

By now, we have introduced all sorts of PMFs for the discrete case. The joint PMF, the conditional PMF-- given an event-- and the conditional PMF of one random variable given another. And we're moving along with the program of defining analogous concepts for the continuous case.

We have already discussed the joint PDF and the conditional PDF, given an event. The next item in our menu is to define a conditional PDF of one random variable, given another random variable. We proceed by first looking at the definition for the discrete case.

A typical entry of the conditional PMF is just a conditional probability, but in different notation. And using the definition of conditional probabilities, this is equal to the ratio of the joint divided by the probability of the conditioning event. Unfortunately, in the continuous case, a definition of this form would be problematic, because the event that Y takes on a specific value is an event that has 0 probability.

And we know that we cannot condition on a 0 probability event. However, we can take this expression as a guide on how to define a conditional PDF in the continuous case. And this is the definition, which just mimics the formula that we have up here. Notice that this conditional PDF-- defined this way-- is well defined, as long as the denominator is a positive quantity.

Let us now try to make sense of this definition. Let us first recall the interpretation of the conditional PDF, given an event, A , that has positive probability. We know that the PDF is used to determine the probability of a small interval. And similarly, the conditional PDF is used to calculate the conditional probability of a small interval given the conditioning event.

We would like to do something similar for the conditional PDF, where we would like to take the event A to be something like the event that Y is equal to some particular value-- little y . But as we said, this is problematic, because this event does not have positive probability. So instead, we can take A to be the event that Y is approximately equal to a certain value.

So we're dealing with a little interval around this value, little y , which in general would be an event of positive probability. And we can try to have a similar interpretation. Let us see how this works out. So what does it mean that Y is approximately equal to some particular value, little y ? We interpret that as follows. We're told that the random variable, Y , takes a value that is within epsilon-- where epsilon is a

small number-- of a given value, little y .

And given this conditioning information, we want to calculate the probability of a small interval. How do we do that? Well, here-- because this, in general, will be a positive probability event-- we can use the definition of conditional probabilities. And it would be equal to the probability of both events happening, divided by the probability of the conditioning event.

What is the probability of both events happening? This is a probability of a small rectangle in xy space. At that rectangle, the joint PDF, has a certain value. And because we're integrating over that rectangle-- and that rectangle has dimensions δ and ϵ -- of that probability, that small rectangle, is approximately equal to this.

Then we need the denominator, which is the probability of the conditioning event. And this is approximately equal to the density of Y evaluated at that point, times the length of the small interval. We cancel the ϵ s. And then we notice that the ratio we have here is what we defined as the conditional PDF. So we get this relation times δ .

So what do we see? We see that the probability of a small interval is equal to a PDF times the length of the small interval. However, because we are conditioning on Y being approximately equal to a certain value, we end up using a corresponding conditional PDF, where the conditional PDF is defined this way. So we now have an interpretation of the conditional PDF in terms of probabilities of small intervals.

Now that we have an intuitive interpretation of the conditional PDF, we can also use it to calculate conditional probabilities of more general events, not just intervals. And we do this as follows. In general, for continuous random variables, we can find the probability that X belongs to a certain set by integrating a PDF over that set. Because here we're dealing with a conditional situation where we're given the value of Y , we use the conditional PDF instead of the true PDF. And this way, we calculate the conditional probability.

Now, the difficulty is that this conditional probability is not a well-defined quantity according to what we did early on in this class. We cannot condition on zero probability events. But we can get the around this difficulty as follows. This quantity is well-defined. And we can use this quantity as the definition of this conditional probability.

And so we have managed to provide definition of conditional probabilities, given a 0 probability event of

a certain type. It turns out that this definition is sound and consistent with everything else that we are doing. But when we're dealing with particular problems and applications, we can generally forget about all of these subtleties that we have been discussing here.

The bottom line is that we will be treating conditional PDFs-- given the value of a random variable, Y -- just as ordinary PDFs, but given the information that this random variable took on a specific value. And in that conditional universe, we will calculate probabilities the usual way, by using conditional PDFs instead of ordinary PDFs.