

module_prob4_lecture8

Fri, 2/18 12:51PM 15:11

SUMMARY KEYWORDS

cdf, pdf, probability, integral, negative infinity, derive, equal, area, integration, point, rectangle, uniform distribution, remember, derivative, calculus, interval, question, stock, suppose, density

SPEAKERS

Sumon Majumdar

S

Sumon Majumdar 00:18

So in this clip I want to explore for the connections between PDF and the CDF. So in particular, if you recall, the CDF has the probability up to a particular point. And the PDF is just a derivative of the CDF. Now, and we know that if we have a CDF, right capital F effects from that we can derive the PDF by taking the derivative. So the question is, can we also go the opposite route? So can we go that? Suppose we start out with a PDF? Can we derive a CDF? So what's the idea? So remember, in the previous clip, I gave you a second way of motivation for how the PDF comes around, basically, by looking at smaller and smaller intervals, and looking at the probabilities connected with those. Now, suppose you have this as your density function. And ask the question, you want to ask the question? What's the cdf of this arbitrary point? Let's say B. So remember, CDF is all the probability of values less than or equal to b? And the way to find it is that you add up this Yeah, the this add up this, add up all of these probabilities up to the point P. Great, and that, if you add all of those up, that will give you the CDF at that particular point. Taking this idea forward, right. And technically, what's happening is that if you have a probability density function or a PDF, and you want to figure out what's the CDF at a particular point B, would you have to look at is the area is this green area, the area under the PDF up to the point so that this area will give you what's the CDF at that particular point. So in the language of calculus, if you want to look at f of b , which is the area that the PDF up to the point B, it will be the integral, the integral of the PDF, starting from the lowest most possible value, which is negative infinity, add up the area all the way from negative infinity, up to B. So it's the integral of f of x from negative infinity to B. That's what gives you the CDF off of this random variable at the point. Now, another way to view this is C in calculus, the integration and differentiation are just opposite sides of the coin. Right? And this is, again, this is what we're doing. Right? That see the PDF is just the derivative of the CDF. So the opposite of that, is that if we want to now figure out what the CDF is, we take the integral of the density function. So let me let me do and so how, so again, here, so the the CDF is just the integral of the PDF from negative infinity to B, right? So if we were to do those areas, right, for the various values of B, the B equal to let's say, point one, b equal to five v equal to six, and in plot all of those up right, you get what would be the CDF and vice versa, if you started with the CDF and took the, the slope of it at the various points, right, and you plotted that you would get the PDF. So you can go from here to here, or you can go from here to here. So when we talk about continuous random variables, we can talk about either its cdf, or its PDF, because if we know one we can

immediately get. So now, if you have the CDF, you can also figure it out, what's the, what's the probability of a particular interval? What's the probability that the random variable lies in the interval a to b . So remember, this is just going to be the probability of less than or equal to b minus the probability of less than or equal to a . So this is equal to $F(b) - F(a)$. And if you view it in terms of areas, it's the total area up to B , and take out or subtract the total area to a . So what's left is just this little bit here. And that will be exactly the probability of the random variable lying in the interval a to b . And again, in the language of calculus, this is going to be integral with the areas the integral, the integral of $f(x)$ between a and b . So let's use this as an example. So remember this example that we did before the uniform distribution. And we derived its PDF, which was $f(x) = 1/10$ for x between zero and 10. And $f(x)$ is zero elsewhere. So now, so it's zero everywhere else, right? And it's $1/10$ in between. So the question now is that suppose I want to derive What's this, the cdf of this uniform distribution? In the earlier example, I had given you this, right. And from here, we derive the PDF by taking the derivative now, going the opposite route and giving you this, and I'm asking you, can you derive me the CDF. So wrong, but again, the CDF is the area under the PDF up to the point B , right, so that's, that'll be the CDF at least. So let's use this concept here. So if I want to figure out $F(b)$, so that we take an arbitrary point like this. So the $F(b)$ will be the area under the PDF up to b , so it's, it'll be the area of this bit. So in this particular case, it's relatively simple because this is just a rectangle. So the base of the rectangle is b times the height, the height is $1/10$. So this will be $b \times 1/10$. Is that this particular area. So the density is going to be beautiful. So remember, this is going to be $F(b) = b/10$ for all values of b between zero and 10. For values of b , which are less than zero, C , the density is zero. So they're the CDF is also going to be equal to zero. For b less than zero. What about greater than 10? So if I want to look at the value, let's say the density at CDF a 12, right? So the CDF at 12 will be the area under the PDF up to 12. So Well, in this part, there is nothing, right? So it will essentially be the area up to this point. So if I'm going to be looking at $f(12)$ it will be $1/10$ times up to the base here will be zero to 10. So times 10. So that's going to be and this is going to be true for all valid Use of b greater than 10. Right? Therefore $F(b)$ is going to be equal to one. So this is how given a PDF, you can use it to derive the CDF. So those of you who like more calculus, right, so you can cast the same thing that if you want to look at $F(b)$, right, it's the integral from negative infinity to B of $f(x)$. Now, $f(x)$ is zero for anything below zero. So I'm going to just therefore write, so the part between minus infinity to zero there, this function $f(x)$ is zero, so I can ignore that. So you can only take the integral from zero up to B . So they're $f(x) = 1/10$ for x between zero and 10. So those of you who struggled with integration, you may look at our module calculus module before, or if you still are, still find it too difficult, you can skip this integration part. If you have understood the concept that I did in the previous slide, just by finding the area, but by looking at the rectangle, that's fine, too. But those of you who are good at handling integration, let's just do this. So it's $1/10$. That's a constant that comes out. So I have d affects between zero and B . Right? So this is just the integral of one times d affects, so the integral of that is just x , right? And we have to evaluate it at zero, and B . So this will be $1/10 \times (B - 0)$. x and B is B , x at zero is zero. So this is equal to $b/10$. This is exactly the answer that we are also caught in the previous slide. Here, right? So that's that was the right. So this here is the CDF for this particular form distribution. And once you have this, right, you can ask lots of questions related to probabilities. So for example, right? Suppose X is the returns on a stock returns on the stock, which has the uniform distribution between zero 10. Right, so this is distributed uniformly between zero, right? And I'm going to introduce this uniform distribution in a later clip a little bit more formally, but right now, right? This is given by this PDF, right? And suppose I want to ask the question, what's the VAT? What's the probability that the stock takes the returns on the stock is less than or equal to five? So that is just the CDF at five? Remember, the CDF is given by $F(b)$ of the tank. So this is going to be just fine for return, which is its probability half that the stock return is less than or equal to five. What if I ask the question,

what's the probability that the stock return is between three and five? Again, to look at the probability of this interval, it's f of five minus f of three. f of five is again five over 10. Remember, f of b is equal to b over 10. Right for values of b between zero and 10. So both of these values are between zero and 10. So I can use this. So f of three is going to be three over 10. So this is going to be $2/10$ point right? So the probability that the returns of the stock are between three and five is 20%. So this is how given a unit given a particular PDF, you can derive the CDF. And once you have the CDF, you can use it to calculate a lot of other things as well. So this here is the connection between PDF CDF so you can go from one to the other and vice versa