

module_prob3_lecture3

Mon, 2/7 10:37AM 12:36

SUMMARY KEYWORDS

probability, random variables, number, x tilde, random, variable, outcome, talk, continuous variable, examples, denote, tilde, values, typically, derived, crimes, similarly, continuous random variables, probability mass function, translated

SPEAKERS

Sumon Majumdar



Sumon Majumdar 00:05

So let's now start with a fresh topic of random variables. Random variables is a term you'll encounter all the time in social sciences. And it means something very simple. So what's a random variable? It's a variable, which takes on numerical values based on the outcome of some random event. So that's the definition. But what it means is something very, very simple, that it is a variable, which takes some values, right? Could be 1, 2, 3, 4, etc. But it's random in the sense that it depends on what happens in a particular event, or a particular experiment. So let me give you some examples. So let's think of suppose the number of forest fires in a province or in a region right now. So the variable could take values either 0, 1, 2, 3 400, so on, it just depends on the number of forest fires, they were in that region in that season. Similarly, it could be the number of crimes or a particular type of crime, the number of murders in a particular city, the number of spectators who attend a particular maybe a baseball game, or a hockey game. So this different, this is random in the sense that the number of spectators one day could be a lot in other days, it could be very few. Now, the day could be something else, right. But it the way the random variable there is the number of spectators. Similarly, the number of children in a family, right, so some families have one, some families have none, some families of five, right, so it's the number of children in a particular in a family. That's the random that's around variable. And all of these are examples of what's called discrete random variables, that means they take only a discrete number of values 0, 1, 2, 3, 4, the number of children cannot be 2.2, but 2.2416. Similarly, the number of spectators are like 1000 1002, and so on, it cannot be 1,000.81. Right. Now, the examples example, is the number of minutes your phone lasts on a single charge, right. So the number of minutes could be 60, number 81, 92. so on. And of course, like, this is something which is based on measure, right? One could talk about a phone, like if you're really, really finicky, and you measure it particularly well, then it could be in 61.9 to 77 minutes or so on it. But typically, what we talk about is minutes is like whole number 61, 62, 65, 70. So the other example is continuous random variables, right? So suppose we talk about the growth rate of an economic, right, so here is the growth rate of the global economics. Now this is a continuous variable, because it can take any number between, like this could be negative numbers could be positive numbers, it could be numbers in between one and two, it could be 1.2785, and so on, right? Similarly, if we're talking about the crime rate, which is the number of crimes divided by the population, now, that may not be a whole

number, that could be anything between one to three, like it could be 1.6666. So on, right. Similarly, profits and losses are the water level in the lake, right? So that's a continuous where these are all continuous variable. And these are random variables, because they differ from season to season, year to year, month to month, and so on. So these are all examples of random variables. Now, so these were all straightforward, right? Now, they could be random variables, which could be derived from the outcome of some event, right? So what's an example? So So see, see your grade in a particular class, right? So Now typically what happens is that you get a numerical grade, let's say you get a grade like 88.5, right? And that gets translated into a letter grade. So if your score is somewhere between 85 to 90, then you get an A. Right? So the so the outcome here, right in terms of the random event is your numerical score, right? Then that gets translated into a letter grade, right? And then you can convert it back into a GPA of 3.9. So that's the that's the numerical score, based on what percentage you got in the, in the course. Right, so this is an example of a derived pair. Other examples of this is, for example, if you look at companies, the profits are based on sale numbers, right? So number of units sold, that could be random, but that translates into certain amount of profits for the company. Or, for example, the commission that you get based on sales, it's the typical commissions are like you get a certain percentage, if your sales are below a certain level, and that percentage goes up when you sell more, and so on. Right, but if you're thinking about the random variable, as the total amount of commission that you get, this is based on some other outcome, which is the number of seats. And I'll give you some solid examples of that. In specific contexts, and later. Now, once you have random variables, now, the what we're really interested in here is the probabilities connected with this, right? What are the probabilities of the different outcomes? So the probability mass function is the probability that the random variable takes a particular fact. Okay, so the probability that- so we typically talk about the random variables, we denote it by number by things like X tilde, U tilde, right? So typically, that denotes that this is a random variable. Now, what's the probability that the random variable X still that takes a particular value small x , right, so this denotes a particular value? So this is called the, this is denoted by P of X , and it's called the probability mass function. So what's the probability that the random variable takes the particular value x small x , right? So typically, we denote by capital X , the random variable and small x the particular value that that random variable takes. So let's do an example of that. Let's say let's say a restaurant has five tables. And x tilde is the number of tables which are occupied. So x tilde, what are the possible values that could take it could take the value 0 1 2 3 4, and five, right, so either no tables are occupied, or all five are occupied, or four are occupied, and so on, right? So this is an example of a discrete random variable. Now, suppose I give you what's like, say, at the lunch hour that these are the probabilities of x tilde taking the different values, the probability that the that it's zero, right, that means no table is occupied, that's point one, the probability that one table is occupied, that's point two, probability that two tables are occupied a point four right? So P of two, that means would be equal to point four. So these different probabilities for the different values of the random variable X tilde, right, they constitute what's called the PMF the probability mass function. So two things to note here, right. So these are probabilities. So see all of these are between zero and one. Right. So they cannot exceed one cannot be less than zero. Secondly, right if you add them all up see if you add these up. So this is point one plus point two, that's point three plus point four, this point seven, plus point two. That's point 9- 95. And another point 0.5, that's one, right? So if you add up all these probabilities, that's going to be equal to one. Because remember, these are the only possibilities that the extent the random variable X tilde can take, are only these six possibilities, right? So if you add up all the probabilities, they cover all the, if they should cover all the possibilities, and if you look at the entire probability of what's called the sample space, right, so that should be equal to one. Now, you can plot these, right? So here are plotted them by histogram, right? So this is the, this gives you the probability of zero tables, the probability of one table is point two of two tables

being occupied is point four, and so on. So this was an example of a random variable and probability mass function. Let's do a one other example. So this is the case where say two dice are rolled. And x tilde is the sum of their outcomes. So what's the possible values of X tilde, so they can be either 2, 3, 4, 5, so on up to 12? Because see, it's the sum of two dice, right? So the minimum is one and one so it could be so the minimum of x tilde is two, and the maximum could be six and six. So the sum maximum is 12. So what's the probability of each of these so this is something you should be able to derive based on what we have done in the previous modules. So I'll stop the clip here. Right? And I would like you to try out deriving this probabilities on your own. And in the next clip, I'll start out by by deriving those probabilities so you can check your answers against mine