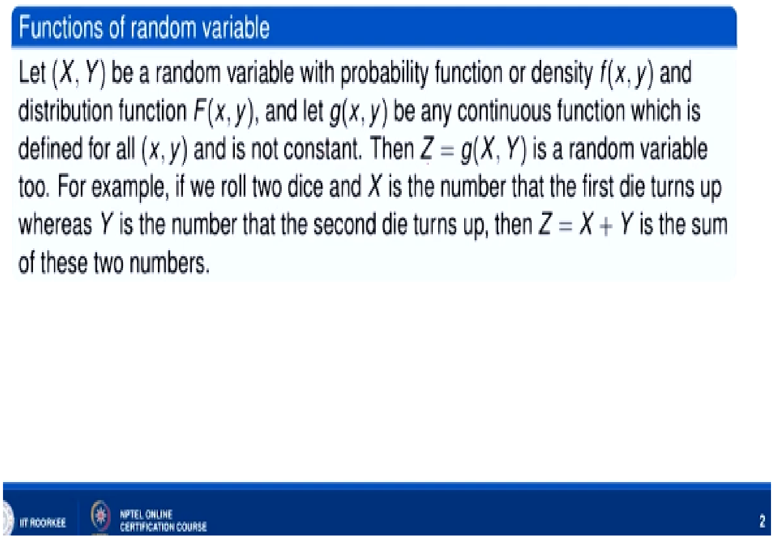
Advanced Engineering Mathematics
Prof. P.N. Agrawal
Department of Mathematics
Indian Institute of Technology – Roorkee

Lecture – 54
Joint Probability Distribution - III

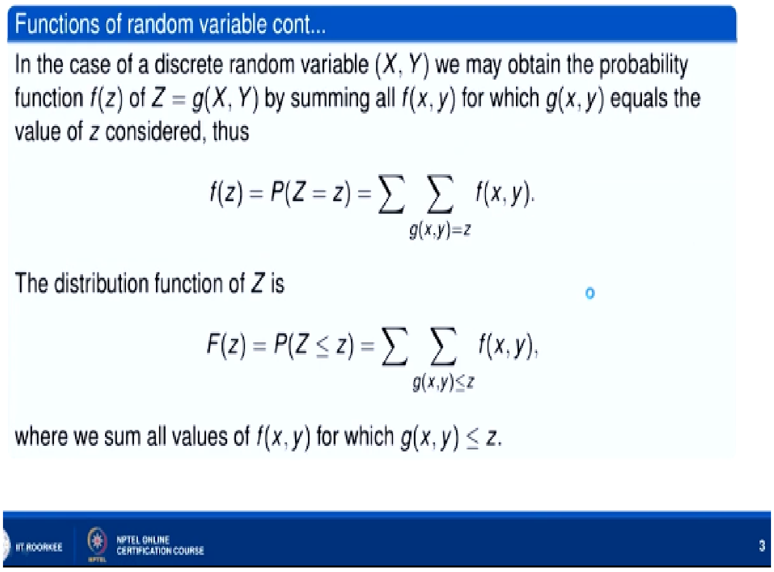
Hello friends welcome to my lecture on joint probability distribution, this is the third and final lecture on joint probability distribution. First we define function of random variable. Let x and y, xy be a random variable with probability function are density fxy and the cumulative distribution function fxy and let gxy be any continuous function which is defined for all xy and is not a constant.

(Refer Slide Time: 00:50)



Then Z = gxy is a random variable 2 okay. For example, if we roll 2 dice, if x is the number then the first dice turns up whereas y is the number that the second dice turns up then Z = x + y is the sum of these 2 numbers.

(Refer Slide Time: 01:12)

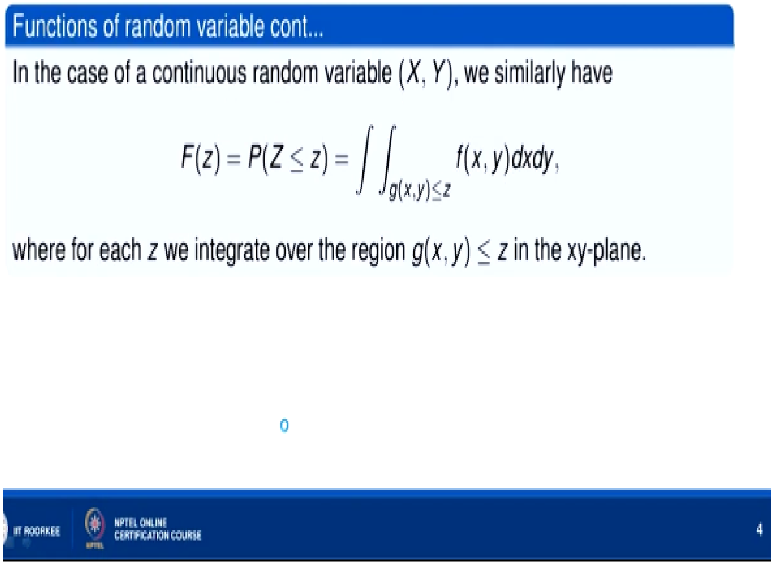


In the case of a discrete random variable xy we may obtain the probability function fz of z = gxy by summing all fxy for which gxy = to the value of z considered okay, that is fz = probability that z takes the value z will be = double sigma fxy where gxy = z, that is sum of the values of x and y becomes = z. We will take the sum over all those values of x and those values of y, where the sum of x and y values give you z okay.

The distribution function of z will be fz = pz <= z where we will sum over all those xy where the sum of the values of x and y are or where the gxy, gxy not necessarily the sum of x and y in the example we have taken gxy as x + y but it is arbitrary here. So gxy must be <= z. So in the case of the probability function fz, we will take the sum over all those pairs of values of xy, such that gxy = z in the case of cumulative distribution function fz, we shall take the sum over all those values of fxy where the pair of values of xy satisfy gxy <= z.

Okay, so where we sum overall values of fxy where gxy is <= z.

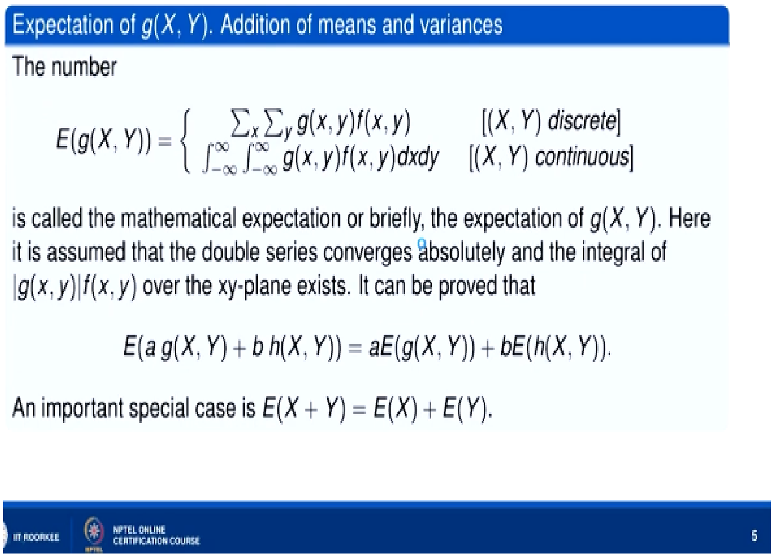
(Refer Slide Time: 02:49)



Now functions of random variable, if you consider here continuous random variable xy okay we earlier discussed discrete case, now let us consider the case of continuous random variable xy. So we similarly have fz = probability that z is <= z, we integrate over all those values of xy which satisfy the inequality gxy <= z okay. So fxy dx dy. So where each z we integrate over the region gxy <= z in the xy plane okay.

And when you take the probability okay the probability here, then we will take a small fz, small fz will be we will integrate over all those xy for which gxy = z okay.

(Refer Slide Time: 03:42)

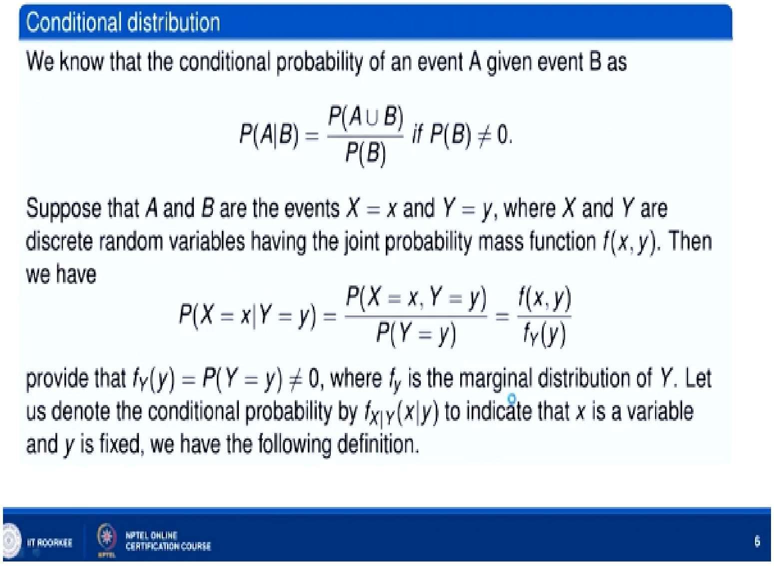


So the number Eg XY okay, we have defined, this is discrete case in the case of continuous random variable we have the distribution function fz, now we have to consider the expectation of gxy. The number expectation of gxy = sigma over x, sigma over y, gxy, fxy when xy is a discrete probability distribution and when gxy is a continuous, joint continuous distribution, we have integral over – infinity to infinity, integral over – infinity to infinity gxy fxy dx dy.

This is called the mathematical expectation or expectation of gxy. We assume here that the double series in the discrete case converges absolutely and the integrals here okay, over the xy plane okay, edges, integral of mod of gxy * fxy over the xy plane edges okay. So it can be shown that expectation of ag xy + b hxy = a times Eg xy + b hxy okay. An important special case here you can take as expectation of ax.

You can take 1 b = 1 and gxy as x hxy as y. So you can say expectation of x + y = expectation of x + expectation of y.

(Refer Slide Time: 05:11)

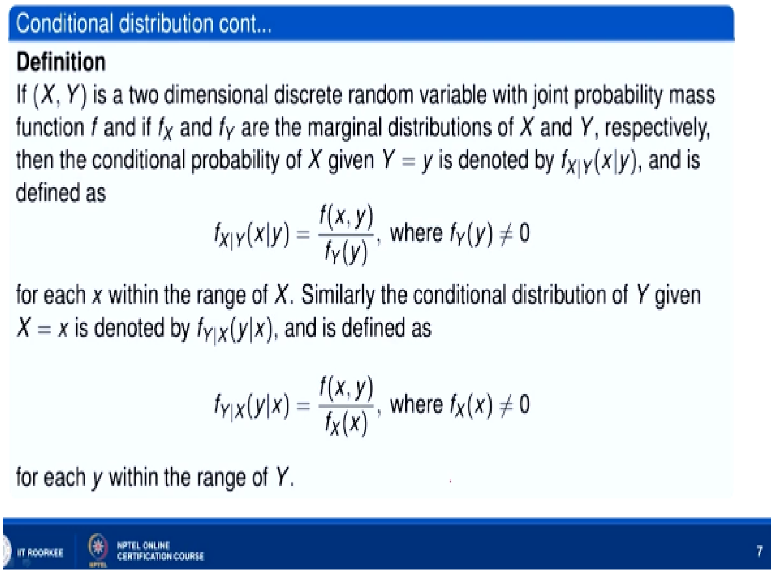


Now we know that the conditional probability of an event A given event B is given by this formula okay. Probability that A given B = probability of A intersection B/probability of B. So if PB is not = 0. Now suppose that A and B are the events X = x and Y = y where x and y are discrete random variables okay, having the joint probability mass function fxy, then we have probability that x takes the value x, y takes the value y = probability that x takes the value x, y takes the value y/probability of that y takes the value y.

And then now probability that x takes the value x, y takes the value y is fxy and probability that y takes the value y is the marginal distribution function of y with respect to the joint distribution okay, provided that the marginal distribution function or marginal density of y with respect to the joint distribution is not 0. So this is the marginal distribution of y. Now let us denote the conditional probability by fx given by x given by okay.

So this can be denoted by fx given by x given by to indicate that x is a variable and y is fixed okay.

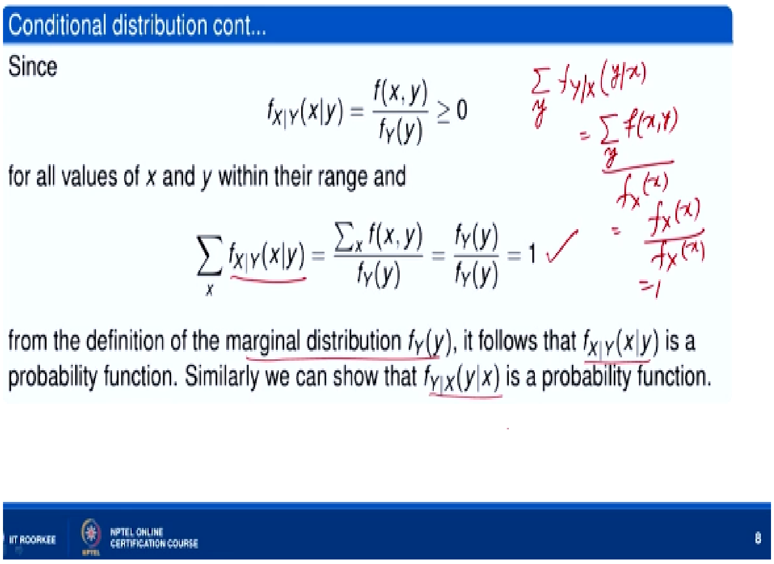
(Refer Slide Time: 06:38)



So we have the following definition. If xy is the 2 dimensional discrete random variable with joint probability mass function f and if fx fy are the marginal distribution functions of x and y, then the conditional probability of x given y = y is denoted by fx given by x given by end is defined as fx given by x given by = fxy over fyy, fyy is the marginal density of f, marginal density of y with respect to the joint distribution.

So for each x within the range of x. Similarly, conditional distribution of y given x = x is denoted by fy given x y given x and is defined as fy given x y given x = fxy over marginal distribution of x with respect to the joint distribution. For each y in the range of y okay. Here we assume that marginal densities fxx and fyy are not 0.

(Refer Slide Time: 07:41)

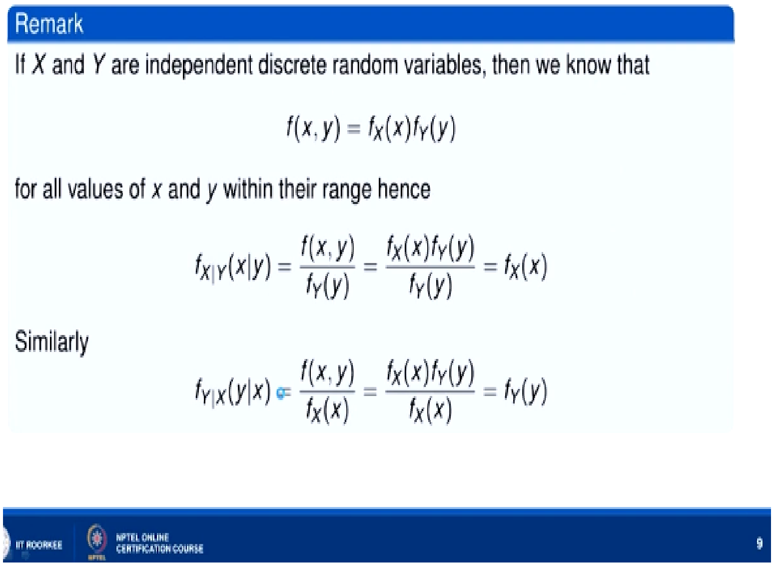


Now since fx given y, x given y = fxy/fyy fxy is >= 0 fyy is also >= 0, this is density function, this is marginal density. So they are both nonnegative and therefore their quotient is nonnegative. This is valid for all xy in the range and sigma fx given by x given by = sigma/x fxy. When you keep y fixed and sum over, when you keep y = y, when you keep y fixed and sum over all the values of x okay.

Then you are getting fyy okay, so fyy/fyy = 1 okay and that is this follows from the definition of marginal distribution and so what happens is that the conditional distribution okay, fx given by x given by is a probability function. This is a probability function, similarly we can show that fy given x, y given x is also a probability function, there we will, if you want to show this then you sum over all y such that fy/x y/x y/x = sigma/y fxy/fxx we will have okay.

Now here we are summing over all y, x is fixed so we get fxx/fxx which is = 1. So both the conditional distributions are probability functions okay. This is conditional probability function of x given y = y, this is conditional probability function of y given x = x.

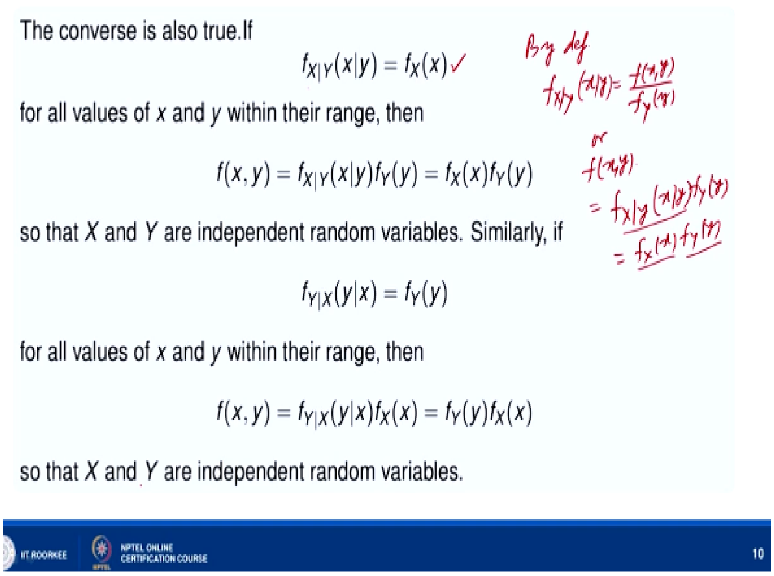
(Refer Slide Time: 09:38)



Now if x and y are independent and they are discrete random variables, independent discrete random variables then we know that fxy = fxx * fyy for all values of xy within their range okay. So now fx given by x given by = fxy/fyy, y definition okay. This is by definition and if x and y are independent by definition, fxy = fxx * fyy. So let us put it here what we get fxx.

So if x and y are discrete random variables and they are independent okay, then the conditional distribution of x given y = y = marginal distribution of x with respect to the joint distribution. Similarly, fy/x y/x = fx/y/fxx = fyy okay. So if x and y are independent discrete random variables then the conditional distribution of x given then the condition distribution of x given y = y = marginal distribution of x with respect to joint distribution and conditional distribution of y given x = x is marginal distribution of y.

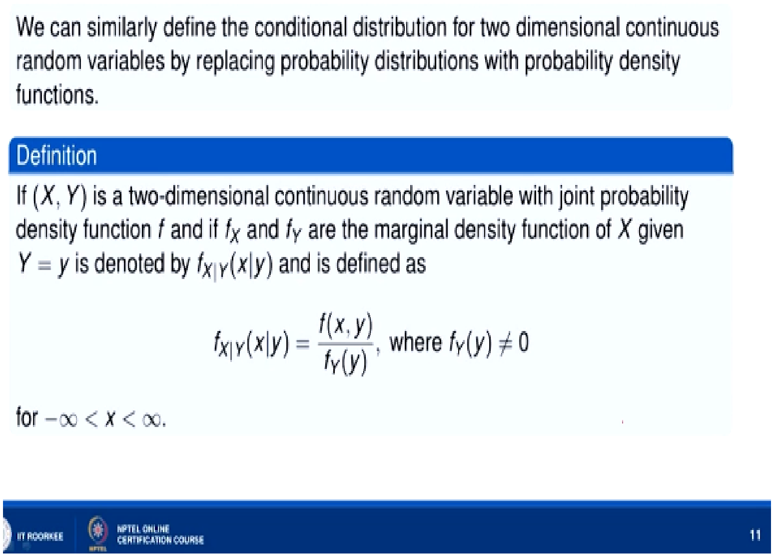
(Refer Slide Time: 10:58)



So the converse is also true, if fx/y, x given y = fxx for all values of x and y within range then fxy = fxx * fyy. How it will follow let us see, fxy = see by definition. By definition fx given y x given y = fxy/fyy okay, so we can say or fxy = fx given y x given y * fyy okay. Now fx given y x given y = fxx okay. So fxx = fxx * fyy and therefore x and y are independent random variables okay.

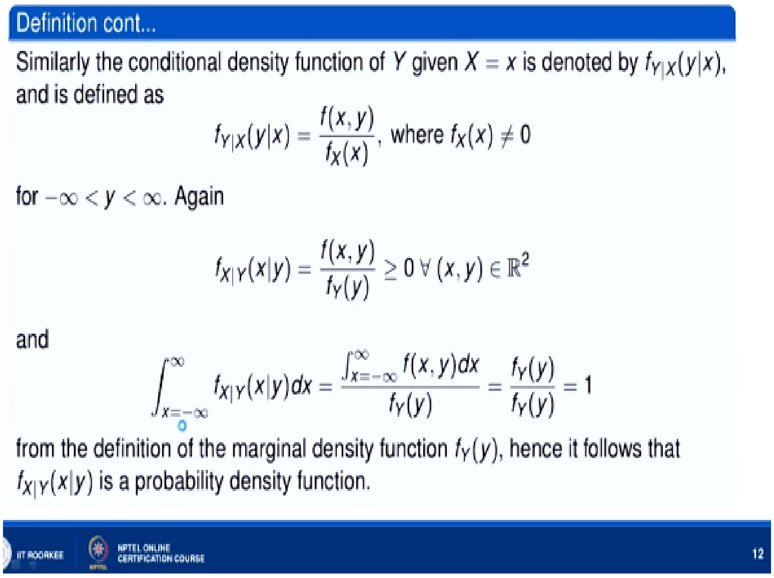
Similarly, if f y/x/given x = fyy then fxy = fy given x/given x fxx = fyy * fxx, so that x and y are independent random variables.

(Refer Slide Time: 12:17)



Now we can similarly define the conditional distribution for 2 dimensional continuous random variables by replacing probability distributions with probability density functions. So if xy is a 2 dimensional continuous random variable with joint probability density function f and if fx and fy are the marginal density function of x given y = y, then fx given y x/y x given y = is defined as fx given y x given y = fxy/fyy where fy y is not = 0 and x grounds to the interval – infinity to infinity.

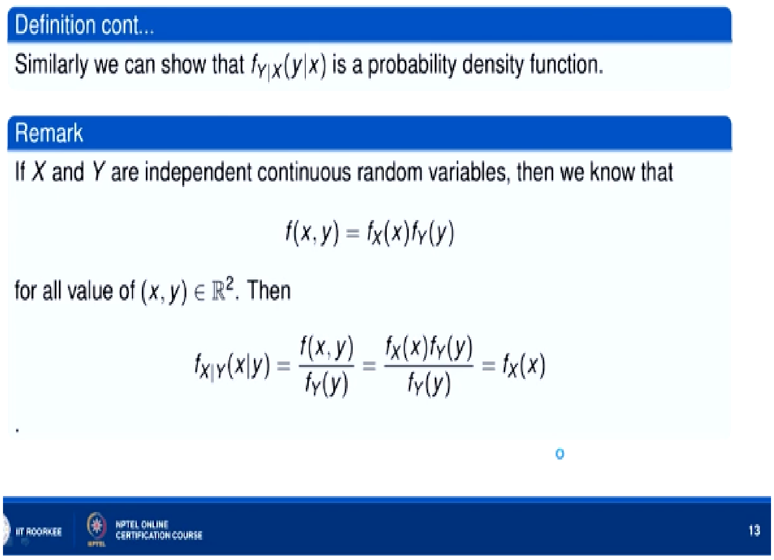
(Refer Slide Time: 12:55)



The conditional density function of y given x = x similarly is defined as fy given x, y given x = fx y/fxx, where fxx is not = 0 and x belongs to the interval – infinity to y belongs to the interval – infinity to infinity. Again fx given by x given by = fxy/fyy is >= 0 for all xy belonging to R square and integral over – infinity to infinity fx given by x given by dx = integral over – infinity to infinity fx/dx and this is independent of x so fyy.

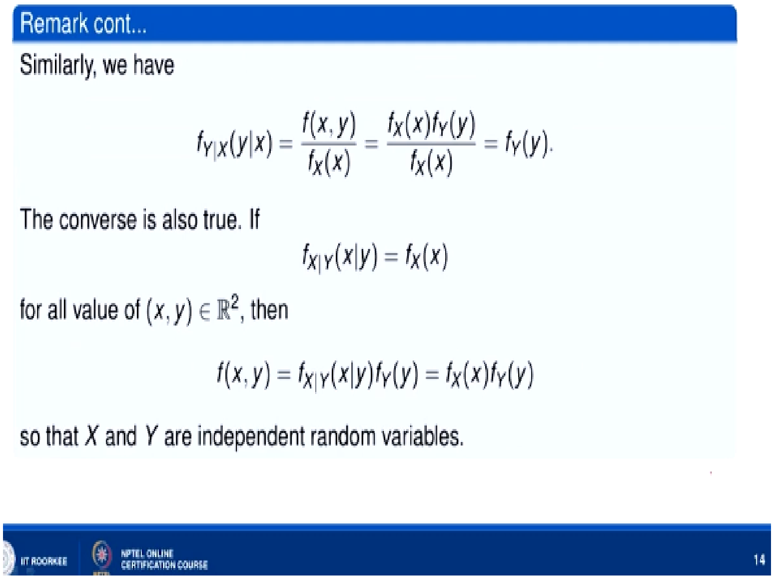
And we get this integral as fyy, so fyy/fyy = 1, so from the definition of marginal density function we find that this value is = 1, so fx given y x given y is a probability density function.

(Refer Slide Time: 13:52)



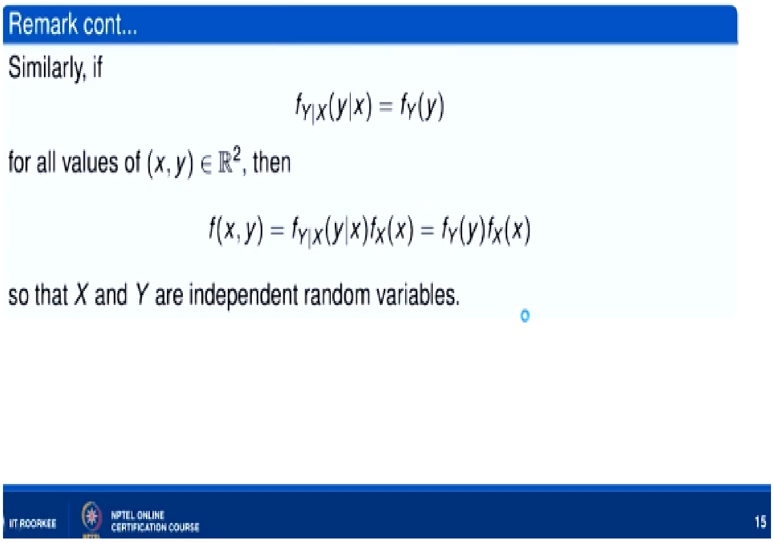
Similarly, we can show that fy given x, y given x is a probability density function. If x and y are independent continuous random variables, then we have fxy = fx given x fxx * fyy for all value of x y belonging to R square and then fx given y x given y by definition is fxy/fyy but fxy = fxx * fyy. So when you put it here you get fx given y x given y = marginal density function of x with respect to the joint distribution.

(Refer Slide Time: 14:27)



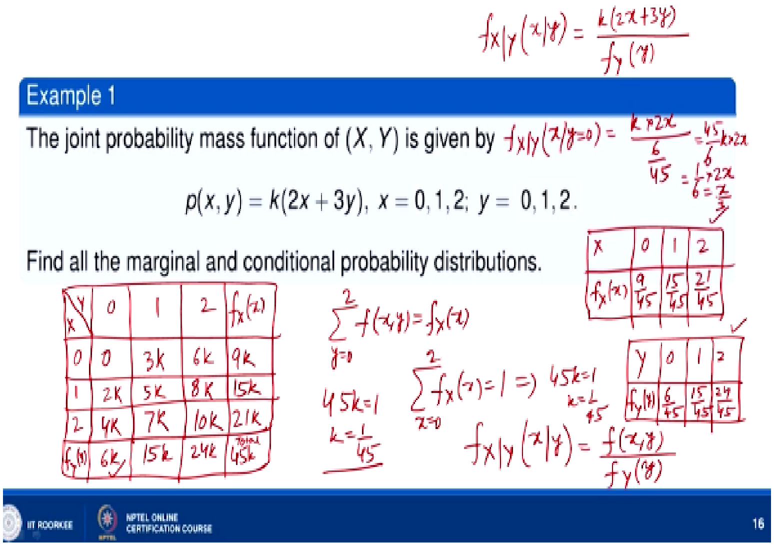
And fy given x y given x = fxy/fxx okay, which is fxx * fyy/fxx, so this is fyy okay. So we have the same result okay like in the discrete case, the converse is also true if fx given y x given y = fxx for all values of xy belonging to R square, then fxy = fx of given y, x given y f yy, this will follow from the definition of conditional distribution and this is = fxx, so we get fxx * fyy and then x and y are independent random variables.

(Refer Slide Time: 15:05)



And similarly we can see for fy given x, y given x, if it is = fyy then fxy = fyy * fxx and so x and y are independent random variables.

(Refer Slide Time: 15:20)



Now let us consider a problem on marginal and conditional probability distributions. So the joint probability mass function f xy is given by pxy = k times 2x + 3y, x = 0, 1, 2, y = 1, 2, 3. Let us find first the joint probability distribution okay. So x varies from 0, 1, 2 and y varies from 0, 1, 2. So let us say here we have x, here we have y and x varies from 0 to 2, okay. So when x is 0, y is 0 okay, p00 = 0.

When x is 0, y is 1, okay we have 3k, when x is 0 and y = 2 we have 6 k, when x is 1, y is 0 we have 2k, when x is 1 okay, y = 1, so we get 3+2 5k okay, and when x is 1, y is 2, so 3 2s are 6, 6 +2, 8k and when x is 2 we have 4 here y =0 okay, so we have 4k. When x is 2 we have 4 here, y = 1 so we have 7 k and when x is 2 we have 4 here, y = 2, so we have 6 here, 6 + 4, 10k okay.

So 6k + 3k, 9k here we have 8 + 5, 13, 13 + 2 15k and we have 10 +7, 17, 17 +4, 21 k. Let us check again, when x is 0, y is 0 we have pxy 0 when x is 0, okay y = 1, we have 3k when x is 0, y = 2 we have 6 k, when x = 1 we have here 2 okay, so y is 0, so we have 2 k, when x is 1, we have 2 here, here we have 1, we have 3 + 2, 5k, when x is 1 y is 2 we have 3 2s are 6 + 2 so 8k.

When x = 2 we have 4 here y is 0 so 4 k, when x is 2, y is 1 we have 7 k when x is 2 y is 2 we have 3 2s are 6 + 4, 10k okay. So this is 15 + 9, 24, 24+21, 45k okay, alright. Thus it is the sum of the columns, so here what is happening, we are getting fxx because when you sum along the row okay, you get the value of fxx because you are summing fxy sigma fxy, you are summing along the row.

So y varies from 0 to 2 okay, so this gives you fxx, okay, so now fyy, so this is 6k, this is 7 +5, 12 +3, 15k and this is 10 + 8, 18 + 6, 24k. So 24 +15, 39 + 6, 45 k. So 45 k is the total okay. Now this is 45 k okay. So 45 k must be = 1 okay and this means that k = 1/45, sigma fxx okay, when x varies from 0 to 2 = 1 okay, so this implies 45k = 1, so k = 1/45 okay. Now we find the marginal probability function.

Marginal distribution of x, so let us form the table of marginal distribution function of x. So xfxx. So x takes the value 0, 1, 2 okay. So when x = 0, we have fxx = 9k that is 9/45 okay. When x = 1, we have 15k, 15/45 and when x = 2 we have 21 k, so 21/45 okay, and we have similarly marginal density function of y. So 0, 1, 2, f/y. So when y = 0 okay, we have 6k okay, so 6/45 okay.

When we have y = 1 it is 15/45 and when y = 2 it is 24/45 okay. So this is the marginal density function of y, this is marginal density function of x. Now let us find conditional probability distributions okay. So conditional probability distributions means we have to find fx given y x given y okay this will be = fxy/fyy okay. So fx given y x over y okay. So f/y, fxy = k times 2x + 3y.

So we have fx given y x given y = k times 2x + 3y, f/y okay, f/y = 6/45, 15/45, 24/45 okay. So we have to do it for y = 1, 2, 3 okay, y = 0, 1, 2, and so on. So fx we have to first find fx given y by x y =0 okay. So y = 0 means we have k times 2x okay, 2x y = 0 means fyy is 6/45 okay. So this is = 45/6 * k * 2x, okay. So 45/6 * k 45k = 1, so we have 1/6 * 2x okay, 1/6 *2x that is x/3 okay.

And similarly we have to do it for y = 1, y = 2 okay. So we have to do it for y = 1, y =2.

(Refer Slide Time: 23:40)

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$$f_{Y|X}(y|x) = \frac{2+3y}{15} \checkmark, \; x=1$$

For Y=1

$$f_{X|Y}(x|y) = \frac{f(x,1)}{f_Y(1)} = \frac{k(2x+3)}{\frac{15}{45}} = 3k(2x+3) = \frac{3}{45}(2x+3) = \frac{1}{15}(2x+3)$$

$$f_{Y|X}(y|x) = \frac{p(2,y)}{f_X(2)} = \frac{k(4+3y)}{\frac{24}{45}} = \frac{4+3y}{24}$$

For Y=2

$$f_{X|Y}(x|y) = \frac{f(x,2)}{f_Y(2)} = \frac{k(2x+6)}{24/45} = \frac{45k(2x+6)}{24} = \frac{2x+6}{24} = \frac{x+3}{8}$$

$$f_{Y|X}(y|x) = \frac{f(0,y)}{f_X(0)} = \frac{3ky}{9/45} = \frac{45}{9} \times 3ky = \frac{y}{3}, \; x=0$$

$f_{X|Y}(x|y)$	$y=0$: $\frac{x}{3}$	$y=1$: $\frac{2x+3}{15}$	$y=2$: $\frac{x+3}{8}$
$f_{Y|X}(y|x)$	$x=0$: $\frac{y}{3}$	$x=1$: $\frac{2+3y}{15}$	$x=2$: $\frac{4+3y}{24}$

$$f_{Y|X}(y|x) = \frac{p(1,y)}{f_X(1)} = \frac{k(2+3y)}{\frac{15}{45}}$$

So fx given by x/y, x given y, okay, we have to do it for this is fxy/fyy okay, let us take y = 1 now okay. So fx given by x given y okay, we have to do it for, this is fxy/fyy okay, let us take y = 1 now okay. So fx given by okay, we have found for y =0, for y = 1 we have to find, so fx1, fy1, fx1 = k times 2x + 3/15/45 k times 2x + 3/15/45 okay. So 15/45 means 1/3 okay, so 3k * 2x + 3 okay.

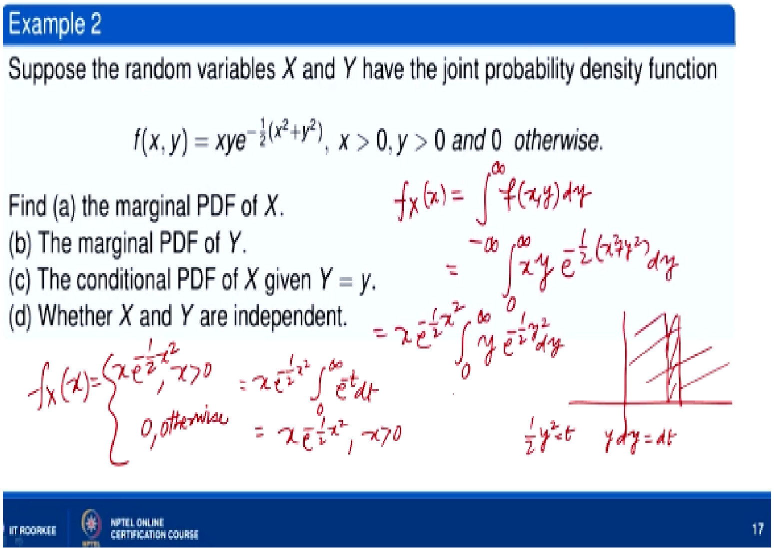
So 3/45, k is 1/45, so 3/45 * 2x +3 okay, so we have 1/15, 2x +3 okay and then fx given by for y = 2 okay. So this if for y = 1, and we have earlier found for y = 0, for y = 2 okay. So fx 2, fy2 okay. This is = k times 2x + 6/24/45 okay. So this is 45k * 2x +6/24 okay, 45k = 1 so we have 2x + 6/24 okay, so we have x + 3/8 okay. So fx given y, x given y okay. For y = 0 we found to be x/3 okay.

For y = 1 we found it to be 2x + 3/15 and for y = 2 we found it to be x + 3/8 okay, similarly we can find fy given x y given x okay. For x = 0, for x = 1, for x = 2 okay. Let us find fy given x, y given x for x = 0. So we have f0y/fx0 okay. So f0y, f0y will be = sorry p0y you can say p0y = x0. So 3ky okay, 3ky/fx0 = 9/45 okay. Now what this is 45/9 * 3ky okay, 45k = 1, 3/9 is 1/3, so it is y/3 okay.

So x = 0 we have found okay, x = 1 if you want. So f y given x y given x will be = f, instead of f actually we have taken p okay. So let us write p, so p 1y/fx1. So how much p1y will be, p1y will be = k times 2 + 3 y/15/45 okay. So this will be = you can say this will be = 45 k = 1. So 2+ 3y/15. So fy given x y given x = 2 + 3y/15 when y = 2 okay, and f/given x/given x can similarly be found for x = 2 okay.

So p2y/fx2, this is =k times 2x + 3 y we had, so 4 + 3y/fx 2 we found to be = 24/45 okay so this is 45k will be = 1. So 4 + 3y/24 okay. So f y given x y given x we find for x = 0 to be y/3, for x = 1 we found it to be =, for x = 0 it was y/3, for x = 1 it was so 2+3y/15 okay and when I take x = 2 y, I wrote y = 2 there, okay this is for x = 1, for x = 2 we got to be 4+3yy 24 okay, so this is how we find the conditional distribution of x given y and y given x okay.

(Refer Slide Time: 30:38)



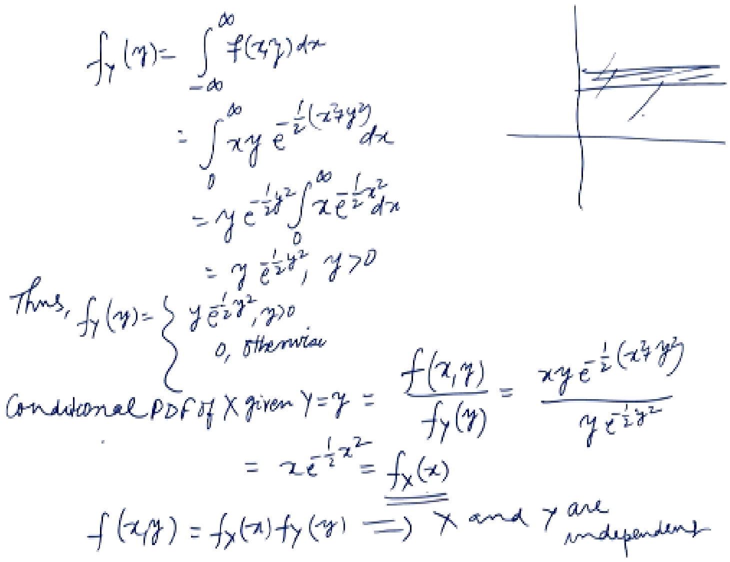
Now suppose we have the random variables x and y, whose joint probability density function is fxy = x * y e to the power -1/2, x square + y square, x > 0, y > 0 and 0 otherwise okay, so when in the first quadrant fxy is given by xy * e to the power -1/2 x * x square + y square. So marginal density function of x let us find first okay. That is we want to find fxx okay. So fxx will be integral over – infinity to infinity fxy dy okay.

So we will have it over first quadrant only, okay, x is > 0, y is > 0, so we have to integrate over in that with respect to y okay. So integral over the first quadrant okay. So we are integrating, when we integrate we here keep x as fixed okay. Probability that x = x we want, y varies okay. So we take a vertical strip. For the vertical strip in the region x is fixed okay, y varies from 0 to infinity.

So 0 to infinity x * y e to the power -1/2, x square + y square dy, okay, so x times e to the power -1/2 x square we can write outside okay, x times e to the power -1/2 x square we can write outside, integral 0 to infinity/times e to the power -1/2 y square we can integrate. Okay so let us make the substitution 1/2 y square = let us say t then we get y dy = dt. So we get here limits remain the same, we get x e to the power -1/2 x square e to the power –tdt, which comes out to be, this integral comes out to be 1 okay.

So we have x e to the power -1/2 x square, where x is > 0 okay and so fxx = x e to the power -1/2 x square when x is > 0 and 0 otherwise okay. So fxx = x e to the power -1/2 x square when x is > 0 and 0 otherwise okay, then marginal pdf of y, marginal pdf density function of y similarly we can find okay. So marginal density function of y we can find to be.

(Refer Slide Time: 33:37)



f/y = integral over – infinity to infinity fxy, we are integrating in this first quadrant okay because there only fxy is nonzero. So we are taking now y fixed okay, y fixed means we have to take a horizontal strip okay. So horizontal strip we have to take in the region like this. So we have to take okay, so this is for where y fixed okay. So this will be = integral over 0 to infinity, xy e to the power – 1/2 x square + y square dx okay.

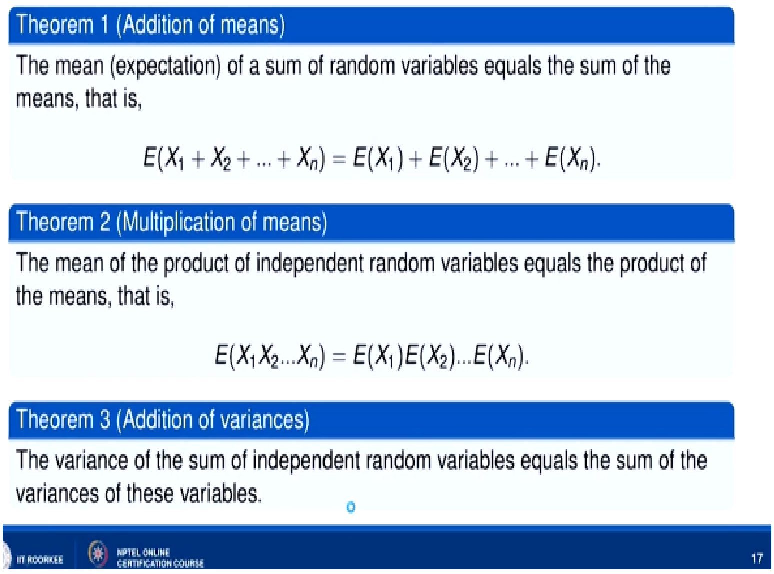
So y e to the power – 1/2 y square we shall write outside and integral 0 to infinity x e to the power -1/2 x square dx, this we have just now seen, this comes out to be 1, so this is y * e to the power -1/2 y square and this is valid when y is > 0, this equal to fyy = this when y is > 0, thus fyy = y e to the power -1/2 y square when y is > 0 and 0 otherwise. Okay, now conditional PDF okay of x given Y =y.

So we have to find this okay. Conditional PDF given Y = y okay, conditional PDF of x given Y = y, so this means that we want the let us go to the definition, conditional okay. So we define the conditional distribution of x given Y = y okay. If xy is a 2 dimensional discrete random variable, this is for the case where fx y/fyy in the case of continuous we have fx/y x/y fx/yy/y okay.

So we come to this one okay. So conditional probability of x given Y = y is fxy okay given Y = y, given Y = y means fyy okay. So this will be = xy e to the power -1/2 x square + y square/fyy that is fyy we have found just now y times e to the power -1/2 y square okay. So this comes out to be x e to the power -1/2 x square okay, x e to the power -1/2 x square which is nothing but you can see this is = fxx okay.

So we can see here that fxy = fxx * fyy okay and therefore x and y are independent random variable okay. So it answers this question also. We have found the conditional probability distribution of x, given Y = y it came out to be fxx okay. So fxx so it turns out that fxy = fx x * fyy and therefore x and y are independent okay.

(Refer Slide Time: 37:30)



So then we have these theorems which are important ones, the mean are expectation of a sum of random variables equal the sum of the means that is expected value of X1 + X2 and so on Xn is expected value of X1 + expected value of X2 and so on expected value of Xn and then the mean n of the product of independent random variable.

Suppose X1, X2, Xn are independent random variables then expectation of their product is product of their expectations and then variance of the sum of independent random variables equals the sum of the variances of those random variables. So that is all in this lecture. Thank you very much for your attention.