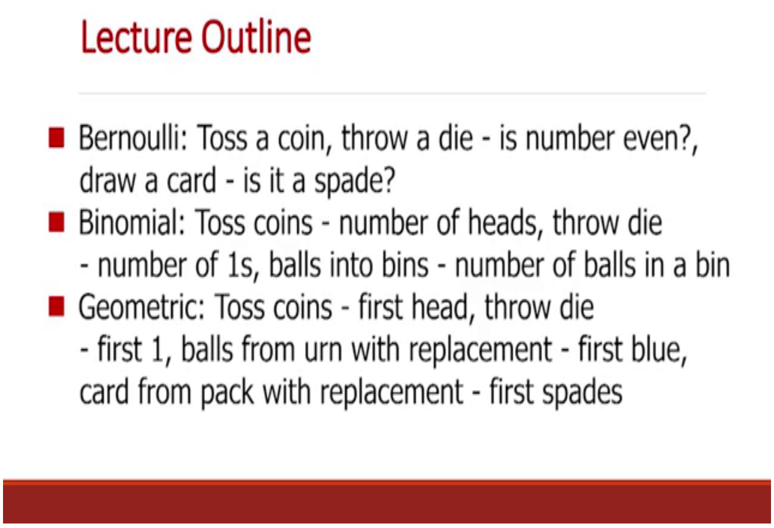
Probability Foundations for Electrical Engineers
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture – 25
Examples: Discrete PMFs

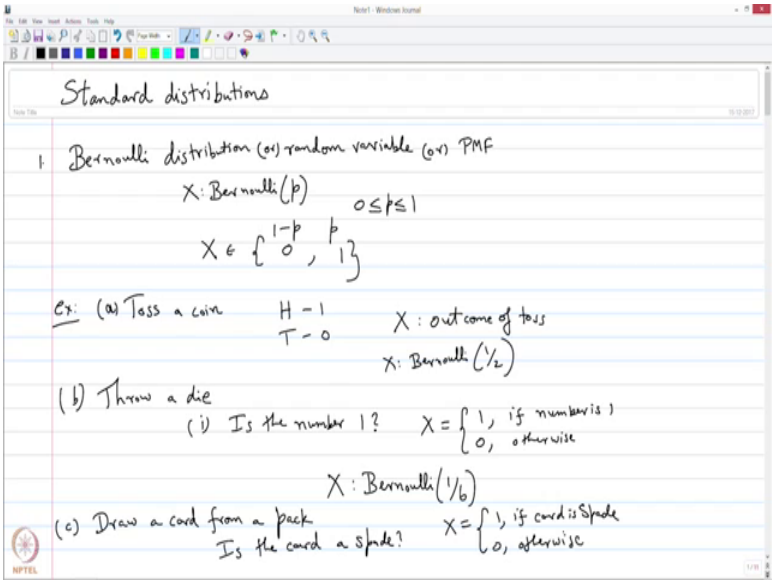
(Refer Slide Time: 00:13)



Welcome to this lecture I hope you are seeing Professor Aravind's lectures on various types of distributions, various types of random variables with different distributions. A few of them are very standard and come repeatedly in various situations. So, these are given some standard names and we call them as specific types of distributions or specific types of random variables or specific types of probability mass functions PMFs. So, if you think about it all of these things are referred to the same quantity which is the same entity which is the probability space and the distribution we are dealing with.

So, one the first example that professor Aravind deals with and I will also deal with here is the Bernoulli distribution or the Bernoulli random variable or the Bernoulli probability mass function. So, these are 3 things which kind of refer to the same thing and we will use this kind of terminology interchangeably.

(Refer Slide Time: 01:09)



So, this lecture is on standard distributions examples of how standard distributions occur in many of our the experiments that we have been doing we have been looking at a few experiments all the time tossing a coin, throwing a die, picking a ball, from an urn throwing balls into bins picking a card from a pack. And Professor Aravind has been talking about the standard distributions or PMFs and we will see how those standard distributions arise in our simple experiments. So, that is going to be this lecture.

So, the first one that we go we will look at as examples of the Bernoulli pdf; Bernoulli PMF sorry Bernoulli distribution. Like I mentioned, I will use these 3 things interchangeably in some sense hopefully you see the distinction between these things and the fact that they refer to the same thing in some sense. So, the distribution is the entire thing the random variable the values it takes the probability with which it takes them the random variable of course, refers to the variable itself the PMF is the probability together they make the distribution that is the way to think of it.

The Bernoulli PMF or random variable if you say X is a Bernoulli there is a parameter p. So, what is this parameter? It is its some number between 0 and 1, p is some number between 0 and 1 and when you say X is a Bernoulli random variable you have to also say what p s. So, p could be a half, half is the most popular value, but you can also have other values you know maybe half is. So, fair value in some sense, but of course, you can also have other values 1.9, 0.999 or 0 or 1 anything is possible. So, when I say X is

Bernoulli random variable what do I mean X takes values in this set 0 1. So, X takes only 2 values. So, I will use 0 or one success or failure or; however, you think of it, but 0 and one in the random variable description and the probability of 1 is p, probability of 0 is 1 minus p.

So, you remember the way I write distributions right. So, particularly discrete distributions I would say X the random variable belongs to this particular set and indicate the probability on top of the value. So, it is a very convenient way to write probability distributions. So, this is the random variable which is Bernoulli it takes the value 0 with probability 1 minus p and takes the value 1 with probability p.

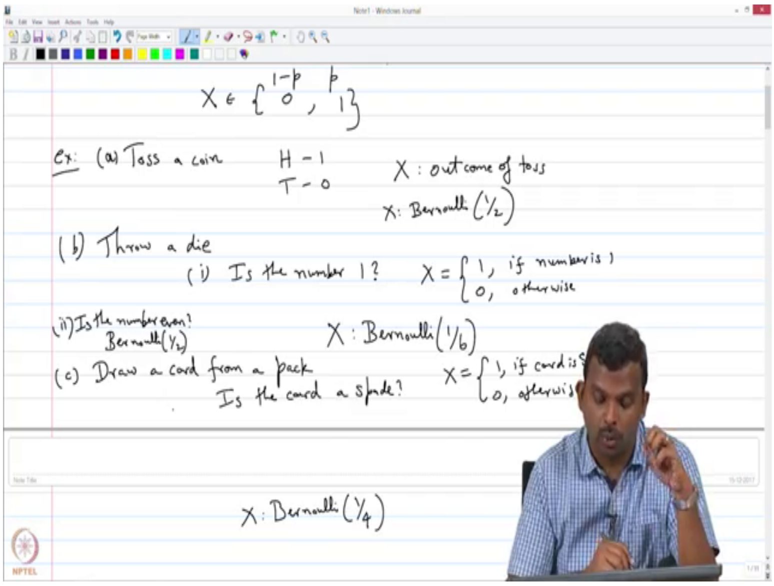
A couple of things to check whenever somebody gives you a distribution is that the total probability probability of everything should add to 1. So, that is the first thing to check the other thing to check is each probability should be between 0 and 1, you cannot have both these conditions violated. So, you should check. So, 0 has probability 1 minus p that is between 0 and 1, 1 has probability p that is between 0 and 1. So, 1 minus p and p together if you add them you get 1. So, this is a valid PMF. So, this is the Bernoulli distribution.

So, what are our examples if you look at the first example and toss a coin it is a very very natural example if you toss a coin the. So, you could say heads this 1 and tails is 0. So, you just think of a mapping like that and your outcome simply becomes the random variable. So, X is. So, X refers to the outcome of the toss that is clearly Bernoulli. So, X is Bernoulli with half probability right. So, probability of heads this half probability of tails is half. Like I said you have to do this simple mapping of heads to one and tails to 0 that does not change the distribution in any significant way. So, tossing a coin has an obvious example.

The other example if you throw a die. So, you can ask a question is the number 1. So, this could be 1 question you can ask. So, the answer to this question could be the result of a random variable. So, you could say X is equal to 1 if the number is 1, 0 otherwise. So, what am I doing here I am throwing a die and I am looking at the number that came out, if the number that came out is one I am thinking of it as success or something and I am assigning one to the random variable X. If the number that came out anything other than 1 2 3 4 5 or 6 I am going to assign a value 0 to the random variable x. So, I have a 2 valued random variable and that is clearly Bernoulli again.

So, now, clearly you can see this X here is Bernoulli and what is the parameter; the parameter is the probability that I get a one when I throw the die that is 1 by 6. So, this is a different Bernoulli distribution it is not 1 by 2 its 1 by 6. So, you can also look at another example you know draw a card from a pack and you can ask the question is the card a spade. So, once again I can assign X is equal to 1 if a card is spade and 0 otherwise.

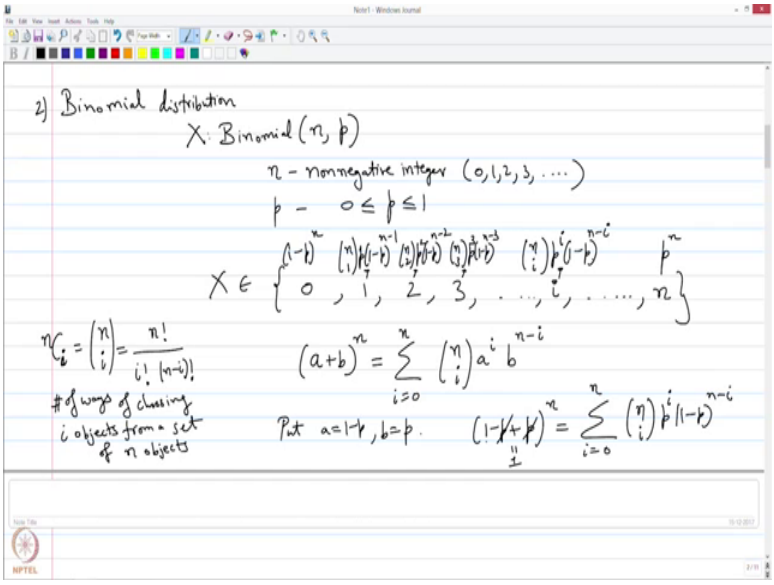
(Refer Slide Time: 06:58)



And you can see that if I do that this X becomes Bernoulli. And what is the probability of 1? It is 1 by 4. So, that is the Bernoulli distribution when you draw a card from a pack of cards so in fact, you can also do something else here I gave a number 1. So, you can ask one more question which is I will put it as 2. So, so you can ask is the number even. So, when you throw a die you can ask a question is did the number you is the number you got even or not. So, even is what 2 4 and 6 are even, 1 3 and 5 are odd. So, if you think about it is the number even is actually a Bernoulli distribution with the probability half right. So, 3 by 6 is the probability that the number is even. So, that is 1 by 2. So, this is these are a few examples of Bernoulli distribution.

And different parameters and how they show up in our standard examples that we have seen. So, far hopefully this is clear. So, Bernoulli is a simple example in some sense.

(Refer Slide Time: 08:14)



The second distribution that process Aravind talks about this binomial distribution this is a very very important distribution as well or binomial PMF or binomial random variable all of these are the same. So, if you say X is binomial you need 2 parameters now a parameter n and a parameter p. So, what is n? n is a positive integer, usually positive I mean you could say n is 0 as well. So, let me just not say positive I will say nonnegative integer. By the way what is the difference between positive and non negative? Positive does not include 0, positive will mean 1, 2, 3, so on; non negative means 0 is also included, 0 comma 1 comma 2 comma 3, so n this could be 0 1 2 3 like that any one of these case.

And what is p? p is my probability indicator. So, that is just 0 less than p less than 1. So, it is some number between 0 and 1. So, for every such n and p, I have a binomial random variable which can be defined. So, what is this binomial random variable? So, it takes values takes it could take n plus 1 values it could be 0, it could be 1, it could be 2, it could be 3, it could be so on till it could also be some arbitrary value, i in the middle and then it could also be n. So, takes values from 0 to n, integer values from 0 to n, 0 1 2 3 4 so on till 1, this I here is some arbitrary integer between 0 and 1 just to show how the probabilities work out.

So, the values that X takes when it has a binomial distribution is easy to write down now what about probabilities, the probability it turns out works like this. So, we need this notation n c 1, n c i which we also denoting n choose i this is n factorial by i factorial times n minus i factorial. What is the meaning of this number? Number of ways of choosing i objects from a set of n objects. So, this is the combinatorial sort of formula n choose i n choose i uses n factorial by i factorial times n minus i factorial number of ways of choosing i objects from a set of n objects. Now, these n choose i will play a role in the definition of probability here it turns out for probability that X equals i is actually n choose i, p power i 1 minus p power n minus i. So, this is the probability.

So, if you put i equals 0, if I put i equals 0 n choose 0 is actually 1, p power 0 is also one. So, you will simply have 1 minus p power n that is the probability that a binomial random variable takes the value 0 and then you have the probability that binomial random variable takes the value 1 its n choose 1, p times 1 minus p power n minus 1 and then you have 2 n choose 2 p squared times 1 minus p by n minus 2, then you have 3 which is n choose 3 p cubed 1 minus p power n minus 3. So, this is for 1, this is for 2, this is for 3 and that is for i and all the way to n, if you go to n you have n choose n p power n 1 minus p power n minus n that actually evaluates to p power n.

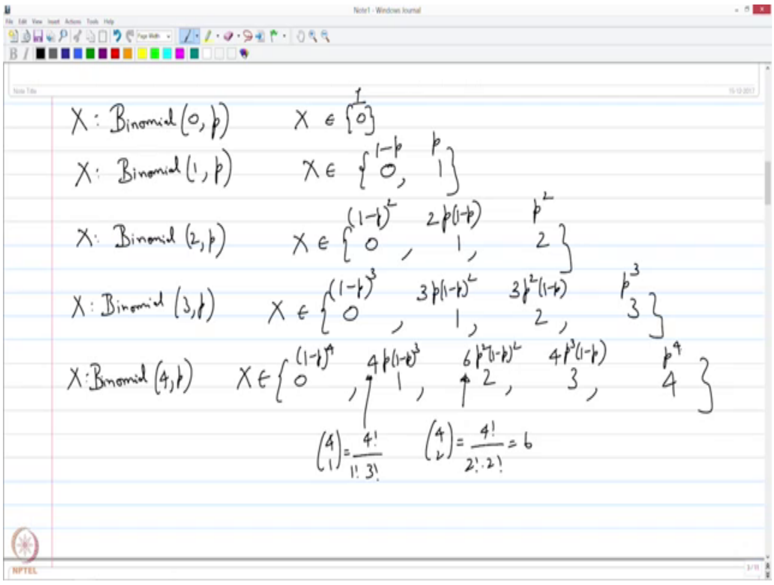
This is the binomial random variable, it has a connection with Bernoulli random variable. In fact, if you repeat the Bernoulli rand; Bernoulli experiment n times independently and identically the number of successes you have number of ones you have is actually binomial distributed. So, that is the relationship anyway. So, it is important to see as well.

So, now, remember this I am claiming is a valid PMF right. So, what is the condition for this to be a valid PMF every probability should be between 0 and 1 and then all the probabilities should add up to 1. In fact, if you think about it if I can show every probability here is positive right. So, that is every probability here is positive that is easy to see right there is no negative going on 1 minus p is also positive. So, everything is positive n choose i is also positive. So, every probability here is positive or non negative and then if I show that they add up to 1 that is enough they have to be less than 1, so all of them are positive and a bunch of positive numbers are adding up to 1 which means each number is also less than 1. So, that is enough to convince yourself that these probabilities will actually add to 1, if you show that they add to 1 then you are done.

So, how do you show something like this adds to 1 it turns out you have to use this binomial expansion formula it is very important that you know this a plus b to the power n. So, some of the standard formula you learn growing up is a plus b squared, you also learn a plus b cubed, maybe some of you will learn a plus b power 4 also. So, all these are standard formulae that you learn now there is an extension to all that formula a plus b to the power n, for n being any integer 1 2 3 4 100 whatever and that formula is the following this is equal to summation i equals 0 to n, n choose i a power i, b power n minus i. So, this is the main big formula you can show for instance if you put n equals 2 you will get your familiar a squared plus 2 a b plus b squared. So, anything you put n equals 3 if you put you will get a cubed plus 3 a b squared plus 3 a squared b plus b cube. So, you will get all those things as special cases of this formula this is the master formula for a plus b power a. So, this is the most important thing here which brings everything together in the binary, for the binomial distribution.

So, here this is true for any a and b. So, you can put a equals p 1 minus p and b equals p, if you do that what happens, to the left hand side you get here 1 minus p plus p whole raised to the power n equals summation i equals 0 to n n choose i p power i 1 minus p power n minus i. Now, what is this guy 1 minus p plus p? p and p will cancel. So, you get 1 itself. So, this is just equal to 1. So, this shows that all these probabilities n choose i p power i 1 minus p by n minus i, i going from 0 to n if you add up all the probabilities you get one all of them are positive or non negative and they add up to 1 which means this is a valid probability mass function. So, this formula for binomial expansion is important to know and it appears in so many forms in probability calculations again and again. So, it is good to know this n choose i very well in some sense what happens to it and all that and these kind of summations involving binomial coefficients and the binomial expansion form it is important to know in many calculations in probability.

(Refer Slide Time: 15:57)



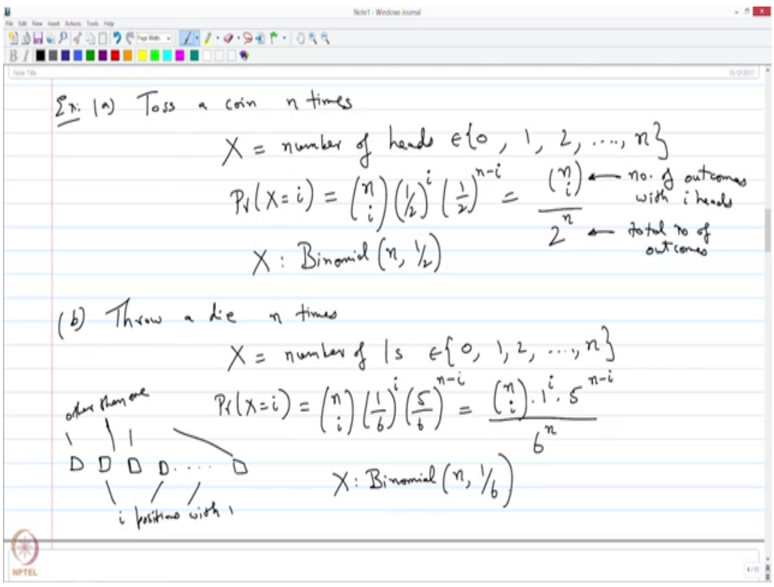
So, this is the binomial distribution. So, let me take a few examples of this and show you how it look if you the simplest case is binomial of 0 comma p. So, what is 0 comma p whatever p you pick, but if n is 0 what will happen, this is just in this case this X which is binomial 0 p will simply take just one value 0. So, if it just takes one value the probability of that has to be one in think about how we put it in. So, this one value it takes probability one it is not at all interesting or you might want to take binomial of 1 comma p. So, in this case X takes values 0 1 and if you put in your formula correctly you will simply get 1 minus p p. So, this is actually Bernoulli right binomial of one comma p is Bernoulli, it is not at all interesting it is you do not get a new random variable if you do it. So, it starts to get interesting if you put binomial of 2 comma p. So, you would get 0 1 2. So, this you will see is 1 minus p squared, this one is 2 p times 1 minus p, this one is p square. So, this is the binomial of 2 comma p.

So, likewise you can do. So, let us let me do it just at least binomial of 3 comma p and after that I will stop and let you visualize what happens in the general case 0 1 2 3, 0 is 1 minus p power 3, 1 is 3 times p 1 minus p squared, 2 is again 3 times p squared times 1 minus p this guy is p power 3. So, these are all the way in which you evaluate the binomial coefficient. So, maybe I will show you just one more example binomial of 4 comma p. So, in which case X takes values 0 1 2 3 4, 0 is with probability 1 minus p power 4 this is 4 times p into 1 minus p cubed, the 6 times p square into 1 minus p squared 4 choose 2 is 6 so how did I get 6 here for instance 4 choose 2 its 4 factorial by 2 factorial into 2 factorial that is 6 you can evaluate it. Likewise how did I get 4 here 4 choose one you can see is actually 1, 4 factorial by 1 factorial into 3 factorial.

I am using the same formula I am not doing some magic here you just use the same formula you get the answer. So, this will be again 4 times b cubed into 1 minus p this is p power 4. So, this is an example this is how the binomial random variable we look if you look at different cases.

So, that is the example. So, I am going to see some situations in which a binomial random variable comes. So, if you look at the first example a toss of coin n times, and I said the random variable X to be number of heads.

(Refer Slide Time: 18:56)



So, it turns out if you look at this number of heads the number of heads could be 0 or it could be one it could be 2 so on till n. So, the values that the X the random variable takes number of heads is from 0 to n now what is the probability that X equals i the number of heads is i the heads could occur in any one of i position chosen from n positions. So, you will have n choose i then the number of heads has to be half right. So, I am sorry half power i times number of tails this is half power n minus i. So, you get the get this probability also. So, clearly from this we see that X is binomial with n and half case, you toss a coin n times the number of heads is going to be a binomial with n comma half.

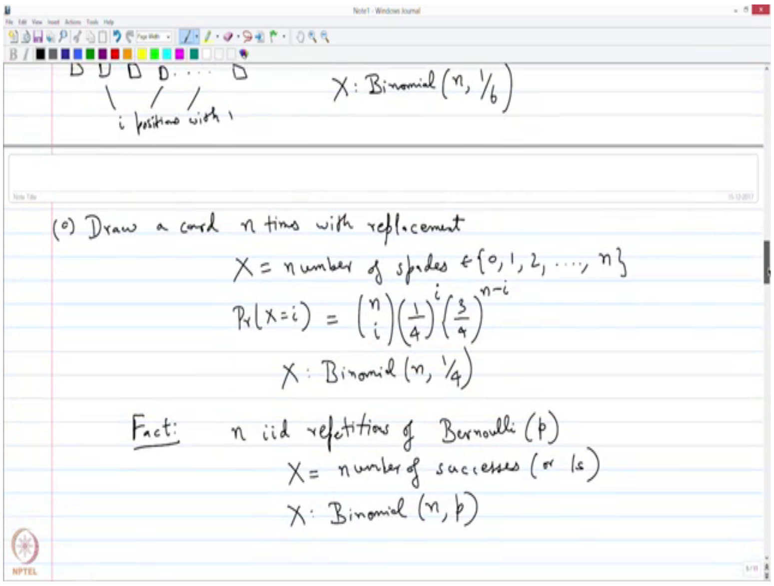
So, hopefully you see this method of how I got this probability that X equals i if you have I heads you have to count the number of possible ways in which I heads can appear so that that is a little bit. So, that is why the n choose i comes about and then you have a 2 power n in the denominator. So, this is just I mean this is just n choose i divided by 2 power n right. So, 2 power n is the total number of possibilities for n tosses I have done this before and what is n choose i, number of possibilities with exactly high heads this is total number, total number of outcomes maybe I should write outcomes here. We have seen this before. So, the n choose i by 2 power n 2 power n I can write as half power i into half power n minus i. So, this comes into that familiar binomial form. So, you have X being binomial with parameter n and a half. So, number of heads that you get when you toss the coin n times is a binomial random variable.

So, I can now repeat this same sort of experiment with. So, anytime you repeat a Bernoulli experiment n times independently you get a binomial distribution. So, for instance you throw a die n times and you let X be the number of 1s, once again this takes values 0 1 2 3 to n and probability of X being i, now there is a little bit more confusing. So, we cannot maybe may be do this in such a simple fashion. So, you should you have to think about the calculation a little bit more, but you can argue that this will be n choose i into one by 6 power i times 5 by 6 power n minus i.

So, how do you write this down? So, for instance the total number of possibility is 6 power n. So, when you throw out dice n times if 6 power n. Now, how many possibilities how many outcomes have exactly i ones? The i ones could occur in any one of the any i of the n position and n throws you have n throws any i of the throws can be one. So, you have n choose i possibilities times. So, for the i guys you have just one power i right. So, one has to occur for the i positions there is no choice for the remaining positions you have it could be any one of the 5 possibilities what are the remaining 5 case 2 3 4 5 6 any one of the 5 could be there hopefully you see that. You have n throw of the dice, there are i position with 1 remaining positions have other than 1. So, how many ways in which you can choose these i positions, n choose i, the i positions themselves have just one possibility all the remaining positions. Other than have to have other than one right other than one how many possibilities are that is 5 possibilities 5 power n minus i. So, that is the same as this case.

So, this clearly shows that X is binomial with parameter n and 1 by 6. Same thing with drawing the cards, but when you do they repeat the experiment iid you have to draw the card with replacement. So, that is the next thing I want to show.

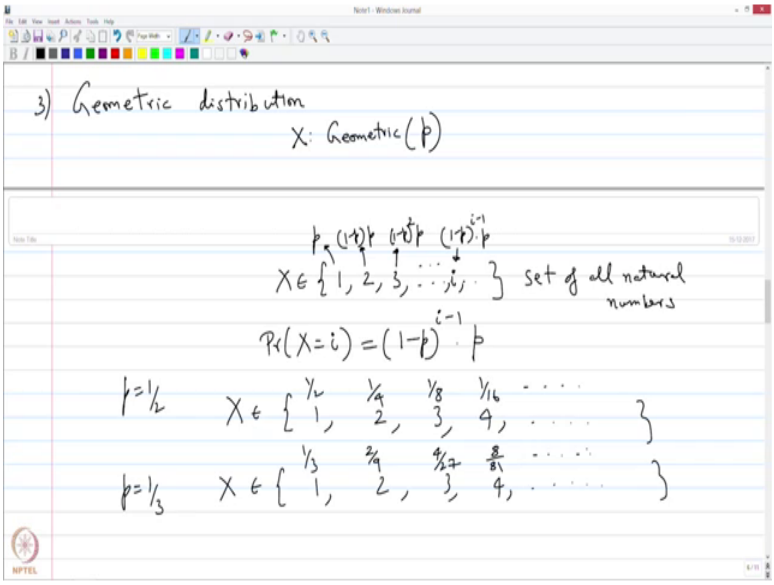
(Refer Slide Time: 23:53)



If you have draw a card n times with replacement, this with replacement is very crucial otherwise you will not get binomial and you let X be the number of spades. Once again this X also belongs to 0 1 2 3 to n and once again with the careful calculation you can show probability that X equals i is n choose i 1 by 4 probability of spades in the i position and 3 by 4 which is probability of non spades in the remaining positions. So, this will show that X is binomial with n and 1 by 4.

So, this is an useful factor remember if you have a Bernoulli distribution you repeated n times independently and count the total number of successes you always get a binomial distribution. So, this is an important fact and you have just shown this by experiment n iid repetitions, repetition of Bernoulli p and you let us set X as the number of successes. Bernoulli is either failure or success, success refers to 1 failure refers to 0. So, the number of success or ones then X is binomial with n comma p. So, this is a very important fact and we have verified this with quite a few examples. So, that is the binomial distribution.

(Refer Slide Time: 25:47)



The next important type of distribution that we will consider is this geometric distribution. So, quite often and many examples and real life situations you have you may not be concerned with the total number of things. That sometimes some sometimes occurs you might want to know the first time an event will occur the success will occur. So, you start the experiment keep repeating it what is the first time that successful occur. So, if you do that you end up getting what is called a geometric distribution.

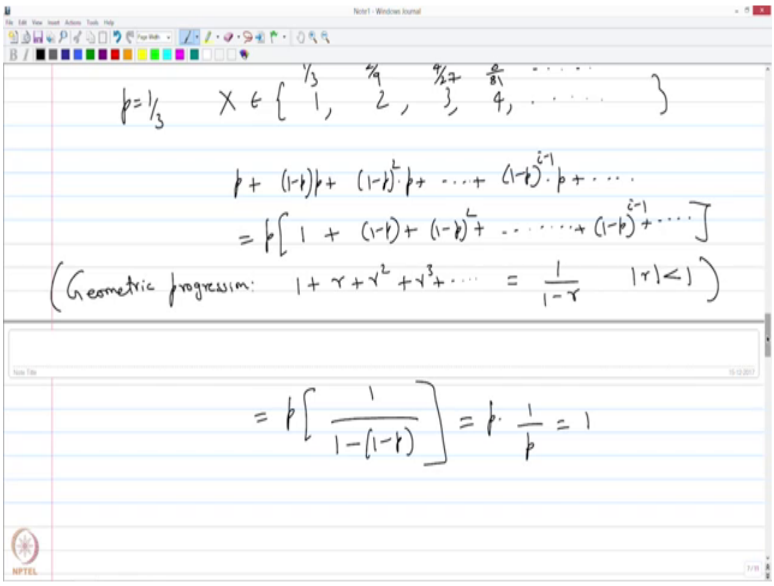
So, let me define the geometric PMF once again. So, the geometric PMF has a just one parameter p it does not have a multiple parameters just p it takes values. So, there are 2 ways in which people write geometric PMF sometimes they start at 1, some people start at 0. So, I am going to start at 1 so on; so on I mean there is no end here so X this is the first step of random variable which is not finite, but Bernoulli is finite as and it takes only 2 values 0 or 1. Binomial is also finite it takes only n values n is a finite number. Geometric can take any large number as its value it could be 1000, 10,000, 100,000 whatever there is no limit to the maximum value that a geometric random variable could possibly take. So, 1 comma 2 comma 3 so on set of all natural numbers.

And what is the probability that X equals i some number here. So, I is some arbitrary number here turns out this is 1 minus p power i minus 1 times p. So, that is the probability that we assign. Remember what is my distribution? Distribution was 2 things all the values that the random variable can take and the probability with which it takes that value. So, what are the values a geometric distribution is going to take 1 comma 2 comma 3 so on what is the probability that the random variable takes a particular value I it is 1 minus p power i minus 1 times p.

So, once again this value is nonnegative it is between 0 and 1 and between its greater than or equal to 0 and we have to check whether all these numbers add up to 1 right. So, remember what is the probability we will write here. So, now, if you write the probabilities on top of these guy this will be just p this will be 1 minus p times p, this will be 1 minus p squared times p, so on till what will be this, this will be 1 minus p power i minus 1 times p and so on. So, this is the probability of a geometric random variable. So, for instance if you take if you take p equals half, X takes values 1 2 3 4 so on, probability of this is half probability of 2 is 1 by 4 right p equals half if we put we will get 1 by 4, this will be 1 by 8, this will be 1 by 16 and so on. The probability keeps on dropping very very fast.

So, you can also take p equals 1 by 3 let us say in which case X will be 1 2 3 4. So, on probability of one is 1 by 3 this guy is what 2 by 3 times 1 by 3, so you have 2 by 9. What about 3? Its 2 by 3 squared times 1 by 3, so it will be 4 by 27. What about 4? Its 2 by 3 power 3 times 1 by 3, so you will have 8 by 81 so on. So, that is how the probability goes. So, particular values of p you can figure it out. So, to do all the probabilities add up to one I have to check that otherwise it will not be a valid PMF, so let us try and do that.

(Refer Slide Time: 30:04)

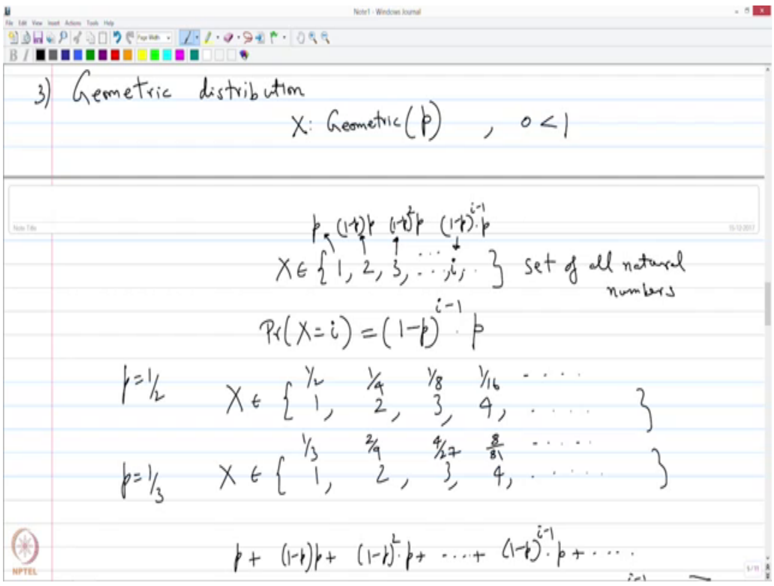


So, you have p time p plus 1 minus p times p plus 1 minus p squared times p plus so on till what is the ith term 1 minus p power i minus 1 times p plus so on now p appears is a common factor here. I can pull it out I have 1 plus 1 minus p plus 1 minus p squared plus so on right the arbitrary term is 1 minus p power i minus 1 it does not matter. So, it goes on and on and on forever.

Now, there is this result on geometric progressions which says that 1 plus r plus r squared plus r cubed plus so on is actually equal to 1 by 1 minus r if absolute value of r is less than 1. So, we will put 0 it does not matter if absolute value of r is less than 1, should be strictly less than 1. So, if you have a summation like 1 plus r plus r squared where r is strictly less than 1, r could be half or 1 by 3 or something then this summation is 1 by 1 minus r. So, that is what it means.

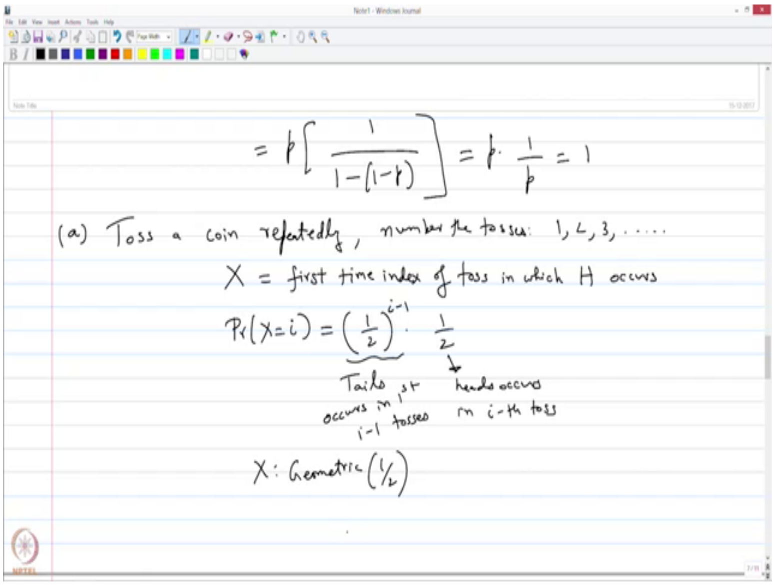
So, if you use this result of geometric progression then the summation becomes p times 1 by what is r here, 1 minus p and I know this 1 minus p is absolute value is less than 1. So, of course, you have the situation I mean for instance p equals 0 as a sort of a bad situation here p is a 0 nothing will really work out you cannot you have to have p being strictly positive in some sense. So, we will assume that. So, we will take p to be greater than 0 between 0 and 1 ok. So, we will take it like that. So, things work out well by source. So, it is 1 by 1 minus r. What is r? 1 minus p.

(Refer Slide Time: 31:50)



So, you notice that this summation has r equals 1 minus p. So, we have p times 1 by 1 minus p which is what p times 1 by p. So, it will cancel and you will get one. So, the summation works or 2 or something that is correct. So, this is the geometric distribution.

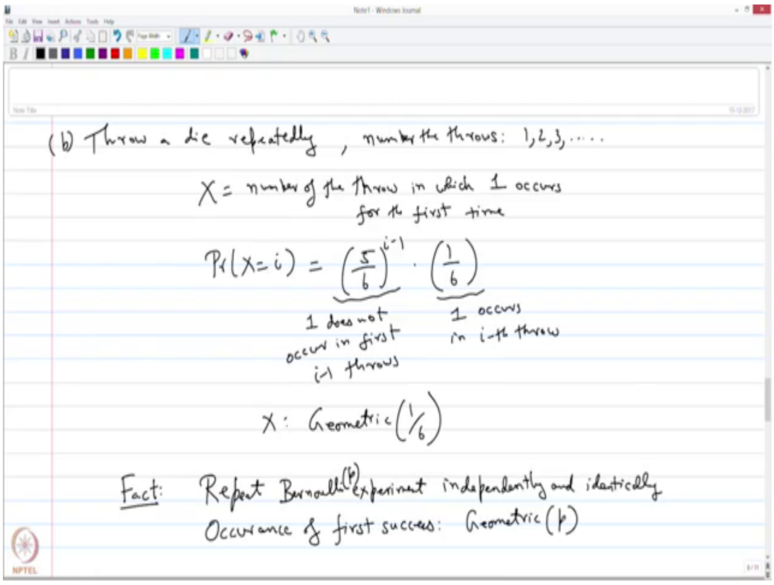
(Refer Slide Time: 32:26)



So, let us once again see some of our examples and I think even Professor Aravind are also mentioned this example if a toss a coin repeatedly and if you set X as first time index also toss a coin repeatedly. So, you also have to number the tosses, you number the tosses as 1 2 3 so on. So, first toss, second toss, third toss so on, so the first toss, second toss, third toss and so on the first time index of toss in which H occurs. So, this is the way in which a geometric random variable solutions show up. So, you have 1 2 3. So, on the first time that H occurs and you can quickly see probability that X equals i with the first time head occurs in the ith toss is half power i minus 1 times half. So, this means. So, this usually means this what, what this means is tails occurs in first i minus 1 tosses right. So, first time head occurs must be in the ith toss. So, in the first i minus 1 tosses you should get tails. So, that is the probability of half i minus 1 and here heads occurs in ith toss.

That is another half. So, this is exactly this. So, you have half power i minus 1 times half. So, X is geometric with half. So, this is hopefully easy to see.

(Refer Slide Time: 34:18)



And then a second example I will say is throw a die repeatedly and again number the throws 1 2 3 so on. So, what is X going to be now? X is the number of the throw in which one occurs for the first time right. So, this is the first I am going to consider success if one comes that is a success and the number of the throw in which one comes is this is going to be my geometric random variable. And here again if you look at probability that X equal to I you will get 5 by 6 power i minus 1 times 1 by 6. So, once again what is this? This is the probability that one does not occur in first i minus 1 throws and then this is the probability that one occurs in ith throw. So, clearly we see that this X is sorry geometric with parameter 1 by 6. So, you can again do the card experiment also you can repeatedly keep drawing cards with replacement from a pack and you will see the first occurrence of spades will be a geometric probable distribution with parameter 1 by 4. So, I am not writing into or in detail.

But in general here is a fact. So, if we repeat a Bernoulli experiment independently and identically, independently occurrence of first success becomes, I should say parameter p here first success is geometric with probe parameter p. So, this is a fact in general this is also true.

So, hopefully these examples give you some simple real life situations where the geometric can the Bernoulli and binomial distributions occur and so I think it is important to know these distributions very very well what kind of values they take, what probabilities with which they take them and how to identify situations in which you can apply them.

Thank you very much, that is the end of this lecture.