

6.041SC Probabilistic Systems Analysis and Applied Probability, Fall 2013 Transcript – Recitation: Using the Central Limit Theorem

Hi. In this video, we're going to do some approximate calculations using the central limit theorem.

We're given that X_n is the number of gadgets produced on day n by a factory. And it has a normal distribution with mean 5 and variance 9. And they're all independent and identically distributed.

We're looking for the probability that the total number of gadgets in 100 days is less than 440.

To start, we can first write this as the probability of the sum of the gadgets produced on each of 100 days being less than 440. Notice that this is a sum of a large number of independent random variables. So we can use the central limit theorem and approximate the sum as a normal random variable.

And then, basically, in order to compute this probability, we'd basically need to standardize this and then use the standard normal table.

So let's first compute the expectation and variance of the sum. So I'm going to actually sum up from 1 to n instead of 100, to do it more generally. So the linearity is preserved for the expectation operator. So this is the sum of the expected value. And since they're all identically distributed, they all have the same expectation, and there are n of them. And so we have this being n times 5.

For the variance of the sum is also the sum of the variances because the independents. And so they're identically distributed to the -- so we have n times the variance of X_i , and this is n times 9.

So now, we can standardize it, or make it 0 mean and variance 1. So to do that we would take these X_i 's, subtract by their mean. So it's going to be 5 times 100 of them, so it's 500 over the square root of the variance, which is going to be 9 times 100 of them, so it's going to be 900. So that's going to be less than 440 minus 500 over square root of 900.

So notice what we're trying to do here is-- notice that the sum of X_i 's is a discrete quantity. So it's a discrete random variable, so it may have a PMF like this. And we're trying to approximate it with a normal density. So this is not drawn to scale, but let's say that this is 440 and this is 439. Basically, we're trying to say what's the probability of this being less than 440, so it's the probability that it's 439, or 438, or 437.

But in the continuous case, a good approximation to this would be to take the middle, say, 439.5, and compute the area below that.

So in this case, when we do the normal approximation, it works out better if we use this half correction. And so, this, in this case, probability, let's call Z the standard normal. And so this is approximately equal to a standard normal with the probability of standard normal being less than whatever that is. And if you plug that into your calculator, you get negative 2.02.

So now, if we try to figure out what this-- from the table, we'll find that negative values are not tabulated. But we know that the normal, the center of normal is symmetric, and so if we want to compute the area in this region, it's the same as the area in this region, above 2.02.

So this is the same as the probability that Z is bigger than 2.02. That's just 1 minus the probability that Z is less than or equal to 2.02, and so that's, by definition, Φ of 2.02. And if we look it up on the table, 2.02 has probability here of 0.9783. And we can just write that in. That's the answer for Part A.

So now for Part B.

We're asked what's the largest n , approximately, so that it satisfies this.

So again, we can use the central limit theorem. Use similar steps here so that we have, in this case, n greater than or equal to 200 plus $5n$. And standardized. So we have n and the mean here-- this is where this comes handy. It's going to be $5n$ and the variance is $9n$. It's greater than or equal to $5n$'s will cancel and you subtract. And then you get 200 over the square root of $9n$.

And we can, again, use the half approximation here, half correction here. But I'm not going to do it, to keep the problem simple. And so in this case, this is approximately equal to the standard normal being greater than probability of the center of normal being greater than or equal to 200 over square root of $9n$.

And so same sort of thing here. This is just 1 minus this. The equal sign doesn't matter because Z is a continuous random variable. And so we have this here.

And we want this to be less than or equal 0.05. So that means that Φ of 200 over square root of 9 has to be greater than or equal to 0.95. So we're basically looking for something here that ensures that this region's at least 0.95.

So if you look at the table, 0.95 lies somewhere in between 1.64 and 1.65. And I'm going to use 1.65 to be conservative, because we want this region to be at least 0.95. So 1.65 works better here.

And so we want this thing, this here, which is going to be 200 over square root of n -- square root of $9n$, to be bigger than or equal to 1.65. So n here is going to be less than or equal to 200 over 1.65 squared, 1 over 9.

If you plug this into your calculator, you might have a decimal in there. Then we just pick n , the largest integer that satisfies this. So we can plug that into your calculator, you'll find that it's going to be 1,632. That's part B.

Last part.

Let n be the first day when the total number of gadgets is greater than 1,000. What's the probability that n is greater than or equal to 220?

Again, we want to use the central limit theorem, but the trick here is to recognize that this is actually equal to the probability that the sum from i equals 1 to 219 of X_i , is less than or equal to 1,000.

So let's look at both directions to check this.

If n is greater than or equal to 220, then this has to be true. Because if it weren't true, and if this were greater than 1,000, then n would have been less than or equal to 219. So this direction works.

The other direction. If this were the case, it has to be the case that n is greater than or equal to 220, because up till 219 it hasn't exceeded 1,000. And so, at some point beyond that, it's going to exceed 1,000 and n is going to be greater than or equal to 220. So this is the key trick here.

And once you see this, you realize that this is very easy because we do the same steps as we did before. So you're looking for this, this is equal to, again, you do your standardization. So this is from 219, and you get 5 times 219 for the mean, and 9 times 219 for the variance, less than or equal to 1,000 minus 5 times 219 over square root of 9 times 219.

Again, you can do the half correction here, make it 1,000.5, but I'm not going to do that in this case, for simplicity. So this is approximately equal to Z being less than whatever this is. And if you plug it in, you'll find that this is negative 2.14.

So in this case, we have this is the probability that Z -- again, we do the same thing-- is greater than or equal to 2.14. And this is 1 minus the probability that Z is less than or equal to 2.14. And that's just Φ of 2.14-- 1 minus Φ of 2.14. And that's-- if you look it up on the table, it's 2.14. It's 0.9838. So here's your answer.

So we're done with Part C as well. So in this exercise, we did a lot of approximate calculations using the central--

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>.