

module_prob1_lecture2

Fri, 1/21 12:30PM 15:14

SUMMARY KEYWORDS

venn diagram, students, surveyed, probability, car, represent, set, owners, events, denoted, intersection, sample, diagrams, talk, venn, union, pfaff, common, professor, remember

SPEAKERS

Sumon Majumdar



Sumon Majumdar 00:04

So, now that we have gotten introduced to the idea of probability, let me introduce some formal terminology. So one of the first things is, which is important, is the set of all possible outcomes. So technically there is a name for it. So this is called the sample space, this is the set of all possible outcomes. So in the context of the die, remember, the set of possible outcomes is 12345, and six, so this here is the sample space. And we typically denoted by S . Now, what we are looking for in terms of probability is some particular occurrence, right. So, technically this is called an event. So this is a collection of outcomes, or a subset of the sample space. So, for example, if you're looking at the events, you want to find out if an even number came up, then that means it's a set of two, four, and six. So, this is a smaller part of the total sample space. So, this is called an event. Similarly, if you wanted to look at all odds, so, this would be a different event, which would encompass one three and five. So, these are sets, this is one set, this is another set, right, the remember, we looked at also the event that the number was greater than four, then that would be a set consisting of just five and six right, and there was some common elements between these two sets, there are some common elements between these two sets. So, oftentimes, when we are representing sets, a common diagrammatic way to do it is using a Venn diagram. So this is something which also professor Pfaff had talked about in an earlier module in the course on sets. But let me just refresh your memory about this. So Venn diagram is a representation of sets. So for example, this here, this rectangle would represent the sample space S . So this would consist of all the numbers from one through six for this particular for die. Now, suppose you want to represent in this the set of events. So the set of events will consist of two, four, and six. So what's outside the set of events is the other numbers, which is one, three, and five. So this is an example of a Venn diagram. So let me do a couple of other examples of Venn diagrams, because you will see that they will help you in calculating probabilities in an easy way. So let's do this example. So suppose there was a survey conducted. And then the survey, 100 people were surveyed 60 of them, they own a car. So this is a survey about car ownership. And they were also asked if they were students. So among the people who were surveyed, 30 were students, and of these students, 15 of them own a car. So how can I represent this in a Venn diagram? So so this is my Venn diagram here. So the rectangle here, this represents all the people who were surveyed. So among them, some were students. Let's at this blue set is the set of students. The green represents car loans. And there is some commonality between them because some

students also own a car. So how can we represent? Let's look at the rates of 15 of the students own car. If you look at this middle section here, which is the common element between students and car owners, so there are 15 of them here. So that means this part, just the plain light blue part, right? So these are the students who do not own cars. So there were 30 students 15 are in this common area. So outside this common area there are another 15 students. What about the car owners? Right? 60 people reported they own the car. Right? So the green is 60 of them. 15 are here. So outside is 45. What about people who are neither students nor car owners? So see if I'd to total up this right, so there are 15 students here 15 students who own a car. So that's 15 plus 15 that's 30. Right there, this 45 car owners were not students are 30 plus 45, that 75. Total 100 people were surveyed. So there are 25 people outside who are neither students nor car owners. So this is an example of a Venn diagram. And how does it help us in calculating probability? Because let me ask this question, right? So if I asked what is the probability of picking someone, right, I'm doing the survey at random, what's the probability of picking someone who is a car owner, but not a student? So if you look at here, go back to the Venn diagram. So there are 45 car owners, who are not students. So if you're looking at the probability, it will be 45. Favorable outcomes, relative to the total number of outcomes, 100 people were surveyed. So it's 45 over 100. So this probability here is going to be 45 over 100, which is .45. Now in the context of this Venn diagram, let me also introduce some other terminology. This is the terminology of sets, right? So let's say let's denote S as the set of students. So this was the blue set in my Venn diagram, let's call C as the set of car owners, this was the green one in my Venn Diagram. So instead, we talk about intersection of sets and union of sets. So the intersection of sets is denoted by this inverted U . So this is again, something which Professor Pfaff talked about before. So intersection is this idea of what are things which are common to both sets, so students, and car owners. So this is this part here, right? The part which represents those students who are also car owners. So it's students and car owners. That's the part which is the intersection of those two sets, $S \cap C$. The other thing that we talk about in four sets is the union of sets. So the union is denoted by U . And what is the union of students and car owners? Is the set of students and the set of car owners. What's the union of that? This represents all the people who are either students or car owners. So in the context of this diagram, what is the union? It's all of this. And all of this, it's the combination of people who could be students, or car owners, or both, right? So that's the union. So that will be $S \cup C$. Again, these are terms which you probably already know and remember, but I just wanted to refresh your memory. So let's do one other example. Right? Just to get the hang of, of this Venn diagrams. So let me extend this survey. That same example, let's say again, 100 people surveyed 60, owned cars. And of them. We also asked the question, how many cars do you own. So these car owners, now 20 owned more than one car. One plus cars. Same thing. We surveyed under 30 students, 15 of them own cars, right? This is the information we already had. Now, we asked them do any of you own more than one car? And five of them they claim that they own more than one car. So how do we represent now this right, because now there are students, there are car owners, and there are also owners who own more than one car. Right, so there are three sets here. One is the blue one, the set of students. One is the light green one, which is the set of car owners. And then a smaller set within the set of car owners of those who own more than one car, the one plus cars, right? So this dark blue set of one plus car owners is within the set of cars. So now let's incorporate the various types of information that we have. So first thing, if you look at the students that 15 of them own cars, and five own one plus cars. So if you look at the P among the students, right, so the blue set, and you look at its intersection with a set of one plus cars, right, so there are five students here who own more than one car. So put those five students, this particular five there. Now 15 students own cars overall, right? So five of them are here. So 10 must be in this segment here. So these are 10 students who own cars, but just one car each. So total, there were 30 students in the in the sample, right? So 10, and five are here, right? So this is 10, this

is five. So that being the remaining 15, must be the other students who don't own any cars. So now let's look at the people who are outside. So who are not students? Right? So this is this part here. Right? So total, if you if you look at the survey, 20 people say that they own one plus cars, right? Well, then there are five are students. So the remaining 15 are non students who own one plus cars. Right? And what about the car owners? Right, so the total there are 60 car owners, so that 10 here, five here, so 15, plus another 15 here. So that's 30. So there are 30 car owners who own just one car and are not students. And what about people who are neither car owners nor students? So remember, the 100 people were surveyed. So if we add up all the people in the set, so this is 15, 5, 20, and then 30, and another 30, 60, and 15, 75. So that means there are 25 people outside. Right? Who are neither students, nor do they own any cars. So once you have this we can talk about different probabilities. So what's the probability in the sample of someone being a car owner? So this, so see who are car owners, the 60 people who are car owners, so 60 out of 100 is the probability of somebody being a car. And you can look at more complicated events, what's the probability of car owners who are not students? Now see, so it's these 30 here, right? Plus this 15 here, because these are the people who do not intersect with the blue set, which is the set of students, right? So these are people who own cars, but are outside the blue set. So there are 45 such people out of the hundred. So the probability of someone in the sample being a car owner, but not a student, is 45 over 100. So that's .45. So this was the idea of Venn diagrams, because they help you represent data and complicated data in a very simple, diagrammatically. And once you have it represented this way, you can look at, you can ask different types of questions and answer them by looking at what the, where the data lies within the Venn diagrams. So that's why these are useful, is a useful tool, especially when we're talking about probability and probability of different sets in union of sets, intersection of sets. So let's stop this clip here and in the next clip pick up on some of the basic axioms of probability.