6.041SC Probabilistic Systems Analysis and Applied Probability, Fall 2013
Transcript – Recitation: Convergence in Probability and in the Mean Part 1

In this exercise, we'll be working with the notion of convergence in probability, as well as some other notion of converge of random variables that we'll introduce later. First type of random variable is xn, where xn has probability 1 minus 1 minus over n to be as 0 and probability of 1 over n to be a 1. And graphically, we see that we have a pretty big mess. 1 minus 1 over n at location 0, and a tiny bit somewhere here, only 1 over n. So this will be the PMF for x.

On the other hand, we have the sequence of random variables, yn. Fairly similar to xn with a slight tweak. The similar part says it also has a very high probability of being at 0, mass 1 over 1 minus n. But on the off chance that yn is not at 0, it has a pretty big value n. So it has probability 1 over n of somewhere out there. So to contrast the two graphs, we see at 0, they have the same amount of mass, 1 over 1 minus n, but for y, it's all the way out there that has a small mass 1 over n. So this will be our Pyn of y.

And for the remainder of the problem, we'll be looking at the regime where the number n tends to infinity, and study what will happen to these two sequences of random variables. In part A, we're to compute the expected value n variance for both xn and yn. Let's get started.

The expected value of xn is given by the probability-- it's at one, which is 1 over n times 1 plus the probability being at 0, 1 over n times value 0. And that gives us 1 over n. To calculate the variance of xn, see that variance is simply the expected value of xn minus expected value of xn, which in this case is 1 over n from the previous calculation we have here.

We take the square of this value and compute the whole expectation, and this gives us 1 over n, 1 minus 1 over n plus the remainder probability 1 over 1 minus n of x being at 0, so 0 minus 1 over n squared. And if we carry out the calculations here, we'll get n minus 1 over n squared.

Now, let's turn to yn. The expected value of yn is equal to probability of being at 0, 0 plus the probability of being at n and times the value n. And this gives us 1. The variance of yn. We do the same thing as before, we have 1 minus 1 over n probability of being at 0, multiplied 0 minus 1 squared, where 1 is the expected value of y. And with probability 1 over n, out there, equal to n, and this is multiplied by n minus 1 squared. And this gives us n minus 1

Already, we can see that while the expected value for x was 1 over n, the expected value for y is sitting right at 1. It does not decrease as it increases. And also, while the variance for x is n minus 1 over n squared, the variance for y is much bigger. It is actually increasing to infinity as n goes infinity. So these intuitions will be helpful for the remainder of the problem.

In part B, we're asked to use Chebyshev's Inequality and see whether xn or yn converges to any number in probability. Let's first recall what the inequality is about. It says that if we have random variable x, in our case, xn, then the probability of xn minus the expected value of xn, in our case, 1 over n, that this deviation, the absolute value of this difference is greater than epsilon is bounded above by the variance of xn divided by the value of epsilon squared.

Well, in our case, we know the variance is n minus 1 over n squared, hence this whole term is this term divided by epsilon squared. Now, we know that as n gets really big, the probability of xn being at 0 is very big. It's 1 minus 1 over n. So a safe bet to guess is that if xn work to converge anywhere on the real line, it might just converge to the point 0. And let's see if that is true.

Now, to show that xn converges to 0 in probability, formally we need to show that for every fixed epsilon greater than 0, the probability that xn minus 0 greater than epsilon has to be 0, and the limit has n going to infinity. And hopefully, the inequalities above will help us achieve this goal. And let's see how that is done.

I would like to have an estimate, in fact, an upper bound of the probability xn absolute value greater or equal to epsilon. And now, we're going to do some massaging to this equation so that it looks like what we know before, which is right here. Now, we see that this equation is in fact, less than probability xn minus 1 over n greater or equal to epsilon plus 1 over n.

Now, I will justify this inequality in one second. But suppose that you believe me for this inequality, we can simply plug-in the value right here, namely substituting epsilon plus 1 over n, in the place of epsilon right here and use the Chebyshev Inequality we did earlier to arrive at the following inequality, which is n minus 1 over n squared times, instead of epsilon, now we have epsilon plus 1 over n squared.

Now, if we take n to infinity in this equation, see what happens. Well, this term here converges to 0 because n squared is much bigger than n minus 1. And this term here converges to number 1 over epsilon squared. So it becomes 0 times 1 over epsilon squared, hence the whole term converges to 0. And this proves that indeed, the limit of the term here as n going to infinity is equal to 0, and that implies xn converges to 0 in probability.

Now, there is the one thing I did not justify in the process, which is why is probability of absolute value xn greater than epsilon less then the term right here? So let's take a look. Well, the easiest way to see this is to see what ranges of xn are we talking about in each case.

Well, in the first case, we're looking at interval around 0 plus minus epsilon and xn can lie anywhere here. While in the second case, right here, we can see that the set of range values for xn is precisely this interval here, which was the same as before, but now, we actually have less on this side, where the starting point and the interval on the right is epsilon plus 2 over n. And therefore, the right hand style captures strictly less values of xn than the left hand side, hence the inequality is true.

Now, we wonder if we can use the same trick, Chebyshev Inequality, to derive the result for yn as well. Let's take a look. The probability of yn minus it's mean, 1, greater or equal to epsilon. From the Chebyshev Inequality, we have variance of yn divided by epsilon squared.

Now, there is a problem. The variance of yn is very big. In fact, it is equal to n minus 1. And we calculated in part A, divided by epsilon squared. And this quantity here diverges as n going to infinity to infinity itself. So in this case, the Chebyshev Inequality does not tell us much information of whether yn converges or not.

Now, going to part C, the question is although we don't know anything about yn from just the Chebyshev Inequality, does yn converge to anything at all? Well, it turns out it does. In fact, we don't have to go through anything more complicated than distribution yn itself.

So from the distribution yn, we know that absolute value of yn greater or equal to epsilon is equal to 1 over n whenever epsilon is less than n. And this is true because we know yn has a lot of mass at 0 and a tiny bit a mass at value 1 over n at location n.

So if we draw the cutoff here at epsilon, then the probability of yn landing to the right of epsilon is simply equal to 1 over n. And this tells us, if we take the limit of n going to infinity and measure the probability that yn-- just to write it clearly-- deviates from 0 by more than epsilon, this is equal to the limit as n going to infinity of 1 over n. And that is equal to 0. From this calculation, we know that yn does converge to 0 in probability as n going to infinity.

For part D, we'd like to know whether the convergence in probability implies the convergence in expectation. That is, if we have a sequence of random variables, let's call it zn, that converges to number c in probability as n going to infinity, does it also imply that the limit as n going to infinity of the expected value of zn also converges to c. Is that true?

Well, intuitively it is true, because in the limit, zn almost looks like it concentrates on c solely, hence we might expect that the expected value is also going to c itself. Well, unfortunately, that is not quite true. In fact, we have the proof right here by looking at yn. For yn, we know that the expected value of yn is equal to 1 for all n. It does not matter how big n gets. But we also know front part C that yn does converge to 0 in probability.

And this means somehow, yn can get very close to 0, yet it's expected value still stays one away. And the reason again, we go back to the way yn was constructed. Now, as n goes to infinity, the probability of yn being at 0, 1 minus 1 over n, approaches 1.

So it's very likely that yn is having a value 0, but whenever on the off chance that yn takes a value other than 0, it's a huge number. It is n, even though it has a small probability of 1 over n. Adding these two factors together, it tells us the expected value of yn always stays at 1.

And however, in probability, it's very likely that y is around 0. So this example tells us converges in probability is not that strong. That tells us something about the random variables but it does not tell us whether the mean value of the random variables converge to the same number.

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.