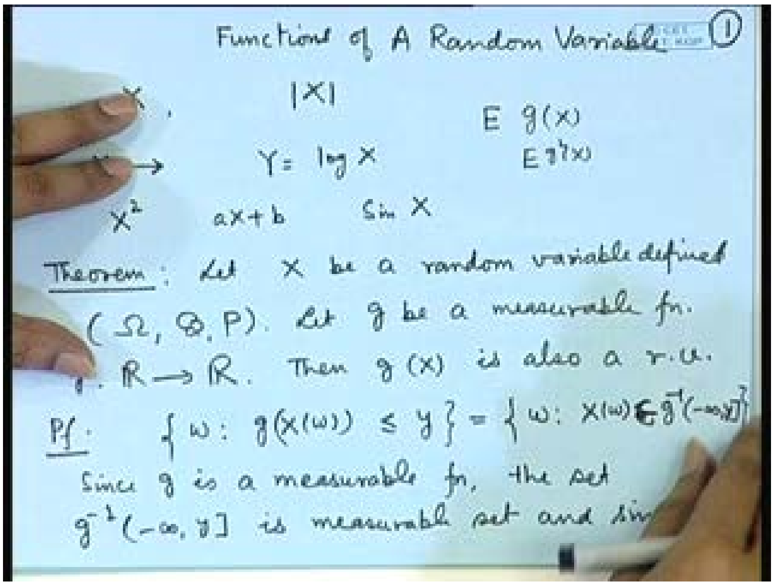
Probability and Statistics
Prof. Dr. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur
Module No. # 01

Lecture No. # 17
Function of a Random Variable

We have discussed the distributions of random variables. So, when we have a sample space and we are interested in certain characteristics arising out of that experiment, such as, we are recording the heights of the individuals, life of equipment, time taken by a sprinter to complete a 100 meter race, etcetera. So, these are the examples of random variables. However, many times we may not be interested directly in the same characteristic, but a function of that characteristic. Consider measurements, where we are recording the errors in the actual measurement. So, the errors may be negative or positive. However, it may turn out that, our losses are dependent upon the absolute value of the error in the recording.

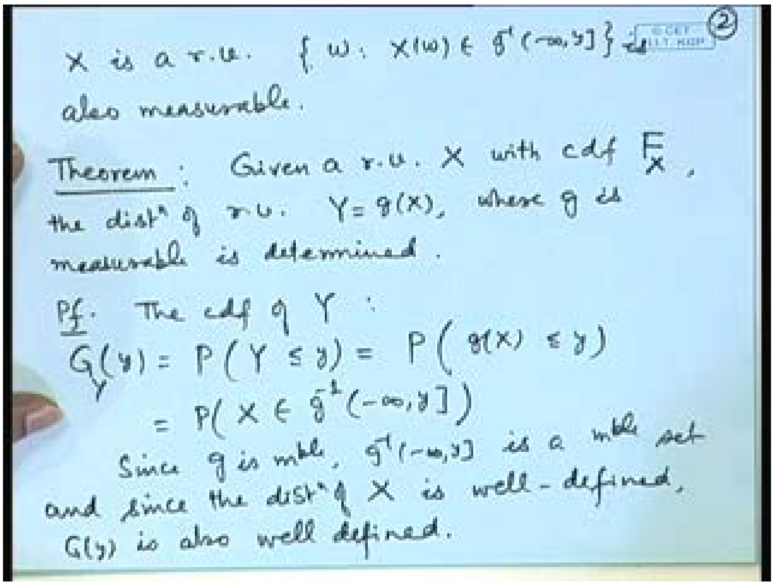
(Refer Slide Time: 01:20)



That means, in place of the random variable X, we may be interested in modulus of X. Suppose X denotes certain astronomical distances; now, if the distances or the numbers are too large, we may be interested in a function of X say log of X to make them scale to our level. In a similar way, sometimes we may be interested in X square, we may be interested in say aX plus b or sine of X, etcetera. Now, it is one thing to consider the characteristics of a function. So, in general, suppose I say g X. We may be interested to look at expectation of g X, expectation of g square X, variance of g X, etcetera. However, we may be interested in the actual distribution or the full probability distribution of g X also.

Now, in order to study the distribution of g X, it is important to firstly ensure that g X is also a random variable. Recall that the definition of a variable says that random variable is a real valued measurable function defined on the sample space. Therefore, if g is also a measurable function on the real line, then Y is equal to g X will also be a random variable. So, we start from this result. Let X be a random variable defined on say omega, B, P. Let g be a measurable function. So, g is a function from R to R. Then, g X is also a random variable. So, to look at the proof of this, we must prove that the set of all those points, such that g X omega is less than or equal to say y is a measurable set for every real y. Now, this is the set X omega belongs to g inverse of minus infinity to y. Now, if g is a measurable function, then g inverse minus infinity to y is a measurable set. And therefore, X omega belonging to that – this is also a measurable set since X is a random variable. So, since g is a measurable function, the set g inverse of minus infinity to y is measurable set.

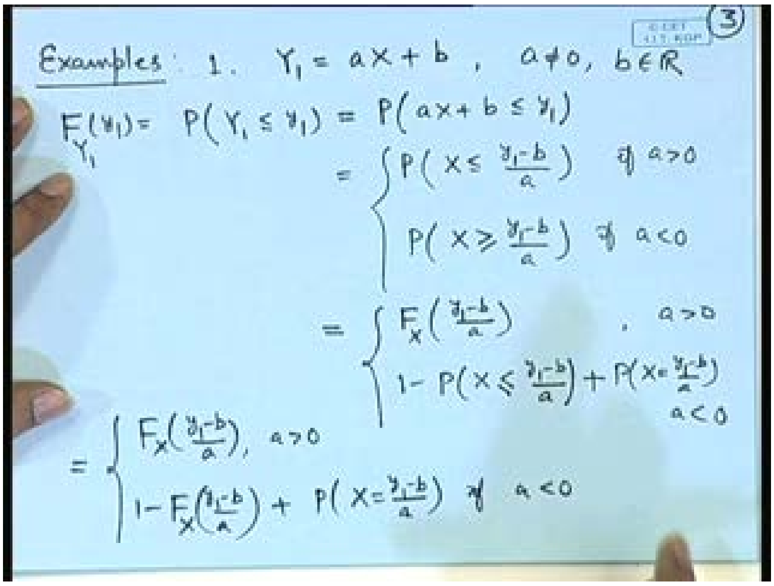
(Refer Slide Time: 05:00)



And, since X is a random variable, the set omega such that X omega belongs to g inverse minus infinity to y, is also measurable. So, in general, we will be considering measurable functions of random variable, so that to ensure that they are also random variable. And then, we can study the probability distributions. Now, we have to ensure that a probability distribution can be found. So, given a random variable X with cumulative distribution function, say capital F X, the distribution of random variable Y is equal to g X, where g is measurable is determined. So, now onwards, whenever we are considering the function g X, then g has to be a measurable function.

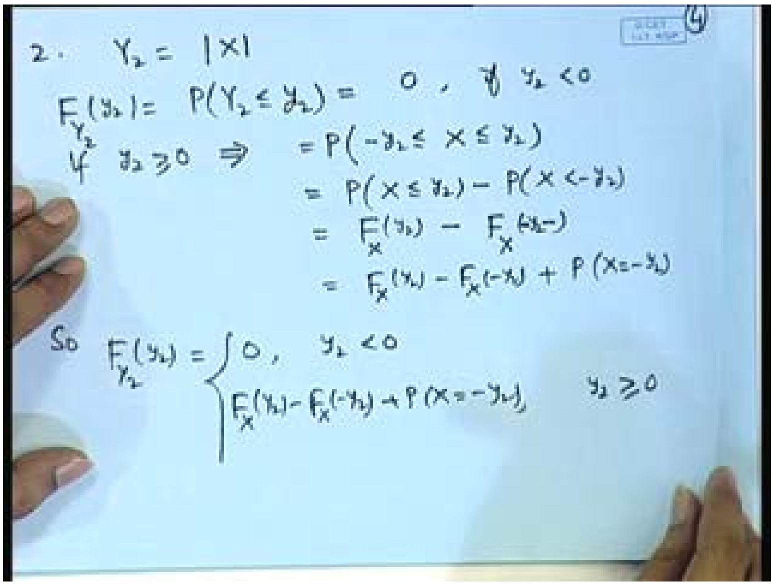
Let us consider the c d f of y. So, let me use the notation say G y, that is, probability of Y less than or equal to y. Now, this is equivalent to probability of g X less than or equal to y, which is equivalent to probability of X belonging to the set g inverse minus infinity to y. Now, g inverse minus infinity to y is a measurable set and the probability distribution of X is well-defined. Therefore, this probability can be determined (Refer Slide Time: 07:04). Since g is measurable, g inverse minus infinity to y is a measurable set. And, since the distribution of X is well-defined, G y is also well-defined. So, the basic existence of the probability distribution of a function of random variable is established. Now, let us look at the practical aspect of it; that means, how do we determine the distribution of a function of random variable. This is a general approach. So, if we use this c d f approach, that means we express the c d f of the function in terms of c d f of X here. So, let us look at this.

(Refer Slide Time: 08:18)



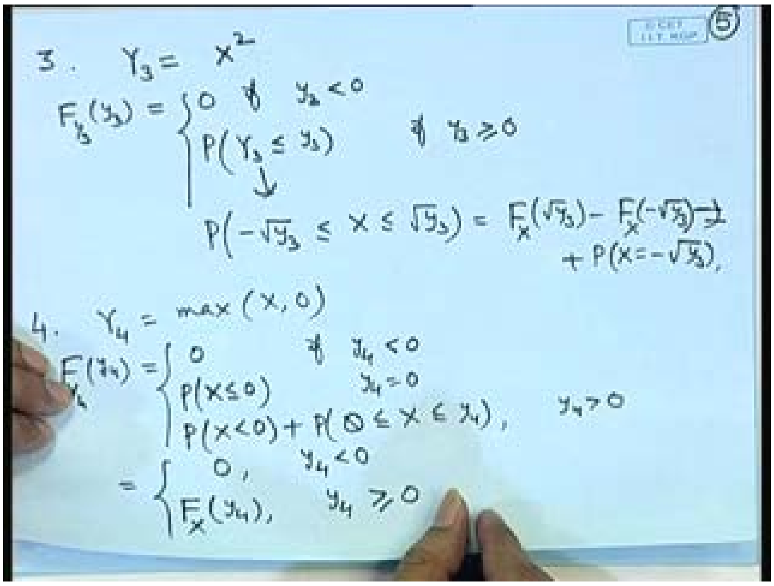
Let us consider a function say Y 1 is equal to say aX plus b. So, if we are looking at the c d f of y 1, here a is a nonzero constant and b is any real. So, this is probability of Y 1 less than or equal to say small of y 1. Now, this is aX plus b less than or equal to y 1. Now, this can be expressed as probability of X less than or equal to y 1 minus b by a if a is positive. And, it will become probability of X greater than or equal to y 1 minus b by a if a is negative. So, notice here that both of these are certain probabilities related to the random variable X. If the c d f of X is well-defined, that means, the capital F X is known, then both of these probabilities are also known; that means, this is equal to for example, F of X at y 1 minus b by a. And, the second term we can write as 1 minus probability of X less than y 1 minus b by a, which we can write as X less than or equal to this plus probability of X is equal to y 1 minus b by a. So, this is equal to F X at y 1 minus b by a if a is positive; and, 1 minus F of X y 1 minus b by a plus probability of X is equal to y 1 minus b by a if a is negative. Therefore, you can see that the c d f of y 1 is well-determined. Another thing that we can note here is that if X is a continuous random variable, then this probability will be 0. So, we will have only these terms here.

(Refer Slide Time: 10:40)



Let us take another example say Y 2 is equal to modulus of X. So, if we look at the c d f of Y 2 – now, notice here that if y 2 is a negative number, then this probability is going to be 0, because modulus of a random variable is always a non-negative quantity. So, this is 0 if y 2 is less than 0. Now, if y 2 is greater than or equal to 0, then we can express this as probability of minus y 2 less than or equal to X less than or equal to y 2. So, this is equal to probability of X less than or equal to y 2 minus probability of X less than minus y 2, which is nothing but the c d f of X at the point y 2 minus the left-hand limit of the c d f at minus y 2; or, we can write it as F X of y 2 minus F X at minus y 2 plus probability of X is equal to minus y 2. Therefore, the c d f of y 2 is expressed as 0 if y 2 is less than 0 and it is F X of y 2 minus F X of minus y 2 plus probability of X equal to minus y 2 if y 2 is greater than or equal to 0. Here (Refer Slide Time: 12:26) you can note that if y 2 is equal to 0, then this term will vanish and this will reduce to probability of X is equal to 0. Also, if X is a continuous random variable, then this term will vanish.

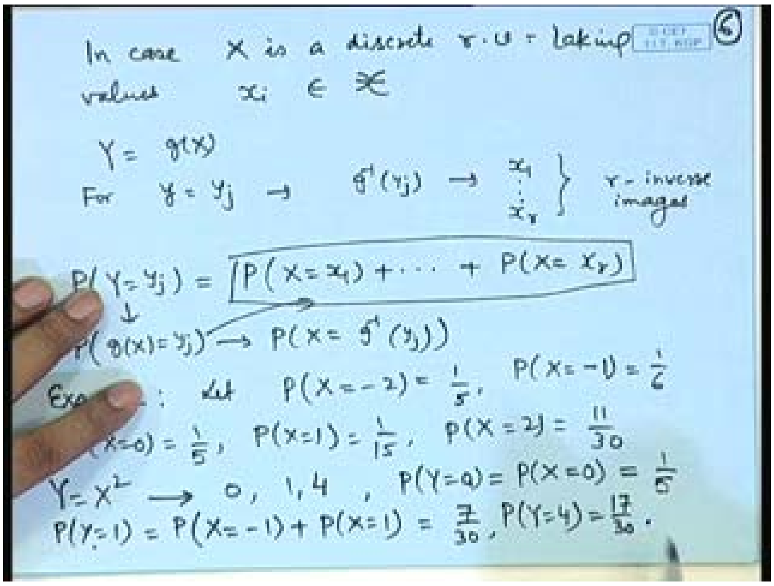
(Refer Slide Time: 12:48)



Let us take say Y 3 is equal to X square. Once again if I consider F of Y 3, then this is 0 if y 3 is less than 0. And, it is equal to probability of Y 3 less than or equal to y 3 if y 3 is greater than or equal to 0. Now, this quantity becomes probability of minus square root y 3 less than or equal to X less than or equal to root of y 3. So, that is equal to F X of root y 3 minus F X of minus root y 3 minus, that is, the left-hand limit at this point, which we can express as this plus probability of X is equal to minus root y 3.

Let us consider say function Y 4 is maximum of X and 0. So, here F of Y 4 – once again you can notice here that this random variable is also no negative; so, this will be 0 if y 4 is negative. If y 4 is equal to 0, then this is simply probability of X less than or equal to 0. And, if I consider y 4 positive, then it is the probability of X less than or equal to y 4. So, that will be X less than 0 plus probability of 0 less than or equal to X less than or equal to y 4. So, if I combine these terms, then this is equal to 0, if y 4 is less than 0; and, it is F X of y 4 if y 4 is greater than or equal to 0. You can consider it as truncation of X at 0. So, this approach, where we find out the inverse in each of the set g X less than or equal to c, that is, g inverse of minus infinity to c and then expressing the c d f of the function g X in terms of the c d f of x. However, many times the function may be quite complicated and we may not be able to express the inverse image set in a proper way. So, we look at the particular approaches, when the random variables are discrete or continuous. And, we may consider the special methods, for example, if I have a one-to-one function or if I have two-to-one function, and if the random variable is discrete, then we may express the probability of a point in terms of the inverse images taken by the random variable. If we have a continuous random variable, then in place of probability mass function, I will have a probability density function; then, there is some method. So, let us develop these methods.

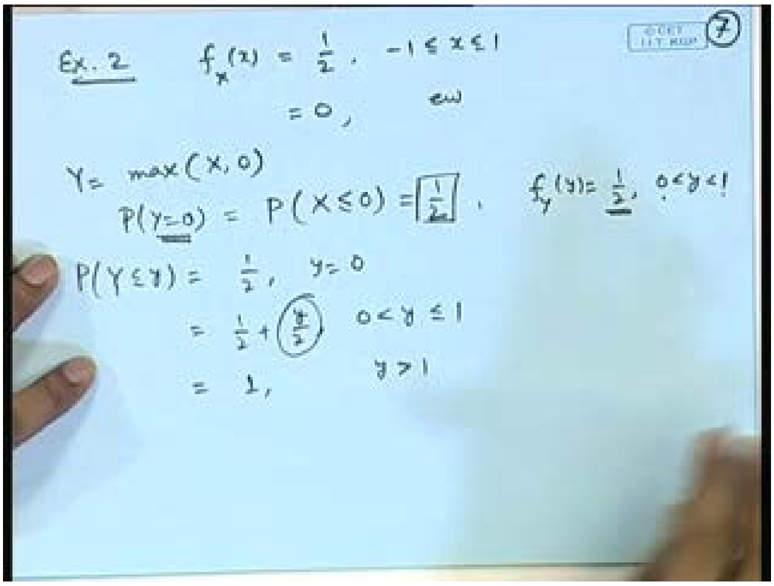
(Refer Slide Time: 16:21)



In case X is a discrete random variable say the random variable is taking values, say x i is belonging to some set script X; Y is equal to g X, is a function. So, we consider that for y is equal to y j, g inverse of y j could be x 1, x 2, x r say; that means, in general, suppose r inverse images are there, then probability of Y is equal to y j – that will be equal to probability of X is equal to x 1 plus probability X equal to x 2 plus probability X equal to x r, because what we are doing is, we are writing it as probability of g X is equal to y j; that is probability of X is equal to g inverse of y j. Now, what are the values of x i, which lead to g X is equal to y j. So, we look at the set of inverse images and add all the probabilities.

As an example, consider probability X is equal to say minus 2, is equal to 1 by 5; probability X is equal to minus 1, is equal to say 1 by 6; probability X equal to 0 is 1 by 5; probability X equal to 1 is say 1 by 15; and, probability X is equal to 2 is say 11 by 30. Let us consider say Y is equal to X square. Now, here X takes values minus 2, minus 1, 0, 1 and 2. So, this Y will take values 0, 1 and 4. Now, probability of Y is equal to 0. Now, this has only 1 inverse image, that is, probability X equal to 0. That will be 1 by 5. If we are looking at probability Y is equal to 1, then there are two inverse images: X equal to minus 1 and X equal to 1. So, the probability is 1 by 6 and 1 by 15, will be added up; we get 7 by 30. In a similar way, probability Y is equal to 4. Now, this will be probability of X equal to minus 2 and probability X equal to 2. So, we will add 1 by 5 and 11 by 30, which is leading to 17 by 30. So, the probability distribution of Y you can see here, is described by probability Y is equal to 0, probability Y is equal to 1 and probability Y is equal to 4. Now, this approach is not directly applicable when we have say mixture random variable or if the function g X is such that the X may discrete, but g X may not be discrete; or, X could be continuous, but g X may not be continuous.

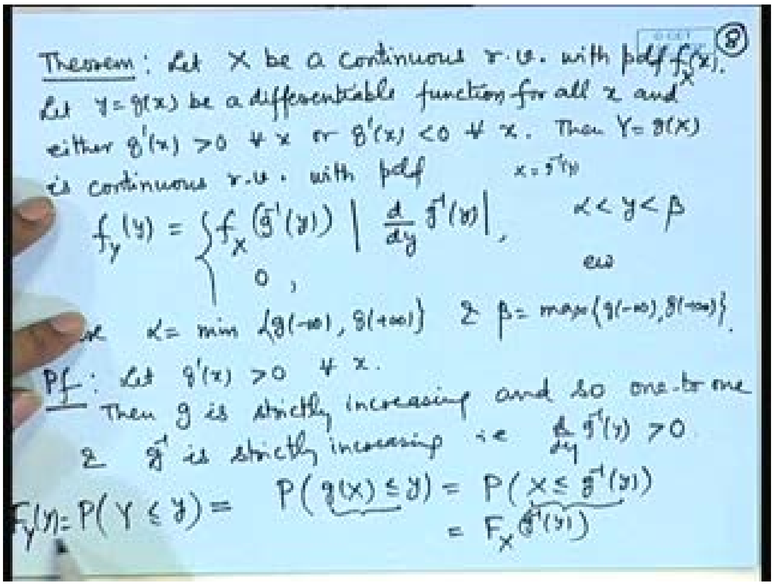
(Refer Slide Time: 19:56)



Let us take this case. Suppose X is a uniform random variable on the interval minus 1 to 1. And, we consider the function say Y is equal to maximum of X and 0. Now, you can see here that this distribution, this random variable is not completely discrete nor continuous, because probability of Y is equal to 0 is half. This is equivalent to probability X less than or equal to 0; that is equal to half. But, if I look at probability Y is equal to half, etcetera, then that is 0, because there after, the random variable is continuous. If X is non-negative, then Y becomes continuous random variable. So, in this case, we can use the formula that we developed for the (Refer Slide Time: 20:59) c d f of Y 4, that is, maximum of X, 0. So, for y 4 less than 0, it is 0. And, thereafter, it is F X of y 4. So, if we use this, we get probability of Y less than or equal to y; it is equal to half for y is equal to 0 and it is equal to half plus y by 2 for 0 less than y less than or equal to 1. And, of course, it is 1 for y greater than 1.

We have then density function here, that is, f Y y is equal to half for 0 less than y less than 1 (Refer Slide Time: 21:40). We have the weight half attached to probability Y is equal to 0; and, in the interval 0 to 1, we have a density function. So, it is an example of mixed random variable. Although the random variable X is a continuous random variable, but the function of that is a mixture random variable. We have certain conditions under which from a continuous random variable, the function is also continuous random variable and the probability density function of the given function Y is equal to g X can be determined using a formula. So, we state it in the following theorem.

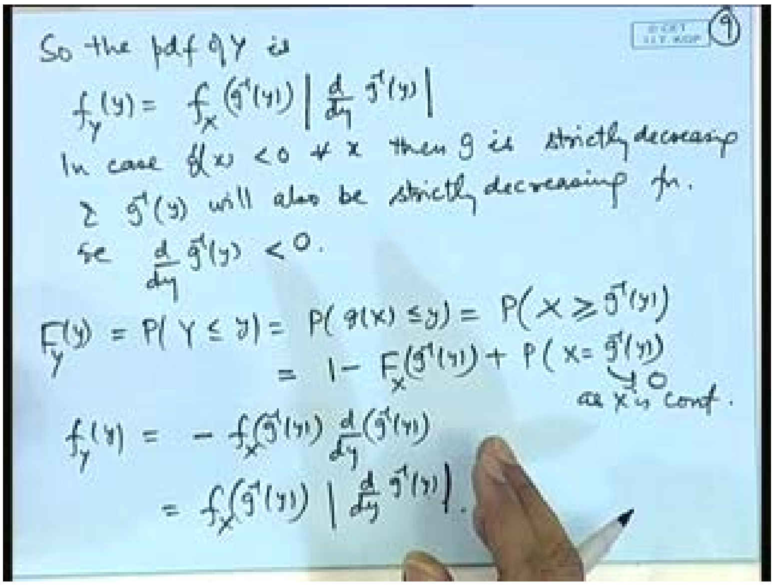
(Refer Slide Time: 22:35)



Let X be a continuous random variable with probability density function say f X X. Let y is equal to g x be a differentiable function for all x and either g prime x is positive for all x or g prime x is negative for all x. Then, Y is equal to g X, is a continuous random variable with p d F given by f Y y is equal to f X g inverse y multiplied by the absolute value of d by d y g inverse y over the range alpha to beta and 0 elsewhere, where alpha is the minimum of the g minus infinity, g of plus infinity. And, beta is the maximum of g minis infinity, g of plus infinity. Notice here the conditions. We are taking the function, such that the function is differentiable everywhere and either the derivative is strictly positive throughout the range or strictly negative. This ensures that the function is strictly increasing or strictly decreasing. Also, it ensures that the function will be a one-to-one function. In that case, there is a direct formula for the determination of the probability density function of Y is equal to g X. It is described in terms of the density function of X itself; that is, in place of X, we substitute g inverse y, which is uniquely determined under the conditions given here. And, we multiply by absolute value of d x by d y, because if y is equal to g x, then x is equal to g inverse y. And, this term is nothing but dx by dy.

Let us consider the proof of this. Let g prime x to be strictly positive for all x. Then, g is strictly increasing; and so, it will be a one-to-one function. And, g inverse will also be strictly increasing, that is, d by d y of g inverse y will be positive. And, this will be differentiable also. So, if we consider the c d f of Y; and, another thing is that the range we have to consider. If x is over certain range; suppose from minus infinity to infinity in general; then, if g is an increasing function, then the minimum value will be g of minus infinity and the maximum value will be g of plus infinity. If g is a strictly decreasing function, then it will be reverse. So, here the range of y will be from alpha to beta, where alpha and beta are defined like this. So, this is equal to probability of g X less than or equal to y; that is, probability of X less than or equal to g inverse y. This is ensured, because g is one-to-one function. So, g inverse is a one-to-one function. Therefore, the regions g X less than or equal to y and X less than or equal to g inverse y are equivalent. That means, whenever g X less than or equal to y is satisfied, X less than or equal to g inverse y is also satisfied. So, this is nothing but the c d f of X at g inverse y. Therefore, the density function of y will be determined by differentiation of capital F Y.

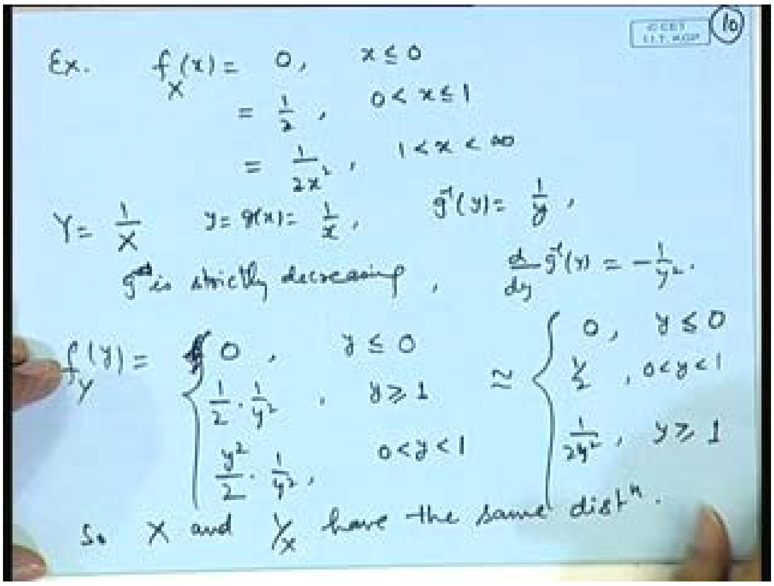
(Refer Slide Time: 27:33)



So, we get the probability density function of Y is f Y is equal to f X g inverse y multiplied by d by d y of g inverse y. Since g inverse y was increasing function, the derivative was positive. Therefore, this is equivalent to absolute value of this. Now, in case g prime x is strictly negative, then g is strictly decreasing and g inverse y will also be a strictly decreasing function; that is, d by d y of g inverse y will be less than 0. So, when we consider probability of Y less than or equal to y, this will be equivalent to X greater than or equal to g inverse y. Since g is a decreasing function, here the event g X less than or equal to y will be equivalent to X greater than or equal to g inverse y. So, which we can write as 1 minus F X of g inverse y plus probability of X is equal to g inverse y.

Now, this term will be 0, because X is continuous. So, the c d f of y is determined in terms of the c d f of X. So, if you differentiate, we get the p d f of y as minus f X g inverse y into d by d y of g inverse y. Since d by d y of g inverse y is negative, minus of this is absolute value (Refer Slide Time: 29:42). So, this is f X of g inverse y multiplied by d by d y of g inverse y. So, you can see that in both the cases, the density function of y is determined as the density function of X at the point X is equal to g inverse y multiplied by the absolute value of d x by d y. So, this theorem is useful when the function g X is a strictly increasing or strictly decreasing function.

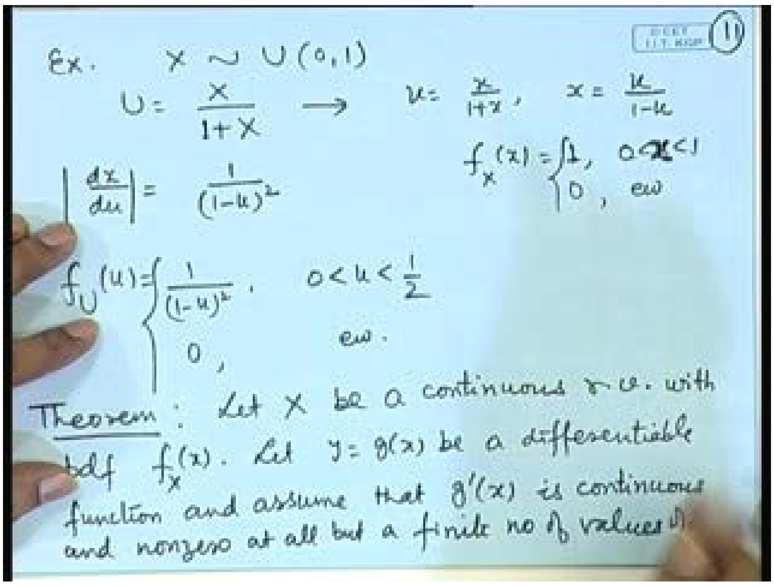
(Refer Slide Time: 30:24)



Let us take an example say the random variable x is having a 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, 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 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 is determined by the density function of x at (Refer Slide Time: 31:43) g inverse y multiplied by the absolute value of the d x by d y term. Now, here it is 0 (Refer Slide Time: 31:51). So, it will remain 0. Whenever x is less than or equal to 0, 1 by x is also less than or equal or 0. When it is half, this remains half multiplied by 1 by y 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, it will reduce to 0 less than y less than 1.

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 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: 33:42)

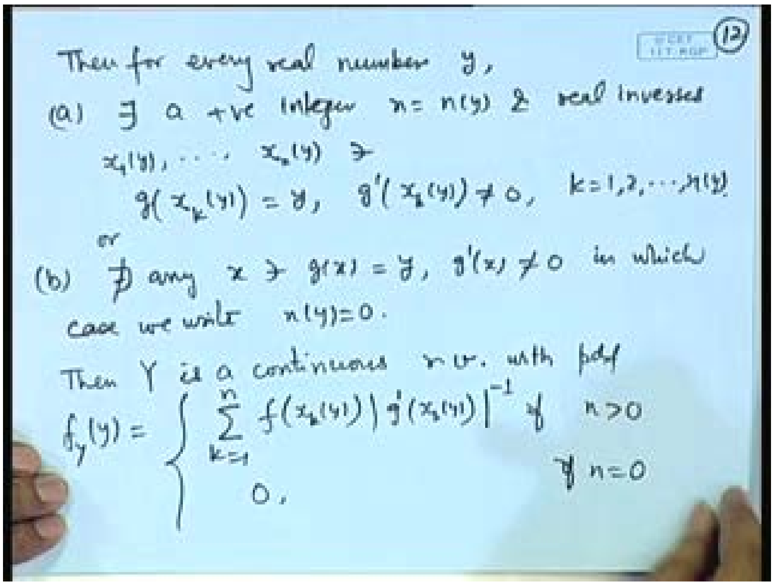


Let us look at another example. Say X follows uniform 0, 1 and we define random variable U is equal to X divided by 1 plus X. Now again, you can see here that this is a one-to-one function. In fact, if u is equal to x by 1 plus x, then the inverse function x is equal to u by 1 minus u. Let us look at the derivative dx by du; 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 minus 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 is 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 one-to-one 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, if there are two inverse images, then we split the region, that is, the domain of X into 2 disjoint regions, such that both of them 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 two inverse images: minus square root y and plus square root y. So, if I consider two 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 part of the domain, the function will be one-to-one. 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 region 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.

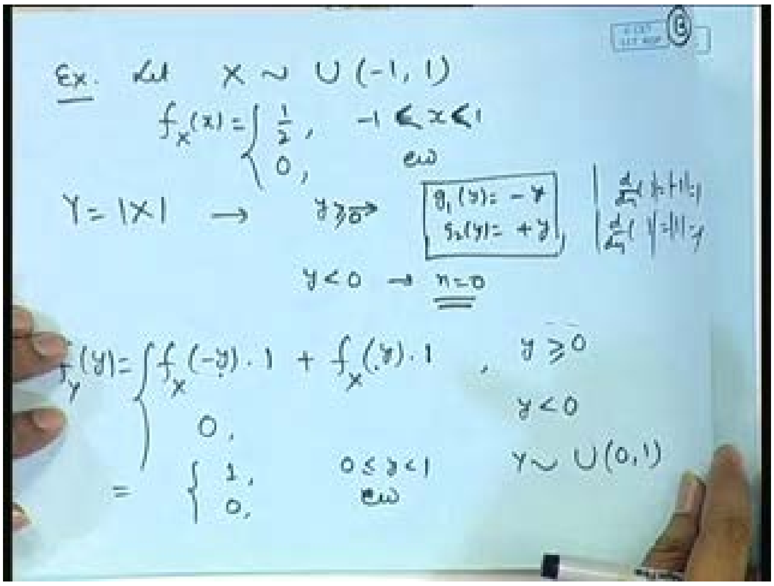
So, let us look at the result here. Let X (Refer Slide Time: 37:13) be a continuous random variable with probability density function say f X x. Let y is equal to g x be a differentiable function and assume that g prime x is continuous and nonzero at all, but a finite number of values of x.

(Refer Slide Time: 38:16)



Then, for every real number y, there exists a positive integer n is equal to say n of y and real inverses say x 1 of y, x 2 of y, x n of y such that g of x k of y is equal to y and g prime of x k of y is not 0 for k equal to 1, 2 up to n of y. Or, there 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 p d f 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.

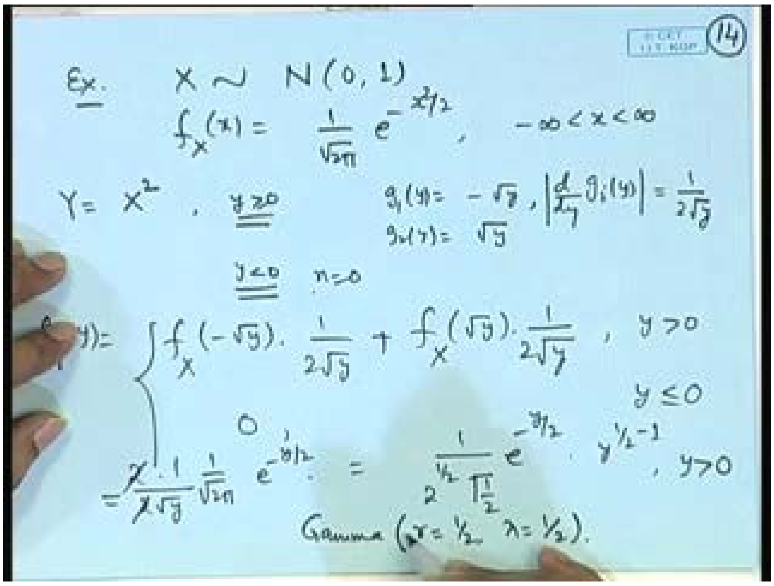
(Refer Slide Time: 40:42)



So, let us look at application of this. I am skipping the proof of this theorem. Let X follow say 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, two inverse images for a given y is there for y positive. If y is negative, n is equal to zero; 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.

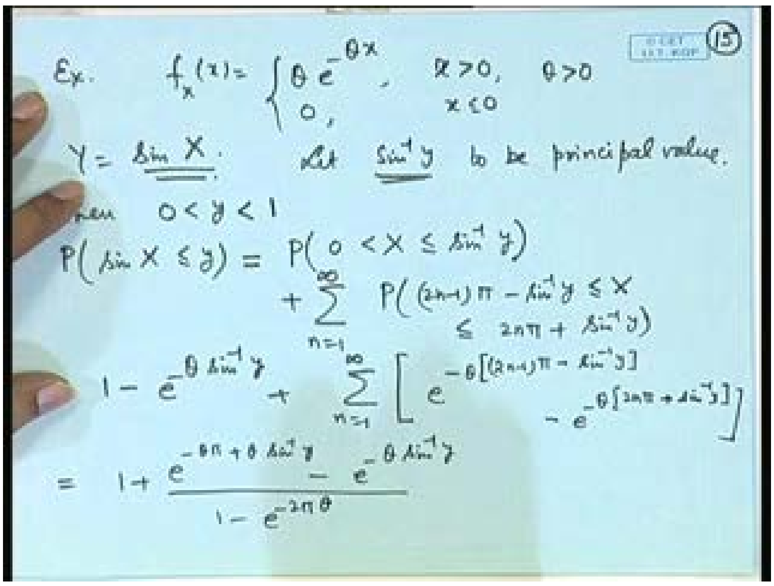
(Refer Slide Time: 43:00)



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 two inverse images; that is, g 1 of y is equal to minus root y and g 2 of y is equal to plus root y. So, if I look at the derivatives, that will be minus 1 by 2 root y or plus 1 by 2 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 at x equal to minus root y multiplied by 1 by 2 root y plus the density function at plus root y multiplied by 1 by 2 root y. 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, 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 into y to the power 1 by 2 minus 1 for y greater than 0. This is nothing but a gamma distribution with r is equal to half and lambda is equal to half. So, you can see here (Refer Slide Time: 45:18) 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 sine 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.

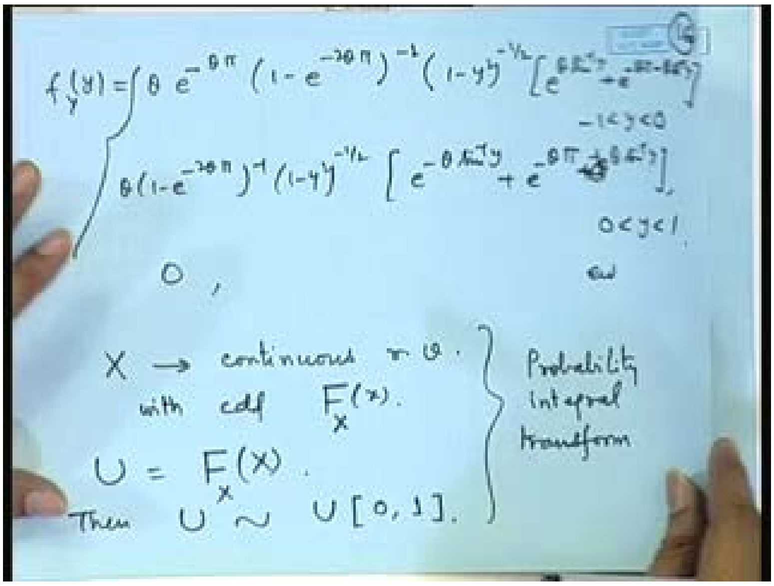
(Refer Slide Time: 46:25)



Let us take one example here say f X x is theta e to the power minus theta x for x greater than 0; that means, it is an exponential distribution with parameter theta. Consider Y is equal to say sine of X. So, we consider sine inverse y to be principal value. So, we consider two cases, because sine 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 sine X is less than or equal to y can be expressed as probability of 0 less than X less than or equal to sine inverse of y, where sine inverse is the principal value plus probability of 2n minus 1 pi minus sine inverse y less than or equal to X less than or equal to 2n pi plus sine 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 sine X less than or equal to y.

Now, this probability using the exponential density function, probability (Refer Slide Time: 48:19) of X lying between 0 to sine inverse y is 1 minus e to the power minus theta sine inverse y plus summation n is equal to 1 to infinity e to the power minus theta 2n minus 1 pi minus sine inverse y minus e to the power minus theta 2n pi plus sine inverse y. If we look at this series here, e to the power sine 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 sine inverse y minus e to the power minus theta sine inverse y divided by 1 minus e to the power minus 2 pi theta.

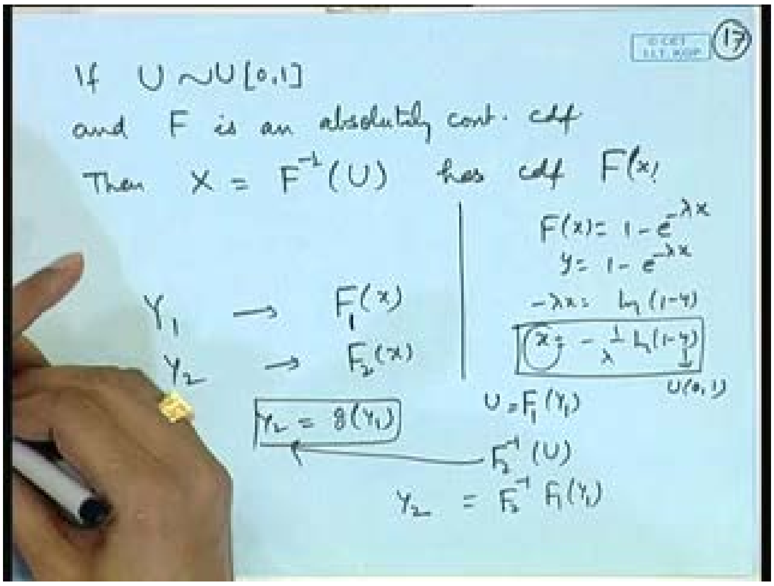
(Refer Slide Time: 49:49)



In a similar way, if y between minus 1 to 0, (Refer Slide Time: 49:40) we can split the regions and evaluate. 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 sine inverse y plus e to the power minus theta pi minus theta sine 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 e to the power minus theta sine inverse y plus e to the power minus theta pi plus theta sine 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 density function of X. If the function is one-to-one, there is a direct formula. If there is a many-one 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 part 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 (Refer Slide Time: 52:01) X is a continuous random variable with c d f, capital F; if we define say U is equal to F X of X, then U is distributed on the uniform interval 0 to 1. So, this is known as probability integral transform.

(Refer Slide Time: 52:48)



The converse of this result is also true; that is, if U is uniform 0, 1 and F is an absolutely continuous c d f, then X is equal to F inverse of U, has c d f F X; basically, F here. Now, this is a 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 c d f F 1, Y 2 is a continuous random variable with some c d f, 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, say 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 said 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, the c d f of that is known, that is, capital F is known, we take F inverse of that. So, suppose I consider F x is equal 1 minus e to the power 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 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.

These transformations 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 distributions, we make use of the transformations. So, especially the 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 number variables is quite important in this sense that to simulate values of various random variables, we make use of these transformations. So, we will proceed to the jointly distributed random variables in the next lecture.