

module_prob4_lecture5

Fri, 2/18 12:51PM 10:36

SUMMARY KEYWORDS

probability, continuous random variable, talk, temperature, random variable, x tilde, values, discrete random variables, variable, continuous, cdf, range, measured, discrete, random, equal, dartboard, takes, similarly, growth rate

SPEAKERS

Sumon Majumdar



Sumon Majumdar 00:06

So let me now talk about continuous variables. In the previous module, we talked about discrete random variables, which take a discrete number values. Now, there is a conceptual subtle difference between continuous, discrete random variables, which makes the study of the continuous random variables conceptually slightly more difficult. So, I'll try to explain what that difference is, and why we need to look at probabilities in particular. So, again, recalling What's the definition of a random variable. So it's a variable, which takes on some numerical values based on the outcome of some random event. And in the last module, we talked about discrete random variables, things like what number comes up in a die, number of children one has, number of cars one owns and things like that. So now we're going to be talking about continuous random variables. So random variable is continuous if it can take any value in some range. So if it can take not just values 0, 1, 2, 3, 4, if it can also take things in between, it can take values. Point 7777, point 888, and so on, right? That would be a continuous random. So in real life, there are lots of examples of a continuous random variable. For example, if you look at the growth rate of an economy or population, or, or COVID infections, right, so the growth rate can take any bow, it can take a value of 1%, it can take 1.2123% were 2.67%, and so on, it's a continuous random. Similarly, crime rate. Or if you look at profits or losses of a company, or the water level in the lake, or the temperature in a lake or temperature in the city, technically, it can take any value within negative positive, maybe water level is only positive, but it can take a height of five meters, height of 5.25 meters, five 5.26 meters, and so on, right, so these are all continuous random there. So sometimes in practice, right, because of measurement limitations, we just talk about discrete values, for example, we talk about temperature is really minus one minus two, or at most minus 2.5. Right, we don't talk about more like temperature of 2.14. But that's just because measurement and rounding off the technically temperatures can be any number. Negative, positive, not just integers things in between. Now, how to talk about probabilities, but before going there, so sometimes we talks about the range of random variable. So if you look at the maximum possible value for random variable, and the minimum possible value of variable, the difference between them as the range. So for example, if you look at temperatures across the world, now the highest possible temperature anywhere is about 70 degrees, right? So I don't think there has been any temperature measured above that. Similarly, if we look at the lowest temperature

in the world, the lowest temperature ever measured was about a minus 100 degrees centigrade sometime, if you look at the Guinness Book of Records, right, so So what's the range of temperature that is measured on Earth? Right? It'll be 70 is the maximum minus minus 100. So that range will be 170. Obviously, if you look at Pluto, the range will be something else and in Mars, something else and so now, let's come to probabilities. And here is where there's a bit of a conceptual difference between continuous and discrete and discrete random variables. It was fair Easy. We looked at what's the probability that X takes a very particular value x ? So what's the chance that die comes up to be three? Right? We know that probability is one sixth. Now for a continuous random variable, there is no number like three, there is a number like three, but see, think about about temperature, right. So temperature can take any values, right? What's the probability that the temperature in your city is exactly right? It's not 2.999. It's not 3.000001. It's exactly three. So that that's impossible to tell. So for a continuous random variable, the probability that the random variable takes a particular value very, very, very specific, particular value, that probability is going to be zero. Here's another example, which makes this point it's supposed to look at a dartboard. What's the probability that you hit exactly this point like that the tip of your Dart hits this exact point, not one micron this way, one micron that way. So that probability, is impossible to tell, right? So that probability is zero. So for a continuous random variable, the probability that the random variable takes a very specific value, that's zero. And this is conceptually different from what we had in discrete random variables. So now, if the probability of a particular value is zero, right, how do we still talk about probabilities here? So an easier way to talk about probabilities here is to instead of thinking about, about hitting a very particular point on the dartboard, we can talk about what's the probability that it hits this area? Or what's the probability that the temperature is less than or equal to three degrees. So this is easier to calculate. Because see, then it doesn't depend so much on measuring exactly that three. 2.99 2.98. All of these are all temperatures which satisfy this. So with continuous random variables, one way to look at probabilities is to look at the probability that the random variable takes values less than or equal to a specific value. Now, if you recall from our discrete random variables case, this probability that the random variable takes values less than a particular x , right, so what's the probability that it takes values less than or equal to x ? This is precisely the CDF or the cumulative distribution function of x . And we denoted that by capital F of x . So the same concept still holds in continuous random variables. So f of x , or the CDF is the probability that the random variable takes values less than or equal to x . And remember, we can use this to not just figure out that it takes values less than or equal to a particular number. We can also look at use this to figure out what's the probability that it takes values in an interval a to b . And the way to do that is first we'll look at what's the probability that it takes values less than or equal to b , right? And then subtract out a probability that it takes values less than or equal to A , then what's left over is this B . So, so, the probability that X takes values between A and B is the probability that it takes values less than or equal to b and minus the probability that it is one was less than or equal to a . So this is $f(b)$ minus $f(a)$. So this gives you so this gives you probability of an interval but still right. So this gives you probabilities of intervals. But now, how can we make get an understanding of that which temperatures are more likely which temperatures are less likely? Or, or things like that right, which water levels in the lake are more likely to occur which are less likely