MITx Video | 6.041x | Spring 2014 | Video transcript | L11-2: A linear function of a continuous r.v.

We now move to the case of continuous random variables. We will start with a special case where we want to find the PDF of a linear function of a continuous random variable. We will start by considering a simple example, and study it using an intuitive argument. And afterwards, we will justify our conclusions mathematically.

So we start with a random variable X that has a PDF over the form shown in this figure so that it is a piecewise constant PDF. We then consider a random variable z, which is defined to be 2 times X. The random variable x takes values between minus 1 and 1. So z takes values between minus 2 and 2.

Now, values of X between minus 1 and 0 correspond to values of Z between minus 2 and 0. The different values of X in this range are, in some sense, equally likely, because we have a constant PDF. And that argues that the corresponding values of Z should also be, in some sense, equally likely. So the PDF should be constant over this range.

By a similar argument, the PDF of Z should also be constant over the range from 0 to 2. And the PDF must, of course, be 0 outside this range, because these are values of Z that are impossible.

Let us now try to figure out the parameters of this PDF. The probability that X is positive is the area of this rectangle. And the area of this rectangle is 2/3. So the area of this rectangle should also be 2/3. And that means that the height of this rectangle should be equal to 1/3.

Similarly, the probability that X is negative is the area of this rectangle, and the area of this rectangle is equal to 1/3. When X is negative, Z is also negative, so the probability of a negative value should be equal to 1/3. And for the area of this rectangle to be 1/3, it means that the height of this rectangle should be 1/6.

So what happened here? We started with a PDF of X and essentially stretched it out by a factor of 2 while keeping the same shape. However, we also scaled it down by a corresponding amount. So 2/3 became 1/3, and 1/3 became 1/6. The reason for this scaling down is because we need the total probability, the total area under this PDF, to be equal to 1.

If we now add a number, let's say 3, to the random variable Z, what is going to happen? The random variable Y now will take values from minus 2 plus 3-- this is plus 1-- all the way up to 2 plus 3, which is plus 5. Values in the range from 1 to 3 correspond to values of Z in the range from minus 2 to 0. These values are all, in some sense, equally likely. So they should also be equally likely here.

And by a similar argument, these values in the range from 3 to 5 should also be equally likely. This rectangle corresponds to this rectangle here. So the area should be the same. And therefore, the height should also be the same. Therefore, the height here should be 1/6. And by the same argument, the height here should be equal to 1/3. So what happens here is that when we add 3 to a random variable, the PDF just gets shifted by 3 but otherwise retains the same shape.

So the story is entirely similar to what happened in the discrete case. We start with a PDF of X. We stretch it horizontally by a factor of 2. And then we shift it horizontally by 3.

The only difference is that here in the continuous case, we also need to scale the plot in the vertical dimension by a factor of 2. Actually, make it smaller by a factor of 2. And this needs to be done in order to keep the total area under the PDF equal to 1.

Let us now go through a mathematical argument with the purpose of also finding a formula that represents what we just did in our previous example. Let Y be equal to aX plus b. Here, X is a random variable with a given PDF. a and b are given constants.

Now, if a is equal to 0, then Y is identically equal to b. So it is a constant random variable and does not have a PDF. So let us exclude this case and start by assuming that a is a positive number. We can try to work, as in the discrete case, and try something like the following.

The probability that Y takes on a specific value is the same as the probability that aX plus b takes on a specific value, which is the same as the probability that X takes on the specific value, y minus b divided by a. This equality was useful in the discrete case.

Is it useful here? Unfortunately not. When we're dealing with continuous random variables, the probability that the continuous random variable is exactly equal to a given number, this probability is going to be equal to 0. And the same applies to this side as well. So we have that 0 is equal to 0. And this is uninformative, and we have not made any progress.

So instead of working with probabilities of individual points which will always be 0, we will work with probabilities of intervals that generally have non-zero probability. The trick is to work with CDFs. So let us try to find the CDF of Y. The CDF of the random variable Y is defined as the probability that the random variable is less than or equal to a certain number.

Now, in our case, Y is aX plus b. We move b to the other side of the inequality and then divide both sides of the inequality by a. And we get that this is the same as the probability that X is less than or equal to y minus b divided by a, which is the same as the CDF of X evaluated at y minus b over a. So we have a formula for the CDF of Y in terms of the CDF of X.

How can we find the PDF? Simply by differentiating. We differentiate both sides of this equation. The derivative of a CDF is a PDF. And therefore, the PDF of Y is going to be equal to the derivative of this side.

Here we need to use the chain rule. First, we take the derivative of this function. And the derivative of the CDF is a PDF, so the PDF of X evaluated at this particular number. But then we also need to take the derivative of the argument inside with respect to y. And that derivative is equal to 1/a. And this gives us a formula for the PDF of Y in terms of the PDF of X.

How about the case where a is less than 0? What is going to change? The first step up to here remains valid. But now when we divide both sides of the inequality by a, the direction of the inequality gets reversed. So we obtain instead the probability that X is larger than or equal to y minus b divided by a. And this is 1 minus the probability that X is less than y minus b over a.

Now, X is a continuous random variable, so the probability is not going to change if here we make the inequality to be a less than or equal sign. And what we have here is 1 minus the CDF of X evaluated at y minus b over a. We use the chain rule once more, and we obtain that the PDF of Y, in this case, is equal to minus the PDF of X evaluated at y minus b over a times 1/a.

Now, when a is positive, a is the same as the absolute value of a. When a is negative and we have this formula, we have here a minus a, which is the same as the absolute value of a. So we can unify these two formulas by replacing the occurrences of a and that minus sign by just using the absolute value. And this gives us this formula for the PDF of Y in terms of the PDF of X. And it is a formula that's valid whether a is positive or negative.

What this formula represents is the following. Because of the factor of a that we have here, we take the

PDF of X and scale it horizontally by a factor of a. Because of the term b that we have here, the PDF also gets shifted horizontally by b. And finally, this term here corresponds to a vertical scaling of the plot that we have. And the reason that this term is present is so that the PDF of Y integrates to 1.

It is interesting to also compare with the corresponding discrete formula that we derived earlier. The discrete formula has exactly the same appearance except that the scaling factor is not present. So for the case of continuous random variables, we need to scale vertically the PDF. But in the discrete case, such a scaling is not present.