

The formula that we just derived for the monotonic case has a nice intuitive explanation that we will develop now.

Suppose that g is a monotonic function of x and that it's monotonically increasing.

Let us fix a particular x and a corresponding y so that the two of them are related as follows-- y is equal to g of x , or we could argue in terms of the inverse function so that x is equal to h of y .

Recall that h is the inverse function, that given a value of y , tells us which one is the corresponding value of x .

Now let us consider a small interval in the vicinity of this x .

Whenever x falls somewhere in this range, then y is going to fall inside another small interval.

The event that x belongs here is the same as the event that y belongs there.

So these two events have the same probability.

And we can, therefore, write that the probability that Y falls in this interval is the same as the probability that X falls in the corresponding little interval on the x -axis.

This interval has a certain length δ_1 .

This interval has a certain length δ_2 .

Now remember our interpretation of probabilities of small intervals in terms of PDFs so this probability here is approximately equal to the PDF of Y evaluated at the point y times the length of the corresponding interval.

Similarly, on the other side, the probability that X falls on the interval is the PDF of X times the length of that interval.

So this gives us already a relation between the PDF of Y and the PDF of X , but it involves those two numbers δ_1 and δ_2 .

How are these two numbers related?

If x moves up by the amount of δ_1 , how much is y going to move up?

It's going to move up by an amount which is δ_1 times the slope of the function g at that particular point.

So that gives us one relation that δ_2 is approximately equal to δ_1 times the derivative of the function of g at that particular x .

However, it's more useful to work the other way, thinking in terms of the inverse function.

The inverse function maps y to x , and it maps y plus $\Delta 2$ to x plus $\Delta 1$.

When y advances by $\Delta 2$, x is going to advance by an amount which is how much y advanced times the slope, or the derivative, of the function that maps y 's into x 's.

And this function is the inverse function.

So this is the relation that we're going to use.

And so we replace $\Delta 1$ by this expression that we have here in terms of $\Delta 2$.

And now we cancel the $\Delta 2$ from both sides of this equality, and we obtain the final formula that the PDF of Y evaluated at a certain point is equal to the PDF of x evaluated at the corresponding point, or we could write this as the PDF of X evaluated at the value x that's associated to that y that's given by the inverse function, times the derivative of the function h , the inverse function.

And this is just the same formula as the one that we had derived earlier using CDFs.

This derivation is quite intuitive.

It associates probabilities of small intervals on the x -axis to probabilities of corresponding small intervals on the y -axis.

These two probabilities have to be equal, and this implies a certain relation between the two PDFs.