

Conditional probabilities are probabilities associated with a revised model that takes into account some additional information about the outcome of a probabilistic experiment. The question is how to carry out this revision of our model. We will give a mathematical definition of conditional probabilities, but first let us motivate this definition by examining a simple concrete example. Consider a probability model with 12 equally likely possible outcomes, and so each one of them has probability equal to  $1/12$ . We will focus on two particular events, event A and B, two subsets of the sample space. Event A has five elements, so its probability is  $5/12$ , and event B has six elements, so it has probability  $6/12$ .

Suppose now that someone tells you that event B has occurred, but tells you nothing more about the outcome. How should the model change? First, those outcomes that are outside event B are no longer possible. So we can either eliminate them, as was done in this picture, or we might keep them in the picture but assign them 0 probability, so that they cannot occur.

How about the outcomes inside the event B? So we're told that one of these has occurred. Now these 6 outcomes inside the event B were equally likely in the original model, and there is no reason to change their relative probabilities. So they should remain equally likely in revised model as well, so each one of them should have now probability  $1/6$  since there's 6 of them. And this is our revised model, the conditional probability law. 0 probability to outcomes outside B, and probability  $1/6$  to each one of the outcomes that is inside the event B.

Let us write now this down mathematically. We will use this notation to describe the conditional probability of an event A given that some other event B is known to have occurred. We read this expression as probability of A given B. So what are these conditional probabilities in our example? So in the new model, where these outcomes are equally likely, we know that event A can occur in two different ways. Each one of them has probability  $1/6$ . So the probability of event A is  $2/6$  which is the same as  $1/3$ .

How about event B. Well, B consists of 6 possible outcomes each with probability  $1/6$ . So event B in this revised model should have probability equal to 1. Of course, this is just saying the obvious. Given that we already know that B has occurred, the probability that B occurs in this new model should be equal to 1.

How about now, if the sample space does not consist of equally likely outcomes, but instead we're given the probabilities of different pieces of the sample space, as in this example. Notice here that the probabilities are consistent with what was used in the original example. So this part of A that lies outside B has probability  $3/12$ , but in this case I'm not telling you how that probability is made up. I'm not telling you that it consists of 3 equally likely outcomes. So all I'm telling you is that the collective probability in this region is  $3/12$ .

The total probability of A is, again,  $5/12$  as before. The total probability of B is 2 plus 4 equals  $6/12$ , exactly as before. So it's a sort of similar situation as before. How should we revise our probabilities and create-- construct-- conditional probabilities once we are told that event B has occurred?

First, this relation should remain true. Once we are told that B has occurred, then B is certain to occur, so it should have conditional probability equal to 1. How about the conditional probability of A given that B has occurred? Well, we can reason as follows. In the original model, and if we just look inside event B, those outcomes that make event A happen had a collective probability which was  $1/3$  of the total probability assigned to B. So out of the overall probability assigned to B,  $1/3$  of that probability corresponds to outcomes in which event A is happening. So therefore, if I tell you that B has occurred, I should assign probability equal to  $1/3$  that event A is also going to happen. So that, given that B happened, the conditional probability of A given B should be equal to  $1/3$ .

By now, we should be satisfied that this approach is a reasonable way of constructing conditional probabilities. But now let us translate our reasoning into a formula. So we wish to come up with a formula that gives us the conditional probability of an event given another event. The particular formula that captures our way of thinking, as motivated before, is the following. Out of the total probability assigned to B-- which is this-- we ask the question, which fraction of that probability is assigned to outcomes under which event A also happens? So we are living inside event B, but within that event, we look at those outcomes for which event A also happens. So this is the intersection of A and B. And we ask, out of the total probability of B, what fraction of that probability is allocated to that intersection of A with B?

So this formula, this definition, captures our intuition of what we did before to construct conditional probabilities in our particular example. Let us check that the definition indeed does what it's supposed to do. In this example, the probability of the intersection was  $2/12$  and the total probability of B was

$6/12$ , which gives us  $1/3$ , which is the answer that we had gotten intuitively a little earlier.

At this point, let me also make a comment that this definition of conditional probabilities makes sense only if we do not attempt to divide by zero. That this, only if the event  $B$  on which we're conditioning, has positive probability. If  $B$ , if an event  $B$  has 0 probability, then conditional probabilities given  $B$  will be left undefined.

And one final comment. This is a definition. It's not a theorem. What does that mean? It means that there is no question whether this equality is correct or not. It's just a definition. There's no issue of correctness. The earlier argument that we gave was just a motivation of the definition. We tried to figure out what the definition should be if we want to have a certain intuitive and meaningful interpretation of the conditional probabilities. Let us now continue with a simple example.