

6.041SC Probabilistic Systems Analysis and Applied Probability, Fall 2013 Transcript – Recitation: Joint Probability Mass Function (PMF) Drill 2

Hey, guys. Welcome back. Today, we're going to do another fun problem, which is a drill problem on joint PMFs. And the goal is that you will feel more comfortable by the end of this problem, manipulating joint PMFs. And we'll also review some ideas about independents in the process.

So just to go over what I've drawn here, we are given an xy plane. And we're told what the PMF is. And it's plotted for you here.

What these stars indicate is simply that there is a value there. But we don't know what it is. It could be anything between 0 and 1. And so we're given this list of questions. And we're just going to work through them linearly together.

So we start off pretty simply. We want to compute, in part a, the probability that x takes on a value of 1. So for those of you who like formulas, I'm going to use the formula, which is usually referred to as marginalization.

So the marginal over x is given by summing over the joint. So here we are interested in the probability that x is 1. So I'm just going to freeze the value of 1 here. And we sum over y .

And in particular, 1, 2, and 3. So carrying this out, this is the P_{xy} of 1, 1, plus P_{xy} of 1, 2, plus P_{xy} 1, 3. And this, of course, reading from the graph, is $1/12$ plus $2/12$ plus $1/12$, which is equal to $4/12$, or $1/3$.

So now you guys know the formula. Hopefully you'll remember the term marginalization. But I want to point out that intuitively you can come up with the answer much faster.

So the probability that x is equal to 1 is the probability that this dot happens or this dot happens or this dot happens. Now, these dots, or outcomes, they're disjoint. So you can just sum the probability to get the probability of one of these things happening.

So it's the same computation. And you'll probably get there a little bit faster. So we're done with a already, which is great.

So for part b, conditioning on x is equal to 1, we want to sketch the PMF of y . So if x is equal to 1 we are suddenly living in this universe. y can take values of 1, 2, or 3 with these relative frequencies.

So let's draw this here. So this is y . I said, already, y can take on a value of 1. y can take on a value of 2. Or it can take on a value of 3. And we're plotting here, P_y given x , y , conditioned on x is equal to 1.

OK, so what I mean by preserving the relative frequencies is that in unconditional world this is dot is twice as likely to happen as either this dot or this dot. And that relative likelihood remains the same after conditioning. And the reason why we have to change these values is because they have to sum to 1. So in other words, we have to scale them up.

So you can use a formula. But again, I'm here to show you faster ways of thinking about it. So my little algorithm for figuring out conditional PMFs is to take the numerators-- so 1, 2, and 1-- and sum them. So here that gives us 4.

And then to preserve the relative frequency, you actually keep the same numerators but divide it by the sum, which you just computed. So I'm going fast. I'll review in a second. But this is what you will end up getting.

So to recap, I did 1 plus 2 plus 1, which is 4, to get these denominators. And so I skipped a step here. This is really $2/4$, which is $1/2$, obviously.

So you add these guys to get 4. And then you keep the numerators and just divide them by 4. So $1/4$, $2/4$, which is $1/2$ and $1/4$. And that's what we mean by preserving the relative frequency. Except so this thing now sums to 1, which is what we want.

OK, so we're done with part b. Part c actually follows almost immediately from part b. In part c we're interested in computing the conditional expectation of y given that x is equal to 1.

So we've already done most of the legwork because we have the conditional PMF that we need. And so expectation, you guys have calculated a bunch of these by now. So I'm just going to appeal to your intuition and to symmetry.

Expectation acts like center of mass. This is a symmetrical distribution of mass. And so the center is right here at 2. So this is simply 2. And if that went too fast, just convince yourselves. Use the normal formula for expectations. And your answer will agree with ours.

OK, so d is a really cool question. Because you can do a lot of math. Or you can think and ask yourself, at the most fundamental level, what is independent? And if you think that way you'll come to the answer very easily.

So essentially, I rephrased this to truncate it from the problem statement that you guys are reading. But the idea is that these stars are unknown probability masses. And this question is asking can you figure out a way of assigning numbers between 0 and 1 to these values such that you end up with a valid probability mass function, so everything sums to 1 and such that x and y are independent?

So it seems hard a priori. But let's think about it a bit. And in the meantime I'm going to erase this so I have more room. What does it mean for x and y to be independent?

Well, it means that they don't, essentially, have information about each other. So if I tell you something about x and if x and y are independent, your belief about y shouldn't change. In other words, if you're a rational person, x shouldn't change your belief about y .

So let's look more closely at this diagram. Now, the number 0 should be popping out to you. Because this essentially means that the 0.31 can't happen. Or it happens with 0 probability.

So let's say fix x equal to 3. If you condition on x is equal to 3, as I just said, this outcome can't happen. So y could only take on values of 2 or 3.

However, if you condition on x is equal to 1, y could take on a value of 1 with probability $1/4$ as we computed in the other problem. It could take on a value of 2 with probability of $1/2$. Or it could take on a value of 3 with probability $1/4$.

So these are actually very different cases, right? Because if you observe x is equal to 3 y can only be 2 or 3. But if you observe x is equal to 1, y can be 1, 2, or 3.

So actually, x , no matter what values these stars have on, x always tells you something about y . Therefore, the answer to this, part d, is no. So let's put a no with an exclamation point. So I like that problem a lot. And hopefully it clarifies independents for you guys.

So parts e and f, we're going to be thinking about independents again. To go over what the problem statement gives you, we defined this event, b , which is the event that x is less than or equal to 2 and y is less than or equal to 2.

So let's get some colors. So do bright pink. So that means we're essentially living in this world. There's only those four dots. And we're also told a very important piece of information that conditions on B . x and y are conditionally independent.

OK, so part e, now that we have this. And by the way, these two assumptions apply to both parts e and part f. So in part e, we want to find out P_{xy} of 2, 2. Or in English, what is the probability that x takes on a value of 2 and y takes on a value of 2?

So determine the value of this star. And the whole trick here is that the possible values that this star could take on are constrained by the fact that we need to make sure that x and y are conditionally independent given B .

So my claim is that if two random variables are independent and you condition on one of them, say we condition on x . If you condition on different values of x , the relative frequencies of y should be the same. So here, the relative frequency, condition on x is equal to 1. The relative frequencies of y are 2 to 1.

This outcome is twice as likely to happen as this one. If we condition on 2 this outcome needs to be twice as likely to happen as this outcome. If they weren't, x would tell you information about y . Because you would know that the distribution over 2 and 1 would be different. OK?

So because the relative frequencies have to be the same and $2/12$ is 2 times $1/12$ this guy must also be 2 times $2/12$. So that gives us our answer for part e. Let me write up here. Part e, we need $P_{xy} 2, 2$ to be equal to $4/12$.

And again, the way we got this is simply we need x and y to be conditionally independent given B . And if this were anything other than 4 the relative frequency of y is equal to 2 to 1 would be different from over here. So here condition on x is equal to 1. The outcome, y is equal to 2 is twice as likely as x is equal to 1.

Here, if we put a value of $4/12$ and you condition on x is equal to 2, the outcome y is equal to 2 is still twice as likely as the outcome y is equal to 1. And if you put any other number there the relative frequencies would be different. So x would be telling you something about y .

So there would not be independent condition on B . OK, that was a mouthful. But hopefully you guys have it now.

And lastly, we have part f, which follows pretty directly from part e. So we were still in the unconditional universe. In part e, we were figuring out what's the value of star in the whole unconditional universe?

Now, in part f, we want the value of star in the conditional universe where B occurred. So let's come over here and plot a new graph so we don't confuse ourselves. So we have xy .

x can be 1 or 2. Y could be 1 or 2. So we have a plot that looks something like this.

And so again, same argument as before. Let me just fill this in. From part e, we have that this is $4/12$. And we're going to use my algorithm again.

So in the conditional world, the relative frequencies of these four dots should be the same. But you need to scale them up so that if you sum over all of them the probability sums to 1. So you have a valid PMF.

So my algorithm from before was to add up all the numerators. So 1 plus 2 plus 4 plus 2 gives you 9. And then to preserve the relative frequency you keep the same numerator.

So here we had a numerator of 1. That becomes $1/9$. Here we had a numerator of 2.

This becomes $2/9$. Here we had a numerator of 4. That becomes $4/9$. Here we had a numerator of 2, so $2/9$.

And indeed, the relative frequencies are preserved. And they all sum to 1. So our answer for part f-- let's box it here-- is that $P_{X,Y}(2, 2)$ conditioned on B is equal to $4/9$, is just that guy. So we're done.

Hopefully that wasn't too painful. And this is a good drill problem, because we got more comfortable working with PMFs, joint PMFs. We went over marginalization. We went over conditioning.

We went over into independents. And I also gave you this quick algorithm for figuring out what conditional PMFs are if you don't want to use the formulas. Namely, you sum all of the numerators to get a new denominator and then divide all the old numerators by the new denominator you computed. So I hope that was helpful. I'll see you next time.

MIT OpenCourseWare
<http://ocw.mit.edu>

6.041SC Probabilistic Systems Analysis and Applied Probability
Fall 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.