

6.041SC Probabilistic Systems Analysis and Applied Probability, Fall 2013 Transcript – Recitation: Probability Bounds

In this problem, we're given a collection of 10 variables, x_1 through x_{10} , where each x_i is a uniform random variable between 0 and 1. So each x_i is uniform between 0 and 1, and all 10 variables are independent. And we'd like to develop a bound on the probability that some of the 10 variables, 1 to 10, being greater than 7 using different methods. So in part A we'll be using the Markov's inequality written here. That is, if we have a random variable, positive random variable x , the probability x is greater than a , where a is again some positive number, is bounded above by the expected value of x divided by a .

And let's see how that works out in our situation. In our situation, we will call x the summation of x_i equal to 1 to 10, and therefore, E of x is simply 10 times E of x_1 , the individual ones, and this gives us 5. Here we use the linearity of expectation such that the expectation of the sum of the random variable is simply the sum of the expectations. Now, we can invoke Markov's Inequality. It says $P(x \geq 7)$ is less than $E(x) / 7$, and this gives us $5 / 7$.

For part B, let's see if we can improve the bound we got in part A using the Chebyshev inequality, which takes into account the variance of random variable x . Again, to refresh you on this, the Chebyshev Inequality says the probability that x deviates from its mean $E(x)$, by more than a is bound above by the variance of x divided by a^2 .

So we have to actually do some work to transform the probability we're interested in, which is $x \geq 7$, into the form that's convenient to use using the Chebyshev Inequality. To do so, we'll rewrite this probability as the probability of $x - 5 \geq 2$ simply by moving 5 from the right to the left. The reason we chose 5 is because 5 is equal to the expected value of x from part A as we know before.

And in fact, this quantity is also equal to the probability that $x - 5 \leq -2$. To see why this is true, recall that x is simply the summation of the x_i 's, the 10 x_i 's, and each x_i is a uniform random variable between 0 and 1. And therefore, each x_i , the distribution of which is symmetric around its mean $1/2$. So we can see that after we add up all the x_i 's, the resulting distribution x is also symmetric around its mean 5. And as a result, the probability of $x - 5 \geq 2$ is now equal to the probability that $x - 5 \leq -2$.

And knowing these two, we can then say they're both equal to $1/2$ the probability $|x - 5| \geq 2$, because this term right here is simply the sum of both terms here and here. At this point, we have transformed the probability of $x \geq 7$ into the form right here, such that we can apply the Chebyshev's Inequality basically directly.

And we'll write the probably here being less than or equal to $1/2$ times, applying the Chebyshev Inequality, variance of x divided by 2^2 . Now, 2 is the same as a right here, and this gives us $1/8$ times-- now, the variance of x , we know is 10 times the variance of a uniform random variable between 0 and 1, which is $1/12$, and that gives us $5/48$.

Now, let's compare this with the number we got earlier using the Markov Inequality, which was $5/7$. We see that $5/48$ is much smaller, and this tells us that, at least for this example, using the Chebyshev Inequality combined with the information of the variance of x , we're able to get a stronger upper bound on the probability of the event that we're interested in.

Now, in part B, we saw that by using the additional information of the variance combined with the Chebyshev Inequality, we can improve upon bound given by Markov's Inequality. Now, in part C, we'll use a somewhat more powerful approach in addition to the Chebyshev Inequality, the so-called central limit theorem. Let's see if we can even get a better bound.

To remind you what a central limit theorem is, let's say we have a summation of i equal to 1 to some number n of independent and identically distributed random variables x_i . Now, the central limit theorem says the following. We take the sum right here, and subtract out its means, which is E of the same summation, and further, we'll divide out, what we call normalize, by the standard deviation of the summation. In other words, the square root of the variance of the sum of x_i .

So if we perform this procedure right here, then as the number of terms in the sums going to infinity, here as n goes to infinity, we will actually see that this random variable will converge in distribution in some way that will eventually look like a standard normal random variable with means 0 and 1. And since we know how the distribution of a standard normal looks like, we can go to table and look up certain properties of the resulting distribution. So that is a plan to do. So right now, we have about 10 variables. It's not that many compared to a huge numbering, but again, if we believe it's a good approximation, we can get some information out of it by using the central limit theorem.

So we are interesting knowing that probability summation of i equal to 1 to 10 x_i greater or equal to 7. We'll rewrite this as 1 minus the probability the summation i equal to 1 to 10, and x_i less equal to 7. Now, we're going to apply the scaling to the summation right here. So this is equal to 1 minus the probability summation i equal to 1 to 10 x_i minus 5. Because we know from previous parts that 5 is the expected value of the sum right here, and divided by square root of $10/12$.

Again, earlier we know that $10/12$ is the variance of the sum of x_i 's. And we'll do the same on the other side, writing it 7 minus 5 divided by square root of $10/12$. Now, if we compute out the quantity right here, we know that this quantity is roughly 2.19, and by the central limit theorem, if we believe 10 is a large enough number, then this will be roughly equal to 1 minus the CDF of a standard normal evaluated at 2.19. And we could look up the number in the table, and this gives us number roughly, 0.014.

Now let's do a quick summary of what this problem is about. We're asked to compute the probability of x greater or equal to 7, where x is the sum of 10 uniform random variables between 0 and 1, so we'll call it x_i . We know that because each random variable has expectation $1/2$, adding 10 of them up, gives us expectation of 5. So this is essentially asking, what is the chance that x is more than two away from its expectation?

So if this is a real line, and 5 is here, maybe x has some distribution around 5, so the center what the expected value is at 5, we wonder how likely is it for us to see something greater than 7? Now, let's see where do we land on the probably spectrum from 0 to 1. Well, without using any information, we know the probability cannot be greater than 1, so a trivial upper bound for the probability right here will be 1.

Well, for the first part we use Markov's Inequality and that gives us some number, which is roughly equal to 0.7. In fact, we got number $5/7$, and this is from Markov's Inequality. Oh, it's better than 1, already telling us it cannot be between 0.7 and 1, but can we do better?

Well, the part B, we see that all the way, using the additional information variance, we can get this number down to $5/48$, which is roughly 0.1. Already, that's much better than 0.7. Can we even do better? And this is the Chebyshev, and it turns out we can indeed do better. Using the central limit theorem, we can squeeze this number all the way down to 0.014, almost a 10 times improvement over the previous number. This is from central limit theorem.

As we can see, by using different bounding techniques, we can progressively improve the bound on the probability of x exceeding 7, and from this problem we learned that even with 10 variables, the truth is more like this, which says that the distribution of x concentrates very heavily around 5, and hence, the probability of x being greater or equal to 7 could be much smaller than one might expect.

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>.