6.041SC Probabilistic Systems Analysis and Applied Probability, Fall 2013 Transcript – Recitation: Inferring a Discrete Random Variable from a Continuous Measurement

Hi. In this problem, we're going to look at how to infer a discrete random variable from a continuous measurement. And really, what it's going to give us is some practice working with a variation of Bayes' rule. So the problem tells us that we have a discrete random variable x with this PMF. It is 1 with probability P, minus 1 with probability 1 minus P, and 0 otherwise. So here is just a diagram of this PMF.

And then we also have another random variable, y, which is continuous. And its PDF is given by this. It's 1/2 lambda e to the minus lambda times the absolute value of y. And so this may look familiar. It looks kind of like an exponential. And in fact, it's just a two-sided exponential. That's flattened by a factor of 1/2. And this is what it looks like, kind of like a tent that goes on both ways.

And then we have a random variable z, which is equal to the sum of x and y. And the problem is going to be figuring out what x is, based on an observed value of what z is. So because x is discrete and y is random-- sorry, x is discrete, and y is continuous, z is also going to be continuous. So our measurement is z, which is continuous. And we want to infer x, which is discrete.

So the problem asked us to find is this what is the probability that x equals 1, given that z is a little z. And you can write this another way, just as a conditional PMF as well. It's the conditional PMF of x, given z, evaluated to 1 conditioned on little z.

All right, so now let's apply the correct variation of Bayes' rule. So remember, it's going to be this, the probability that x equals 1, or the PMF of x evaluated to 1, times the-- you flip this conditioning. So now it's going to be a conditional PDF of z, since z is continuous. It's going to be a conditional PDF of z, given x, evaluated at some little z condition on x being 1.

And the bottom is the conditional PDF-- or sorry, just the regular PDF of z. And of course, we can rewrite this denominator. Remember, the denominator is always just-- you can use the law of total probability and rewrite it. And one of the terms is going to be exactly the same as the numerator.

So one of the ways that z can be some little z is it's in combination with x being equal to 1. And the probability of that is exactly the same thing as the numerator. And the other way is if x is equal to negative 1. And that gives us this second term.

All right. And now let's just fill in what all these different terms are. So with the PMF of x evaluated at 1, that is just P. What is the conditional PDF of z, given that x is equal to 1? Well, that takes a little bit more work.

Given that x is 1, then z is just going to be-- so if x equals 1, then z is just y plus 1, which means that you can just imagine taking y-- this is what y is, the distribution of y-- and just adding 1 to it, which, in this diagram, would amount to shifting it over by one. So now, it would look like this, the distribution. And algebraically, all you would do is just change this thing in the absolute value to y minus 1. That amounts to shifting it over to the right by one.

All right. So what is that? That's just 1/2 lambda, 1/2 lambda, e to the minus lambda y-- sorry, not y, z-- z minus 1. And the denominator, well, the first term is going to be exactly the same. It's just also P 1/2 lambda e to the minus lambda z minus 1.

What about the second term? The second term, first we need to figure out what is the PMF of x evaluated at a negative 1. Or in other words, what's the probability that x is negative 1? That is given to us by the PMF. It's 1 minus P.

And then the second part is, what is the conditional PDF of z, given that x is negative 1? Well, we can just do the same sort of trick here. If x is negative 1, then z is just y minus 1.

In which case, the PDF of z would just look like this. You're shifted to the left by one now. And now what you have to do is change this into a plus 1. So this conditional PDF would be 1/2 lambda e to the minus lambda z plus 1, absolute value of z plus 1.

All right, so this looks pretty messy. And we can try to simplify things a little bit. So we can get rid of these 1/2 lambdas. And then we can multiply the numerator and the denominator by the same term. Let's multiply it by e to the lambda absolute value of z minus 1.

So what we're going to do is try to cancel out some of these exponential terms. So that will cancel out this top term. So all we have in the numerator now is just P. It will also cancel out this exponential in the denominator. And then we'll have to change this here, because it'll have an extra e to the lambda absolute value of z minus 1.

All right, now let's rewrite this. And what we get is plus 1 minus P e to the minus lambda absolute value of z plus 1 minus absolute value of z minus 1. OK, so that is pretty much as far as you can go in terms of simplifying it.

And now the question is, are we comfortable with this answer? And it helps always to try to interpret it a little bit, to make sure that it makes intuitive sense. And one way to do that is to try to-- some of the limiting cases of what some of the parameters can be.

So in this case, the parameters are P and lambda. So P is the parameter related to x. And lambda is the parameter related to y. So let's try to see if it makes sense under some limiting cases.

The first one we want to think about is when P goes to 0. So if P goes to 0, what happens to our answer? Well, the numerator is 0, this is 0, this is 1. But it doesn't matter, because the numerator is 0. So in this case, this would go to 0.

Now does that make sense? Well, what does that mean when P goes to 0? When P goes to 0, that means that the probability that x is equal to 1 is 0. So even without thinking about y or z, there is already a 0 probability that x is equal to 1.

Now this whole calculation, what we found is, well, if I had some more information, like what z is, does that help me find out what the probability of x being 1 is? Well, no matter what z tells me, I know for a fact that x can't be 1, because P is 0. So this posterior, or this conditional probability, should also be 0, because there's just no way that x can be 1. So in this case, this formula does check out.

Now let's think about another case where P goes to 1. If P goes to 1, that means that X is for sure going to be 1. And it can't be anything else. In which case, what does our formula tell us?

Well, this numerator is 1. This term is 1. 1 minus 1 is 0. So the second term gets zeroed out, and the answer is just 1/1 is 1.

So what does this tell us? This tells us that if I know beforehand that x is for sure equal to 1, then, if I now give myself more information and condition on what I observe for z, that shouldn't change anything else. I should still know for sure that x is equal to 1. So the probability of this conditional probability should still be equal to 1. And it does, so our formula also works in this case.

Now let's think about lambda. What about when lambda goes to 0? Well, when lambda goes to 0, that's a little harder to visualize. But really, what would happen is that you can imagine this distribution getting shallower, shallower and shallower, lower and lower, so that it's like it is kind of flat and goes on forever.

And so what this tells you is that, basically, y-- this is the distribution y-- so when lambda goes to 0, that tells you that y has a really flat and kind of short distribution. And so what does our formula tell us in this case? Well, when lambda goes to 0, this exponent is equal to 0.

And so e to the 0 is 1. So we get P over P plus 1 minus P, which is just 1. So the answer here, our formula will give us an answer of P.

So what does that tell us? That tells us that, in this case, if lambda goes to 0, then our posterior probability, the probability that x equals 1 conditioned on z being some value, conditioned on our continuous measurement, is still P. So the prior, or the original probability for x being equal to 1 is P. And with this additional continuous measurement, our guess of the probability that x equal to 1 is still P. So it hasn't changed.

So basically, it's telling us that this additional information was not informative. It didn't actually help us change our beliefs. And so why is that?

Well, one way to think about it is that, because the distribution of y looks like this, is very flat and it could be anything, then, if you observe some value of z, then it could be that that was due to the fact that it was x equal to 1 plus some value of y that made z equal to that value. Or it could have just as equally been likely that x equal to negative 1 y equals to some other value that made it equal to z.

And so, essentially, it's z-- because y has a shape, it can be likely to take on any value that complements either x being equal to 1 or x equal being to negative 1, to make z equal to whatever the value it is that you observe. And so because of that, in this case, y is not very informative. And so this probability is still just equal to P.

Now the last case is when lambda goes to infinity. And now we have to break it down into the two other cases now. The first case is when-- lets write this over here-- when lambda goes to infinity.

The first case, it depends on what this value is, the sine of this value. If this value, the absolute value of z plus 1 minus the absolute value of z minus 1, if that's positive, then, because lambda goes to infinity and you have a negative sign, then this entire exponential term will go to 0. In which case, the second term goes to 0. And the answer is P/P, or is 1. And so if absolute value of z plus 1 minus absolute value of z minus 1 is greater than 0, then the answer is 1.

But in the other case, if this term in the exponent, if it's actually negative, if it's negative, then this negative sign turns to a positive, and lambda goes to infinity. And so this term blows up, and it dominates everything else. And so the denominator goes to infinity. The numerator is fixed at P, so this entire expression would go to 0.

OK, so now let's try to interpret this case. Let's start with the first one. When is it that absolute value of z plus 1 minus absolute value of z minus 1 is greater than 0? Or you can also rewrite this as absolute value of z plus 1 is greater than absolute value of z minus 1. Well, when is that case?

Well, it turns out, if you think about it, this is only true if z is positive. If z is positive, then adding 1-- let me draw a line here, and if this is 0-- if z is positive, something here, adding 1 to it and taking the absolute value-- the absolute value doesn't do anything-- but you will get something bigger. Where subtracting 1 will take you closer to 0, and so because of that, the absolute value, the magnitude, or the distance from 0 will be less.

Now if you're on the other side, adding 1 will take you-- if you're on the other side, adding 1 will take you closer to 0. And so this magnitude would be smaller. Whereas, subtracting will take you farther away, so the absolute value actually increased the magnitude. And so this is the same as z being positive. And so this is the same as z being negative.

So what this tells you is that, if z is positive, then this probability is equal to 1. And if z is negative, this probability is equal to 0. Now why does that make sense? Well, it's because when lambda goes to infinity, you have the other case. Essentially, you pull this all the way up, really, really far, and it drops off really quickly.

And so when you take the limit, as lambda goes to infinity, effectively, it just becomes a spike at 0. And so, more or less, you're sure that y is going to be equal to 0. And so, effectively, z is actually going to be equal to x, effectively.

And because of that, because x can only be 1 or negative 1, then, depending on if you get a z that's positive, then you know for sure that it must have been that x was equal to 1. And if you get a z that's negative, you know for sure that it must have been that x was equal to negative 1. And so because of that, you get this interpretation.

And so we've looked at four different cases of the parameters. And in all four cases, our answer seems to make sense. And so we feel more confident in the answer.

And so to summarize, this whole problem involved using Bayes' rule. You start out with some distributions, and you apply Bayes' rule. And you go through the steps, and you plug-in the right terms. And then, in the end, it's always helpful to try to check your answers to make sure that it makes sense in some of the limiting.