# 6.041SC Probabilistic Systems Analysis and Applied Probability, Fall 2013
# Transcript – Recitation: Joint Probability Mass Function (PMF) Drill 1

Welcome back guys. Today we're going to work on a problem that tests your knowledge of joint PMFs. And we're also going to get some practice computing conditional expectations and conditional variances. So in this problem, we are given a set of points in the xy plane. And we're told that these points are equally likely.

So there's eight of them. And each point has a probability of 1/8 of occurring. And we're also given this list of questions. And we're going to work through them together.

So in part a, we are asked to find the values of x that maximize the conditional expectation of y given x. So jumping right in, this is the quantity we're interested in. And so this quantity is a function of x.

You plug-in various values of x. And then this will spit out a scalar value. And that value will correspond to the conditional expectation of y conditioned on the value of x that you put in. So let's see, when x is equal to 0, for instance, let's figure out what this value is.

Well, when x is equal to 0 we're living in a world, essentially, on this line. So that means that only these two points could have occurred. And in particular, y can only take on the values of 1 and 3.

Now, since all these points in the unconditional universe were equally likely, in the conditional universe they will still be equally likely. So this happens with probability 1/2. And this happens with probability 1/2.

And therefore, the expectation would just be 3/2 plus 1/2 which is 4/2, or 2. But a much faster way of seeing this-- and it's the strategy that I'm going to use for the rest of the problem-- is to remember that expectation acts like center of mass. So the center of mass, when these two points are equally likely, is just the midpoint, which of course is 2.

So we're going to use that intuition on the other ones. So I'm skipping to x is equal to 2 because 1 and 3 are not possible. So when x is equal to 2, y can only take on the values of 1 or 2. Again, they're equally likely. So the center of mass is in the middle which happens at 1.5 or 3/2.

Similarly, x is equal to 4. We're living in this conditional universe, where y can take on of these four points with probability 1/4 each. And so again, we expect the center of mass to be at 1.5 or 3/2. And this quantity is undefined otherwise.

OK, so we're almost done. Now we just need to find which value of x maximizes this. Well, let's see, 2 is the biggest quantity out of all of these numbers. So the maximum is 2. And it occurs when x is equal to 0.

So we come over here. And we found our answer. x is equal to 0 is the value, which maximizes the conditional expectation of y given x.

So part b is very similar to part a. But there is slightly more computation involved. Because now we're dealing with the variance and not an expectation. And variance is usually a little bit tougher to compute.

So we're going to start in the same manner. But I want you guys to see if you can figure out intuitively what the right value is. I'm going to do the entire computation now. And then you can compare whether your intuition matches with the real results.

So variance of x conditioned on a particular value of y, this is now a function of y. For each value of y you plug in you're going to get out a scalar number. And that number represents the conditional variance of x when you condition on the value of y that you plugged in.

So let's see, when y is equal to 0 we have a nice case. If y is equal to 0 we have no freedom about what x is. This is the only point that could have occurred.

Therefore, x definitely takes on a value of 4. And there's no uncertainty left. So in other words, the variance is 0.

Now, if y is equal to 1, x can take on a value of 0, a value of 2 or a value of 4. And these all have the same probability of occurring, of 1/3,

And again, the reasoning behind that is that all eight points were equally likely in the unconditional universe. If you condition on y being equal to 1 these outcomes still have the same relative frequency. Namely, they're still equally likely.

And since there are three of them they now have a probability of 1/3 each. So we're going to go ahead and use a formula that hopefully, you guys remember. So in particular, variance is the expectation of x squared minus the expectation of x all squared, the whole thing squared.

So let's start by computing this number first. So conditioned on y is equal to 1-- so we're in this line-- the expectation of x is just 2, right? The same center-of-mass to argument.

So this, we have a minus 2 squared over here. Now, x squared is only slightly more difficult. With probability 1/3, x squared will take on a value of 0.

With probability 1/3, x squared will take on a value of 4. I'm just doing 2 squared. And with probability 1/3, x squared takes on a value of 4 squared or 16.

So writing down when I just said, we have 0 times 1/3 which is 0. We have 2 squared, which is 4 times 1/3. And then we have 4 squared, which is 16 times 1/3.

And then we have our minus 4 from before. So doing this math out, we get, let's see, 20/3 minus 12/3, which is equal to 8/3, or 8/3. So we'll come back up here and put 8/3.

So I realize I'm going through this pretty quickly. Hopefully this step didn't confuse you. Essentially, when I was doing is, if you think of x squared as a new random variable, x squared, the possible values that it can take on are 0, 4, and 16 when you're conditioning on y is equal to 1. And so I was simply saying that that random variable takes on those values with equal probability. So let's move on to the next one.

So if we condition on y is equal to 2 we're going to do a very similar computation. Oops, I shouldn't have erased that. OK, so we're going to use the same formula that we just used, which is the expectation of x given y is equal to 2. Sorry, x squared minus the expectation of x conditioned on y is equal to 2, all squared.

So conditioned on y is equal to 2, the expectation of x is 3. Same center of mass argument. So 3 squared is 9.

And then x squared can take on a value of 4. Or it can take on a value of 16. And it does so with equal probability.

So we get 4/2, 4 plus 16 over 2. So this is 2 plus 8, which is 10, minus 9. That'll give us 1.

So we get a 1 when y is equal to 2. And last computation and then we're done. I'm still recycling the same formula. But now we're conditioning on y is equal to 3. And then we'll be done with this problem, I promise.

OK, so when y is equal to 3 x can take on the value of 0. Or it can take on the value of 4. Those two points happen with probability 1/2, 1/2. So the expectation is right in the middle which is 2. So we get a minus 4.

And similarly, x squared can take on the value of 0. When x takes on the value of 0-- and that happens with probability 1/2-- similarly, x squared can take on the value of 16 when x takes on the value of 4. And that happens with probability 1/2.

So we just have 0/2 plus 16/2 minus 4. And this gives us 8 minus 4, which is simply 4. So finally, after all that computation, we are done. We have the conditional variance of x given y.

Again, we're interested in when this value is largest. And we see that 4 is the biggest value in this column. And this value occurs when y takes on a value of 3. So our answer, over here, is y is equal to 3.

All right, so now we're going to switch gears in part c and d a little bit. And we're going to be more concerned with PMFs, et cetera. So in part c, we're given a random variable called r which is defined as the minimum of x and y.

So for instance, this is the 0.01. The minimum of 0 and 1 is 0. So r would have a value of 0 here. Now, we can be a little bit smarter about this.

If we plot the line, y is equal to x. So that looks something like this. We see that all of the points below this line satisfy y being less or equal to x. And all the points above this line have y greater than or equal to x.

So if y is less than or equal to x, you hopefully agree that here the min, or r, is equal to y. But over here, the min, r, is actually equal to x, since x is always smaller.

So now we can go ahead quickly. And I'm going to write the value of r next each point using this rule. So here, r is the value of y, which is 1.

Here, r is equal to 0. Here r is 1. Here r is 2. Here r is 3.

Over here, r is the value of x. So r is equal to 0. And r is equal to 0 here.

And so the only point we didn't handle is the one that lies on the line. But in that case it's easy. Because x is equal to 2. And y is equal to 2. So the min is simply 2. So with this information I claim we're now done. We can just write down what the PMF of r is.

So in particular, r takes on a value of 0. When this point happens, this point happens, or this point happens. And those collectively have a probability of 3/8 of occurring.

r can take on a value of 1 when either of these two points happen. So that happens with probability 2/8. r is equal to 2. This can happen in two ways. So we get 2/8. And r equal to 3 can happen in only one way. So we get 1/8.

Quick sanity check, 3 plus 2 is 5, plus 2 is 7, plus 1 is 8. So our PMF sums to 1. And to be complete, we should sketch it. Because the problem asks us to sketch it.

So we're plotting PR of r, 0, 1, 2, 3. So here we get, let's see, 1, 2, 3. For 0 we have 3/8. For 1 we have 2/8. For 2 we have 2/8. And for 3 we have 1/8.

So this is our fully labeled sketch of Pr of r. And forgive me for erasing so quickly, but you guys can pause the video, presumably, if you need more time. Let's move on to part d.

So in part d we're given an event named a, which is the event that x squared is greater than or equal to y. And then we're asked to find the expectation of xy in the unconditional universe. And then the expectation of x times y conditioned on a.

So let's not worry about the conditioning for now. Let's just focus on the unconditional expectation of x times y. So I'm just going to erase all these r's so I don't get confused.

But we're going to follow a very similar strategy, which is at each point I'm going to label what the value of w is. And we'll find the expectation of w that way.

So let's see, here, we have 4 times 0. So w is equal to 0. Here we have 4 times 1.

w is equal to 4. 4 times 2, w is equal to 8. 4 times 3, w is equal to 12.

w is equal to 2. w is equal to 4. w is equal to 0. w is equal to 0.

OK, so that was just algebra. And now, I claim again, we can just write down what the expectation of x times y is. And I'm sorry, I didn't announce my notation. I should mention that now.

I was defining w to be the random variable x times y. And that's why I labeled the product of x times y as w over here. My apologies about not defining that random variable. So the expectation of w, well, w takes on a value of 0. When this happens, this happens or that happens. And we know that those three points occur with probability 3/8.

So we have 0 times 3/8. I'm just using the normal formula for expectation. w takes on a value of 2 with probability 1/8. Because this is the lead point in which it happens, 2 times 1/8.

Plus it can take on the value of 4 with probability 2/8, 4 times 2/8. And 8, with 1/8 probability. And similarly, 12 with 1/8 probability.

So this is just algebra. The numerator sums up to 30. Yes, that's correct. So we have 30/8, which is equal to 15/4. So this is our first answer for part d.

And now we have to do this slightly trickier one, which is the conditional expectation of x times y, or w conditioned on a. So similar to what I did in part c, I'm going to draw the line y equals x squared. So y equals x squared is 0 here, 1 here. And at 2, it should take on a value of 4.

So the curve should look something like this. This is the line y is equal to x squared. So we know all the points below this line satisfy y less than or equal to x squared. And all the points above this line have y greater than or equal to x squared.

And a is y less than or equal to x squared. So we are in the conditional universe where only points below this line can happen. So that one, that one, that one, that one, that one and that one. So there are six of them.

And again, in the unconditional world, all of the points were equally likely. So in the conditional world these six points are still equally likely. So they each happen with probability 1/6.

So in this case, the expectation of w is simply 2 times 1/6 plus 0 times 1/6. But that's 0. So I'm not going to write it.

4 times 2/6 plus 4 times 2/6 plus 8 times 1/6, plus 12 times 1 over 6. And again, the numerator summed to 30. But this time our denominator is 6. So this is simply 5.

So we have, actually, finished the problem. Because we've computed this value and this value. And so the important takeaways of this problem are, essentially, honestly, just to get you comfortable with computing things involving joint PMFs.

We talked a lot about finding expectations quickly by thinking about center of mass and the geometry of the problem. We've got practice computing conditional variances. And we did some derived distributions. And we'll do a lot more of those later.