

In this important segment, we will develop a method for finding the PDF of a general function of a continuous random variable, a function  $g$  of  $X$ , which, in general, could be nonlinear. The method is very general and involves two steps.

The first step is to find the CDF of  $Y$ . And then the second step is to take the derivative of the CDF and then find the PDF. Most of the work lies here in finding the CDF of  $Y$ . And how do we do that?

Well, since  $Y$  is a function of the random variable  $X$ , we replace  $Y$  by  $g$  of  $X$ . And now we're dealing with a probability problem that involves a random variable,  $X$ , with a known PDF. And we somehow calculate this probability. So let us illustrate this procedure through some examples.

In our first example, we let  $X$  be a random variable which is uniform on the range from 0 to 2. And so the height of the PDF is  $1/2$ . And we wish to find the PDF of the random variable  $Y$  which is defined as  $X$  cubed. So since  $X$  goes all the way up to 2,  $Y$  goes all the way up to 8.

The first step is to find the CDF of  $Y$ . And since  $Y$  is a specific function of  $X$ , we replace that functional form. And we write it this way. So we want to calculate the probability that  $x$  cubed is less than or equal to a certain number  $y$ . Let us take cubic roots of both sides of this inequality.

This is the same as the probability that  $X$  is less than or equal to  $y$  to the  $1/3$ . Now, we only care about values of  $y$  that are between 0 and 8. So this calculation is going to be for those values of  $y$ . For other values of  $y$ , we know that the PDF is equal to 0. And there's no work that needs to be done there. OK.

Now,  $y$  is less than or equal to 8, so the cubic root of  $y$  is less than or equal to 2. So  $y$  to the  $1/3$  is going to be a number somewhere in this range. Let's say this number.

We want the probability that  $X$  is less than or equal to that value. So that probability is equal to this area under the PDF of  $X$ . And since it is uniform, this area is easy to find. It's the height, which is  $1/2$  times the base, which is  $y$  to the  $1/3$ .

So we continue this calculation, and we get  $1/2$  times  $y$  to the  $1/3$ . So this is the formula for the CDF of  $Y$  for values of  $y$  between 0 and 8. This completes step one.

The second step is simple calculus. We just need to take the derivative of the CDF. And the derivative is

$\frac{1}{2}$  times  $\frac{1}{3}$ , this exponent,  $y$  to the power of minus  $\frac{2}{3}$ . Or in a cleaner form,  $\frac{1}{6}$  times  $\frac{1}{y}$  to the power  $\frac{2}{3}$ .

So the form of this PDF is not a constant anymore.  $Y$  is not a uniform random variable. The PDF becomes larger and larger as  $y$  approaches 0. And in fact, in this example, it even blows up when  $y$  becomes closer and closer to 0. So this is the shape of the PDF of  $Y$ .

Our second example is as follows. You go to the gym, you jump on the treadmill, and you set the speed on the treadmill to some random value which we call  $X$ . And that random value is somewhere between 5 and 10 kilometers per hour. And the way that you set it is chosen at random and uniformly over this interval. So  $X$  is uniformly distributed on the interval between 5 and 10.

You want to run a total of 10 kilometers. How long is it going to take you? Let the time it takes you be denoted by  $Y$ . And the time it's going to take you is the distance you want to travel, which is 10 divided by the speed with which you will be going. So the random variable  $y$  is defined in terms of  $x$  through this particular expression. We want to find the PDF of  $y$ .

First let us look at the range of the random variable  $Y$ . Since  $x$  takes values between 5 and 10,  $Y$  takes values between 1 and 2. Therefore, the PDF of  $Y$  is going to be 0 outside that range. And let us now focus on values of  $Y$  that belong to this interesting range. So  $1 < y \leq 2$ .

And now we start with our two-step program. We want to find the CDF of  $Y$ , namely, the probability that capital  $Y$  takes a value less than or equal to a certain little  $y$  in this range. We recall the definition of capital  $Y$ . So now we're dealing with a probability problem that involves the random variable capital  $X$ , whose distribution is given to us.

Now, we rewrite this event as follows. We move  $X$  to the other side. This is the probability that  $X$  is larger than or equal after we move the little  $y$  also to the left-hand side.  $X$  being larger than or equal to  $\frac{10}{y}$ .

Now,  $y$  is between 1 and 2.  $\frac{10}{y}$  is going to be a number between 5 and 10. So  $\frac{10}{y}$  is going to be somewhere in this range. We're interested in the probability that  $X$  is larger than or equal to that number.

And this probability is going to be the area of this rectangle here. And the area of that rectangle is equal

to the height of the rectangle-- now, the height of this rectangle is going to be  $1/5$ . This is the choice that makes the total area under this curve be equal to 1-- times the base. And the length of the base is this number 10 minus that number. It's  $10 - 10/y$ .

So this is the form of the CDF of  $Y$  for  $y$ 's in this range. To find the PDF of  $Y$ , we just take the derivative. And we get  $1/5$  times the derivative of this term, which is minus 10, divided by  $y$  squared. But when we take the derivative of  $1/y$ , that gives us another minus sign. The two minus signs cancel, and we obtain  $2$  over  $y$  squared.

And if you wish to plot this, it starts at 2. And then as  $y$  increases, the PDF actually decreases. And this is the form of the PDF of the random variable  $y$ . This is the form which is true when  $y$  lies between 1 and 2. And of course, the PDF is going to be 0 for other choices of little  $y$ .

So what we have seen here is a pretty systematic approach towards finding the PDF of the random variable  $Y$ . Again, the first step is to look at the CDF, write the CDF in terms of the random variable  $X$ , whose distribution is known, and then solve a probability problem that involves this particular random variable. And then in the last step, we just need to differentiate the CDF in order to obtain the PDF.