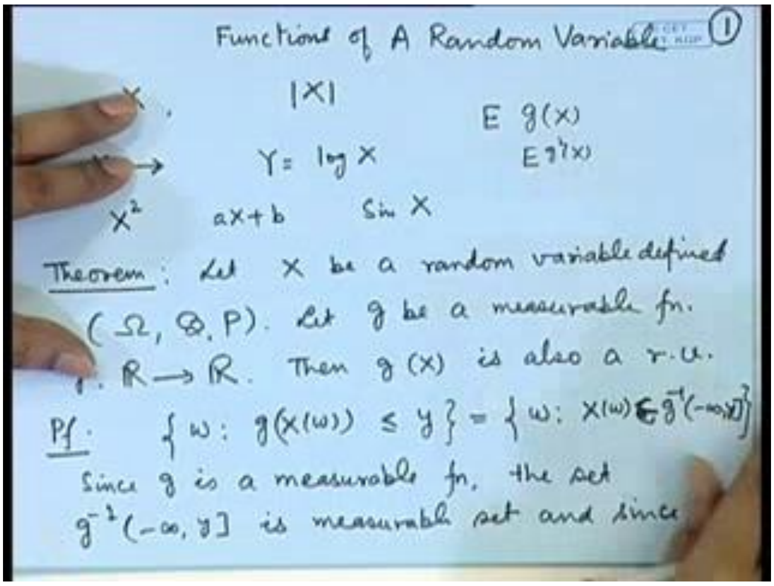
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 33
Function of a Random Variable – I

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 individual's, 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 loses are dependent upon the absolute value of the error in the recording.

(Refer Slide Time: 01:19)

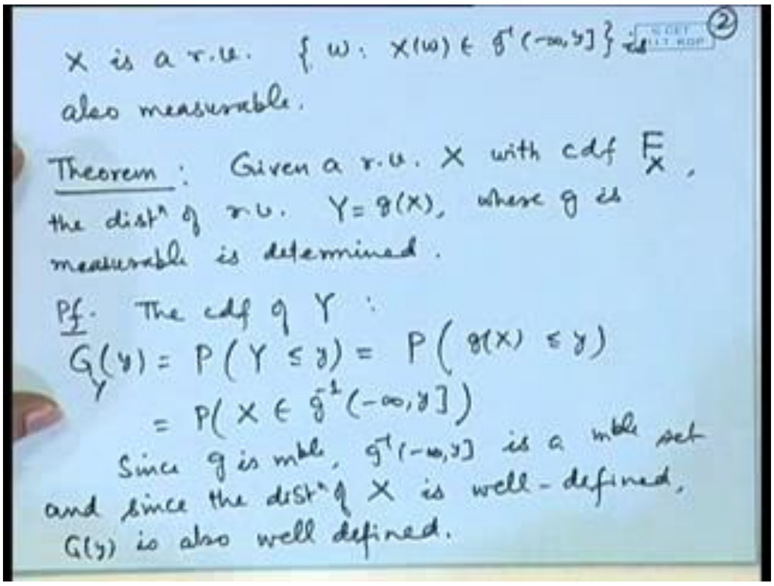


That means in place of the random variable X we may be interested in X square modulus of X. Suppose X denotes certain astronomical distances, now if the distances are the numbers are too large we may be interested in a function of X say log of X to make them scale to our level. In a similar way sometimes we may be interested in X square we may be interested in say a x plus b or sin of X etcetera.

Now, it is one thing to consider the characteristics of a function. So, in general suppose I say g x. So, we may be interested into 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 random 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 is will also be a random variable.

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 less than or equal to or you can say 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 random variable. So, since g is a measurable function the set g inverse of minus infinity to y is a measurable set.

(Refer Slide Time: 04:56)



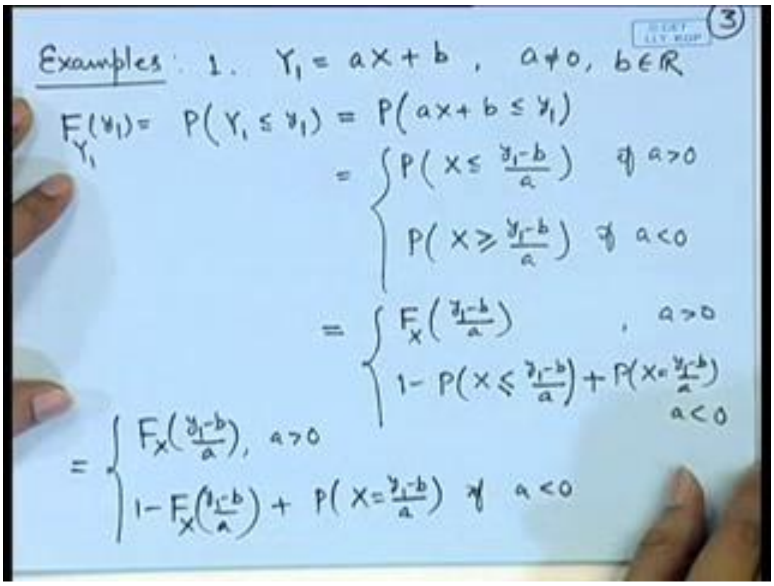
And since X is a random variable the set omega such that X omega belongs to g inverse minus infinity to y is also measurable.

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. So 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, the distribution of random variable y is equal to g x where g is measurable is determined. So, now onwards whenever we are considering a function g x then g has to be a measurable function.

Let us consider the cdf of y. So, let me use the notations 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. 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 this; that means, how do we determine the distribution of a function of a random variable? This is the general approach. So, if we use this cdf approach; that means, we express the cdf of a function in terms of the cdf of X here. So, let us look at this.

(Refer Slide Time: 08:23)

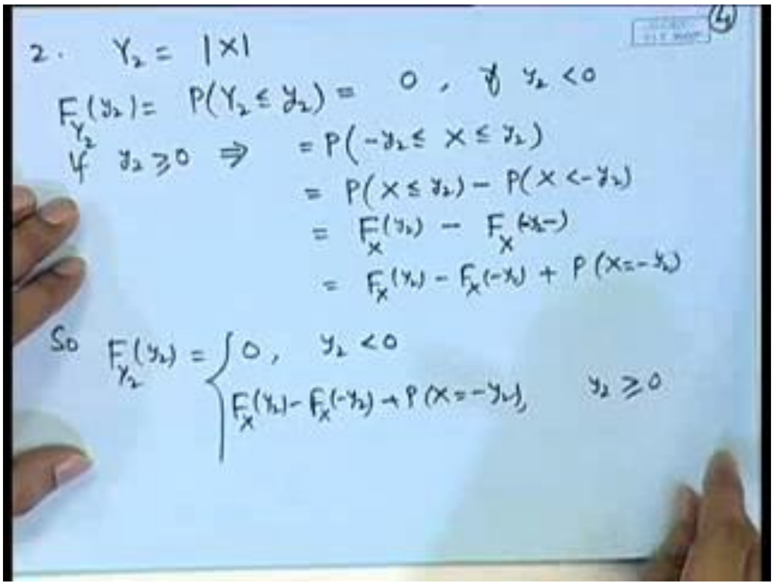


Let us consider a function say y 1 is equal to say a x plus b. So, if we are looking at the cdf of y 1, so, here a is a non 0 constant and b is an areal. So, this is probability of y 1 less than or equal to say a small of y 1, now this is a x plus b less than or equal to y 1. Now this can be expressed 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.

Notice here that both of these are certain probabilities related to the random variable X, if the cdf of X is well defined; that means, the capital F; X is the known then both of these probabilities are also known. That means, this is equal to for example, F of x 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 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 cdf of y 1 is well determined.

Another thing that we can note here is that if X is a continuous random variable then probability will be 0. So, you will have only these terms here.

(Refer Slide Time: 10:37)

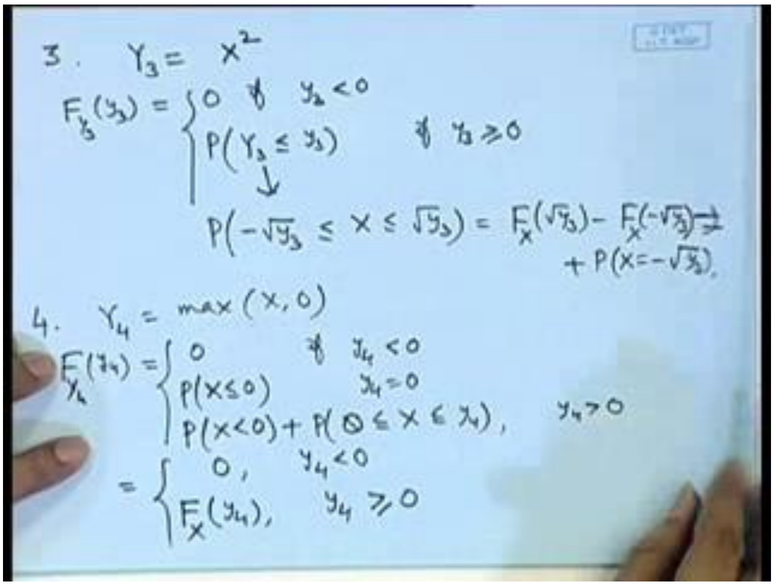


Let us take another example say Y 2 is equal to modulus of X. So, if we look at the cdf of y 2, now notice here that if Y 2 is a negative number then this probability is going to be 0 because modulus of a random variable is always a non negative quantity. So, this is

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 of minus y 2 which is nothing but the cdf of X a the point y 2 minus the left hand limit of the cdf 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 cdf 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 F x plus probability of X minus y 2 if y 2 is greater than or equal to 0. Here you can note that if y 2 is equal to 0 then this term will vanish and this will reduced to probability of X is equal to 0 also if y 2 if X is a continuous random variable then this term will vanish.

(Refer Slide Time: 12:47)



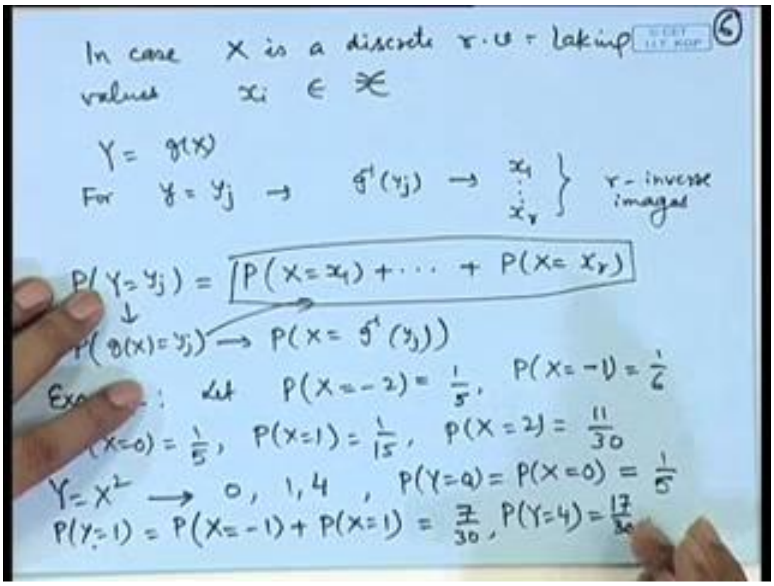
Let us say y 3 is equal to X square once again if I consider F of y 3 at then this is 0 if y 3 is less than 0 and it is equal to probability of y 3 less than or equal to y 3 if y 3 is greater than or equal to 0. Now this quantity becomes probability of minus square root y 3 less than or equal to X less than or equal to root of y 3. So, that is equal to f of root y 3 minus F 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 a function y 4 is maximum of X and 0. So, here F of y 4 once again you can notice that this random variable is also non-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 a probability of X less than or equal to y 4. So, that will be X less 0 plus probability X 0 less than or equal to X less than or equal to y 4, see if I combine these terms then this is equal to 0 if y 4 is less than 0 and it is F of y 4, if y 4 is greater than or equal to 0 you can consider it as truncation of X at 0.

This approach where we find out the inverse image of the set g x less than or equal to c that is g inverse of minus infinity to c, and then expressing the cdf of the function g x as the in terms of the cdf 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 1 to 1 function or if I have a 2 2 1 function and if the random variable is discrete then we express the probability of a point in terms of the inverse image 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 and then there is some method.

So, let us develop these methods.

(Refer Slide Time: 16:23)



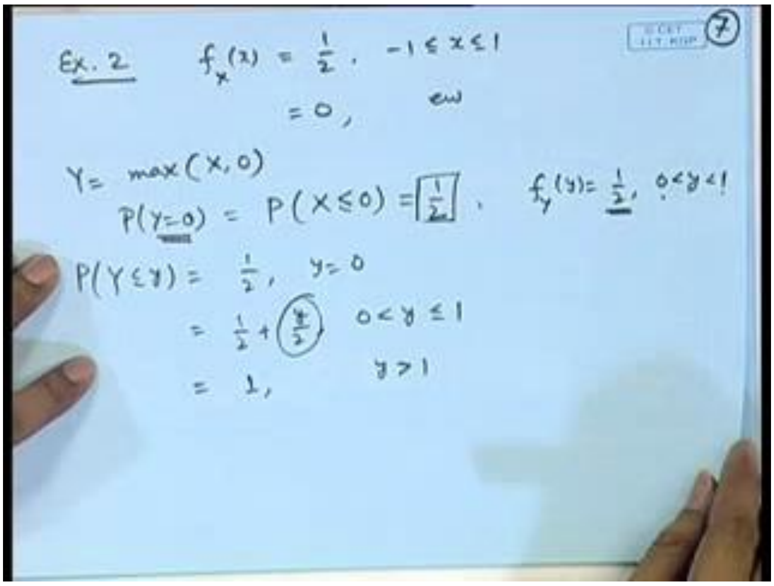
In case X is a discrete random variable say the random variable is taking values say X i’s belonging to some set a script X y is equal to g x is a function. So, we consider that for y is equal to y j g inverse of y j could be X 1 X 2 X R say; that means, in general suppose

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's which lead to g x is equal to y j. So, we look at the set of the 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 is equal to 0 is 1 by 5 probability of X is 1 is say one by 15 and probability X is equal to 2, we say 11 by 30, let us consider 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 is equal to 0 that will be one by 5 if we are looking at probability y is equal to 1 then there are 2 inverse images X is equal to minus one and X equal to one. 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 is 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 one and probability of y is equal to 4, now this approach is not directly applicable when we have a say mixture random variable or if the function g x is such that the X may be discrete, but g x may not be discrete or X could be continuous, but g x may not be continuous.

(Refer Slide Time: 19:53)

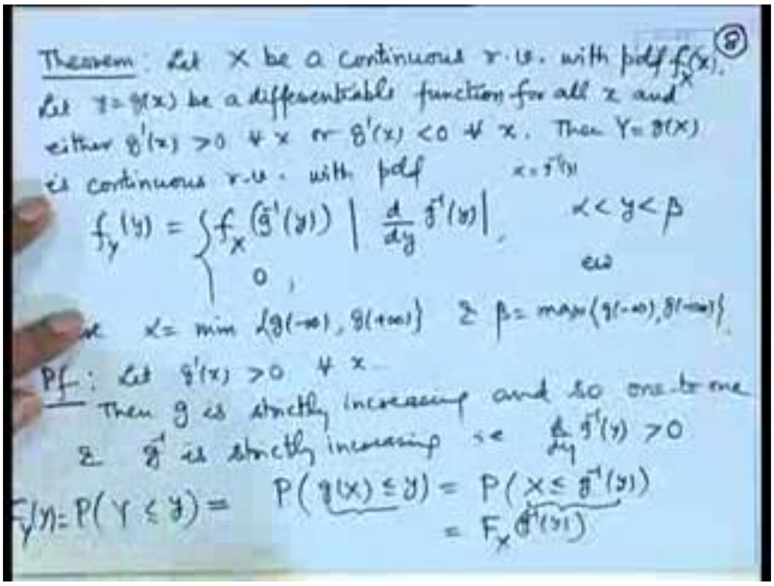


Let us take this case suppose X is a uniform random variable on the interval minus 1 to 1 and we consider the function say y is equal to maximum of X and 0. Now you can see here that this distribution this random variable is not completely discrete nor continuous because probability of y is equal to 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 a continuous random variable. So, in this case we can use the formula that we developed for the cdf of y 4 that is maximum of X 0.

For y 4 less than 0 it is 0 and there after it is f 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 one and of course, it is one for y greater than 1. So, we have then density function here that is F y is equal to half for 0 less than y less than 1 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 a example of mixed random variable all though 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 a continuous random variable and the function the probability density function of the given function y is equal to g x can determined using a formula.

So, we state it in the following theorem.

(Refer Slide Time: 22:36)



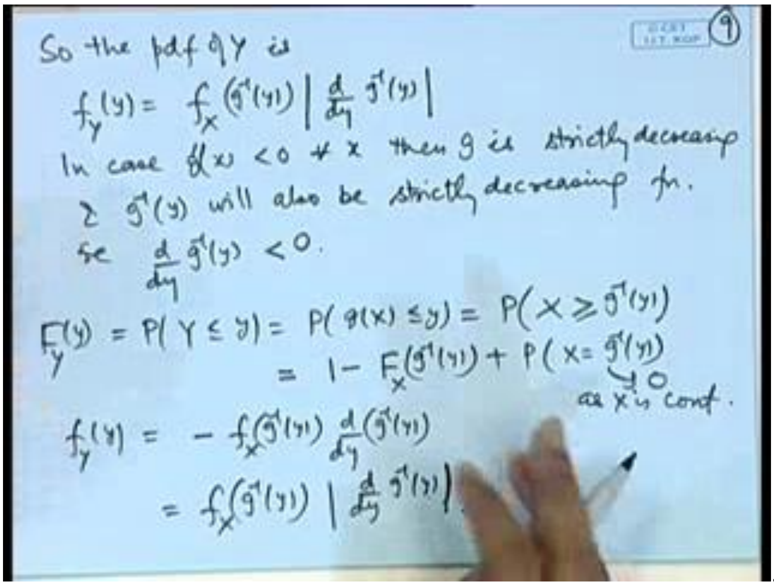
Let X be a continuous random variable with probability density function say F x, let y is equal to g x be a differentiable function for all X and either g prime X is positive for all X or g prime X is negative for all X then y is equal to g x is continuous random variable with pdf given by F y is equal 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 notice here the conditions we are taking the function such that the function is differentiable everywhere and either the derivative is strictly possible throughout the range or strictly negative.

This ensures that the function is a 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 d X by d y.

Let us consider the proof of this. So, let g prime X to be strictly positive for all X then g is a strictly increasing and. So, it will be a one to one function and g inverse will also be is 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 cdf 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 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 reversed.

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 a 1 to 1 function. So, g inverse is a 1 to 1 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 satisfied X less than or equal to g inverse y is also satisfied. So, this is nothing, but the cdf 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)



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 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 a 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 g.

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 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 cdf of y is determined in terms of the cdf of x.

So, if you differentiate we get the pdf 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 so minus of this is the absolute value. 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 the 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 a strictly decreasing function.