

# module\_prob3\_lecture5

Mon, 2/7 10:37AM 14:28

## SUMMARY KEYWORDS

probability, cumulative distribution function, random variable, cdf, temperature, sales, tilde, profits, tables, equal, values, called, derived, calculate, depends, restaurant, point, ice cream parlor, ice cream, continuous distributions

## SPEAKERS

Sumon Majumdar

---



Sumon Majumdar 00:04

Since the last example, I showed you how the distribution of one random variable can be derived from data of another. In particular, we looked at how the profits of restaurant depends on the number of restaurant tables. So let me do one more examples on the slides, again from business. So, a business's profits probably depends on a lot of other factors trade. So for example, it could depend on the demand from consumers, it could depend on supply chain problems, it could depend on the weather, right, so the lots of different random factors that could have an impact on businesses sales and profits. So let's do another example. Along those lines. So in particular, consider an ice cream parlor, who sells depends on the temperature. And the temperature is given by a random variable  $t$  tilde. And here is the probability mass function of  $T$  tilde. So the temperature can take values less than 20 degrees, it's the probability of that happening is point five, the probability that the temperature is between 20 and 25 is point five, that between 25 and 30.2. Now this ice cream parlor sells depends on the temperature, as the temperature goes up, more people are likely to demand ice cream. So in particular, if the temperature is below 20 degrees, the sales are zeros. So this is another random variable as tilde, right, which gives the sales of this ice cream parlor. So if the temperature is below 20, the sales are zero. So these are sensitive people. So they only buy ice cream if the temperature is off 20. Now if the temperature is between 20 and 30, then the sales is the moderate 30 units of ice cream. If the two bridges between 30 and 40, that's when the sales come up to 60. And when the temperature crosses 40, everyone is Bryan ice cream, and then says 200. So what we have here is the probabilities of the temperature. So how can we use that to do to get the probability mass function of the sales random variable, which is? So first thing know that the seals here can take only four possible values either at zero 3060, or 100. And what we want to figure out is the probability connected with each of these. So let's do that. Right, so let's look at sales equal to zero. So this only happens when the temperature is below 20. And what's the probability of that? That's point one, five, here. So that's the probability therefore, that the sales will be zero. When can the sales be 30? The sales will be 30, when the temperature is between 20. So what's the probability of that so see, temperature between 20 and 30? It's either this or this? Right? So if we add those two up, so point one, five, and point two, that's point three, five. So that's the probability that the temperature is between 20 and 30. And then for the sales. Similarly, for looking for, what's the probability that the sales are going to be 60,

that's here. And this is going to only happen when the temperature is between 30 and 40. So when is temperature between 30 and 40, either this or this. So if we add these two probabilities up, point two and point two five, so that's point four five, that's the probability of the of sales being 60. What about the remaining the sales of 100 That only happens when the temperature is more than 40? And the probability of that is 0.05. So what we have done is we have derived this random variables from this room. And we have used the basic distribution on  $T$  tilde to derive the probability mass function That's the sales. So this is how you can derive the probability of one random variable from that. Okay, let me now move on to another very important concept. This is another way of presenting probabilities. And this is called the cumulative distribution function. This is going to be super, super important in our next module on continuous distributions of continuous discrete variables. But let me introduce you to the concept. So, so far, we have looked at what's the probability that the random variable takes a particular value,  $x$ . So, that was called the probability mass function of the PN. That's what we have been derived so far. Now, here, the idea is that, I asked the question, what's the probability that this random variable takes values less than or equal to small  $x$ ? So what's the probability that  $X$  tilde takes values less than or equal to small  $x$ ? Or another way of saying what's the probability that the value is that the random variable the staking is app no small  $x$ . So this is called the cumulative distribution function. That particular probability is called the cumulative distribution function, or CDF  $H$ . And typically, this is denoted by this capital letter  $F$  of  $x$ . And how do we calculate this, this cumulative distribution? So see, what we're interested in is what's the probability that this  $x$  tilde, right, so this is my random variable, it takes values less than or equal to. So what we do is that we look at what are the possible values that it can take less than or equal to  $x$ , and look at the probability connected with this and this and this and this, add all these up, that will give you the overall probability that the random variable takes values less than two. So succinct way of writing it using the summation notation that we had done in a previous module, the  $C$ , this is going to be the summation, that means the sum of the probabilities, sum of  $p$  of  $y$ , and for all those  $y$ 's, which are less than or equal to  $x$ . So this is let's say, one, one, let's call it  $y$  one,  $y$  two,  $y$  three. So we're looking for all those  $y$ 's which are less than or equal to  $x$ . And adding up the probabilities connected with this probability connected to this probability connected this and so on, right? So if we can add all of these up, right, so that will be the cumulative probability up to  $x$ , or what's called the cumulative distribution function. So let's apply this to an example. So this was our example one about a restaurant with five tables. And these were the, these were the probabilities of no table being occupied, one table being occupied two tables. So now what we want to calculate there is the CDF for the cumulative distribution function. So what we first want is to look at  $f$  of  $z$ . So what's the probability that the number of tables occupied is less than or equal to  $C$ ? So the only case there is zero, right? That's the only way the number of tables can be less than or equal to zero. So this is just going to be point. Now, the next one, which is what's so let's look at  $f$  one, right? So this is the probability that the number of tables is less than or equal to one right number of tables occupied is less than or equal to one. So what are the possibilities? Either the number of tables is zero, or occupied is zero, or the number of tables occupied is one So what we need to do is to add up this probability and this probability. So point one and point two. So that's good. And this is how we progress if we look at our two, right? So if you're looking here, right, so this is, what's the probability that the number of tables occupied is two or less, right? So it can either be two, which is this, or it can be one, or it can be  $z$ . If we add these three things up, it's point 4.2 1.1. That's. And you see what we're doing right? So if we're to do the next one, right,  $f$  of three,  $C$ , up to here, point seven is adding up this three. Now, if we were to look at what's the probability that the number of tables is occupied is three or less, so we already have 0.1. And two, we have up to here. Now we just add to that this, and that becomes 0.9 So this will be CDF at three, right? So this will be  $f$ . The next one here, this will be  $f$  of four, that's the cumulative distribution function at four. So what we just need to do is to this, we

need to add this one. So that will be point five. And then here, we ask the question,  $f$  of five, what's the probability that the number of tables occupied is five or less? So we already know the answer for formulas. Now we just add one more, this one to that. So that becomes point nine, five plus 0.05. So this is my cdf, this is my CDF in the index. Now, let's do one more, right. So this was this was the profits of that restaurant. And we have the PMF. This is the probability of this is the probability of profits being minus 100. This is the profits of profit. This is the probability of profits being minus 50. And so on, if we want to calculate the CDF, which is this. So less than or equal to 100, that's going to be point one. Less than or equal to 50, we add a point one and point two, that's point three. Less than or equal to zero, we have to add a point 1.2 And point four, right, so point one and point we already know a point three, so we just have to add this to seven. Next one for to get the CDF at this point, we just need to add this to this. So this will become point nine, next one, this to this, this, five, and this to this. So every step, we are just adding the extra probability, and that's giving us the CD. Right? So if we look at this point here, right,  $f$  of 100. So what is that? So this is the probability that the profits of this restaurant will be less than or equal. And we've calculated that to be is naught point nine, five. So that means 95% chance the profits of this restaurant will be less than or equal to. So this is the basic idea of CDF or cumulative distribution function. Let me stop the clip here. And in the next clip, I'll be picking up on some properties.