

In this segment, we develop the inclusion-exclusion formula, which is a beautiful generalization of a formula that we have seen before. Let us look at this formula and remind ourselves what it says.

If we have two sets, A_1 and A_2 , and we're interested in the probability of their union, how can we find it? We take the probability of the first set, we add to it the probability of the second set, but then we realize that by doing so we have double counted this part of the diagram. And so we need to correct for that and we need to subtract the probability of this intersection. And that's how this formula comes about.

Can we generalize this thinking, let's say, to the case of three events? Suppose that we have three events, A_1 , A_2 , and A_3 . And we want to calculate the probability of their union. We first start by adding the probabilities of the different sets. But then we realize that, for example, this part of the diagram has been counted twice. It shows up once inside the probability of A_1 and once inside the probability of A_2 .

So, for this reason, we need to make a correction and we need to subtract the probability of this intersection. Similarly, subtract the probability of that intersection and of this one. So we subtract the probabilities of these intersections. But, actually, the intersections are not just what I drew here. The intersections also involve this part.

So now, let us just focus on this part of the diagram here. A typical element that belongs to all three of the sets will show up once here, once here and once there. But it will also show up in all of these intersections. And so it shows up three times with a plus sign, three times with a minus sign, which means that these elements will not be counted at all. In order to count them, we need to add one more term which is the probability of the three way intersection.

So this is the formula for the probability of the union of three events. It has a rationale similar to this formula, and you can convince yourself that it is a correct formula by just looking at the different pieces of this diagram and make sure that each one of them is accounted properly. But instead of working in terms of such a picture, let us think about a more formal derivation. And the formal derivation will use a beautiful trick. Namely, indicator functions.

So here is the formula that we want to establish. And let us remind ourselves what indicator functions are. To any set or event, we can associate an indicator function. Let's say that this is the set A_i . We're

going to associate an indicator function, call it X_i , which is equal to 1 when the outcome is inside this set, and it's going to be 0 when the outcome is outside.

What is the indicator function of the complement? The indicator function of the complement is 1 minus the indicator of the event. Why is this? If the outcome is in the complement, then X_i is equal to 0, and this expression is equal to 1. On the other hand, if the outcome is inside A_i , then the indicator function will be equal to 1 and this quantity is going to be equal to 0.

If we have the intersection of two events, A_i and A_j , what is their indicator function? It is X_i times X_j . This expression is equal to 1, if and only if, X_i is equal to 1 and X_j is equal to 1, which happens, if and only if, the outcome is inside A_i and also inside A_j .

Now, what about the indicator of the intersection of the complements? Well, it's an intersection. So the associated indicator function is going to be the product of the indicator function of the first set, which is 1 minus X_i times the indicator function of the second set, which is 1 minus X_j .

And finally, what is the indicator function of this event? Here we remember De Morgan's Laws. De Morgan's Laws tell us that the complement of this set-- the complement of a union-- is the intersection of the complements. So this event here is the complement of that event. And, therefore, the associated indicator function is going to be 1 minus this expression.

And if we were dealing with more than two sets-- and here we had, for example, three way intersections-- you would get the product of three terms. And if we had a three way union, we would get a similar expression, except that here we would have, again, a product of three terms instead of two. So now, let us put to use what we have done so far.

We are interested in the probability that the outcome falls in the union of three sets. Now, an important fact to remember is that the probability of an event is the same as the expected value of the indicator of that event. This is because the indicator is equal to 1, if and only if, the outcome happens to be inside that set. And so the contribution that we get to the expectation is 1 times the probability that the indicator is 1, which is just this probability. Now, the indicator of a three way union is going to be, by what we just discussed, 1 minus a product of this kind, but now with three terms.

Let us now calculate this expectation by expanding the product involved. We have this first term, then, when we multiply those three terms together, we're going to get a bunch of contributions. One

contribution with a minus sign is 1 times 1 times 1. Another contribution would be minus minus-- that's a plus-- X_1 times 1 times 1. And similarly, we get a contribution of X_2 and X_3 .

And then we have a contribution such as X_1 times X_2 times 1. And if you look at the minus signs-- there are three minuses involved-- so, overall, it's going to be a minus. Minus X_1 times X_2 . And then there is going to be similar terms, such as $X_1 X_3$ and $X_2 X_3$. And, finally, there's going to be a term X_1 times X_2 times X_3 . There's a total of four minus signs involved, so everything shows up in the end with a plus sign.

So the probability of this event is equal to the expectation of this random variable here. We notice that the ones cancel out. The expected value of X_1 for an indicator variable is just the probability of that event. And we get this term. The expected value of X_2 and X_3 give us these terms.

The expected value of X_1 times X_2 . This is the indicator random variable of the intersection. So the expected value of this term is just the probability of the intersection of A_1 and A_2 . And, similarly, these terms here give rise to those two terms here. Finally, X_1 times X_2 times X_3 is the indicator variable for the event A_1 intersection A_2 intersection A_3 . Therefore, the expected value of this term, is equal to this probability. And, therefore, we have established exactly the formula that we wanted to establish.

Now this derivation that we carried out here, there's nothing special about the case of three. We could have the union of many more events, we would just have here the product of more terms, and we would need to carry out the multiplication and we would get cross terms of all types involving just one of the indicator variables, or products of two indicator variables, or products of three indicator variables, and so on.

And after you carry out this exercise and keep track of the various terms, you end up with this general version of what is called the inclusion-exclusion formula. So the probability of a union is-- there's the sum of the probabilities, but then you subtract all possible probabilities of two way intersections. Then we add probabilities of three way intersections, then you subtract probabilities of four way intersections, and you keep going this way alternating signs until you get to the last term, which is the probability of the intersection of all the events involved.

And this exponent here of n minus 1 is the exponent that you need so that the last term has the correct sign. So, for example, if n is equal to 3, the exponent would be 2, so this would be a plus sign, which is

consistent with what we got here. So this is a formula that is quite useful when you want to calculate probabilities of unions of events. But also, this derivation using indicator functions, is quite beautiful.