

We have already worked through some examples in which X was a random variable with a given PDF, and we considered the problem of finding the PDF of Y for the case where Y was the function x cubed or the function of the form a/X . What both of these examples have in common is that Y is a monotonic function of X .

In this case, Y is increasing with X . In this case, Y was decreasing with X . It turns out that there is a general formula that gives us the PDF of Y in terms of the PDF of X in the special case where we're dealing with a monotonic function. So, let us assume that g is a strictly increasing function. And what that means is that, if x is a number or smaller than some other number x' , the value of g of x is going to be smaller than the value of g of x' . So, when you increase the argument of the function, the function increases.

To keep things simple, we will also assume that the function g is smooth, in particular that it is differentiable. Then we have a diagram such as this one. Here is x , and y is given by a function of x . It's a smooth function, and that function keeps increasing. Now, because of the assumptions we have made on g , we have an interesting situation.

Given a value of x , a corresponding value of y will be determined according to the function g . But we can also go the other way. If I tell you a value of y , then you can specify for me one and only one value of x that gives rise to this particular y . So, the function g takes us from x 's to y 's, but you can also go back the opposite way from y 's to values of x .

And the mapping that takes us from y 's to x 's, this is the inverse of the function g . And we give a name to that inverse function, and we call it h . So, h of y is the value of x that produces a specific value y . Let us now move on with the program of finding the PDF of Y . We will follow the usual two step procedure. And the first step is to find the CDF of Y .

So we fix some little y , And we want to find the probability that the random variable y takes a value in this range. When does this happen? For Y to take a value in this range, it must be the case that X takes a value in this range here. Values of X smaller than this particular number result in values of Y that are less than or equal to this particular number.

So, we can rewrite the event of interest in terms of the random variable X and write it as follows. We need to have x less than or equal to h of y . But this is just the CDF of X evaluated at h of y . We now carry out to the second step of our program. We take derivatives of both sides and we find that the PDF of Y is equal to the derivative of the right hand side, the derivative of the CDF is a PDF.

And then the chain rule tells us that we also need to take the derivative of the term inside here with respect to its argument. And this is a general formula for the PDF of a strictly increasing function of a random variable X . How about the case of a decreasing function? So, let us assume that g now is a strictly decreasing function of X .

So, we might have a plot for g that looks something like this. What happens in this case? We can start doing a calculation of this kind. But now, how can we rewrite this event? The random variable Y will take a value less than or equal to this number y . When does this happen? When the value of g of x is less than y .

And that happens for x 's in this range. So, this is the set of x 's for which is the value of g of x is less than or equal to this particular number y . So the event of interest in that case is the event that X is larger than or equal to h of y , which is 1 minus the probability that X is less than h of y . Because X is a continuous random variable, we can change this inequality to one that allows the possibility of equality.

And so this is 1 minus the CDF of X evaluated at h of y . Now we take the derivatives of both sides and we find the PDF of Y being equal to, there's a minus sign here, then the derivative of the CDF, which is the PDF. And finally, the derivative of the function h . Now in this case, g is a decreasing function of x . So when x goes down, y goes up.

When x goes up, y goes down. This means that when y goes up, x goes down. So it means that the inverse function h is going to be also monotonically decreasing. Since it is decreasing, it means that the slope, the derivative of the function h is going to be either 0 or negative. And so minus a negative value gives us the absolute value of that number.

So we can rewrite this by removing this minus sign here, and putting an absolute value in this place. Of course, in the case where g is an increasing function, when x goes up, y goes up. This means that when y goes up, x goes up. So h in that case would have been an increasing function, so this number here would have been a non-negative number, and so it would be the same as the absolute value.

So using these absolute values, we obtain formulas that are exactly the same in both cases of increasing and decreasing functions, and so our final conclusion is that in either case, the PDF of Y is given in terms of the PDF of X times the derivative of this inverse function.

Let us now apply the formula that we have in our hands for the monotonic case to a particular example, where y is the square of X , and where X is uniform on the interval 0 to 1. So the function g , in our case, the function g is the square function. Now, you could argue here that this function is not monotonic, so how can we apply our results?

On the other hand, the random variable X takes values on the interval from 0 to 1, and therefore the form of the function g outside that range does not concern us. Over the range of values of interest, the function g is a monotonic function. So, what is the correspondence? y is going to be equal to x squared.

That's the g of x function. And when that happens, we have the relation that x is going to be the square root of y . This tells us that the inverse function, h of y , which tells us what is the particular x associated with a given y , the inverse function takes the form square root of y . So now we can go ahead and use the formula.

The density at some particular little y where that little y , belongs to the range of values of interest, x things values between 0 and 1, so y also takes values between 0 and 1. So over that range, the density of Y is the density of X , which is uniform, therefore it is equal to 1, times the derivative of the square root function. And the derivative of the square root function is $1/2$ times the square root of y .

As you can see, the amount of calculations involved here are rather simpler compared to what we would have to do if we were to go through our two step program and work with CDFs. All that you need to do is essentially to identify the inverse function that given a y produces x 's, and write down the corresponding derivative.