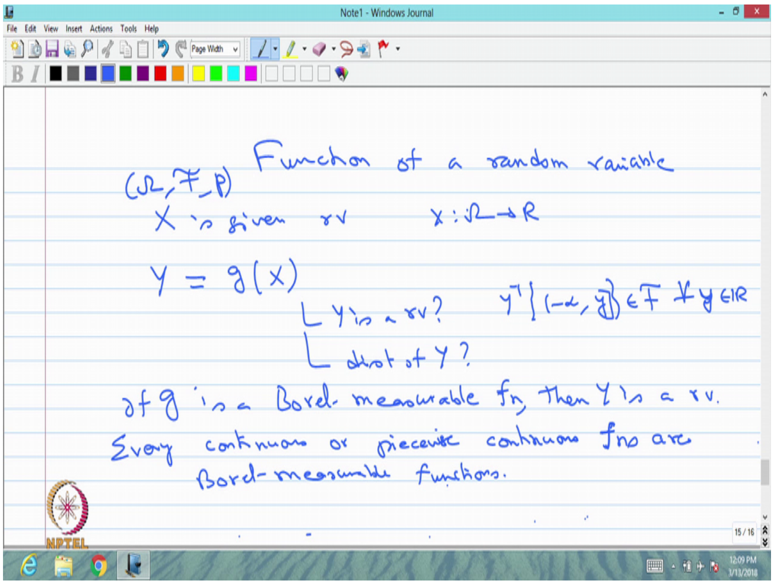
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture - 11

In week 2, we are discussing the random variable and we started with the definition of the random variable. Then, we discussed the distribution function, then we have related the distribution function with the random variable with the probability in the form of, we get the CDF; CDF of the random variable.

And by seeing the CDF, we classify the random variable or we get the types of random variable that is a discrete type or continuous type or mixed type random variable.

(Refer Slide Time: 00:36)



So, now we are moving into the function of random variable. Function of a random variable. Sometimes, we started creating 1 random variable and then we come to the conclusion oh we need another random variable, what to do?

So, we can always create a new random variable from the scratch that is our real valued function from omega to R satisfying the condition, then we can conclude that is a random variable

But the easy thing is suppose you are able to relate the already found one random variable with the new random variable in some form, then you can find the new random variable and it’s a distribution through the existing already the random variable.

So, the existing random variable, we use the letter capital X is given and we are interested to create a new random variable that is Y. Instead of creating a new real valued function and find out we make a relation in the form of earlier random variable as a g of X; where g is the function from r to r.

I will repeat X is the given random variable that is defined from omega to R. Already we know it is a random variable and we know the CDF of the random variable. From the CDF, we know it is a discrete type or continuous type or mixed type random variable. We are interested to find the distribution of another random variable that is Y; for that we identify the relation that is g of X.

Now, the first question is whether the way we make a relation g of X, whether that is going to be a random variable? That means, whether Y inverse of minus infinity to small Y that is belonging to the same probability space? So that means, we have a given probability space omega F P and in this probability space X is a given random variable.

And now, the first question is whether Y is a random variable. That means, whether it satisfies Y inverse of minus infinity to small Y sorry, closed interval that is belonging to F for all Y belonging to (Refer Time: 03:21). If it is a random variable, then the second question is what is the distribution of the random variable Y?

I repeat the issue; we have a probability space, we have a one random variable and we are interested to find the distribution of the other random variable for that we make the relation g of X. Therefore, first question is whether Y is a random variable and then, if Y is a random variable what is the distribution of Y?

First, we can answer the first question whether Y is a random variable. Whenever X is a random variable and g is Borel-measurable function; then if g is a Borel-measurable function, then the Y is a random variable. Now, what is the Borel-measurable function in the measure theory course, one can study if you have a real valued function; the inverse image of a Borel set is belonging to the Borel set.

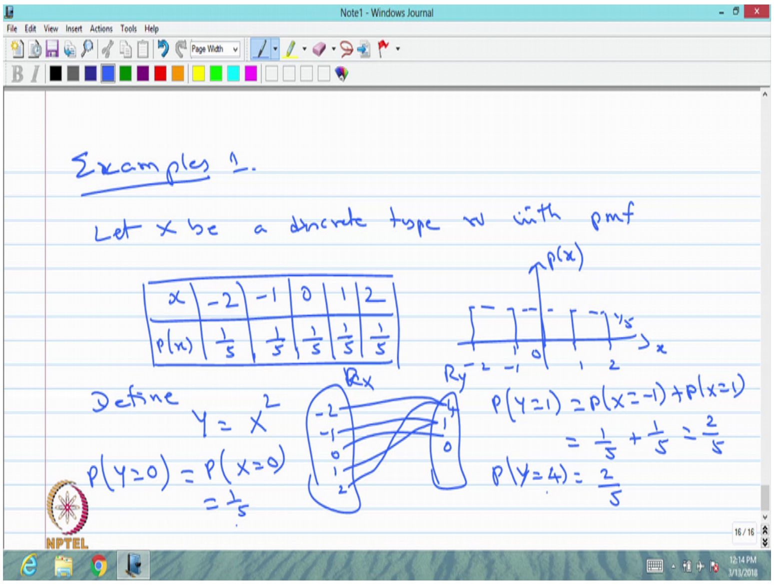
Then, we can conclude the given real valued function is a Borel measurable function. As far as this course is concerned, we do not need to worry about the Borel measurable function and so on. We can take it every continuous or piecewise continuous functions or Borel measurable function. So, we can make sure whether the g is going to be a continuous or Borel continuous or piecewise continuous function; therefore, it is going to be a Borel measurable function. Therefore, we it is a random variable.

So, we can use this concept every continuous or piecewise continuous; one should know the definition of piecewise continuous. So, every continuous or piecewise continuous functions are Borel measurable functions. Therefore, the first question is answer. So, as far as this course is concerned, we always give g such a way that Y is going to be a random variable.

Now, the question is how to find the distribution of Y. Distribution means in general, it is CDF of the random variable. If you know that it is a discrete type random variable, if you find the probability mass function of that, that is also called distribution function. If you know that it is a continuous type random variable, then if you find the probability density function of the random variable; then that is also called the distribution function.

So in general, distribution means cumulating distribution function of the random variable. Otherwise, it could be a probability mass function or probability density function based on the random variable is discrete type random variable or continuous type random variable. Now, we will start find out the distribution of function of random variable with the few cases.

(Refer Slide Time: 07:05)



So, let us start with the examples through that we will study the distribution of functions of function of random variable. The first example, let X be a discrete type random variable with the probability mass function given for different values of x, what is a p of x? So, this way also one can give.

Suppose, x takes the value minus 2, minus 1, 0, 1 and 2; what is the probability mass at those points? Then, we are defining the probability mass function; that means, other than this points the values are 0 and if you add this is going to be 1. Therefore, this is probability mass function.

So, let us give some values such a way that it is going to be probability mass function. So, we are going for the example to discuss function of random variable. Therefore, I am just going for the easy example in which the probability mass function at the point. So, this is a probability mass function at the point minus 2, minus 1, 0, 1 and 2; all the values are 1 by 5 fine.

Now, I am going to define new random variable or new real valued function which is Y is equal to X square is a easiest function. So, the possible x values are minus 2, minus 1, 0, 1 and 2. The way we defined Y is equal to X square, now the random variable Y is going to be a discrete type; because for different values of X, the Y values are going to be either 0 or 1 or 4 because Y is equal to X square. Therefore, the way I have made the relation Y is equal to X square and x is a discrete type random variable.

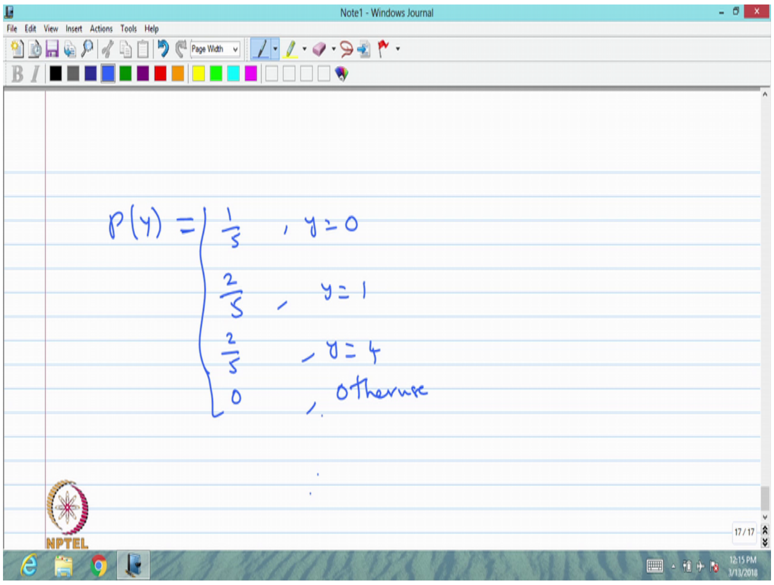
I can very well say the random variable Y is of the discrete type. Therefore, I can go for either CDF of the random variable Y or I can go for the probability mass function of Y. We can make a nice the possible values of x that is minus 2, minus 1, 0, 1 and 2 and this is mapped with the possible values of Y that is 0 is mapped with 0 and minus 1 and plus 1 is mapped with 1 and minus 2 and plus 2 is mapped with 4. So, this is the mapping from x to y..

That means, if I want to find out the probability of Y takes a value 0 that is same as I have to go for what is the possible outcome or what is the possible values of x, which gives the value Y is equal to 0. So, X equal to 0 will give the value Y is equal to 0. Therefore, probability of Y is equal to 0 is same as probability of X equal to 0 and I know that probability of X equal to 0 is 1 by 5; therefore, this is 1 by 5.

Similarly, I can find probability of Y is equal to 1 that is same as either probability of X takes a value minus 1 or probability of X equal value 1; both will give the value Y is equal to 1. Therefore, probability of Y is equal to 1 is nothing but it is a 1 by 5 plus 1 by 5; therefore, it is 2 by 5.

Similarly, you can find probability of Y takes a value 4 that is x takes a value is a minus 2 as well as x takes the value plus 2; those probabilities we will put together, give the probability of Y is equal to 4. Therefore, that is again 2 by 5.

(Refer Slide Time: 11:55)

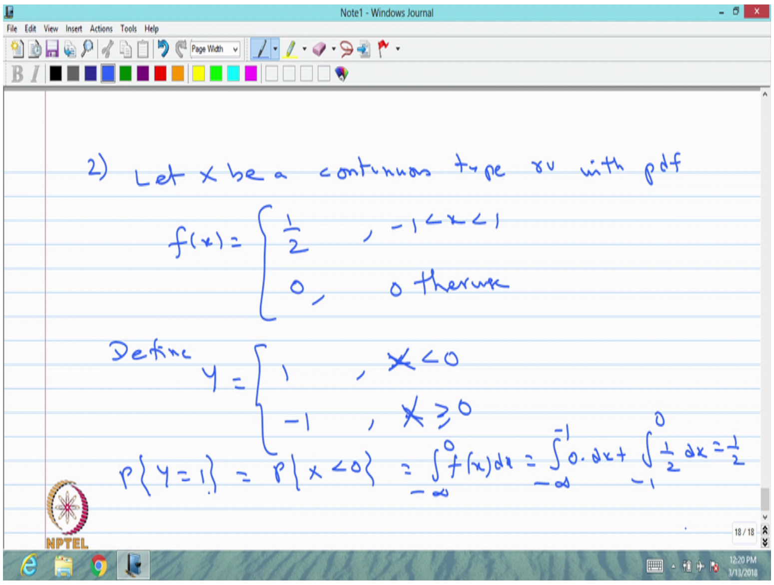


Therefore, I can make probability mass function for Y that is takes a value 1 by 5, when y is equal to 0 and 2 by 5, when y takes a value 1 and 2 by 5, when y takes a value 4 and 0 otherwise. That means, from discrete type random variable and Y is equal to X square gives the again discrete type random variable.

Because it has the mass at the 0, 1 and 4; if you add all the probability masses, it is going to be 1. So, this is the probability mass function and if draw the CDF, then it is 0; till 0, at 0 it is 1 by 5 height. Then, it is 1 by 5 till 1, at the point 1 it becomes 3 by 5. Then, 3 by 5 till 4 and at the point 4 onwards, it becomes 1.

Therefore, this CDF has 3 jumps. So, this discrete type random variable whose probability mass function is probability of Y takes the value 1 by 5, 2 by 5, 2 by 5 at those points otherwise it is 0. So, this is the easiest example in which we are getting discrete random variable into discrete random variable.

(Refer Slide Time: 13:21)



Then second example is let x be a continuous type random variable with since it is a continuous type random variable, I am going to give the probability density function of the random variable x that is takes a value 1 by 2 between the interval minus 1 to 1; otherwise it is 0. Whenever the problem is given, you can verify whether it is a correct. Probability density function, it is always greater than or equal to 0; that is a first property.

Second property, if you integrating the whole interval minus infinity to infinity; it has to be 1. So, if we check minus infinity to minus 1, the probability density function is 0. From minus 1 to 1, it is 1 by 2. So, if you integrate minus 1 to 1, 1 by 2 you will get 1 and integration from 1 to infinity, again the probability density function is 0; therefore, it is 0. So, the whole interval minus infinity to infinity, the probability density function that integration is 1; therefore, it is a probability density function.

Now, I am going to define new random variable. Why I am saying the random variable? That means, we are saying it is a Borel measurable function; therefore, it is a random variable. So, we do not need a question about the whether it is a random variable or not we started with the random variable.

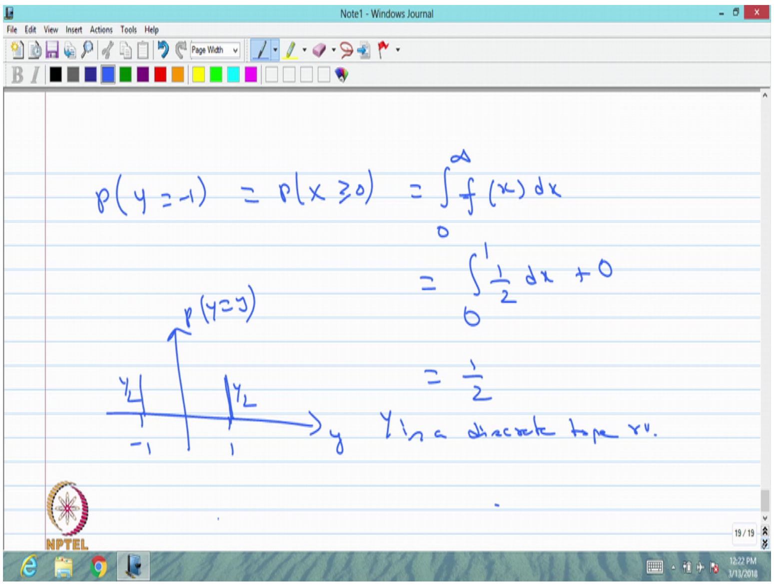
So, define a random variable Y which takes the value 1 when x is lesser than 0. It takes a value minus 1, when x takes a value greater than 0. I am defining a new random variable which takes the value 1, when x is less than 0; minus 1 when x is greater than 0. I can write capital letter also. So, capital X is a random variable. So, when the random variable takes a value less than 0, it is a 1; when it is greater than 0, it is a minus 1. I can use greater than 0 also.

So, now you can find probability of Y takes a value 1, that is same as the probability of x is less than 0, that is same as since x is a continuous type random variable x is less than 0 or less than or equal to 0; both are one and the same and it is same as minus infinity to 0 and the probability density function..

And the probability density function is greater than 0 between the interval minus 1 to 0; therefore, it is minus infinity to minus 1 it is 0 plus minus 1 to 0 f of x d of x here it is 1 by 2 takes. Therefore, you can integrate and you can get the value and this value is going to be 1 by 2.

The way we define Y takes a value 1 for x is less than 0. So, probability of Y is equal to 1 that is same as probability of x is less than 0.

(Refer Slide Time: 17:01)

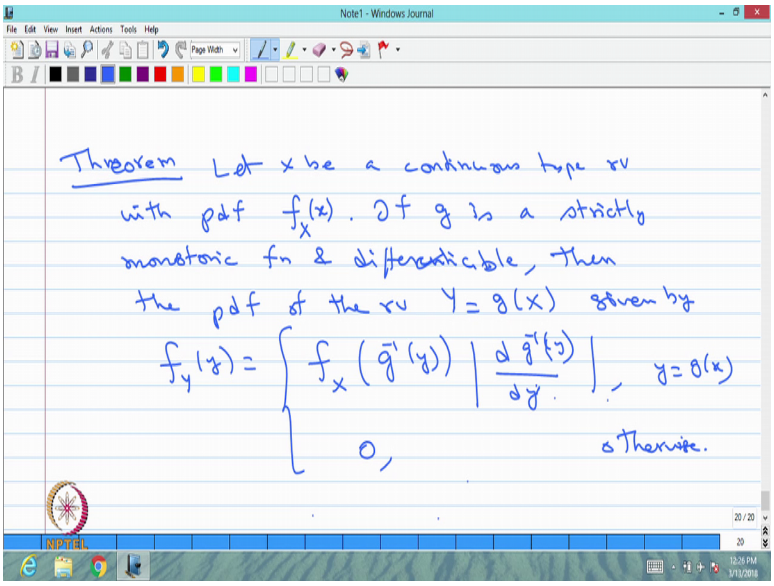


Similarly, the probability of Y takes a value minus 1 that is same as probability of x is greater than or equal to 0; that is same as integration from 0 to infinity f of x dx that is same as the again the f of x is greater than 0 between the interval 0 to 1; whereas, 1 to infinity in the density function is a 0. Therefore, it is a 0 to 1 f of x is 1 by 2 the other quantity is 0. So, if you simplify you will get the value 1 by 2.

Therefore, it has the mass the probability mass at the point minus 1, 1 by 2 and at the point 1, it has another 1 by 2 and if you add both the masses it becomes 1. Therefore, this is a discrete type random variable. X is a discrete type random variable. The first example X is a discrete type, Y is also discrete.

Now the X is a continuous type random variable. The way we defined the function Y, the Y is a discrete type random variable whose probability mass function is 1 by 2 at the point 1 and minus 1 otherwise it is 0.

(Refer Slide Time: 19:07)



Now, we are going for one general result as a theorem and this theorem will be useful whenever you want to find out the probability density function of a continuous type random variable, whenever x is also continuous type. So, I am going to give the theorem first; then we will give the proof of the theorem followed by 1 example, then I will conclude.

Let x be a continuous type random variable with probability density function is small f of x. If g is a strictly monotone function and differentiable, then the probability density function of the random variable capital Y as a g of capital X is given by see the theorem, you can directly write the probability density function of the random variable y in terms of the probability density function of x.

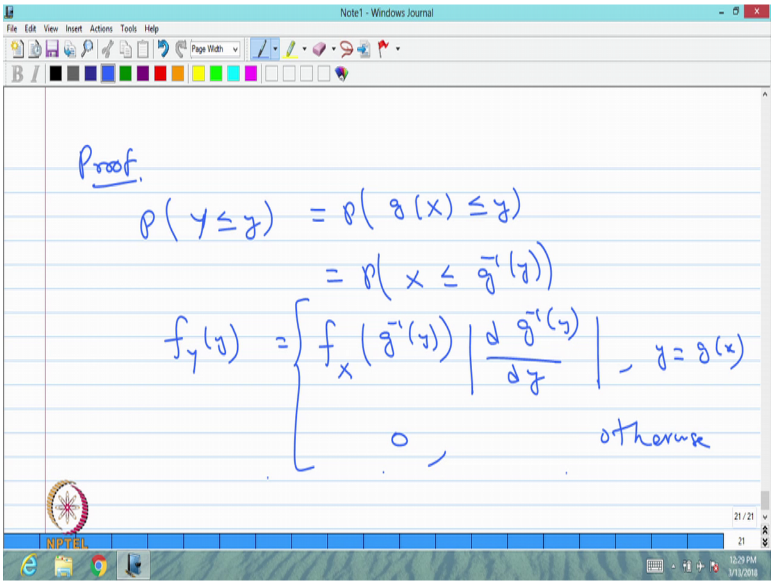
I can rewrite this as the suffix x. I am using capital small f for all the probability density function by writing suffix x or suffix y, we know that we conclude it is a probability density function of x; it is a probability density function of y.

So, the probability density function of y, we can write it in the form of probability density function of x by replacing x by g inverse of y; not only that by multiplying the absolute of derivative of g inverse of y with respect to y.

So, this is going to be greater than 0 whenever y takes a value g of x; whenever y is not going to take the value g of x, it is going to be 0. That means, by just substituting x by g inverse of y in the probability density function of x, you will not get the probability density function of y unless otherwise you multiply the absolute of derivative of g inverse of y with respect to y.

That means, you recall the probability density function has a 2 properties; it is going to be greater than or equal to 0 and the integration is going to be 1. By multiplying this absolute quantity, the integration is going to be 1. Therefore, this multiplication absolute of derivative of g inverse of y with respect to y, that is called normalizing constant.

(Refer Slide Time: 23:01)



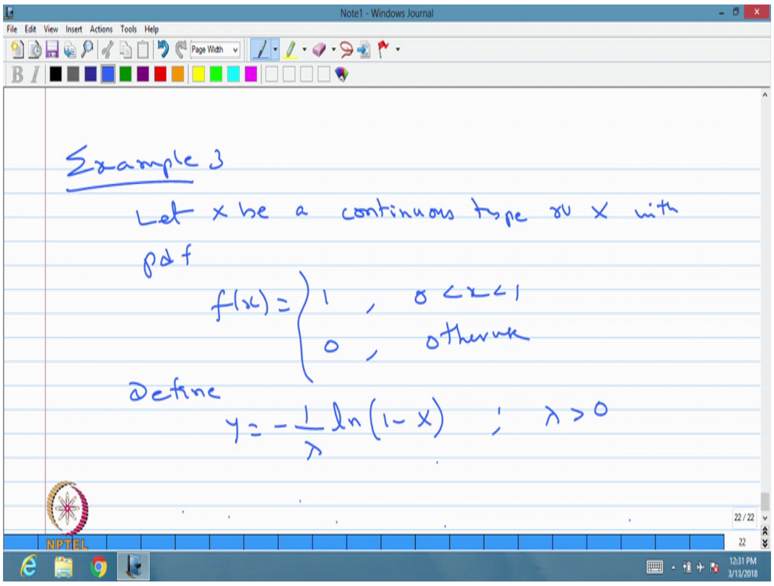
You can prove, you can prove. I will give the proof for the one part since I am saying strictly monotonic it could be increasing or decreasing. So, we will do the one part, then similarly one can do the other part. So, you can find probable CDF of the random variable y correct that is nothing but the y is replaced by g of capital x less than or equal to small y; whenever I write capital letter; that means, it is a random variable whenever I write the small letter; that means, it is a variable values.

So, I am directly finding the CDF of the random variable y. I am replacing y by g of x; that is same as the probability of x is less than or equal to g inverse of y. This is valid only g is a strictly monotonic function. In this case it is a strictly increasing function. Once I know the CDF, I can find the probability density function by differentiating both side with respect to y. By differentiating CDF with respect to y, I can get the probability density function of y; that is a small f suffix capital y. That is nothing but once you do the differentiation in the right hand side with respect to y, you can use the chain rule.

Therefore, it is a the probability density function evaluated at the point g inverse of y; then differentiate the g inverse of y with respect to y. So, this is valid when g is strictly increasing function. Suppose, g is decreasing function, then also you can do the similar calculation. Then you will get the values with the negative sign, then you can take the negative first.

So, you can go for absolute of this in general whether it is a strictly increasing or strictly decreasing, you can take absolute of this derivative terms and substitution the probability density function of x with the g inverse of y will give the probability density function of y whenever y is equal to g of x, otherwise it is 0. You can go for example for how to apply this theorem.

(Refer Slide Time: 25:51)



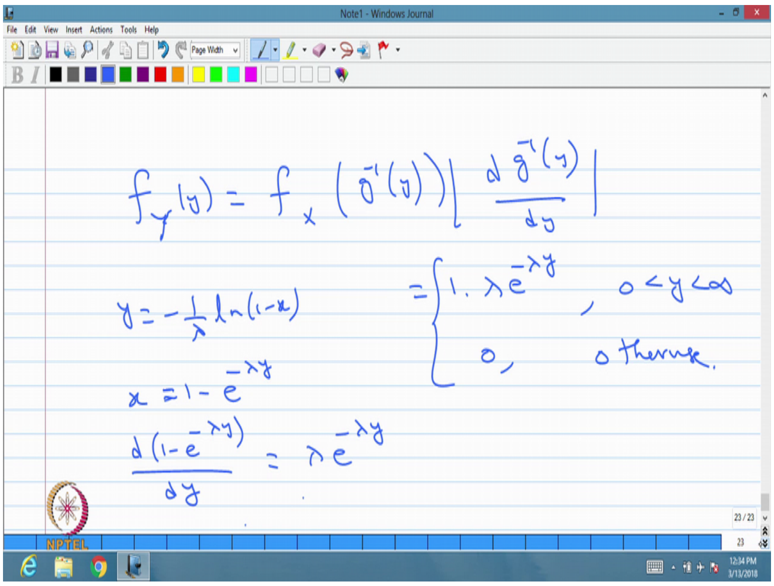
So, easiest example, example number 3 because we have already discussed 2 examples.

The third example, let x be a let x be a continuous type random variable with the probability density function f of x takes a value 1 between 0 to 1 otherwise it is 0; that means, it is a constant probability density between the interval 0 to 1, otherwise its 0. If you integrate you will get the value 1, then greater or equal to 0; therefore, this a probability density function.

Now, we are defining a new random variable Y is equal to minus 1 by lambda l n of 1 minus x. I am defining new random variable Y is equal to minus 1 by lambda l n of 1 minus x. You can prove it is a continuous function; therefore, it is a Borel measurable function; therefore, it is a random variable.

Here, the lambda has to be strictly greater than 0. So, x is a continuous type random variable with the probability density function 1 between the interval 0 to 1; 0 otherwise and Y you define it as a minus 1 divided by lambda l n of 1 minus x. Now you can verify whether this theorem can be applied. x is a discrete x is a continuous type random variable, y is this function; it is a strictly increasing function therefore, and differentiable also. So, therefore, we can apply the theorem and you can directly get the probability density function.

(Refer Slide Time: 27:57)

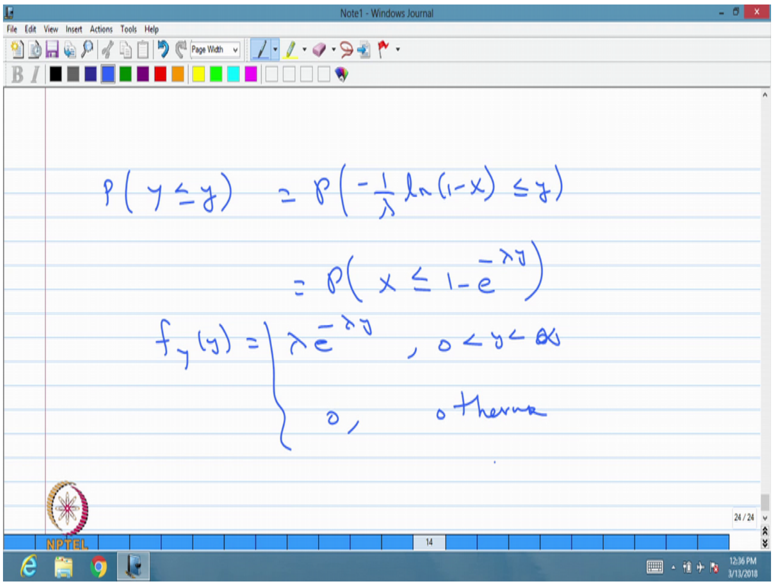


So, the probability density function of y that is a probability density function of x by replacing x by g inverse of y. Then derivative of g inverse of y with respect to y in (Refer Time: 28:15). So, we have small y is equal to 1 by lambda l n of 1 minus x. So, you can find x, you can find x. So, the x is going to be x is going to be 1 minus e power minus lambda times y; that is a g inverse of y. So, you can find the derivative of 1 minus e power minus lambda y with respect to y, you will get that is lambda times e power minus lambda y.

So, you need g inverse of y as well as you need derivative. Therefore, this is going to be the probability density function substituted x by g inverse of y, but since the probability density function is a constant between the interval 0 to 1. Therefore, it is going to be again 1 and the derivative with absolute that is e power lambda times e power minus lambda y and if you do the little home work when x lies between 0 to 1, you will get y lies between 0 to infinity. Therefore, the probability density function is between 0 to infinity; one times so one can be avoided. So, 0 otherwise.

So, whenever x takes the value 0 to 1, y takes the value 0 to infinity in that the probability density function is lambda times e power minus lambda y, otherwise it is 0. So, since this theorem is satisfied, we are able to get the probability density function directly. You can find the CDF of the random variable y directly without using the theorem also. By using the theorem we got the probability density function.

(Refer Slide Time: 30:29)

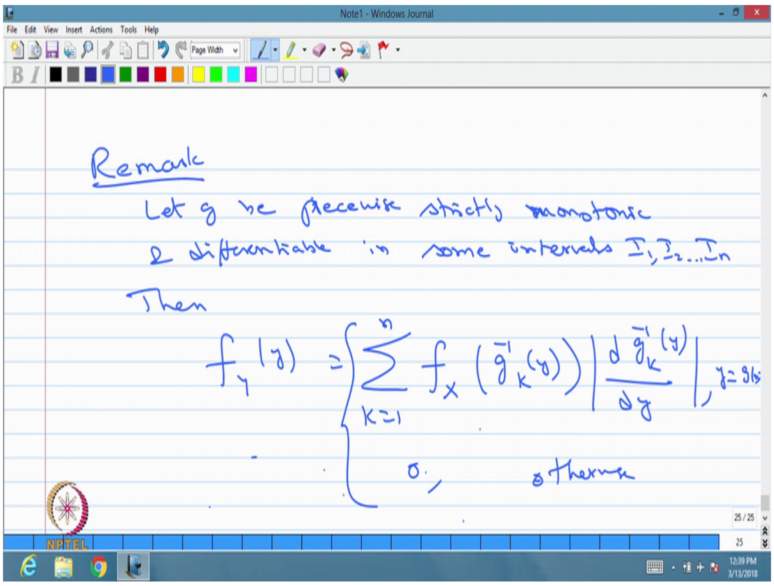


But the another method you can directly compute what is the CDF of the random variable y; that is same as probability of minus 1 by lambda l n of 1 minus x less than or equal to small y. Then you can do the little simplification, you will get probability of x is less than or equal to 1 minus e power minus lambda y. If you do the little simplification, you will get a probability of x is less than or equal to 1 minus e power minus lambda y.

Therefore, I can go for probability density function by differentiating both side that is going to be lambda times e power minus lambda y when Y is lies between 0 to infinity. I am skipping in between some steps that you can work out separately. So, there are 2 ways; either you can apply the theorem or you can find the CDF first by differentiating you can get the density function.

Sometimes, it may be a not strictly increasing or strictly decreasing; in some interval the function may be monotonically increasing, some interval it may be monotonically decreasing, but still you can use the theorem.

(Refer Slide Time: 31:54)



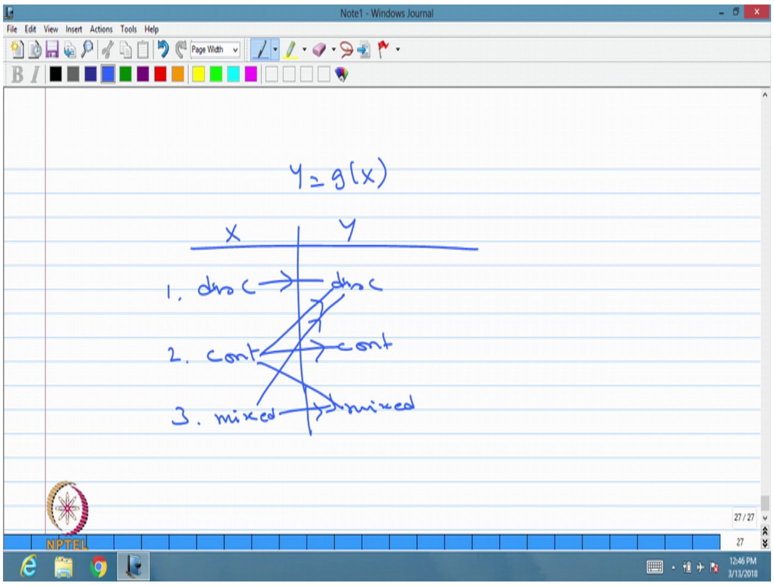
So, I am coming back to the remark of the theorem. Whenever you have a issue in which the function is the function g be a piecewise strictly monotone and differentiable in some intervals. Earlier in the whole interval, it is strictly increasing or strictly decreasing. Now it is piecewise strictly monotonic and differentiable in some intervals. We label this intervals as the I 1, I 2 and so on suppose I n intervals.

Then, we can find the probability density function of the random variable y. Then, we can find the probability density function of y as sum of n intervals in which you get what is the sum of f x at different k g inverse; then find out the absolute of derivative of g inverse of y, the kth one with respect to y. So, this is going to be the probability density function which is greater than 0 whenever you get the points which is y is equal to g of x, otherwise it is 0.

So that means, there is a possibility you may have a function in which it is increasing or decreasing or decreasing or increasing or it may be in many intervals. Then, you can add all the intervals corresponding density function, if you sum it up; then that is going to be the probability density function of y.

So, we have seen many examples in which discrete to discrete, continues to continuous or continuous to discrete and so on. So, the way we create the random variable y accordingly you will end up with the different types of the random variable for y for a given random variable x.

(Refer Slide Time: 35:05)



So, in conclusion whenever X is given and you are interested to find out the distribution of Y as a function of x in the g, in conclusion we can go for the table form when X is of discrete type. I am just writing disc discrete type or X is continuous type random variable or X is of the mixed type random variable; the way we defined the Y is equal to g of x, the random variable Y that is also of the form discrete or continuous or mixed. You can think of many functions in your mind, then you can make a line whether you will get a discrete to discrete or discrete to continuous or mixed and so on. So, in 3 in to 3, 9 out of 9 which possibility is possible that you can make out.

So, discrete to discrete is possible, I have given the example also. Y is equal to X square example where X is a discrete; therefore, Y is also discrete. From discrete you can go for from the discrete type of random variable, you cannot go for the continuous you cannot go for the mixed type random variable. Whereas, if you have a continuous type random variable by default you can always make a function Y is equal to X or something.

Therefore, from continuous to continuity, it is easy. From continuous to discrete is possible, I have given 1 example. X is a continuous type random variable, Y is equal to 1 and minus 1. So, that sorry Y is yes Y is equal to 1 and minus 1 that example is the continuous to discrete type example.

Since, continuous to continuous and continuous to discrete is possible; therefore, continuous to mixed is also possible because the continuous means the unit mass is distributed over the interval, over the real line. So, you can make a many to one function; therefore, some density can be a mass at some points and whereas, the few density you can keep density as it is. Therefore, continuous to mixed is possible.

I will repeat, from continuous type random variable mixed type random variable is possible because the unit mass can be transform into density in some interval and masses at some points. Therefore, it is a mixed type random variable because continuous to continuous is possible, continuous to discrete possible; therefore, continuous to mixed is also possible

Now, we come to the third example third type when X is a discrete type; when X is a mixed type random variable mixed means it has the density as well as mass. So, by making a many to one function, all the density you can put it mass at some points. Therefore, mixed to discrete is possible. Very important observation mixed to continuity is impossible, mixed to continuous type random variable is impossible because mixed has mass at some points and density between some interval. So, mixed to continuous is not possible. Whereas, by default by 1 to 1 mapping, you can always have mixed to mix.

Therefore, whatever you do the different problems in the distribution of function of random variable, at the end you will end up with, these are all the only possible ways you will get the different type of random variable for Y and once you know the distribution of x, you can find the distribution of Y in the form of a CDF or if you know that it is a discrete type random variable in the form of probability mass function, if you know Y is continuous type random variable in the form of probability density function.

With this, we are completing the random variable with the 3 topics; one is definition and the CDF and the second one is a types of random variable and the third topic is distribution of function of random variable.