

module_prob1_lecture4

Fri, 1/21 12:30PM 13:37

SUMMARY KEYWORDS

probability, intersection, rule, venn diagram, february, customers, snow, addition, nonfiction, axiom, cold, fiction, buying, nonfiction book, area, ottawa, disjoint, book, figure, bought

SPEAKERS

Sumon Majumdar

Sumon Majumdar 00:04

Let me now talk about another rule of probability. And something which is equally important. This is called the addition rule. What this gives you is the probability of the union of two sets A and B is probability of A plus probability of B minus the probability of A intersection B. So let me show it to you in the context of a Venn diagram, what is the set that what does this rule actually telling you, it's actually pretty simple rule once you get the hang of it. So if you want to find out the probability of the union of A and B, so here is a and b. So if we look at the union of A and B, it's this area. So what this rule is telling you is that if I want to find the probability of this union, that is the probability of A plus the probability of B. But C, once I've done probability of B, and I've done probability of a lead, so this middle area, I have double counted this, right? I've counted once in a once in B. Right? So if I want to get rid of one of those double accountings, I need to take, you need to subtract this area out. Okay? And what is this, this is A intersection B. So probability of A plus probability of B minus probability of A intersection B. That's the probability of the union of A and B. Now, I want to relate this to something which you've seen before. So remember, one of the axioms of probability was that if A and B are disjoint, when A and B are disjoint. That means they have nothing in common. So that is probability of A plus probability of B. Remember, that was the third axiom of probability. Now here, this introduces this new term minus P of A intersection B. If A and B are disjoint, then in that case, there's nothing in common. So this set here is an empty set. And therefore the probability of this is zero. Right, so that, therefore, in the previous axiom, this part didn't even exist. But when sets are not disjoint, when they have something in common, that's when this formula comes in. So probability of A union B is probability A, plus probability of B minus the probability of the intersection. Let's now apply this. So here's an example of a customer's visiting a bookstore. Let's say this particular day, 200 customers visit this bookstore. So 75 bought of nonfiction book, 100 bought a fiction book and 25 bought books. So there are some customers who which bought bought two books, others 75 Others bought just a single nonfiction, or a single fiction, some didn't buy any. So how can we so it's easy to represent this by a Venn diagram? Right? So here's the set of people who bought fiction. Here's the set of people who bought nonfiction. So this part here, the intersection the common part, right, so there are 25 people there who bought both fiction and nonfiction. So how many just bought a fiction so see 100 bought a fiction. So 25 are here. So 75 are here. Now 75 bought a nonfiction book, of which 25 are already here. So these 50 just bought a single nonfiction book. So there

were a total of 200 customers. So if we want to try and figure out how many customers didn't buy any book at all, so let's add these guys up. So this is 75. This is 25. That's 100 plus 50. That's 150. Right? So the 150 customers within this area So that means the remaining 50 People are outside this area out of the total of 200. So these 50 people didn't buy any book. So now let's try to use this in the context of probability. Let's suppose you're trying to figure out what's the probability of a customer buying a book. So here, the customer could either be buying a fiction, or he could be buying a nonfiction. So what you're interested in is what's the probability that the customer buys a book, whether it's fiction, or nonfiction. So this will be equal to by the addition rule is equal to the probability of someone buying a fiction book. So 75 out of sorry, 100 out of 200, bought a fiction book. So this is 100 over 200 plus P of B, so B is nonfiction. So what's the probability of someone buying a nonfiction, so that's plus 75. Water nonfiction out of 200 minus the probability of A intersection B. So A intersection B is the people who bought both the fiction as well as the nonfiction. Right? How many such customers are there, there are 25. So this is minus 25 200. Right, so we putting the addition rule to use, so this is P of A plus P of B minus P of A intersection B. So they have the same denominator that's 200. So let's, so this is 100 plus 70, 175, minus 25. So that becomes 150. So this is point seven, five. So the probability of a customer buying a book, some book, right, fiction, or nonfiction, is 75 .75, or 75%. You can also so this is using the addition rule to figure this out. Of course, you could also do it much easier here. So see, see if you added up the total number of people who could buy a book so that 75 plus 25, 100 plus 50, 150. So the probability of someone buying a book is 150, or 200. But you could also do it this way. But see this, the addition rule is essentially putting this this idea of the Venn diagram again, if you go back to how we actually derived this, it's putting the idea of this Venn diagram that is this plus this, and there's a double counting. So we should take that out by just putting that more formally. So once you get the idea of this, right, the rest of it, so don't get blown away by looking at what's this big formula. What it is, is just putting this simple idea of the Venn diagram into more formal language. So let's, let's do one more example. Right. So this is an example of the Winterfest in Ottawa. So you know, in Ottawa, there's always a winter fest somewhere around February. So suppose you're trying to think about organizing the festival in Ottawa on February 5, and in fact that that's the date of the Winterlude, beginning of the Winterlude. In, I think the next year 2022. And what you're worried about is whether it will snow or be very cold that day. So typically, what you may want to do is you look at the data of the past 10 years, right? And this is what I did, I looked up and Weather Underground, and looked at the past 10 years of data for auto for February 5 for that particular day. And notice saw, so it snowed, two out of those 10 years on that particular date. Was it very cold. So I looked at temperatures below minus 10 degrees centigrade, right. So five of those years it was below. So it was below 10 degrees centigrade. And how many of those years did it both snow and be very cold be less than 10 degrees, it was one out of 10. So this is the data I collected. And what I want to know is what's the probability that this year on that day of the first that is on February 5, there could be snow, or be very cold. So what we're interested in is what's the probability of there being snow, or very cold. So it's P of A, so p of snow, which is step two over 10 plus P of B, which is very cold. So that's five over 10. minus, so there's two minus the probability of the intersection that is both snow and very cold. That's one over 10. So these all have the same denominator. So that's 10. So it's two plus five, seven minus one, that's six over 10. Or .6. Right? So based on this, on this my research of last 10 years, it seems that on on February 5, what's the probability of there being snow or being very cold is point six? And in fact, you could do the opposite. What's the probability of the day of being good weather in the sense of what's the probability of there being no snow? What's the opposite. So this is the being snow or being very cold. So the opposite of no snow and temperature greater than minus 10? Right. So this is the compliment event. So what we can use is the complement rule to figure out this probability, this will be one minus the probability of snow or very cold. So we already figured this out. So that's point six. So

this will be one minus point six, which is point four. So this is the type of things that the, the addition rule is very useful for is to figure out what's the probability of two things happen, right? Like one thing, or the other thing? What's the probability of them happening? It's probability of the first plus probability of the second minus the probability of the common. So please do remember this column, this addition rule, and sometimes it see once you have the probability of this event, you can now use the complement rule to figure out the probability of the complement event or the opposite event. So let's stop this clip here. And in the next clip, I'm going to begin with a clicker question, which is going to be based on this addition rule. So please do review it