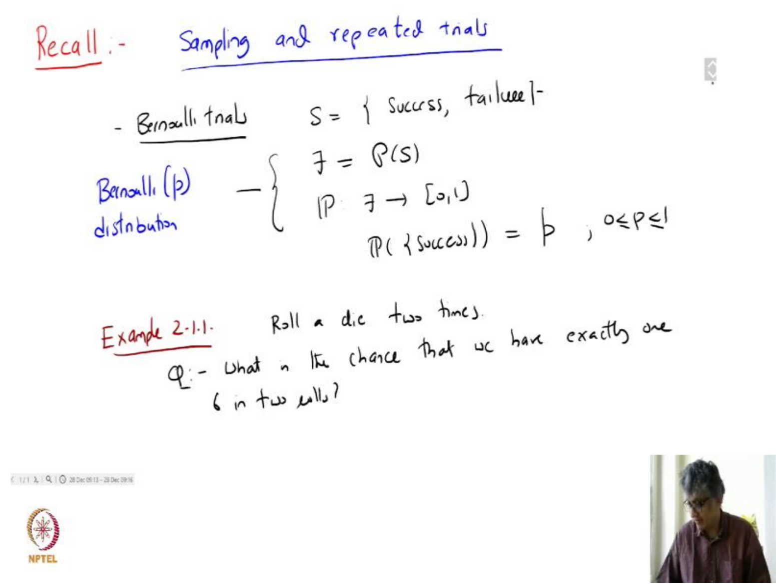
Introduction to Probability – With Examples Using R
Professor Siva Athreya
Theoretical Statistics & Mathematics Division
Indian Statistical of Institute, Bangalore
Lecture 9
Sampling and Repeated Trails – Part 01

(Refer Slide Time: 0:21)



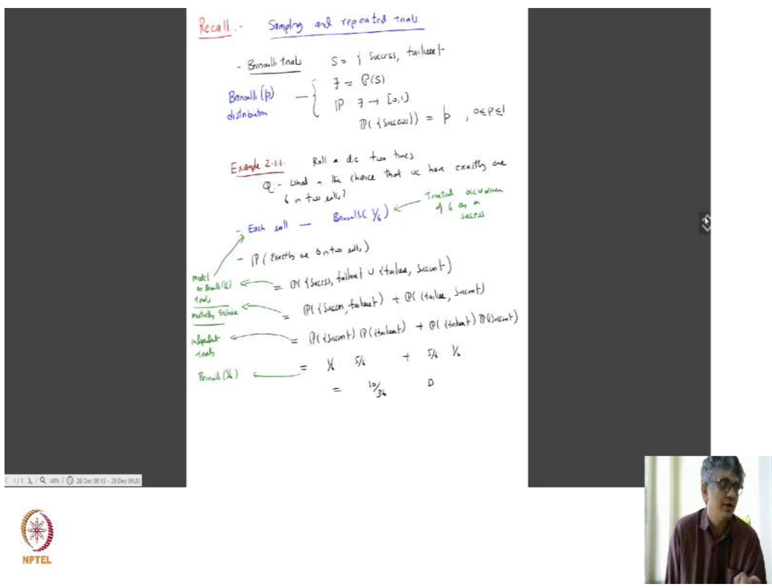
So, I will begin by recalling what we were doing last time. So, what we were doing last time was this idea of sampling and repeated trials, we discussed sampling and repeated trials and the first thing we discussed was this idea of Bernoulli trials. So, and here the motivation was that you have an experiment very interested in a particular event in the experiment. And what you would like to do is you if the event happens you declare the experiment of success, it does not happen they event you did an exp1ntial failure.

So, what happens is that your sample space is reduced to just two possibilities, success and failure and then the event space is just the power set of S that is all possible thing and the probability as a function from 0 to 1 is defined completely the moment you define the probability of success. And you define the probability that is equal to p the p is the number between 0 and 1, if it is 0 means that there is no success, always a failure and 1 means it is always success.

And this is what we would call as a Bernoulli p distribution. So, this is what we discussed last time and then I said that this if you repeat this thing again you have independent Bernoulli trials and I had d1 this example for you last time, I said that, so I have d1 this example which I have called as 2.1.1 here we roll a die two times.

And the question I asked you was that was a little different, I asked you what is the chance that the exactly 1 six and 2, okay what is the chance that we have exactly one 6 and two rolls and there way we did was we had two methods, one was to outline all the 36 outcomes and then write down this event as our chosen event and description of it and every output is equally likely and then you divide by just the number of elements in this event divided by the total number of elements that is one way of doing it.

(Refer Slide Time: 3:33)



But the other way of doing it was that we said that the way to do it was you set up each role as a Bernoulli a one sixth experiment because that is the success probability is 1 6, so the way what this was you treated a occurrence of 6 as success and then so that would imply that the probability success is one six, probability failure is one six.

And the question you are interested in is that the probability of exactly one six in two rolls, exactly one 6 and two rolls and this we wrote down as probability of success of failure only one 6 and the other 1 was failure or failure success for failure and success this is what could happen in the two roles.

And then what would happen is that this is simply a mutually exclusive but the property of probability you could just write this as success, so it is success and failure plus the probability of failure and success, continuously exclusive events and by the axioms of the property probability you could show this.

And this now use the fact that independent, so you get probability of success times probability of failure and plus here you get probability of failure times probability of success and since now I am in a Bernoulli one six experiment I know this is one six, this is five sixth, this is five six and this is one six and so I get 10 by 36.

So, just to sort of reiterate what all we use, we use the fact here that we did two things, we modeled each role as a Bernoulli experiment that is what we need to get this and then at this step we use the fact that they were modeled as Bernoulli one sixth trials so and here we use the fact that of independence, so here we are not independent so I am sorry for that, here we use the fact that they are mutually exclusive, so they are disjoint event so you could add them up.

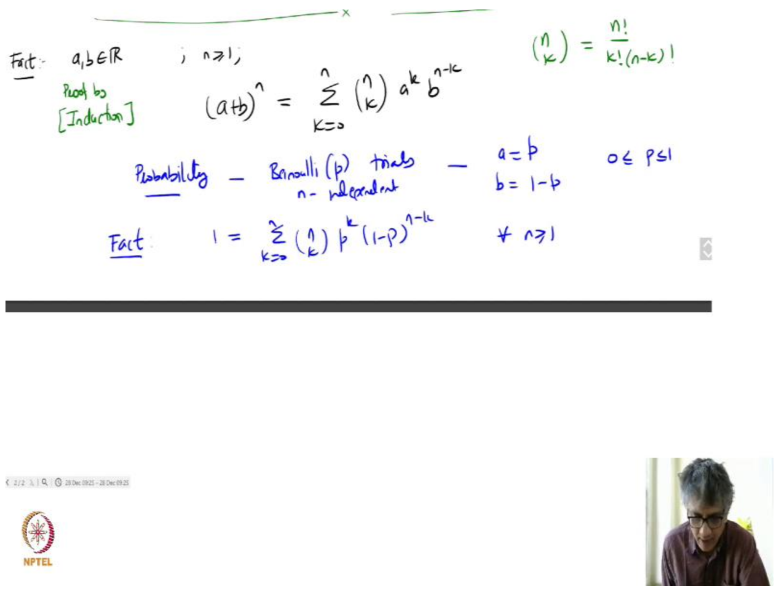
And here we use the fact that they are independent and here I use the fact that they are Bernoulli. So, in some sense this is 1 way of understanding this problem. So, you could convert, so the way we did this was we said that we had an event of interest, the event of interest is exactly 1 6 in two roles so we said that last time is that I am interested let us say a 6 occurs in an experiment, so that is my success and if 6 does not occur that is my failure, so that is what each roll is defined.

Then I said that if I want to do two independent Bernoulli trails because I am throwing 1 role another role exactly 1 6 and 2 rolls means that I have a success and failure and or a failure and success that is the exact event that I am interested in success failure or failure success in my two trials.

And then I discuss independence and will exclusive to complete the problem. So, exact sample space for two rolls would be the tuples, failure, success-success, failure-failure and success-failure, and that would be your next sample space for your element of interest. But I did that calculation last time and here I am just computing this problem.

So, now I will, I will now try to generalize this. So, now I will try and say if I am performing this role 10 times let us say I cannot keep writing down my sample space and again again and doing so. So, now I would like to do something a little different. So, I will just recall so this probably you have seen in your 12th class substantially like.

(Refer Slide Time: 8:38)



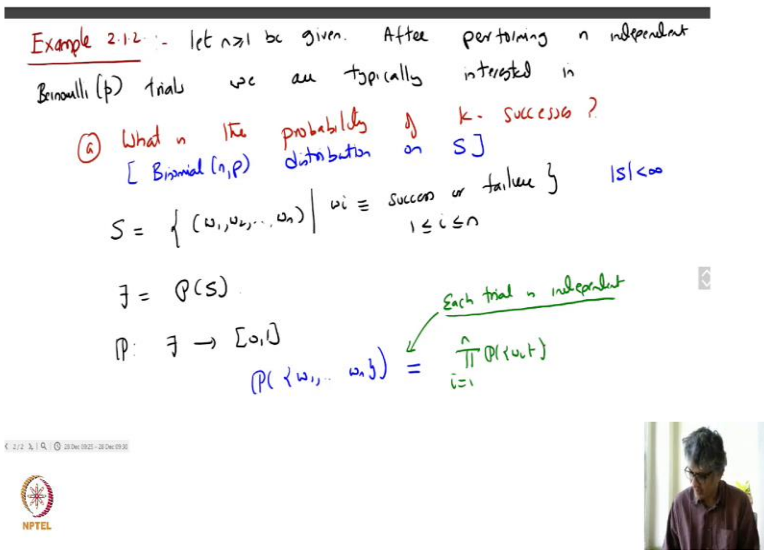
So, if I have any two real numbers a and b in R this is a fact let us say then for any integer n equal to 1 then I know that a plus b to power n is going to be equal to the sum from k equal to 0 to n, n choose k times a power k b power n minus k, where n choose k was just that the number of ways of choosing k people from n people that is just n factorial by a k factorial and then n minus k, very good.

So, now this is something that you should know how to prove, so I do not know if you have done this in analysis or writing a mathematics. So, how does one prove a formula like this, so you should, if you have not practiced it, just practice the proof by induction, so that is how you prove this so it is proved by induction.

So, this you must know that for n equal to 1 it is trivially true, n equal to 2 you assume that n equal to k or n equal to m let us say and then n equal to plus 1 you try and show it. So, this fact I will assume that you know, so in probability it is kind of very, in probability it appears in a very natural way, this formula which you will see in probability associated with Bernoulli pre-trials with n Bernoulli pre-trial let us say n independent of them, n independent trials.

And here you should use the formula with a equal to p and b equal to 1 minus p, so I need to use this formula, I need, so in that case I know that, that case I know 1 is equal to r on this side and on this side I will have k equal to 0 to n, n choose k p power k and 1 minus p power n minus k, this is true for all n (())(11:07), so I will use this fact. So, I will require this fact.

(Refer Slide Time: 11:39)



So, now I will now do an example and then by which I will illustrate this idea of n Bernoulli trials in the example, see the example a 2.1.2 so here what we will do is, we will perform, we will take let n be equal to 1 p given some fixed number given to us then what I do is I perform n independent Bernoulli p trails.

So, after performing n independent Bernoulli p trial we are typically interested in the following questions, like in the previous example I performed it twice, interested in the following questions. What are the questions we are interested in? So, one question is a is, what is the probability of k successes?

Let me mark the question in a different color, so the first question we are interested in, what are the probability of k successes? So, the first question is this. So, I perform an experiment n times n Bernoulli trials trial each outcome is successful failure, I am interested in what is the probability of k successes. So, this is called the binomial n, p distribution, so I will, this outcome space is called the binomial n, p distribution on S, so I will discuss what S is.

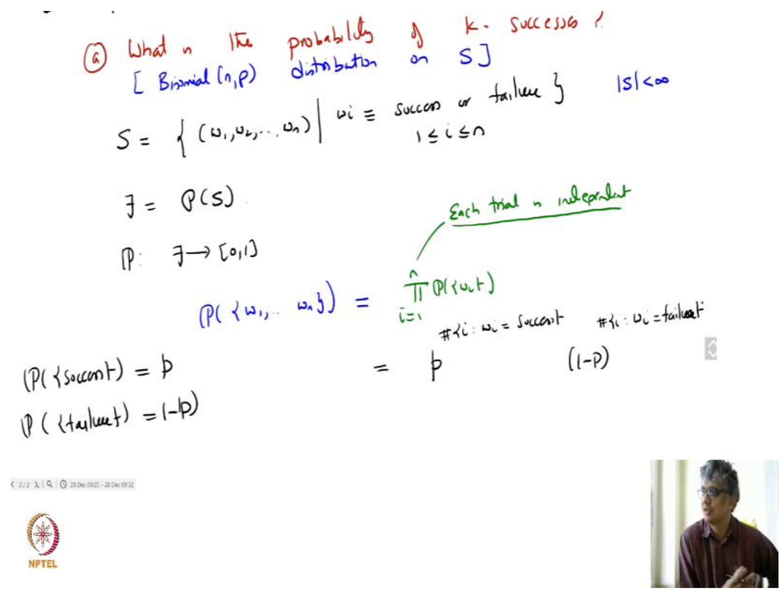
So, what is S, I will discuss in a sample. So, what is happening now? Each trial you are throwing out a probability, it is either a success or a failure and success probability is p and failure probability is 1 value. So, what you do is you let S to be the set of all tuples that is omega 1, omega 2 up to omega n, such that each omega i is a success or a failure, this is for 1 less than equal to i, less than equal to n.

So, your experiment you perform n times, each time you get a success or a failure this is something that you keep in mind (())(14:49). And then again, your event space is again equal to the power setup. So, your probability is the function from f to 0, 1 and what is the probability going to be, now your probability means it is a finite, S is finite so number of, so notice that S is finite, so you each place your either only two outcomes, either a successor of failure so there are two possibilities each position, so you can easily see that this space is finite.

Now, one needs to understand how to assign a probability. So, now already I have shown you before if I know the probability of any single outcome then I am done. So, now if I know the probability of any single outcome, a single outcome is 1 omega 1, omega 2, omega n but each omega n is either a success or a failure. How do I do this now?

Now, each trial is independent, so each trial is independent, each trial is independent, each trial is independent. Now, this is just going to be the product of each, is going to be the product of i equal to 1 to n, the probability of each omega because that all that are independent so each event, this event is independent, so therefore the probability of a 1 and a two and a n is going to positive.

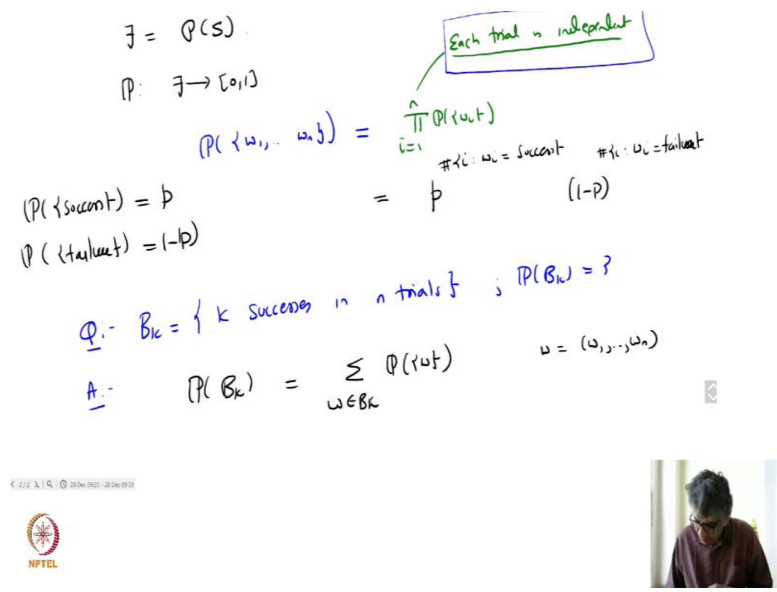
(Refer Slide Time: 16:32)



So, now this is also easy to see. So, now we know that probability of success is p, probability of a failure is 1 minus p. So, immediately 1 can try and write this probability out, the probability of this product is just going to be p to the power the number of i such that omega i is equal to success.

And 1 minus p the power the number of i is such that omega i is a failure. So, let me just make this little bit better, yes, 0, 1 and the this and here number of i omega i is a failure that is how you assign the probability, let us see if it comes my internet is slow… So, now this is how you compute the probability of every unit, every outcome, so this exactly fine, now this defines the probability on this (())(18:39).

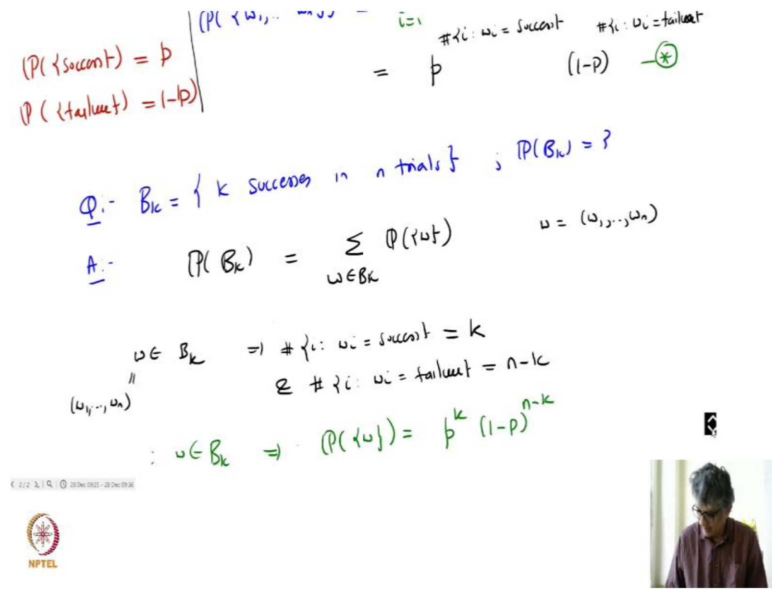
(Refer Slide Time: 19:00)



Now, what I can to do is, I would like to sort of now see, what the question of interest, our question of interest is what is the probability of k successes? So, how do I do that let us see. So, I have to have let us say, let us say I denote b sub k, let us say b sub k is the event that there are k successes in n trails, that is there are k successes in n trails that is the event of intersection.

And I want to know the question is asking, the question is asking B sub k is this, what is probability of B sub k? So, here I have used independence, (())(19:34). So, here is the answer, that is the answer. So, now you already know because p is a probability, the probability of B sub k is the same as the sum from omega in B sub k the probability of omega where omega is omega 1 up to omega n.

(Refer Slide Time: 20:19)

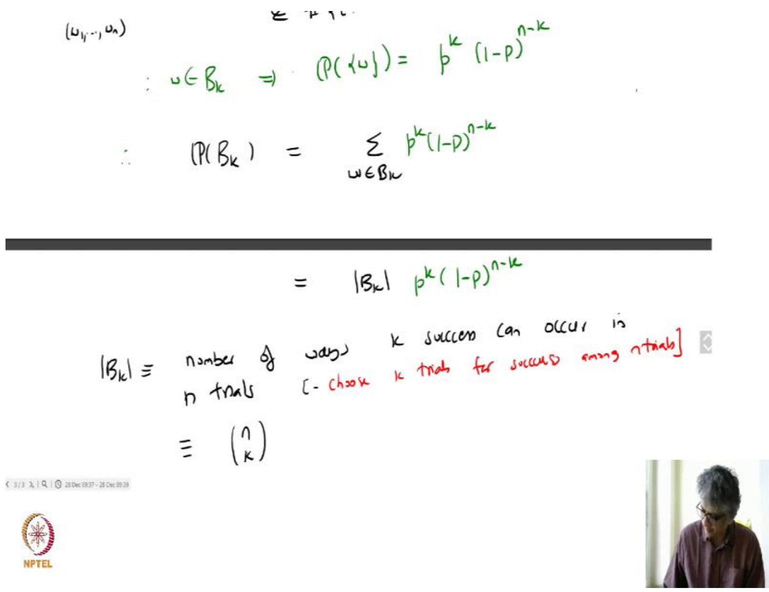


Now, you also know that probability of omega is just given by this formula this formula here let me make this up red perhaps (())(20:19). So, now I know the probability of any omega is just p times the sum number of i's omega i success and number of i omega S frequently. Now, you note if omega is in k, what do we know there are k successes that means you know two things the number of i such that omega i is equal to success because p omega is in Bk, Bk has k successor this got to be equal to k.

And the number of i such that omega i is a failure that is going to be equal to n minus k because total number of people are n so omega is equal to omega 1 up to omega n so there are n people and although it is k of their success and the n minus k of them are fake. So, therefore from this let us call this equation star, so from star, so therefore omega is in Bk sorry, omega is in Bk will imply a probability of the single outcome omega is going to be equal to p to the power the number of times success k and 1 minus p to the power n, so p power k and 1 minus p to the power n minus k.

Because if you look at the above formula there it is p times the number of y successes omega i equal to success and number of S omega i is equal to failure. So, our success is k and failure is n minus k, very good. So, probability each outcome is p power k n minus k, it is the same for every omega in Bk, so it is a constant because it does not depend on how the ordering is, it just depends the total number of successes and total number of failures.

(Refer Slide Time: 22:59)

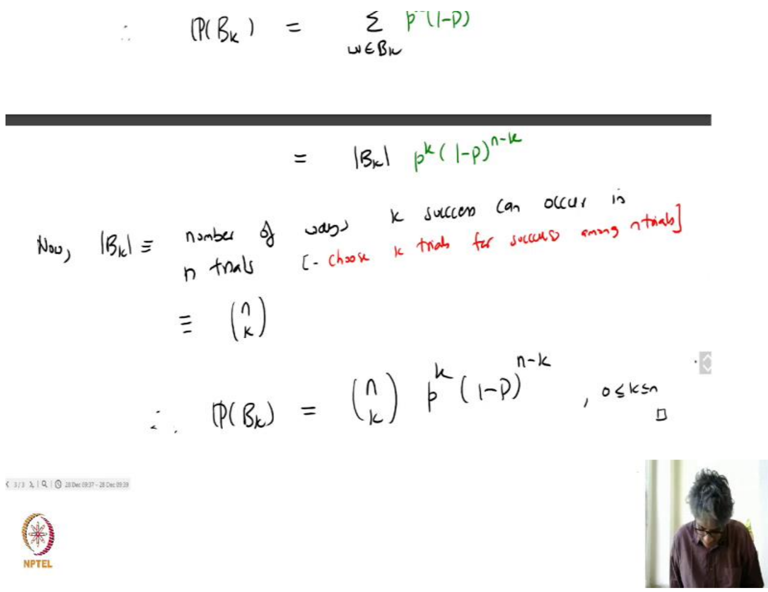


Therefore, now probability of Bk is going to be the sum over omega in B k and for each B k, each probability omega I have this formula, so it is p power k, so it is just p power k into 1 minus p to the power n minus k. So, it is I am adding the same thing again and again. So, what will happen now that is the same as saying this is going to be equal to the size of B k times this expression bracket p power k sorry p power k into 1 minus p to the power n minus k.

So, now I have reduced the problem of computing the probability to just understanding what the size of B power, but size of B k you all know, so B k means is just R, you are performing n trials out of which k is a success and rest have to be failures. So, size of B k is just a counting problem, it is just the number of ways k successes can occur in n trails.

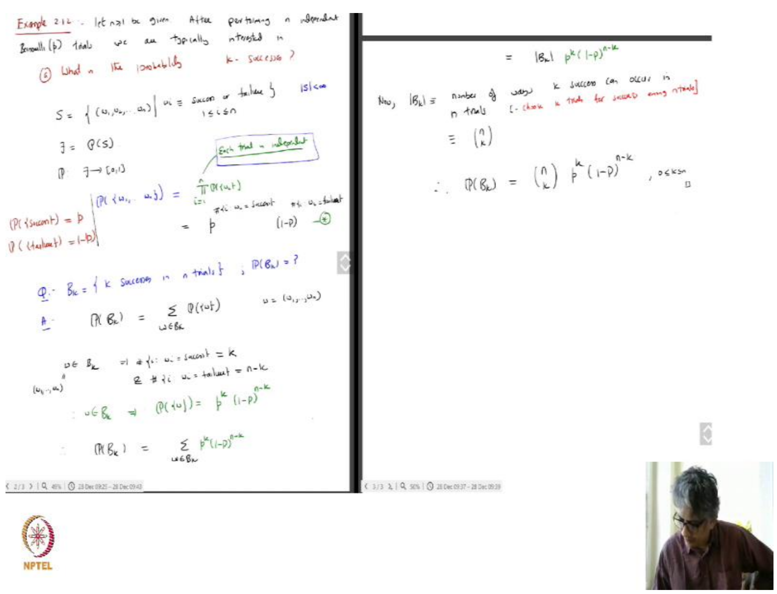
But this is a well-known, this is just all you have to do is choose the spots for the k successes and everybody else is a failure. So, all you have to do is choose just same thing as saying, you choose k spots or k trials for success among n trials. So, that you know is from 12th class or earlier you know that is same as n choose k.

(Refer Slide Time: 25:12)



So, therefore, so now therefore the probability of B sub k which is your quantity of interest is just going to be n choose k p power k and 1 minus p to the power n minus k. So, that is your, that is your answer that is right and this is true for any k between 0 and n.

(Refer Slide Time: 25:58)



Let me recap a little bit, so what I do, I had n trials, I was interested in the probability of having k successes in among n trails. And then the way I did was that I said let S equal to omega 1 to omega n be the outcomes in each trial, so each trail has out comes success or a failure and then I put this probability on it which defined probability of omega 1 to omega n, the trial are independent so that means it is a plot of the probabilities, I am using independent state clearly.

So, once I did that, I get a broader problem, but I know the probability of each omega i is given by either a p or a 1 minus p depending on the success or a failure that means this expression right here is just p times the number of times your successes and 1 minus p to the power number of times you have failure.

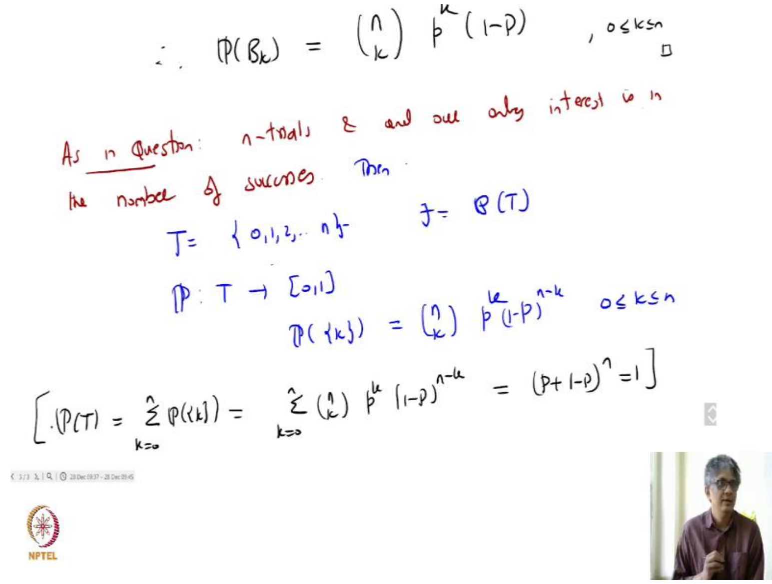
So, that is the expression for probability of every out, now since mod S is finite, we already seen before if I know the probability of every outcome that is enough to define the probability. Why is that, because I do again, I look at the chance that I have k successes in n trails, I call it event as Bk but probability of B k is going to be just by the mutual exclusiveness property summation of omega Bk probability omega, but omega is just a tuple omega 1 through omega.

Now, omega and B k implies, omega and B k means what that means B k is the event there are k successes in n trials that means you have omega 1 to omega n, k of them have to be success and n minus k that have to be failure, it does not matter which order they come, it is not a total number. And that defines your B sub k.

So, that means omega in B k, probability of omega from this star is going to be B power k n 1 minus (())(28:15). So, let me go back probability of B k is going to be sum over omega p power k 1 minus p to the power n minus k and then that is going to be the same as size of Bk because every term is constant so it comes out it is size of Bk times p power k 1 minus p to the power n minus k.

So, mod of B k is just the number of ways of k success can occur in n trials that is just choosing k spots among n trails and then just to get n choose k. So, probability of B k is n choose k p power k 1 minus p to the power n minus k. So, I made 1 small error so I do not want to, I do not want to call this the binomial, those little binomial, I just want to bring this back here, let me erase it and bring it back.

(Refer Slide Time: 29:40)



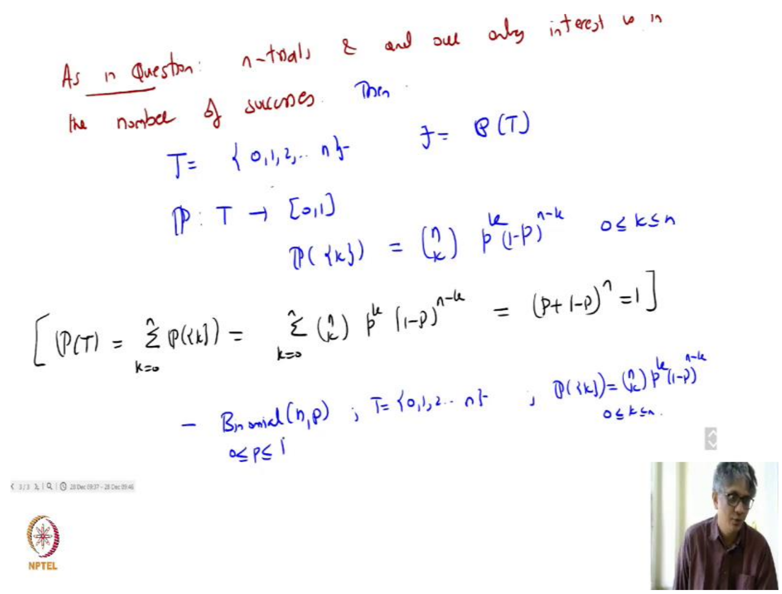
So, now let me discuss what a binomial is in a second, so now, let me go back to the original screen and close this class with that. So, this calculation showed you how to compute the chance of k successes in n trails. Now, what I do is suppose I am only interested in the number of successes, not in, not anything else they expected.

So, what I will do is, so as in question, as in question, so we have n trials and our only interest is in the number of successes, we do not know, we do not care about how they occur or what is happening otherwise we are only interested in number success. So, then the way you do it is you redefine your sample space in the following way.

Then you could think of sample space as T letus say, T as 0, 1, 2, up to k, up to n. Let F be the power set of T that is all possible and then you let the probability of T to 0, 1 given by the probability of an outcome k is going to be just like here, so it is going to be k do not know from successes this is going to be n choose k p power k 1 minus p to the power n minus k for in 0 and n.

Now, I know the legitimate probability because so if you have a finite set sample space you know that your legitimate probability if the sum of the probabilities over all outcomes become 1 but that is easy to see we know that the summation over k probability of p sub k, sorry as p of S, p of T in this case is going to be equal to this is from 0 to n, let us come over here and this n choose k p power k 1 minus p power n minus k and by our earlier expression of binomial expansion this is just p plus 1 minus p to the power n and that is just 1. So, this is a legitimate probability.

(Refer Slide Time: 32:11)



And this is called the binomial np distribution, so this is called the binomial, this is called the binomial np distribution, so when I say binomial np that means my sample space S is just 0, 1, 2, up to n and the probability of every outcome k is just going to be equal to something called this is T, it is for this notation and it is just n choose k p power k 1 minus p power n minus 1.

So, the binomial np distribution models understanding the number successes in n (())(32:44) where the probability is, success in each trial is given by p, so p means 0, so this is the second distribution we have learned now, or third distribution, first was uniform, then was Bernoulli p, now binomial n.