that means, you just think of a this is going to be capital Y takes a value small y. Then look for what are all the possible joint probability mass function when Y is equal to small y. You collect those joint probability mass and you make it addition, then normalize it that becomes the conditional distribution; that means, this conditional probability mass function is nothing, but find out the joint probability mass function divided by the probability of Y is equal to small y by normalizing this value is going to be between 0 to 1.

The same way one can visualize the conditional distribution of y given x is equal to x that is conditional distribution of Y given X as function of y treating small x as a constant. So, this is called conditional probability mass function of Y given X takes a value small x. So, this also can be visualized by making X is equal to x. So, this is a line is equal to small x. So, you collect all the possibilities of the probabilities when X is equal to x from the joint probability mass function, you normalize it therefore, the conditional distribution of Y given X is equal to x you may get it.

So, this can be visualized only for 2 dimensional random variable and the same concept can be extended to any n dimensional random variable.

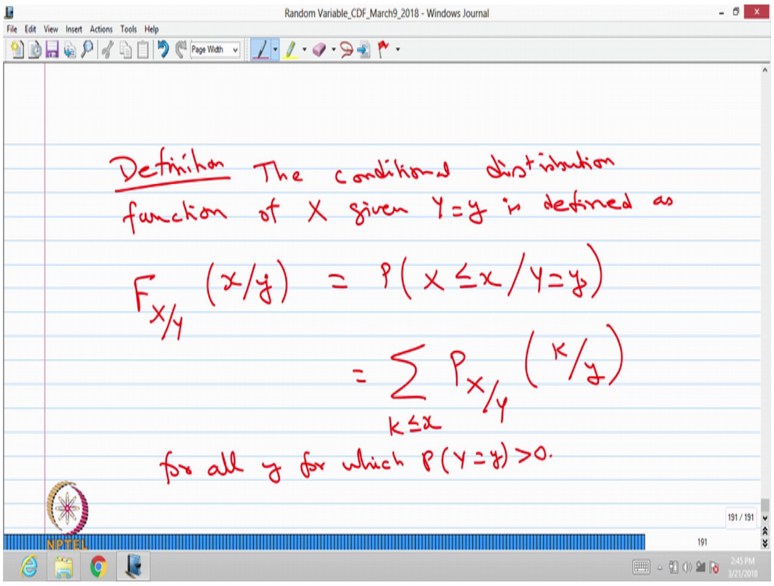
(Refer Slide Time: 10:59)



For example, suppose I have a discrete type random variable n random variables all are going to be a discrete type random variables. One can define conditional probability mass function of Xi given few X values X i minus 1. So, this is going to be a function of a small xi given that all other values are already taken some values.

So, this is going to be a conditional distribution of a Xi given X 1 takes a values small x 1, X 2 takes a value small x 2 and so on capital X i minus 1 takes a value X i minus 1. Now I am going for the simple example of a how sorry, now I will go for the conditional distribution function as a.

(Refer Slide Time: 12:20)



I am going to give the definition of conditional distribution of conditional distribution function of the random variable X given the other random variable takes a value small y, that is defined as earlier we have given conditional probability mass function, now we are going to give conditional distribution.

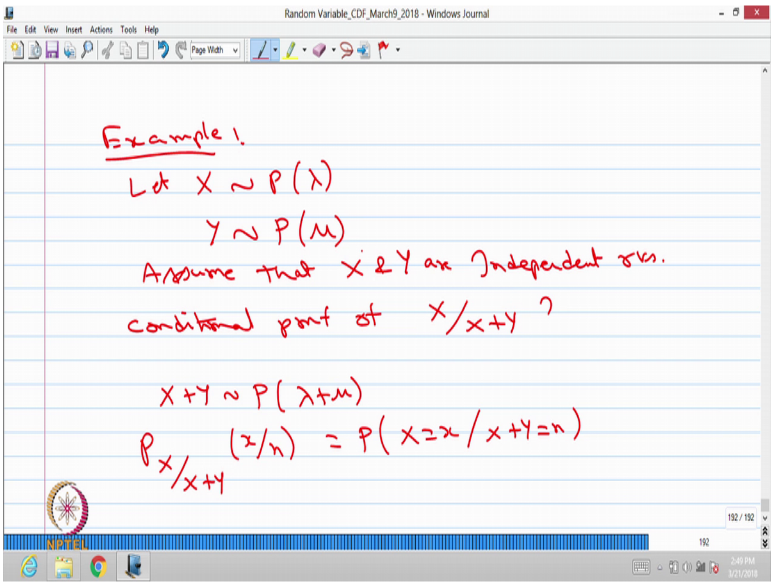
That is the CDF therefore, it is a capital F in the conditional form therefore, X slash Y it is not X divided by Y. Whenever I use a slash; that means, the other random variable already taken some value. Again this is also a function of x and y, but you have to treat y as a constant. That is same as the probability of X is less than or equal to x given the other random variable takes a value you know, that is same as already we have defined probability of X takes the value x given Y takes a value small y

Now, we are finding the conditional distribution function of the random variable X given Y takes the small y therefore, we got X is less than or equal to x that is same as from the conditional probability mass function of a various values of you can treat some K given small y where K has to be less than or equal to x. So, this is a conditional probability mass function when capital X takes a value k and capital Y takes a value small y given condition.

Summing over K less than or equal to x that will give probability of X less than or equal to x given Y takes a value small y that is called conditional distribution of a x given y takes a value small y. This is also true whenever for all y for which the probability Y is equal to small y has to be strictly greater than 0 otherwise, the conditional probability mass function itself not well defined therefore, you cannot get the conditional distribution function.

So, in the first definition we have explained how to get the conditional probability mass function when two random variables are of the discrete type. In the second definition we have given when both the random variables are of the discrete type, how one can represent conditional distribution function of x given Y takes a value small y. We will go for one easy example in which you can discuss the conditional distribution again we take only two random variables. So, the same concept can be extended to many random variables.

(Refer Slide Time: 16:11)



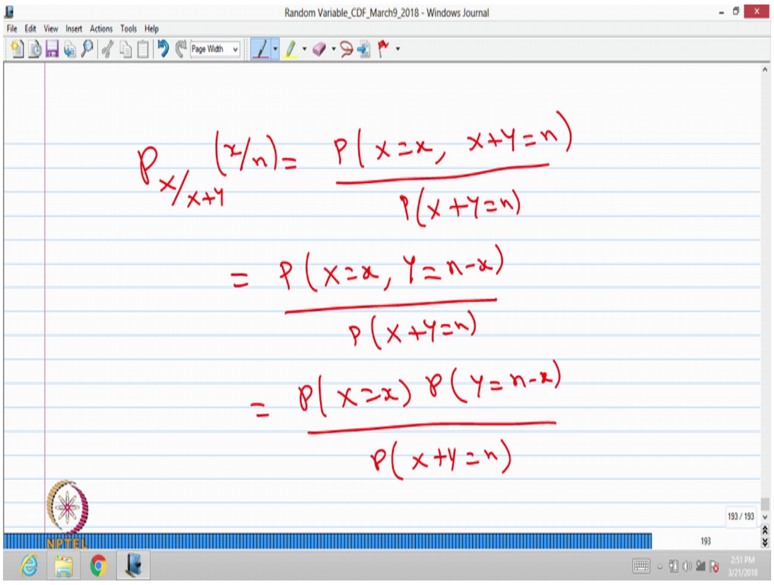
So, as example we give. So, example 1 let X be Poisson distributed with the parameter lambda and Y be again Poisson distributed with the parameter mu and I make the assumption assume that X and Y are independent random variables ok.

We will find out what is the conditional probability mass function of X given plus Y. Earlier in the definition we have discussed the conditional probability mass function of one random variable given another random variable, but here conditional distribution of X given X plus Y; that means, first you should know what is the distribution of X plus you know, then we have to go for finding the conditional probability mass function of X given X plus Y.

Since X is a discrete type variable, Y is also discrete type random variable we know that X plus Y is also going to be a discrete type random variable therefore, we can go for finding a conditional probability mass function of x given X plus Y that is a question. We know that X plus Y is going to be again Poisson distributed with the parameter lambda plus mu, we got this result from the earlier exercise earlier examples of finding the distributions of a functions of several variables several random variables.

So, we know that when X is Poisson Y is also Poisson the summation is also going to be a Poisson distributed by the reproductive property also. So, we know the distribution of X plus Y is Poisson distribution, we will go for finding what is the probability mass function of X given X plus Y as a function of x given n we treat this value is going to be n, X plus Y is going to be n that is same as the conditional probability mass function of X takes a value small x given that X plus Y takes a value small n.

(Refer Slide Time: 19:29)

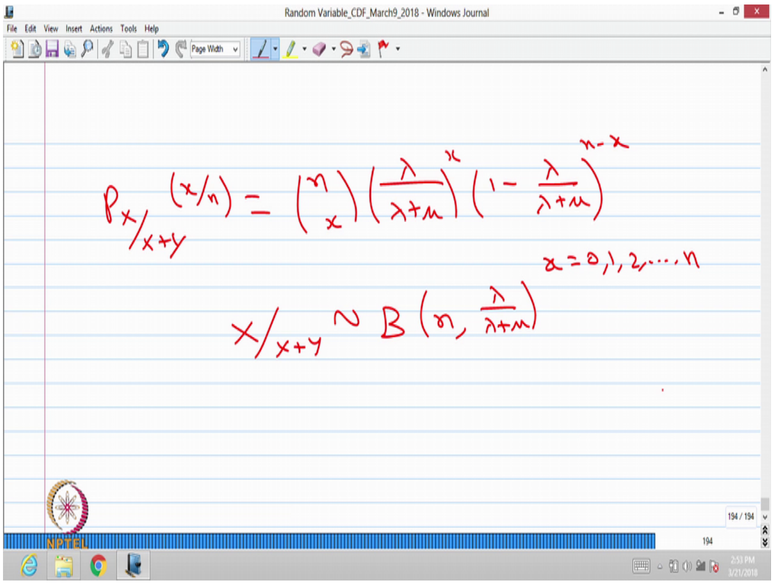


That is same as that is same as the left side is conditional probability mass function of X plus Y as a function of x and n that is same as the probability of X takes a value small x and X plus Y takes a value n divided by the probability of X plus Y takes a value n by using the definition of conditional probability mass function. That is same as the probability of X takes a value small x since X takes a value small x the other one y is going to take the value n minus x divided by the probability mass function of X plus Y takes a value n.

Already we made the assumption the random variables X and Y are independent therefore, the numerator joint probability mass function is the product of a probability mass functions. Now, we can substitute that the probability mass function for the random variable X and Y similarly X plus Y you can substitute the probability mass function of x we know that the random variable x is Poisson distributor you can substitute.

Similarly, y is also Poisson distributed with the parameter mu we can substitute the probability mass function and in the denominator x plus y that is also Poisson distributor with the parameter lambda plus mu. So, you can substitute the probability mass function of the denominator also.

(Refer Slide Time: 21:36)



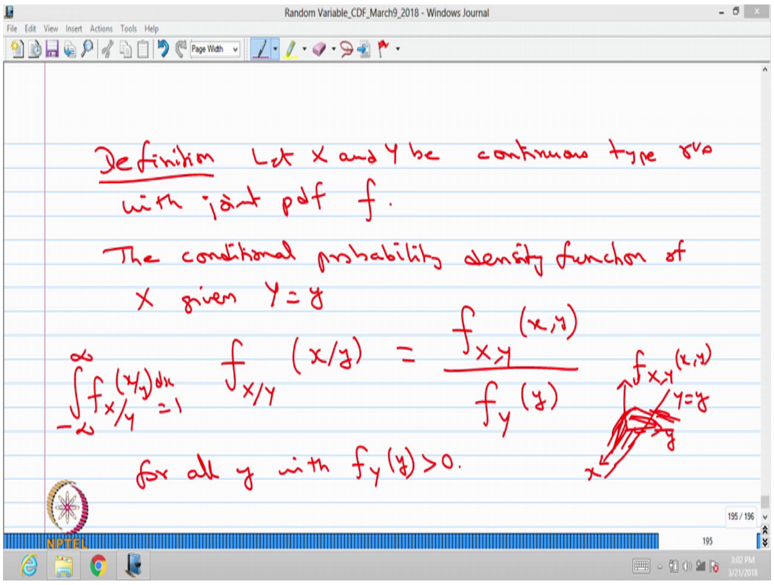
You can substitute and you can do the simplification you can get the answer it is ncx lambda divided by lambda plus mu power x, 1 minus lambda divided by lambda plus mu whole power n minus x.

This is the conditional probability mass function of X given X plus Y as a function of X and n, n to be treated as a constant. So, here the possible values of a x or 0 or 1 or 2 and so on till n. By seeing the probability mass function of this conditional distribution of X given X plus Y you can conclude X given X plus Y that follows binomial distribution with the parameters n comma here the P is lambda of divided by lambda plus mu.

From the Poisson distribution lambda is strictly greater than 0 mu is strictly greater than 0 therefore, lambda divided by lambda plus mu that is lies between open interval 0 to 1. Therefore, you can conclude this follows binomial distribution with the parameters n comma lambda divided by lambda plus mu. It is a very important result from a Poisson distribution, the summation is also going to be a Poisson distribution if they are independent by the reproductive property, where as the conditional distribution over 1 random variable given sum of this 2 random variables, that is binomial distribution.

Now, we will move into the conditional distribution for the continuous type random variables. Let me start with the definition of a conditional probability density function definition.

(Refer Slide Time: 24:09)



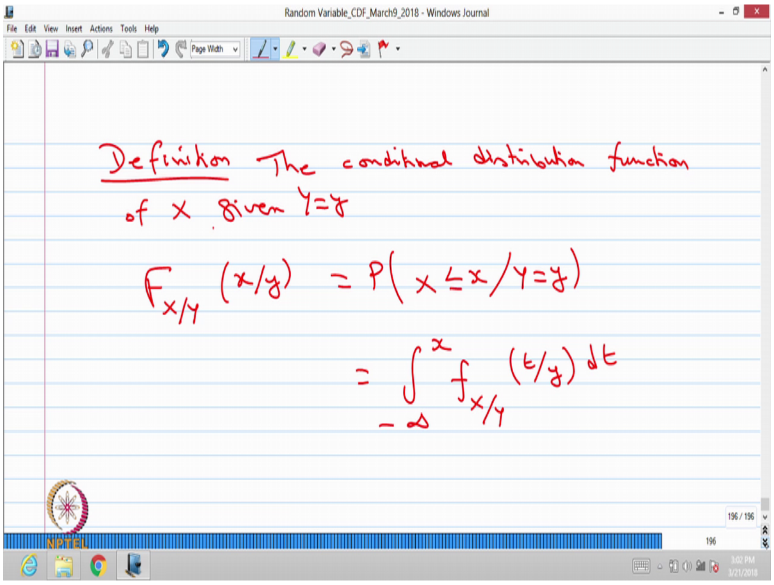
Let X and Y be continuous type random variables with joint probability density function with the joint probability density function small f; the conditional probability density function of the random variable X given the other random variable takes a value small y that is defined as small f. I am using a small f for probability density function, capital F for the CDF small f, but in the suffix I am going to use the notation X slash Y; that means, it is the conditional probability density function of given Y.

Again this is also going to be a function of x and y, but you have to treat y as a constant. This is same as the joint probability density function divided by the marginal distribution of y or probability density function of y. For all y with the probability density function at that point small y has to be greater than 0.

You we know that the probability density function will be greater than or equal to 0, but when you define the conditional probability density function you have to make sure that the denominator does not vanish. That is f y small y has to be strictly greater than 0, this ratio is going to be the probability density function. From the probability density function one can get the probability of any interval. So, here we are getting a probability density function.

Now, we are going for the same way conditional distribution as a next definition.

(Refer Slide Time: 26:50)



That is the conditional distribution function of the random variable X given the other random variable takes a value small y that is capital F. This is in the conditional distribution that is X given Y this also a function of X given small y, that is same as the probability of X takes value less than or equal to X given Y takes a value small y.

That is same as since both are continuous type random variable, that is same as integrate from minus infinity to small x, the conditional probability density function of X given Y with the variable t given small y with respect to dt. That means, by integrating the conditional probability density function, you will get conditional distribution function. The way I have explained conditional probability mass function, one can visualize the conditional the way I explained the conditional probability mass function, one can visualize conditional probability density function also. Always the joint probability density function is a surface over the X Y plane; that means, this function is always greater than or equal to 0 and double integration over the region in which it is going to be greater than 0, that volume has to be 1.

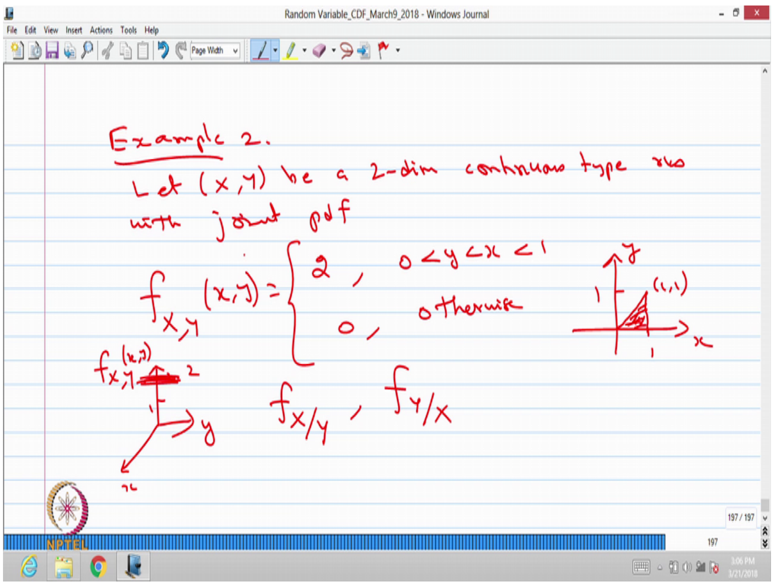
So that means, if you think of a some surface over the xy plane; that is the joint probability density function whenever you go for Y is equal to small y; that means, you are just cutting a one a plane; that means, the surface and the plane Y is equal to small y that will make a one cut. So, you will get a where the plane Y is equal to small y cut at the surface, you will get the sum sort of curve.

That curve by with a value of f y of small y, you are diminishing the curve or enlarging the curve based on the values going to be less than 1 or greater than 1. So, that the area below that area below the curve is going to be one; that means, this is a probability density function. So, minus infinity to infinity f X given Y with respect to x that is going to be 1; The conditional probability density function is nothing, but intersecting the surface with the plane Y is equal to small y by multiplying a 1 divided by f suffix y small y means you are either enlarging or diminishing the curve.

So, that the area is going to be 1; that means, the 1 divided by f y small y is a normalizing constant for the conditional probability density function of x given y is equal to small y. This satisfies the both the properties that is always going to be greater than or equal to 0 and integration over dx is going to be 1 similar to the discrete type.

So, whenever you go for the conditional distribution you are going for a again getting the same property of probability mass function or probability density function based on the random variable discrete or continuous. Whereas, here the conditional distribution function of X given Y is equal to small y when both the random variables are of continuous type. This integration from minus infinity to X which is same as the usual way of one dimensional random variable instead of the probability density function you are using a conditional probability density function of X given Y.

(Refer Slide Time: 32:09)



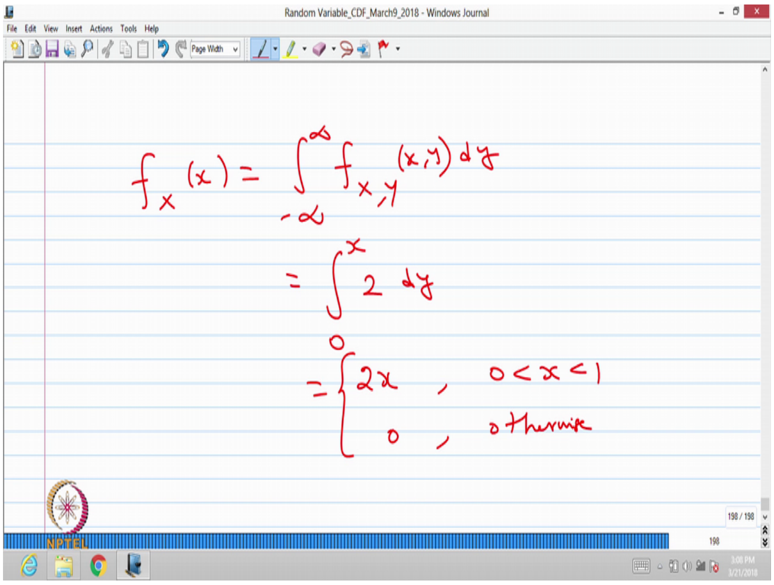
Now, we will go for another example to explain for the continuous type random variables. Already we have given one example for discrete type random variable now we will explain the same concept for a continuous type random variables. Let X comma Y be a 2 dimensional continuous type random variables with the joint probability density function is given by is a function of x and y that takes a value 2 when y is lies between 0 to x and x is lies between y to 1 0. Otherwise, this is a joint probability density function of a 2 dimensional continuous type random variables. You can verify if you do the double integration from minus infinity to infinity dx dy will be1.

But before that we will find out what is the region in which the joint probability density function is greater than 0 that is 2. So, we can make a x versus y, you shade the region in which y is lies between 0 to x, x is lies between y to 1 so; that means, 1 1 that is 1 comma 1. So, you draw a line that is y is equal to x therefore, this is a shaded region. So, in this region the joint probability density function is 2.

That means x axis, y axis the joint probability density function is going to be at the height of a 2 it has the some sort of a plane at the height of 2 over the region xy in this triangle part fine. So, this is the joint probability density.

Our interest is to find out the conditional probability density function of X given Y as well as the conditional probability density function of Y given X. For this problem we are interested to find the conditional probability density function of X given Y similarly Y given X. To do this first you should know what is the marginal distributions of X and Y then only you can go for it. For first one you should know the marginal distribution of Y for the second problem we should know the marginal distribution of X. So, let us find both the marginal distribution of X and Y, then we will go for finding the conditional probability density function of X given Y and Y given X.

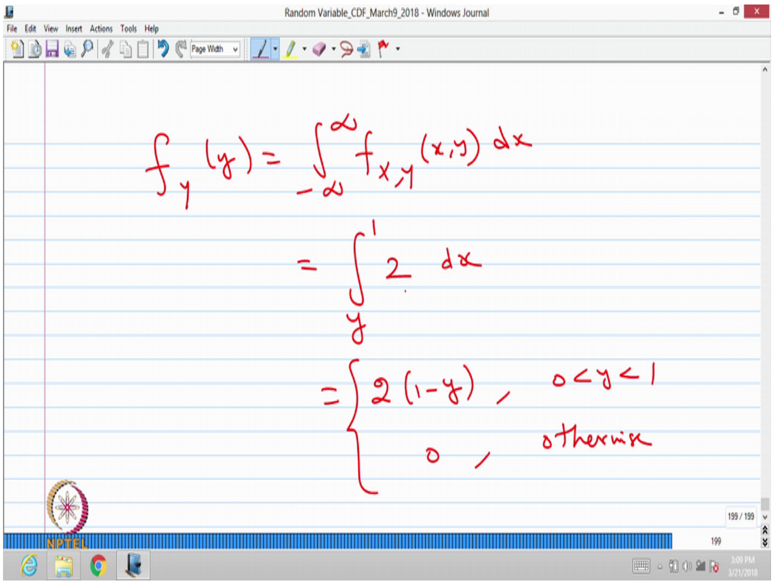
(Refer Slide Time: 35:56)



The marginal probability density function of X that is nothing, but the integration from minus infinity to infinity, the joint probability density function with respect to y correct. So, this is same as you should go and see the joint probability density function value is 2 between 0 less than Y less than X less then 1 and you want integration with respect to Y. Therefore, the integration is 0 to x the value is 2 with respect to y therefore, when you simplify si you will get the answer is 2 x and this is going to be 2 x when x lies between 0 to 1.

So, the probability density function is going to be 2 x when x is lies between 0 to 1 0 otherwise, similarly you can do the marginal distribution of y.

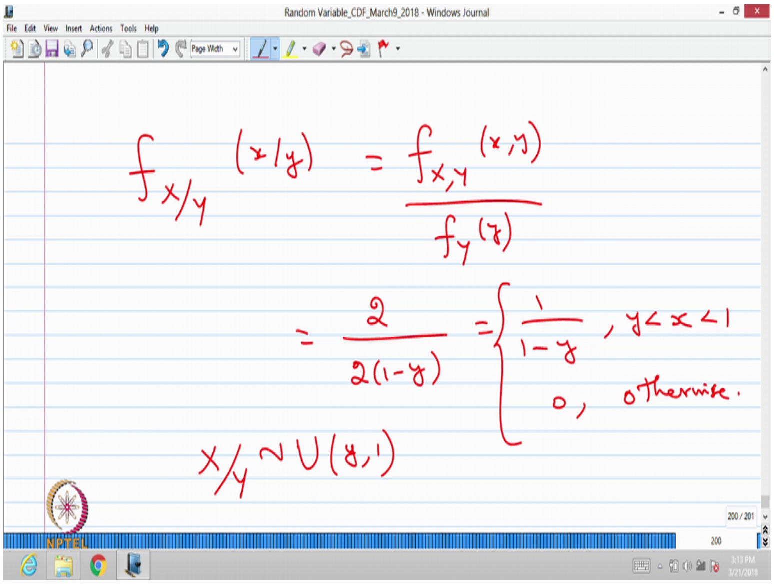
(Refer Slide Time: 37:12)



That is the probability density function of y, that is going to be minus infinity to infinity the joint probability density function of a x and y with respect to x, that is same as. Now, you see the interval again it is 2 between 0 less than y less than x 0 sorry less than 1 therefore, the integration with respect to the value is 2 integration with respect to x and x is between the interval y to 1. When you do the simplification you will get the answer 2 times 1 minus y, and this is going to be the probability density function is 2 times 1 minus y when y is lies between 0 to 1

So, this is the marginal distributions of x and y, now we will go for finding the conditional distributions.

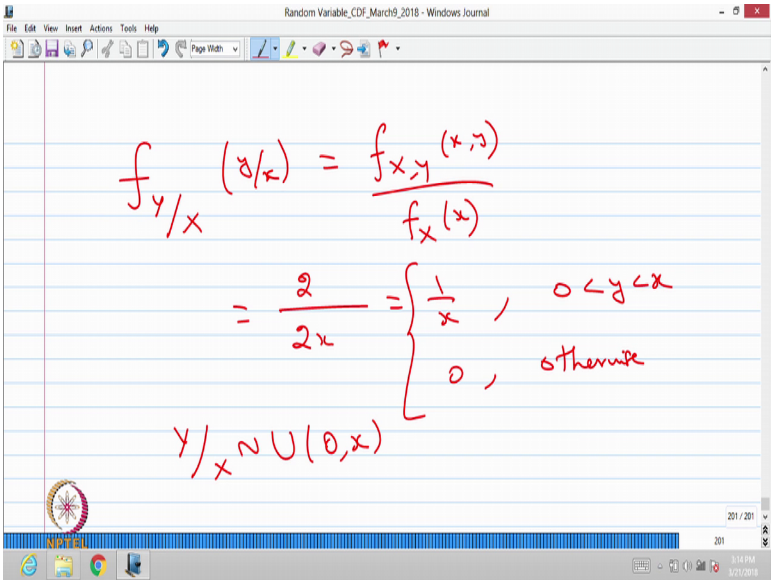
(Refer Slide Time: 38:27)



The first one we are going for finding the conditional probability density function of X given the random variable capital Y takes a value small y, that is same as the joint probability density function divided by the probability density function of y. The joint probability density function is 2 between that interval and if the denominator we have already got the answer 2 times 1 minus y.

So, when you simplify you will get 1 divided by 1 minus y. So, this is the conditional probability density function of x given y. So, you have to treat y as a constant here. So, the value is going to be 1 divided by 1 minus y when x lies between y to 1. When y x takes a value y to 1 the conditional probability density function of x given y that is 1 divided by 1 minus y 0 otherwise.

(Refer Slide Time: 39:51)



Similarly, you can go for finding conditional probability density function of Y given X by treating small x as a constant. Again I write the same definition joint divided by marginal of X. So, this is same as the joint is 2 and marginal is 2 x. So, the simplification will give 1 by x. So, this is going to be 1 by x when y lies between 0 to x 0 otherwise. In this you have to treat x as a constant, y as a variable y lies between 0 to x and the conditional probability density function of y given x is 1 by x, conditional probability density function of x given y that is 1 divided by 1 minus y, where x lies between y to 1.

Here you have to treat a y as the constant. By seeing the probability density function you can say the conditional distribution of x given y that is continuous type uniform distribution between the interval between the interval y to 1 you have to treat y as a small y as a constant.

So, the conditional distribution of X given capital Y takes a value small y, that is continuous type uniform distribution with the parameters or with the interval with the intervals Y to 1. Similarly here the conditional distribution of Y given X that follows continuous type uniform distribution between the interval 0 to X, here you have to treat x as a constant small x as a constant.

So, we started with the joint distribution we started with the joint distribution and we are finding the conditional distribution after finding the marginal distributions of a individual random variables. So, this is a very easy problem in which you are applying the definition and you can feel what could be the distribution of the conditional distributions.