We will now go through a sequence of examples that illustrate the different types of questions that we usually answer using a normal approximation based on the central limit theorem.

In general, one uses these approximations to make statements of this type.

That the probability of the sum of n, i.i.d.

random variables being less than a certain number, that this probability is approximately equal to some other number.

Notice that this statement involves three parameters, a, b, and n, and you can imagine problems where you are given two of these parameters, and you're asked to find the third.

And this gives us the different variations of the questions that we might be able to answer.

So we will go through examples of each one of these variations.

Our setting will be as follows.

We have a container, and the container receives packages.

Each package has a random weight, which is an independent random variable that's drawn from an exponential distribution with a parameter 1/2.

And we load the container with 100 packages.

We would like to calculate the probability that the total weight of the 100 packages exceeds 210.

For example, 210 might be the capacity of the container.

Since we will be using the central limit theorem, we will have to work with the standardized version of Sn in which we subtract the mean of Sn and divide by the standard deviation of Sn.

And to do that, we will need to know the mean and the standard deviation.

Now for an exponential, the mean is the inverse of lambda and the standard deviation is also the inverse of lambda and so we know what these quantities are.

Then, the next step is to take this event here and rewrite it in a way that involves this random variable.

So what we would do is that we take the original description of the event, subtract from both sides of this inequality this number n times mu.

In this case, n is 100, mu is 2.

So we subtract 200, divide by this quantity, square root of 100 is 10 times sigma, this gives us 20.

And we do the same on the other side of the inequality.

This is just an equivalent representation of the original event, but we have here is the probability that this standardized version of Sn is larger than or equal to this number, which is 0.5.

And at this point, we can use the central limit theorem approximation to say that this probability is approximately the same if we use a standard normal instead of Zn.

Now for a standard normal, we can calculate probabilities in terms of the CDF that's given in the table.

But here, we have the probability that Z is larger than something, not smaller than something.

The CDF gives us the probability that Z is less than something.

This is easy to handle.

This probability is 1 minus the probability that Z is less than 0.5, which is 1 minus the CDF of the standard normal evaluated at 0.5.

And at this point, we look up the normal table, the standard normal table, and value for an argument of 0.5.

The corresponding value is this one, so we obtain 1 minus 0.6915, which evaluates to 0.3085.

And this is the answer to this particular problem.

In the next example, we ask a somewhat different question.

We fix again the number of packages to be 100, but we're given some probabilistic tolerance.

We allow the packages, their total weight, to exceed the capacity of the container.

But we don't want that to happen too often, we want to have only 5% probability of exceeding that capacity.

How should we choose the capacity of the container if we want to have this kind of a specification?

So we proceed as follows.

We want this number, 0.05, to be approximately equal to this probability.

But now, we take this event and rewrite it in terms of the standardized random variable.

That is, we start from both sides of the inequality and subtract n times mu, which is 200, and then divide by the standard deviation of Sn, which is this quantity and which is 20, exactly as in the previous example.

And now, this random variable, Zn, is approximately a standard normal.

So we're asking for the probability of the standard normal is larger than or equal to something which, using the argument as in the previous example, is 1 minus the CDF of the standard normal, evaluated at this particular value.

Now, what this tells us is that this quantity here, the value of the CDF, should be equal to 1 minus 0.05.

So this quantity here should be 0.95.

What does this tell us about the argument of the CDF?

We can look at the table and try to find somewhere an entry of 0.95.

And we find it either here or there.

We could choose either one, or we might decide to split the difference and say that we get the value of 0.95 when the argument is 1.645.

And so we conclude that in order for this to be 0.95, we need a minus 200 divided by 20 to be equal to 1.645.

And then we solve for a and we find that a should be 232.9.

And we can choose the capacity of our container this way.

Our next example is a little more challenging.

Here, we will fix a and b and we will ask for the value of n.

Here's a type of question that has this flavor.

We are given the capacity of our container.

We want to have small probability of exceeding that capacity.

How many packages should you try to load?

What is the value of n for which this relation will be true?

So we proceed, as usual, by taking this event and rewriting it in a way that involves the standardized version of Sn.

So we need to subtract n times mu, which in this problem is 2 times n.

We subtract it from both sides of the inequality.

And then we divide by the square root of n times sigma, which is 2.

So we divide both sides of the inequality by this number.

Once more, this event here is identical to the original event, but now it is expressed in terms of the standardized version of Sn.

This is a random variable that's approximately a standard normal, so once more, we're talking about the probability that the standard normal exceeds a certain value, and by the central limit theorem, this is approximately equal to 1 minus the standard normal CDF, evaluated at this particular value here.

Now we want this quantity to be approximately equal to 0.05, which, once more, means that this quantity should be 0.95 and arguing as before that we try to find 0.95 in the standard normal table.

And this tells us that the argument of the normal CDF should be equal to 1.645.

Here, we get an equation for n.

Unfortunately, it is a quadratic equation.

However, we can solve it.

And after you solve it, numerically or using the formula for the solution of quadratic equations, you find the value of n that's somewhere between 89 and 90.

Now, n is an integer, so you could choose either 89 or 90.

If you want to be conservative, then you would set n to the smaller value of the two and set n to be 89.

Our last example is going to be a little different.

Here's what happens.

We start loading the container, and the container has a capacity of 210.

Once we load the package and we see that the weight has exceeded 210, we stop.

Let N be the number of packages that have been loaded, and this number is random.

If you're unlucky and you happen to get lots of heavy packages, then you will stop earlier.

We would like to calculate, approximately, the probability that the number of packages that have been loaded is larger than 100.

Now, this problem feels a little different.

The reason is that N is not the sum of independent random variables and so we do not have a version of the central limit theorem that we could apply to N. What can we do?

Well, we try to take this event and express it in terms of the Xi's.

And here's how we go about it.

What does it mean that we loaded more than 100 packages?

This means that at the time we were loading the 100th package, we didn't stop.

And this means that at that time, after we loaded the 100th package, the weight had not exceeded 210.

So the event that we're dealing with here is the same as the event that the first 100 packages have a total weight which is less than or equal to 210.

But now we're back into a problem that we know how to solve.

And the way to solve it is to take this random variable, standardize it-- this actually is essentially the same calculation as in our very first example-- and we will get the standard normal CDF evaluated at 210 minus the mean of this random variable, which is 200, divided by the standard deviation of this random variable, which is 20.

So we're looking at phi of 0.5, which we look up at the standard normal table, and has a value of 0.6915.

So this was our last example, and these four examples that we worked through cover pretty much all of the types of problems that you might encounter.

Of course, sometimes it might not be entirely obvious what kind of problem you are dealing with.

You may have to do some translation from a problem statement to bring it in the form that we dealt with at this point.

But once you bring it into a form where you can get close to applying the central limit theorem, then the steps are pretty much routine, as long as you carry them out in a systematic and organized manner.