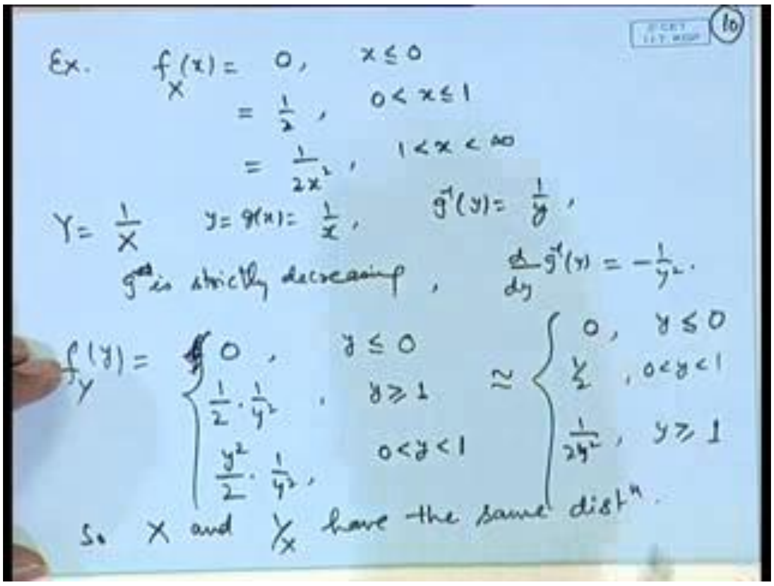
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 34
Function of a Random Variable – II

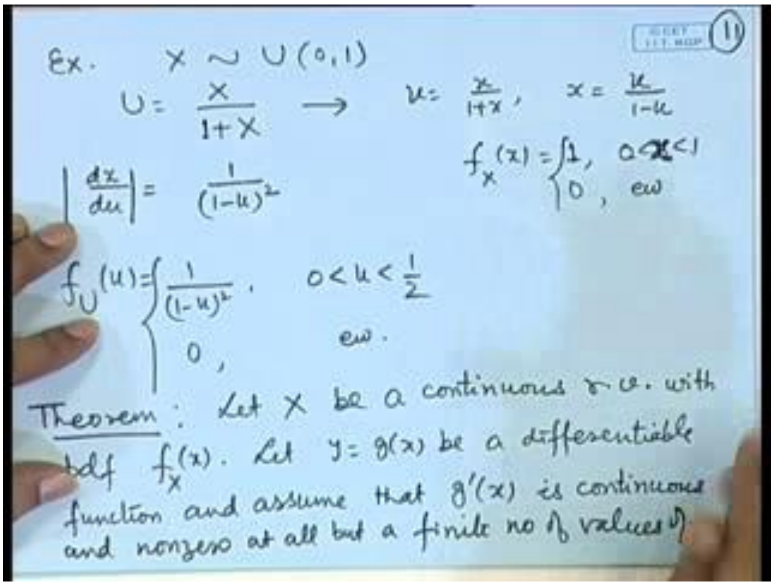
(Refer Slide Time: 00:23)



Let us take an example. Say the random variable X is having the density 0, for X less than or equal to 0 it is equal to half, for 0 less than X less than or equal to 1 and it is equal to 1 by 2 x square for 1 less than x less than infinity and consider the function say Y is equal to 1 by X. So, here g x function is 1 by x. So, g inverse function is also same if I write this as y. So, y is equal to 1 by x. So, x will be equal to. So, if I look at this, this is a strictly decreasing function g is strictly decreasing and g inverse is also a strictly decreasing. So, d by d y of g inverse y is equal to minus 1 by y square. The density function of y then it is determined by the density function of X at g inverse y multiplied by the absolute value of the d X by d y term. Now here it is 0, so it will remain 0, whenever x is less than or equal to 0, 1 by x is also less than or equal to 0. So, when it is half this remains half multiplied by 1 by square.

Now, the range 0 less than x less than 1 is translated to y is greater than or equal to 1 and in the third region it is y square by 2 multiplied by 1 by y square, when x is greater than 1 will reduced to 0 less than y less than 1. So, after simplification this is equal to 0 for y less than or equal to 0, half for 0 less than y less than 1. 1 by 2 y square for y greater than or equal to 1; notice here that these f X x and f Y y they resemble. So, for X less than or equal to 0 it is 0 here it is 0, here it is 0 less than x less than or equal to 1 then it is half and when x is greater than 1 it is 1 by 2 x square. So, that is also satisfied here except that equality at the end point, but that hardly matters because if I put y is equal to 1 here, the value is half and here the value at the y is equal to 1 is half and here also. Actually at the end points because it is a continuous random variable, the values will be immaterial. So, basically X and Y, X and 1 by X have the same distribution here.

(Refer Slide Time: 03:42)



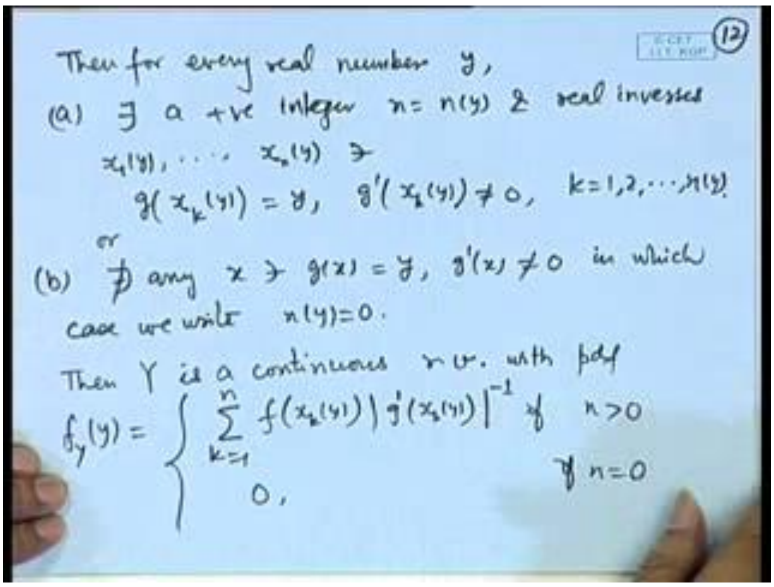
So, let us look at another example, say X follows uniform 0, 1 and we define random variable U is equal to X divided by 1 plus X. Now again you can see here that this is a 1 to 1 function. In fact, if u is equal to x by 1 plus x, then the inverse function is equal to u by 1 minus u; let us look at the derivative d x by d u that is equal to 1 by 1 minus u whole square. So, here the density function of X is 1 for 0 less than u less than 1. So, the density function of u is obtained 1 into 1 by 1 minus u square, what will be the range when this is 0 less than x less than 1. So, when X is 0 this is 0 when X is 1 then this half. It is a strictly increasing function, so 0 less than u less than half, and it is 0 elsewhere; many times the function g X may not be a 1 to 1 function it may be a many one function.

For example y is equal to x square, y is equal to modulus x. So, in that case we see that the region that is from r to r the functions domain and range. So, what we do? We look at the inverse image for a given y and so if there are 2 inverse images, then we split the region that is the domain of x into 2 disjoint regions, such that both of them match map g from each part of the domain to the full range for example, you consider y is equal to x square, now from minus infinity to infinity this maps to 0 to infinity.

Now, for a given y which is positive; I have 2 inverse images, minus square root y and plus square root y. So, if I consider 2 portions of the domain that is minus infinity to 0 and 0 to infinity, both are mapped by this mapping to 0 to infinity. So, the idea here is that in each domain each part of the domain the function will be 1 to 1; that is if you are considering only one inverse image say root y or minus of root y then the function is either increasing or decreasing, we calculate the density in each regions separately and add this gives the density function of the continuous random variable in case the function y is equal to g x is a many valued function.

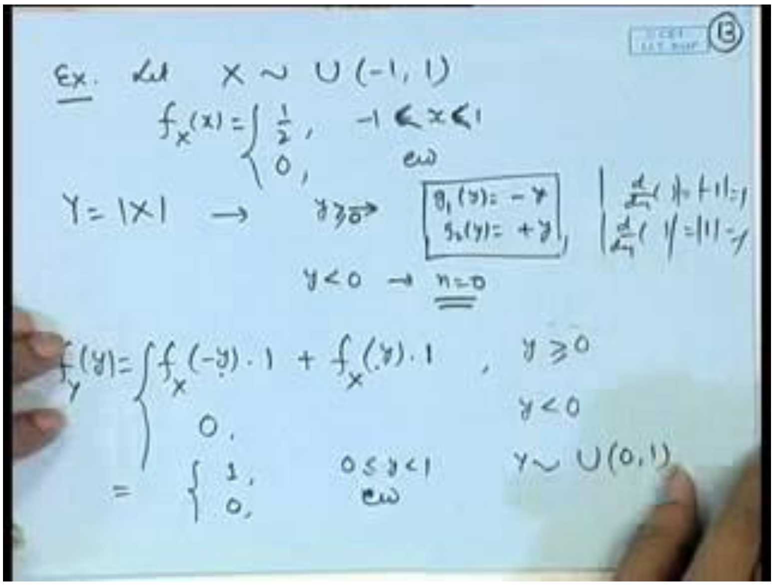
So, let us look at the result here let X be a continuous random variable with probability density function say f X, let y is equal to g x be a differentiable function and assume that g prime x is continuous and non zero at all, but a finite number of values of x.

(Refer Slide Time: 08:16)



Then for every real number y, there exists a positive integer n is equal to say n y, and real inverses say x 1 y, x 2 y, x n y such that g of x k y is equal to y, and g prime of x k y is not 0, for k equal to 1, 2 up to n y or they are does not exist any x such that g x is equal to y, g prime x is not 0, in which case we write n y is equal to 0; then y is a continuous random variable with pdf given by sigma f of x k y, g prime x k y inverse k equal to 1 to n, this is if n is greater than 0, n means n y and it is equal to 0 if n is 0. So, the idea is that we consider n separate regions of the domain such that each region maps to the range of g, calculate the density in each area and sum over all such areas that gives the density function of y.

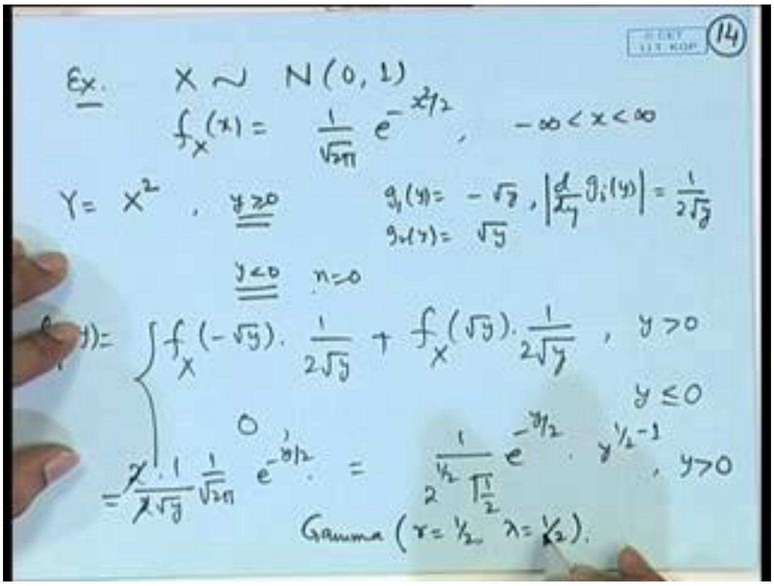
(Refer Slide Time: 10:40)



So, let us look at application of this I am skipping the proof of this theorem; let X follows a uniform distribution on minus 1 to 1. So, the density function is half for minus 1 less than or equal to x less than or equal to 1 it is 0, consider the function say Y is equal to modulus X. So, here for a given y we have g 1 of y is equal to minus y and g 2 y is equal to plus y. So, 2 inverse images for a given y is there for y positive. If y is negative n is equal to 0; that means, there is no inverse image, and the derivatives here you can see d by d y of this is equal to minus 1 and d by d y of this is plus 1. So, if you take absolute values then both are 1. So, the density function of y is the density function of X at minus y into 1 plus the density function at plus y into 1 for y greater than 0, if you want you can include equal to 0 also that does not make any difference, and it is equal to 0 for y less than 0. So, this is half plus half that is equal to 1.

Now, when we say y is greater than or equal to 0, here it will reduce to the region half plus half for minus 1 to 1. So, it is becoming modulus of x between 0 to 1 and it is equal to 0 for y outside 0 to 1. So, you can see that y follows uniform 0, 1. If X is minus 1 to 1 uniform distribution, then modulus of X is uniform distribution on the interval 0 to 1.

(Refer Slide Time: 12:58)

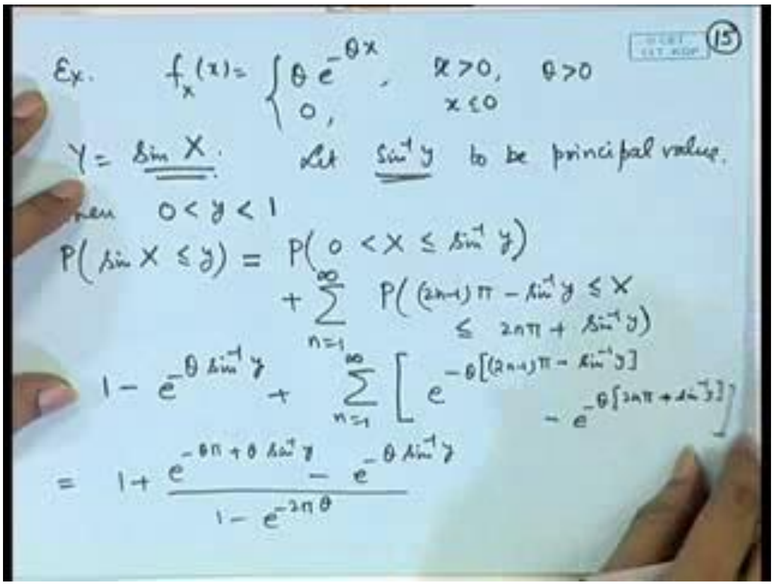


Let us consider say X follows normal 0, 1 distribution. So, the density function is 1 by root 2 pi e to the power minus x square by 2, consider the function say Y is equal to X Square. So, now, for a given y which is non negative, we have 2 inverse images; that is g 1 of y is equal to minus root y and plus root y. So, if I look at the derivatives that will be minus 1 by 2 root y or plus 1 by root y. So, if I take absolute value it is reducing to 1 by 2 root y; for y less than 0 there is no inverse image. So, the density function of y is the density function at X is equal to minus root y multiplied by 1 by 2 root y, plus the density function at plus root y multiplied by.

So, now you see the value will be 1 by root 2 pi e to the power minus y by 2, and in the second term also same terms will be coming because minus root y and plus root y both will give x square is equal to y. So, the 1 by 2 root y, 1 by 2 root y term is coming so it will become 2 times 1 by 2 root y that is equal to 1 by 2 to the power half gamma half, e to the power minus y by 2, y to the power 1 by 2 minus 1; this is nothing but a gamma distribution with r is equal to half and lambda is equal to half. So, you can see here the square of a standard normal random variable is a gamma random variable.

Now, sometimes it may happen that in place of finite number of inverse images we have infinite number of inverse images that may happen in cases such as periodic functions which are trigonometric functions such as sin function or cos function etcetera. So, the theorem can be extended here in place of a finite number of inverses we have infinite number of inverses. So, we again split the domain into infinite number of distinct regions such that each of them is mapping to the full range of g, calculate the density in each region by the same formula that is f x of g inverse y multiplied by the absolute value of d by d y of g inverse y, and add in all the regions.

(Refer Slide Time: 16:25)

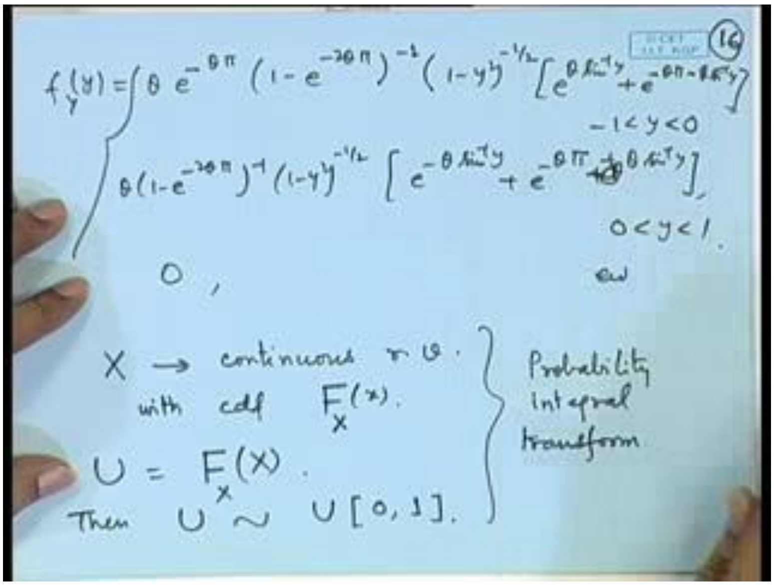


Let us take one example here say f x is theta e to the power minus theta x, for x greater than 0; that means, it is a exponential distribution with parameter theta, consider Y is equal to say sin of X. So, we consider sin inverse y to be principle value. So, we consider 2 cases; because sin lies between minus 1 to 1, so we consider case 0 to 1 and minus 1 to 0. So, if we take 0 to 1 then probability that sin X is less than or equal to y can be expressed as probability of 0 less than X less than or equal to sin inverse of y, where sin inverse is the principle value plus probability of 2 n minus 1 pi minus sin inverse y less than or equal to X, less than or equal to 2 n pi plus sin inverse y; n is equal to 1 to infinity.

So, this takes care of all the infinite number of distinct regions, each of which map to the region sin X less than or equal to y. So, now, this probability using the exponential density function, probability of X lying between 0 to sin inverse y is 1 minus e to the power minus theta sin inverse y, plus summation n is equal to 1 to infinity, e to the power minus theta, 2 n minus 1 pi minus sin inverse y, minus e to the power minus theta 2 n pi minus sin inverse y plus sin inverse y.

If we look at this series here, e to the power sin inverse y terms can be separated out and the remaining terms become geometric sums and infinite geometric series can be added. So, this is simplified to 1 plus e to the power minus theta pi, plus theta sin inverse y minus e to the power minus theta sin inverse y divided by 1 minus e to the power minus 2 pi theta. In a similar way if y is between minus 1 to 0, we can split the regions and evaluate.

(Refer Slide Time: 19:47)

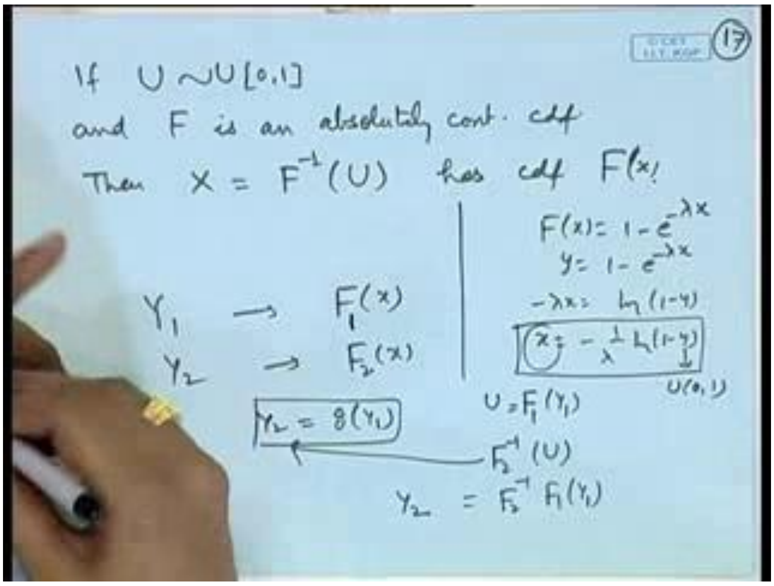


So, after carrying out the calculations, the density of y can be obtained after differentiation as theta e to the power minus theta pi, 1 minus e to the power minus 2 theta pi, inverse 1 minus y square to the power minus 1 by 2, e to the power theta sin inverse y, plus e to the power minus theta pi minus theta sin inverse y, this is for minus 1 less than y less than 0 and it is equal to theta 1 minus e to the power minus 2 theta pi inverse, 1 minus y square to the power minus 1 by 2,0 e to the power minus theta sin inverse y, plus e to the power minus theta pi plus theta sin inverse y, this is for y between 0 and 1 and of course, it is 0 for all other values of y, because y will lie between minus 1 to 1.

So, if the transformation y is equal to g x is such that the random variable y is equal to g x is also continuous, then the probability density function of y can be determined in terms of the density function in x, if the function is 1 to 1 there is a direct formula if there is a many 1 function then we have to split the domain into disjoint sets such that each part of the domain maps to the full range, calculate the inverse image in each of them and utilize that to find out the density function of y in each parts separately and then sum.

There is one important result here which connects all the continuous distributions, which is known as the probability integral transform; this basically says that if X is a continuous random variable with cdf capital F, if we define say U is equal to F of X then u is distributed on the uniform interval 0 to 1.

(Refer Slide Time: 22:47)



So, this is known as probability integral transform the converse of this result is also true that is if U is uniform 0, 1 and F is an absolutely continuous cdf then X is equal to F inverse of U has cdf F x basically F here.

Now, this is the very crucial result, first thing is that it connects all the continuous random variables. So, if Y 1 is a continuous random variable with some cdf F 1, Y 2 is a continuous random variable with some cdf then there exists a function g, such that Y 2 is equal to say g of y 1 and the distribution of y 2 will be given by this. So, basically what we can do is that we can consider F 1 of Y 1 that is a U and then we consider F 2 inverse of U then that will be this. So, basically F 2 inverse of F 1, Y 1 is equal to Y 2.

In the modern age of simulations this result is quite useful. So, in some practical problem we may be interested to generate the values of a random variable which is a having say exponential distribution. So, we will use a pseudo random number generator to generate numbers uniformly between 1 to say some n, and then we can consider division by n to make it a uniform random variable between the values of the uniform random variable on the interval 0 to 1.

Now, if we are having the exponential distribution. So, the cdf of that is known that is capital F is known, so, we take f inverse of that. So, suppose I consider F x is equal to 1 minus e to the power minus lambda x. So, if y is equal to 1 minus e to the power minus lambda x, the inverse function can be obtained. So, minus lambda x is equal to log of 1 minus y. So, x is equal to minus 1 by lambda log of 1 minus y. So, if y's are the uniform random variables on 0 to 1 then if we consider log of 1 minus y and minus 1 by lambda then x i's will be exponential distributed random variables with parameter lambda.

So, this transformation play extremely important role in the simulation of random variables, because we can use some pseudo random number generator to generate uniformly distributed random values and now for any other distribution we make use of the transformations. So, especially their probability integral transform is extremely useful in this and also we have the relationships between various other continuous distributions. So, the discussion on the distribution of the function random variables is quiet important in this sense that to simulate the values of various random variables we make use of these transformations. So now, we will proceed to the jointly distributed random variables in the next lecture.