Okay let's start with the last chapter, or the last lesson, not chapter. So what we're going to do is we're going to start setting up the idea of a, of a distribution.

Now when I talk about the distribution, I mean a function that's going to describe a discrete random variable. And we're going to start with, with a really simple example called a Bernoulli distribution. Now I'm going to recognize here that things are going to look very, very mathematical here. And I get that the notation can be a little daunting. But we're not going to move super-fast through this. We're going to, kind of, take it one step at a time. But in order to start defining some of the facets of a distribution, we have to start kind of basic and start full force using the terminology.

So let's take look at what's called a Bernoulli Distribution or Bernoulli probability distribution. We're going to note that when we're talking about parameters, we're talking about quantities that make up a distribution, that distribution depends on. So a parameter is really not all that far off from what we think about in terms of the inferential side, where a parameter describes a population. [Its] kind of the same thing here. Parameters describe a distribution which the distribution describes the population. Ok, so, so it's really not that long of a leap to get there.

And so let's take a look. So a Bernoulli probability distribution can take on one of two values. It can take on usually it's called a success or a failure. But oftentimes we'll note this as being 0 or 1. So 1 usually represents a success. 0 usually represents a failure. And it's just a way for us to have a little bit of notation to save you some writing space. So, when we have this expression here P of X given alpha that's what it actually represents. That's how we read this. So we read this as the probability of X given Alpha so that the semicolon there represents, is read as, given. And what we do is we list the parameters of the distribution after the semicolon, and then the variable before the semicolon. Now some families will have multiple parameters and actually the binomial distribution, which we'll see here soon, will have multiple parameters. Well let's just look at setting up the Bernoulli probability distribution and kind of dig into that a little bit.

So let's consider example one, that's setup to model the probability of drawing a card from a standard deck of cards. So just kind of here, off to the side let's talk about a standard deck of cards, and what, what it's made up of. So decks of cards, they have four of what we're going to call suits. So those are hearts, diamonds, and then there are clubs and then spades. And the hearts and diamonds are red colored. The spades and clubs are black colors, they are red and black cards. And then in each suit there are 13 base cards, or face value cards. So it goes 2 through 10. There's no 1. And then there's a jack, queen, king and an ace. Sometimes people put ace the at the beginning because ace can sometimes represent a 1, but that's kind of getting into a specific, you know, a couple of specific games like blackjack and poker, where an ace can be low, a 1, or high as an11. But we'll kind of ignore that for now.

So how many cards are there? This is a basically a sample size calculation. So, we have two choices, right. We have the suit and then we have the value. So the suit, there are four choices the value there are 13. So we multiply those two together and there are 52 cards. And that's what we expect.

So with that knowledge in hand, let's scroll back over here to example one. And let's suppose we want to model the, the scenario of drawing a card, and drawing a heart card from a standard deck of playing cards. What would the Bernoulli distribution look like?

So first thing I'm going to do is I'm going to say that drawing a heart card is what I'm going to consider a success. And not drawing a heart part is what I'm going to consider a failure. So with that in mind let's consider. What does it take to draw hearts card? So, so lets say probability that we draw a heart card. And so we know there are 52 cards in the deck. And if we choose at random, then we're not putting any kind of a weight on one type of card over another. So we treat all 52 cards as equally likely of being selected. So there are 13 heart cards. And so that gives us 1/4th. Or, we could write this as a decimal 0.25. Now just a real quick note, this is going to be alpha. This is going to be our parameter, so this is going to be alpha and this particular distribution.

So let's write what the PMF will look like. So the way we're going to write this is, I'm going to say this is the probability of X, given alpha being equal to 0.25. And then this is a piecewise function, so it's equal to 1 minus alpha which is going to be 0.75. And this is if X is 0, and, 0.25 if X is 1 and it's going to be 0 otherwise. So basically that last line tells me that I can't have anything other than 0 or 1. And 1, in this case, is going to represent a heart card. And a 0 is going to be a failure which is going to be a non-heart [card]. Okay so we have a success and failure which we denote by 1 and 0 respectively. And we're just saying, oh if I draw a card, the probability of 1 given that alpha is 0.25 is 0.25 and the probability of 0, not drawing a heart card, is 0.75. That's pretty much it. That's, that's most of the setup of it.

So that being said, let's move on to the idea of a binomial experiment, where this comes from and why we started with a Bernoulli distribution. So binomial experiments our experiments that have four conditions. Now I split some of these conditions out. We could probably do a little rearranging and write this as maybe three conditions, but let's just leave it as four for a moment.

So what are the conditions? The first one is that the experiment has n trials so we're going to repeat some process n times and we fix that in advance. So I say, hey we're going to roll a die five times. And so I set the number of times I'm going to repeat a process ahead of time. The second condition is that the trials are identical. So my experiment doesn't change from trial to trial. And each trial can result in to one of two categories. It can either be a success which we usually denote by an S or a failure which we'll often denote by an F. The third condition is, the trials are independent. And when we talked about independence before, it means that the outcomes of any trial are not affected, or do not affect, any other trial. So they're all kind of in their own little world. But we're doing a collection of these trials. And the last one is that the probability of success, which we're going to denote P, by P, is constant from trial to trial, it doesn't change. And that kind of gets captured here as well, that trials are identical. But it's, it's worth kind of calling that out.

Okay. Now let's talk about the idea of a binomial PMF, a binomial distribution. It's going to be a distribution that essentially models a scenario where all four of these conditions are met. And so what we can do is, we could kind of talk about the derivation of the formula a little bit. It looks very daunting but keep in mind, please, that we're actually not going to do this by hand. This is going to be the first lesson in which we start looking at R. Now there's going to be labs, and things that we work with, but we're going to actually find that at this point, starting here, it's going to be convenient to use some sort of software to do our calculations. rather than doing them by hand Not saying that doing it by hand is beyond anyone's ability here but, we've got the technology, we might as well use it. So let's do that. So binomial PMF. So we're going to have X be a random variable. And what it's going to do is it's going to count the number of successes in n trials. So I'm going to run a trial, So many times. I'm going to define something to be a success in that trial. I'll know the probability that I observe a certain number of successes in n trials.

So here is the formula. It's kind of, it's kind of yucky, I'll admit that from beginning. So there are a few pieces here and we'll we'll kind of look at them in more detail. But there's this term N X that's read as N choose X. We'll delve into that here in a moment. P is the probability of success, and we're saying we have x successes. So we want to have probability of success X times. Probability of failure is the complement. So we're going to have that 1 minus P times. And the remaining n trials will be that. So we'll multiply it by itself and minus X times. That's the number of failures. So note that we have X successes and minus X failures. You add those up you're back to n. And so I can take on any value between 0 and n inclusive. We could have no successes. 1 success, 2 successes all the way up to all of our n trials be successes. So that term there that N and X that looks like a fraction but missing that horizontal line is what's called a binomial coefficient, kind of related to this, but we could talk about that for a long time if we wanted to.

This is read as N choose X. So if I write N and X, we read this as N choose X. And what it represents is the number of ways for me to choose X out of the N objects without regard to the order in which they're chosen. And the formula is given here. It's N factorial over N minus X factorial over times X factorial where factorial N factorial is just N times N minus 1 times N minus 2 all the way down to 1 and 0 factorial is 1. So just a real quick example of this. Again, we're going to kind of wiggle our fingers over this. But let's say I wanted to find the value of 6 choose 2.

So I have six objects let's just say 1, 2, 3, 4, 5, 6. How many ways can I pick two of these? So I can pick the 5 and the 6. And in this case the 5 and 6 would be the same as 6 and 5. I don't care about which order they come in. I just care about what I get. So how many ways are there for me to choose 2 out of 6 items? Well, if we follow the formula here, it's going to be 6 factorial over 6 minus 2 factorial times [whoops, thinking a little ahead there] times 2 factorial. So let's just rewrite this. This is going to be 6 factorial times 4, or sorry, divided by 4 factorial times 2 factorial. So let's see 6 factorial is 6 times 5 times 4 times 3 times 2 times 1. 4 factorial is 4 times 3 times 2 times 1. 2 factorial is 2 times 1. And so we'll see that we have common factors in the numerator and denominator here. First, we have 4s, 3s, 2s and those 1s go away. This 1 kind of just goes away anyway. And so what we're left with here is 6 times 5 is 30 Over 2, or 15 ways. So there are actually 15 ways for me to choose 2 out of these 6 objects. And we could list them. It's not really that difficult. But, you know, with 15 we have 1 and 2, 1 and 3, 1 and 4, 1 and 5, 1 and 6. And then we can say, we can have 2 and 3. I don't need to care about 2 and 1 because 2 and 1 is covered in the 1 the 2. Then we have 2 and 4, 2 and 5, 2 and 6 And that could be a 3 and 4, 3 and 5, 3 and 6 4 and 5, 4 and 6 and then 5 and 6. So it's 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 and that works out.

Okay so that was just a quick demonstration. Again we're not really going to come back to this. We'll talk about it a little bit really, really late in the next course. But that's about all I'm going to say about the binomial coefficient for now. The next video I'm going to start looking at using R. Doing calculations in R, and just kind of how we work with that. So if you haven't yet, please make sure that you can log in to AWS and start up R studio. It's a little difficult that it's kind of windows, within windows, within windows, if I do the AWS version. So I'm going to just use the installation of R Studio on my local computer and use that. But it'll be the same. It's just that looks a little bit cleaner if I use my local version. So we will meet back up in the next video.