

We will now go through an example that brings together all of the concepts that we have introduced. We have a stick of length  $l$ . And we break that stick at some random location, which corresponds to a random variable,  $X$ .

And we assume that this random variable is uniform over the length of the stick. So its PDF has this particular shape. And for the PDF to integrate to 1, the height of this PDF must be equal to  $1/l$ .

Then we take the piece of the stick that we are left with, which has length  $X$ , and we break it at a random location, which we call  $Y$ . And we assume that this location  $Y$  is uniformly distributed over the length of the stick that we were left with. What does this assumption mean?

It means that if the first break was at some particular value,  $x$ , then the random variable  $Y$  has a conditional distribution, which is uniform over the interval from 0 to  $x$ . So the conditional PDF is uniform. A conditional PDF, like any other PDF, must integrate to 1. So the height of this conditional PDF is equal to  $1/x$ .

Are  $X$  and  $Y$  independent? No. One way to see it is that if you change little  $x$ , the conditional PDF of  $Y$  would have been something different. Whereas if we have independence, all the conditional PDFs have to be the same when you change the value of little  $x$ .

Another way to see it is that if I tell you that  $x$  is 0.5, this gives you lots of information about  $Y$ . It tells you that  $Y$  has to be less than or equal to 0.5. So the value of the random variable  $X$  gives you plenty of information about the other random variable. And so we do not have independence.

Notice that in this example, instead of starting with a full description of the random variables in terms of a joint PDF, we use a marginal PDF and then a conditional PDF to construct our model. Of course, with these two pieces of information, we can reconstruct the joint PDF using the multiplication rule. The marginal is  $1/l$ . The conditional is  $1/x$ . So the joint is equal to  $1/lx$ .

But for which values of  $x$  and  $y$  is this the correct expression? It is correct only for those values that are possible. So 0 has to be less than  $y$ , less than  $x$ , less than  $l$ . This is the range of values that are possible in this particular experiment. And we can visualize those values.

They are those that correspond to this shaded triangle here.  $x$  and  $y$  are less than or equal to  $l$ . And  $y$  has to be less than or equal to  $x$ . If you try to visualize the joint PDF, notice that since it only depends on  $x$  not on  $y$ , if you fix a value of  $x$  and you look at the slice of the joint PDF, the value of the joint PDF is going to be a constant on that slice.

On this slice, it's going to be another constant, actually a bigger one. On that slice, an even bigger constant. And actually, this constant is bigger and bigger and goes to infinity as we approach 0. Of course, the fact that the slice is constant is just a reflection of the fact that the conditional PDF is constant over the range of values that the random variable can take.

Let us now continue with some calculations. Let us find the marginal PDF of  $Y$ . How do we do it? Since we have in our hands the joint PDF, we can find the marginal by integrating the joint. And in our case, the joint is equal to  $1/x$ . And we integrate over all  $x$ 's.

Now, what is the range of the integration? If we fix a certain value of  $y$ , the joint PDF is actually 0 in this region and in that region. So we should only integrate over  $x$ 's that correspond to this interval. What is that interval?

It's the interval that ends at  $l$ . And because this is a line of slope 1, this value here is also  $y$ . So we integrate over an interval where  $x$  ranges from  $y$  to  $l$ . In fact, this is just the range of  $x$ 's that are possible for a given value of  $y$ .  $x$  must always be larger than or equal to  $y$ .

Now, the integral of  $1/x$  is a logarithm. And using this fact, we can evaluate this integral. And it's  $1$  over  $l$  times the logarithm of  $l$  over  $y$ . For what  $y$ 's is this a correct expression? Well, it makes sense only for those  $y$ 's that are possible in this experiment. And that's the range from 0 to  $l$ .

When  $y$  is equal to  $l$ , we have the logarithm of 1, which is equal to 0. So the value of the PDF is 0 here. As  $y$  decreases, this ratio increases and goes to infinity. So the log of that also blows up to infinity. And we get a shape of this form, where the function that we're dealing with goes to infinity as we approach 0.

Is this a problem having a PDF that blows up to infinity? Not really. As long as the area under this PDF is equal to 1, it's still a legitimate PDF. And blowing up to infinity is not an issue.

Let us now calculate the expected value of  $Y$ . One way of doing this is by using the definition of the

expectation. It's the integral of  $y$  times the density of  $y$ , which is  $1/l$  times the log of  $l$  over  $y$ . And the range of integration has to be those values for which we have a non-zero density.

So we integrate from 0 to  $l$ , which are the possible values of the random variable  $Y$ . This is an integral that's pretty messy. One can actually integrate it using integration by parts. But the calculation is a bit tedious. So let us look for an alternative and more clever approach.

The idea is to divide and conquer. We're going to use the total expectation theorem, where we're going to condition on  $X$ . The total expectation theorem tells us that the expected value of  $Y$  is the integral over all possible values of the random variable  $X$ , which is from 0 to  $l$ .

The density of  $X$ , which is  $1/l$ , times the conditional expectation of  $Y$  given that  $X$  is equal to some little  $x$ . And we integrate over all  $x$ 's. Why is this simpler? When we condition on  $X$  taking a specific value,  $Y$  has a uniform distribution between 0 and  $x$ . And therefore, this conditional expectation is the expectation of a uniform, which is  $1/2$  the range of that uniform.

So we obtain the integral from 0 to  $l$ .  $1/l$  times  $x$  over 2,  $dx$ . And finally, that's an integral that we can evaluate easily. Or we can think even in a simpler way. This expression here is the density of  $x$ . This is  $x$  itself.

So the integral of this times  $x$  gives us the expected value of  $X$ . And there's only a factor of  $1/2$  that's left out there. So we obtain  $1/2$  the expected value of  $X$ . But now,  $X$  itself is uniform on an interval that has length  $l$ . And therefore, the expected value of  $x$  is  $l$  over 2.

And so we get the final answer, which is  $1/2$  times  $l$  over 2, which is  $l$  over four. This answer makes intuitive sense. If we break a stick once, the expected value or what we're left with is half of what we started with. But if we break it once more, then we expect it on the average to be cut by a factor again of  $1/2$ . And so we expect to be left with a stick that has length  $1/4$  of what we started with.

So this example is a particularly nice one, because we used all of the concepts that we have introduced-- marginal PDFs, joint PDFs, conditional PDFs, and the relations between them, as well as expectations, calculations of expectations, and conditional expectations, as well as the total probability theorem.