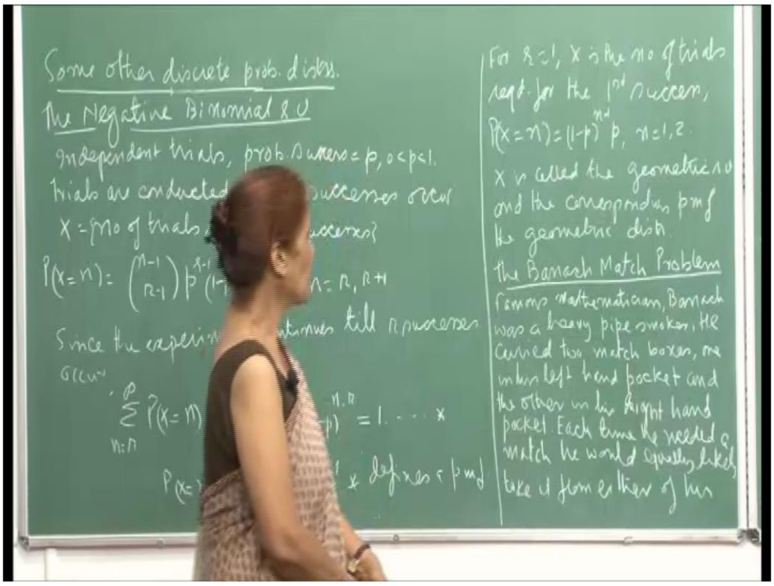
Introduction to Probability Theory and its Applications

Prof. Prabha Sharma
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture - 7
Discrete Random Variables and Their Distributions

(Refer Slide Time: 00:15)



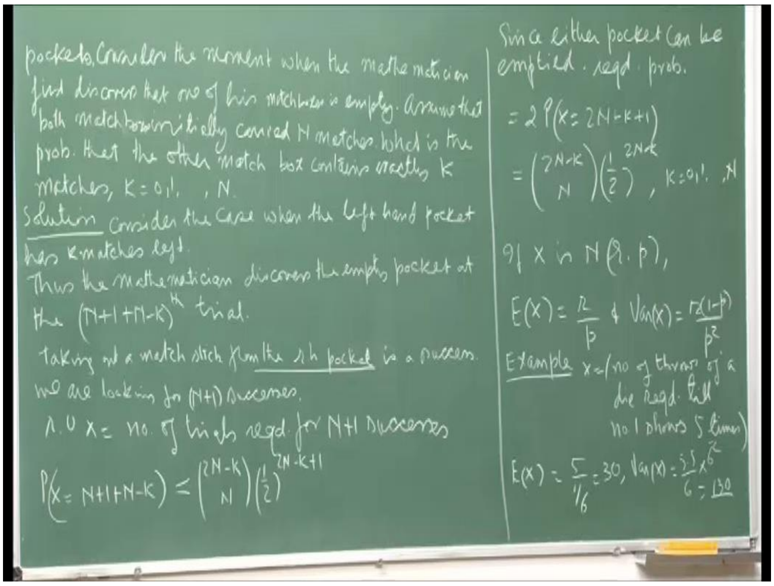
So, I will be taking about some other discrete probability distributions. So, the next one is negative, the negative binomial random variable, and this name is very suggestive. We consider the binomial distribution, which where the random variable, was representing the number of successes, where the number of trials is fixed. So, you perform n trials, probability of success is p, and then we asked for the probability of r successes. Now here it is the reverse, what we are saying is that independent trials are performed. Probability of success is p, p between 0 and 1. Trials are conducted till r success is occur. So, now it is the opposite; that mean now you go and conducting the trials, till r success is have occurred, and so x will be the number of trials required for r successes. In the binomial case, the trials were fixed, and you have to then ask for the probability r successes. Here the number of successes is fixed, and you are saying, what are the numbers of trials that you are required, so that r success is take place.

So, here the random variable is the, number of trials. So, probability x equal to n and it is very simple to write it down, because so that means, see you will stop your experiment, the moment you hit the r x success; so that means, up to n minus 1 trials. If I want r successes occur in n trials, then up to n minus 1 trials r minus 1 successes should have occurred, and therefore the probability of that is n minus 1 choose r minus 1. Here we are using the binomial concept, and then p r minus 1 r minus 1 successes and 1 minus p n minus r failures, and then the last one is the success. So, therefore, the total number of successes add up to r, but you are here wanting r minus 1 successes to occur anywhere up to n minus 1 trials, the last trial must be… So, the n'th trial in this case must be a success. So, your n can vary from r r because you want r successes. So, at least r trials have to be conducted. So, therefore, n is r r plus 1 and so on, and this number can go on up to infinity.

Now, since the experiment continues till r successes occur. So, therefore, when you add up n equal to r to infinity probability x equal to n, this must add up to 1, and therefore, this is a combinatorial argument, to say that what we have defined here, is a probability math's function. Analytically also 1 can show, that this sum will equal one, but then you require little more mathematics. So, therefore, we just satisfy ourselves, by giving this combinatorial argument, that we will continue the trials still a success occur; the r success occur. And also probability x equal to n is non negative for n varying from r r plus 1 n and so on, so this defines a p m f, so this is a valid p m f: When r is equal to 1; that means, if you just looking for the first success, and x is the number of trials required for the first success, then obviously, you put r equal to 1, so that becomes 1 and so this is 1 minus p n minus 1 into p; that means, the first n minus 1 trials must end up in failure.

So, therefore, 1 minus p rise to n minus 1, and the moment you hit a success you stop. So, 1 minus p rise to n minus 1 into p and n can vary from 1 onwards. x is now called the geometric random variable, and the corresponding p m f is the geometric distribution. Now, interesting application of a negative binomial random variable, and the distribution. So, this is known as the Banach match problem. Banach was a very famous mathematician, and he was a heavy pipe smoker. So, he does not waste time in looking for the match box and then lighting up a match, to light his pipe. He would carry 2 match boxes; one is in his left hand pocket, and the other one in the right. So that wherever he puts his hand, he gets a match box and he lights his pipe. Therefore, that is how he was dependent on smoking his pipe. So, each time he needed a match, he would equally likely take it from either of his pockets.

(Refer Slide Time: 05:21)



So, it was equally likely that he would put his hand in the right hand pocket or on the in the left pocket, so probability of that is the same. And now consider the moment when the mathematician first discovers, that 1 of his match box is empty. So, now, assume that both the match boxes initially carried N matches. So, we are asking what is the probability that the other match box contains exactly k matches. So, that means, initially both had capital n number of match sticks, in the match boxes, and then when he discovers that 1 of the match boxes is empty, and the other 1 is contains exactly k matches. So, this is what we have to find out the probability of this event. So, let us say that, consider the case when the left hand pocket has k matches left; that means, right hand pocket is empty, and left hand pocket has, the match box has k matches left.

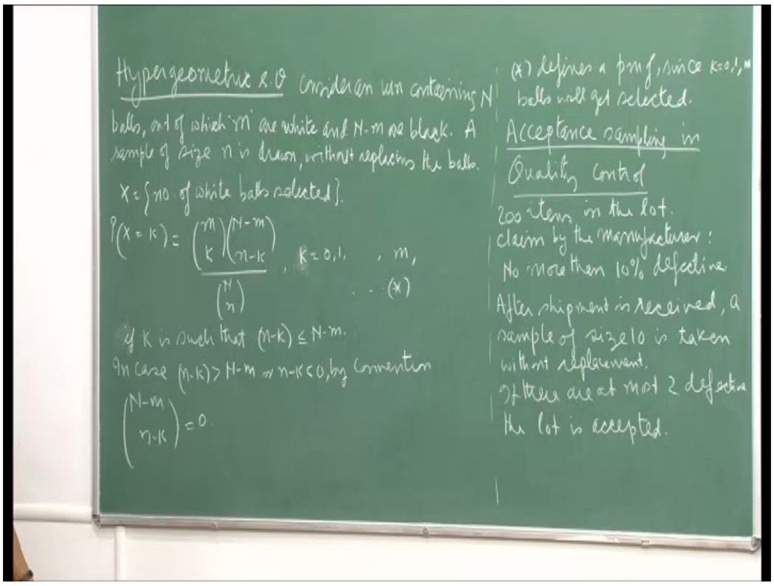
Here again I should say when the left hand pocket match box, is having k matches left, so I have left out that match box; that is the mathematician discovers the empty pocket at the. see he emptied the n match sticks in the right hand pocket, and here he emptied n minus k, because k are left in the left hand pocket, and then on the n plus 1 nth, when he put his hand in the n plus 1 nth, or when he takes out the match box at the 1 plus nth time. I mean from the right hand pocket, then he discovers it is empty. So, the total numbers of trials are n plus n minus k plus 1, so exactly situation of a negative. So, taking out a match stick from the right hand pocket as a success. So, we will (( )) is a success. So, let us say that. So, we will say that, the right hand pocket, taking out or putting the hand in your right hand pocket, and of course, when you put as long as the match box has a stick, you are also taking out stick match stick. So, therefore, whichever way you want to put it, but anyway taking out a match stick from the right hand pocket is the success.

Here looking for the n plus one th success, but where actually he discovers that it is empty. So, that a consequence of the match box getting empty, but essentially what we are saying is that putting his hand in the right hand pocket is a success, and putting his hand in the left hand pocket is a failure. So, random variable x which is equal to the number of trials required for n plus 1 successes. So, this is the exactly the case for a negative binomial distribution, and what we are saying here is, that this is actually; that means, this is happening, when you have add this many trials n plus 1 plus n minus k and therefore, by our argument, see the last one he discovers that it is; that means, at the 2 n minus k plus one'th trial, he discovers that the left right hand pocket match box does not have any sticks left in it. So, this would be 2 n minus k, if you n minus 1 r minus 1 and then 1 by 2 rise to 2 n minus k plus 1, so this will be the probability.

But then since, either of the pockets could have emptied first. So, therefore, what we have saying is twice this. So, the required probability is, since either pocket can be emptied, required probability is twice of that, and so when you multiply by 2 this 1 disappears the required probabilities 2 n minus k choose n into 1 by 2 rise to 2 n minus k. Now, again these results I am just giving you without, because handling this thing, requires lot of mathematics, you will not do it. So, I will just simply say that, expected value of negative random variable, which has parameters r and p; that means, r number of successes are required, and p is the probability of success; that is r by p, and the variance of a negative binomial r comma p random variable is r into 1 minus p upon p square.

So, for example, if you are throwing of a die; x is the random variable which is the equal to the number of throws of a die, required till number 1 shows 5 times. So, your r is 5 here, and p, because die we are assuming is fair die, so probability of each number showing up is 1 by 6. So, probability of number 1 showing up is 1 by 6. So, our p is 1 by 6 r is 5 and therefore, by this these 2 formulae, expected value of x is 5 upon 1 by 6 which is 30, and where is; that means, at least the expected value, the expected number of trials throws of the die required, so that number 1 shows 5 times is 30, and the variance is again by this formula comes out to be 130, 25 into 6. No this one 50. So, this is 150. So, that is about negative binomial and geometric distribution.

(Refer Slide Time: 11:27)



So, another discrete random variable which is quite useful, is hyper geometric random variable. So, I will demonstrate this through an example, and define this variable and corresponding distribution. So, consider an urn containing N balls, out of which m are white, and n minus m are black. Now a sample of size n is drawn, without replacing the balls; that is important, the experiment is conducted without replacing the balls. So, I keep taking out the balls and put them aside. So, now if x is the random variable, which counts a number of white balls selected. So, sample of size n is drawn. Total numbers of white balls are m. So, now, we look at the probability of the number of white balls being equal to k. So, k can obviously vary from 0 1 to m, because there are m white balls. So, the probability would be. Now here you see, we are using the multinomial distribution.

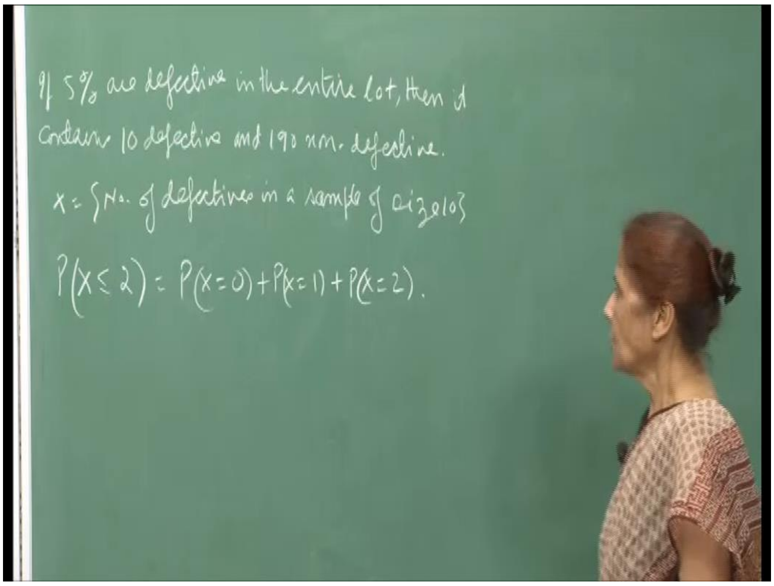
So, out of m white balls, you want to select k, and out of the remaining N minus m black balls, you are selecting n minus k black balls. So, your total sample size is n balls, and the total number of ways of selecting small n balls from N balls is n choose n. So, this we have already gone through, when we were talking about the counting procedures. So, this gives you the, total number of possible ways in which you can select n balls out of N balls, and the number of ways which you can select k white balls from m white balls, and n minus k black balls from N minus m black balls. now here of course, this is meaning, when k such; that n minus k is less than or equal to n minus m, because. See, there is some connection between m k and n.

So; obviously, if my number m is very large, then I can select k number of balls, but then n minus k becomes negative, or n minus k is, so that means, it has to be less than or equal to n minus m, but then see, by convention if n minus k turns out be greater than n minus m, or n minus k is less than 0, then by convention we say that this n minus m choose n minus k 0. So, therefore, this has a meaning. So, we do not have to worry, because then there will be the probability would be. If k such that n minus k is greater than n minus m, or n minus k is less than 0, then these probabilities will be 0, and so there will be no math's attached with those values of k. So, this is your this thing, and again; since when we are drawing out a sample of n balls, white ball may appear or may not appear, and a white ball numbering 1 2 3 up to m, may appear may not appear.

So, this takes care of all possible cases, and therefore, this is a valid p m f. So, that means, what we are saying is, that summation probability x equal to k, k varying from 0 to m is equal to 1, and also these probabilities are all non negative. So, therefore, this is a valid p m f. So, this exercise you must do every time, we define a random variable and its corresponding probability math's function. Now, let me just show you an example of, where we use, where we make use of hyper geometric random variable and its distribution, acceptance sampling in quality control. So, what you do is, of course, I have give you the numbers here are small, but usually the numbers are much bigger than what I am using here. So, suppose 200 items in the lot; some instrument or something is being delivered by a manufacturer, the whole lot is of size 200, and the claim by the manufacturer is, that no more than 10 percent are defective. So, this is the claim.

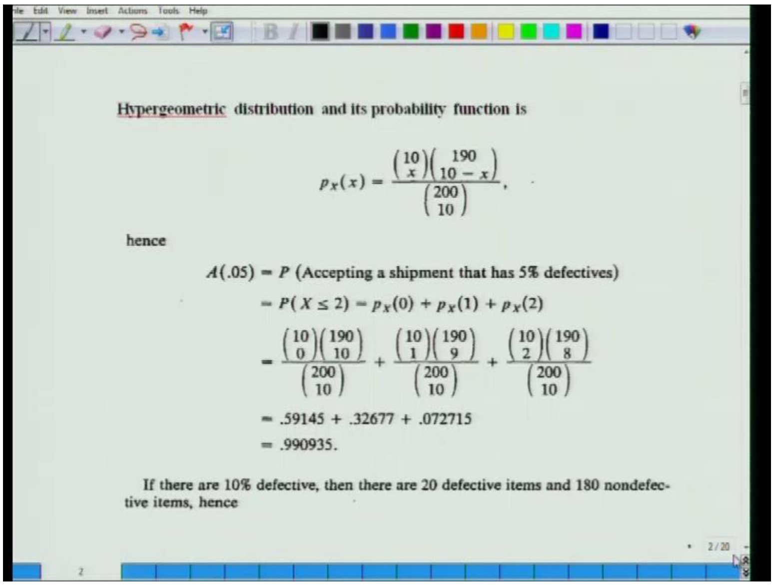
Now; obviously, people do not have time and energy and man power, to actually inspect all the 200 items, and usually this number is very big. So, what the practice is, and that is why it is acceptance sampling. So, what you do is, after the shipment is received, a sample of size 10 is taken. Again the numbers are all, just for convenience sake, but usually there will be more realistic numbers. So, anyway a sample of size 10 is taken without replacement, and if there are at most 2 defective, the lot is accepted. So, you just at random choose a sample of size 10, from this whole lot of 200 items, and then you inspect those 10 items, and you have taken out the sample by without replacement. So, you inspect those and then in that sample of size 10, if you find 0 1 or 2 defective you will accept the whole lot, and you will say that it is ok, and then if there more than 2 defective, then you will reject the lot. So, this is what you call acceptance sampling in quality control.

(Refer Slide Time: 17:33)



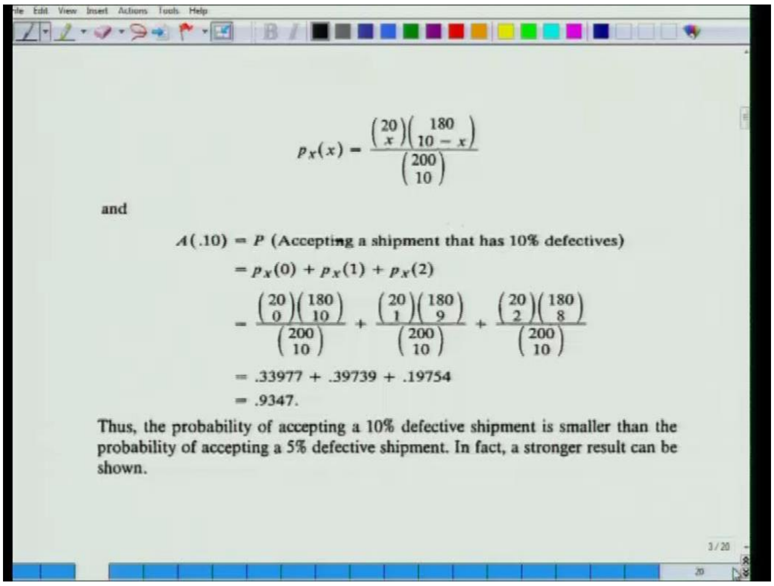
So, let me just show you simple you know combinational exercise here I will do. So, that means for example, if 5 percent are defective in the entire lot, then it contains 10 defective and 190 non defective. So, here the claim by the manufacturer is that no more than 10 percent are defective. So, let me consider the case when 5 percent are defective in the entire lot, then it contains 10 defective and 1, because 5 percent of 200 is 10. So, 10 are defective and 190 are non-defective. So, now if you want to compute the number of defectives in a sample of size 10, x is this. Then you want to compute the probability let x is less than or equal to 2, which means probability x equal to 0 plus probability x equal to 1 plus probability x equal to 2. So, this is what you want to compute, and I will just show you the calculations here.

(Refer Slide Time: 18:36)



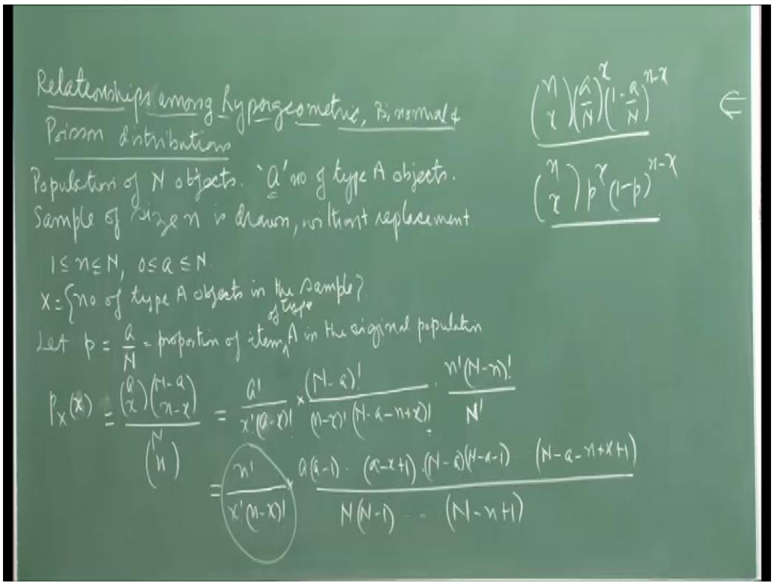
So, what I am showing you here the probability, of number of defectives being x. So, that will be, this is the case when we are, say assuming that 5 percent are defective in the whole lot. So, then it is 10 choose x and then 190 are non defective. From the non defective am choosing 10 minus x, and then divided by 200 choose 10. So, this is a hyper geometric probability of choosing x defective; the samples size containing x defective, our sample size is 10. Hence probability of accepting a shipment that is 5 percent defective. So, we are saying, if at most 2 are defective in the sample size of 10. So, that probability would be probability x equal to 0 x equal to 1 and x equal to 2, which I have written down here, and if you look at the numbers. These are the 3 probabilities you add them up, and they come out to be 0.99 0935; that means, the probability of accepting the whole lot is 0.99. Now, if there are 10 percent defective, then there are 20 defective in the whole lot and 180 non defective.

(Refer Slide Time: 20:01)



So, in that case, the hyper geometric probability of x defective in your sample of size 10, would be 20 choose x ,180 choose 10 minus x, because you are taking a sample of 10 and divided by 200 choose 10, so this will be the probability of having x defective in your sample of size 10, which is taken without replacement. So, therefore, in that case the probability, of accepting the shipment that is 10 percent defective. Again, we want number of defective in the sample, to be not more than 2, so it can be 0 1 and 2. So these are the numbers, and here the probability is 0.9347. So, this is less than the probability that we attain for, when the shipment had 5 percent defective. That is the probability of accepting a 10 percent defective shipment is smaller, then the probability of accepting a 5 percent defective shipment. So, obviously, the less the defective, the more the chance of accepting the shipment, because the probability of getting 2, at most 2 defective in a sample of size, 10 will be smaller, if 5 percent are defective and when 10 percent are defective, this is what it is saying, and therefore, you can experiment with other values of the number of defective items in the whole lot, and then you can compute the probabilities accordingly.

(Refer Slide Time: 21:34)



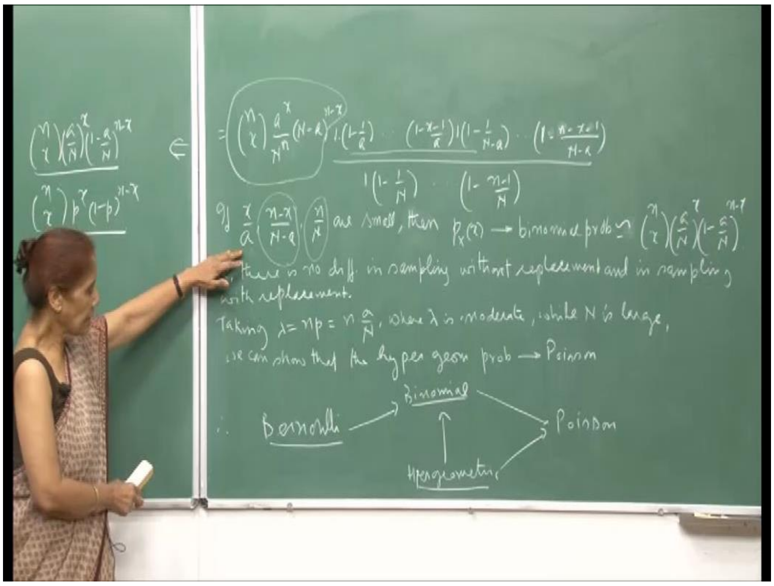
So, relationships among hyper geometric, binomial and Poisson distributions. So, let me show you. I have already shown you, how binomial will approximate to Poisson, when n is large and the number of trials is large, and your n p converges to moderately small number. So, now here let me show you the inter connection between all these 3; in fact, 4 discrete random variables that we have discussed so far. So, now, again we say that population is an objects, and a number of a number of type a objects are there in the population, so the others are not a type. And the sample of size n is drawn without replacement. So, sample of size n is drawn without replacement, where of course, n has to be between 1 and N, and a has to be between 0 and N, and while discussing the hyper geometric series, I told you that even if this number is bigger than this, and by convention this is 0, so there will be no mass.

So, we are really do not have to worry about the values of, as to what the, when I am writing this. hu sorry This is x. So, now, if I want to, and let p be a by n. This is the proportion of item, I am not using it now, I will be using it later on anyway. This is the proportion of items, item of type a, in the original population. So, I am denoting this by p, which I will make use of later on. Now, this probability; that is your sample size of n, contains x objects of type A. So, that will be choosing from a num x. a is the total number of type a objects, you want to select, so your sample should have x of them. So, this is a choose x, then remaining the population n minus a are the other kind of objects, from this you are choosing n minus x other kind of objects divided by the total ways in which you can choose a sample of size n from N. Now just write down these expressions.

So, this will be a factorial divided by x factorial a minus x factorial. This will be n minus a factorial divided by n minus x factorial n minus a minus n plus x factorial. And then this is the denominator, so this flips over, and you get here in the numerator n factorial n minus n factorial divided by n factorial. So, let us simplify this expression what I will do is, this n factorial comes here, then x factorial and n minus x factorial. This I put together and you can see that am heading towards the binomial distribution, and then you see here, a factorial divided by a minus x factorial. So, the terms after a minus x plus 1 will cancel out. So, you will be left with a a minus 1 a minus x plus 1, from here and this is gone.

And similarly here this many terms will go away. So, you will because left with n minus a n minus a minus 1 up to n minus a minus n plus x plus 1. One more from here, that will be up to that many terms. And then here similarly; n factorial I have used. Here now here n minus n factorial, so those terms will cancel out and you will be left with n n minus 1 up to n minus n plus 1. So, this is ok from here to here, this is fine. Now what I am doing is. So, this term this I can write as n choose x, then from here if I remove a from each of this, and this is a rise to x, because this is a a minus 1 a minus x plus 1. So, x minus 1 is comes from here and this is the accept a, so a raise to x this comes out. Then similarly if you remove n minus a from here, from each of them, from each of these terms n minus a, so that will also be in number n minus x, because this is n minus x plus 1, so when you include this point you will get this. So, this is 1, and then this is this, and similarly here it was, would be 1.

(Refer Slide Time: 26:30)



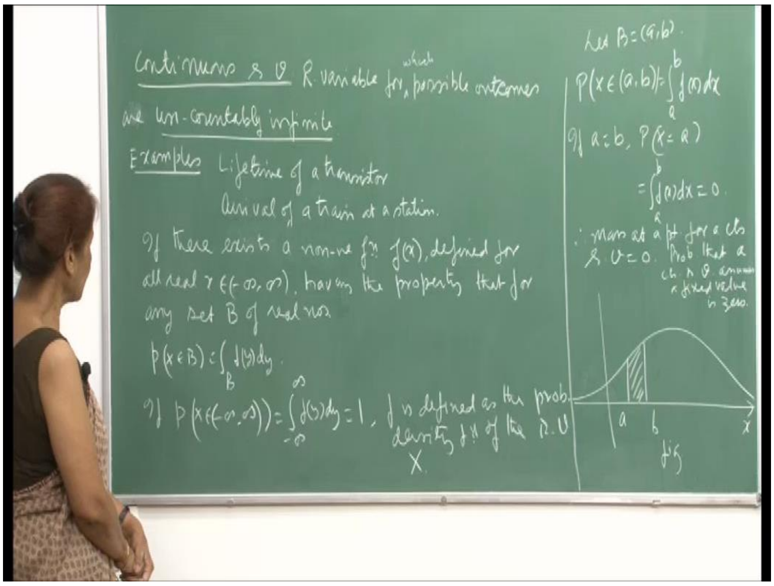
So, 1 minus 1 upon n minus a, and this is what you had, and then n rise to n we have taken, because here, this is minus n plus 1. So, you take out n from each of them and divide correspondingly. So, you will have N rise to small n, that is this here. Now, what I will do is, I should have written down this step here, what I am doing is, I will write this as, this is this. So, I will write this as maybe I will write it here. So, this is n c x, then a by n raise to x, out of this and then into divide by n. So, 1 minus a by n raise to n minus x, so all this, is equal to this, which is your binomial probability of choosing x items from n; that means, your number of defective, or number of type a items, from the sample size n x this, is probability is that. Now, this you see here, what we are saying is that x by a small; that means, a is a big number, large number; that means, the number of type a objects in your total population, is large.

Then this number is also small, and this number is also small; that means, your N is very large, in such a situation you can see that, all these numbers go to 1 all of them, and this also goes to 1. So, this whole reduces to approximately number 1. So, this probability hyper geometric probability, of choosing particular type of objects, in your sample, that probability reduces to the binomial. Now I should have, I did not properly point it out here. See when we conduct the binomial experiment, we say that, the trials are occurring independently, and then you keep counting number of successes. So, in other words in this situation, the binomial experiment would be, when you are replacing the balls, because getting a white ball is a success, and so in the binomial situation you replaced back the ball that you have taken out. So, each time the probability remains the same, of getting a white ball, which is equal to a by n; see here a by n p. So, this actually comes out to be this p raise to x 1 minus p rise to n minus x. So, this is what I got.

So, that means, the difference between a hyper geometric and binomial is, that you know for small values of the population size, it is without replacement the hyper geometric, but if you make the population size is large, and as I said all these 3 numbers should be small, in that case the hyper geometric reduces to is approximated by the binomial probability. And I have already shown you that binomial can be approximated by Poisson, where we were saying the same thing, that this should be small. So, in this case n into p, our p is a by N.

So, this number should be moderately small, then we say that for capital n being large. Then we say that the binomial can be approximated by Poisson, or the binomial probability goes to Poisson, and here I have shown you that a hyper geometric goes to binomial, and by the same argument that I have shown you here, when I talk about, when I take lambda to be n into a by n, then you can show by again manipulating the terms; that the hyper geometric will go to Poisson, already in the earlier lecture I have shown you, that Bernoulli is binomial 1 p. So, relationship is there that when you have n, when you add up n Bernoulli random variables, you get the binomial. So, this diagram shows you the relationship between the, various discrete distributions that we have discussed so far.

(Refer Slide Time: 31:26)

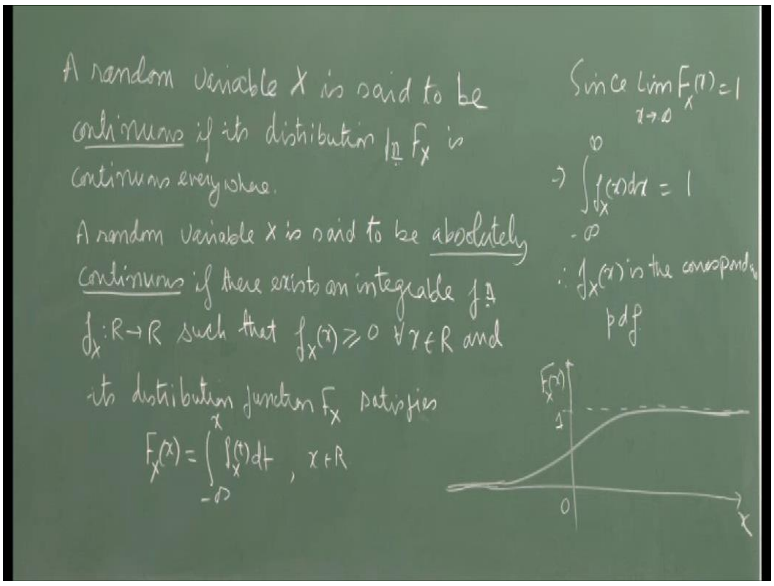


Let us now look at the other types of random variables, which are continuous random variables. So, far we looked at discrete random variables, and their special cases. Now I want to describe continuous random variables, and then again we will look at the special cases of continuous random variables. So, essentially, these are random variables of which possible outcomes are uncountable infinite. So, here you cannot count, and therefore, for example, if you take a subset in r 2, then the number of points, is uncountable, and similarly if you take an interval on the real line, and you consider all possible real numbers, that is a uncountable set.

Examples are lifetime of a transistor, because you do not know when exactly a transistor will fail, but see you might say that, because you have finite clock, so you can say, it failed at this, you can actually say that the lifetime is finite, but then it depends on your counting system. I mean as fine as you make it then you know the lifetime, you can treat this as a uncountable infinite; the arrival of a train at station and so on. So, 1 can go on adding list to this, and as we go through the topic, we will come across so many continuous random variables. Now, one way to define a continuous random variable would be, that suppose there is a non negative function f x, define for all real x on the real line minus infinity to infinity, having the property that for any set b of real numbers.

The probability that x belongs to b, is integral of the function, this non negative function f by d y over b. So, here by our definition we are saying that if x belongs to the whole real line, then the probability of that will be integral minus infinity to infinity, and f by d y d y is 1. So, we are putting this condition, and therefore, by definition f is now known as the probability density function, as a post to probability math's function, because now this is a continuous case. So, we differentiate between the continuous and discrete by. So, in this case the function is probability density function, and for the discrete case, we called it probability math's function. So, this is for the random variable x. So, if there is a function like this, and if a it satisfies these conditions. So, it is a non negative function, then we say that f is the probability density function of the continuous random variable x.

(Refer Slide Time: 34:31)

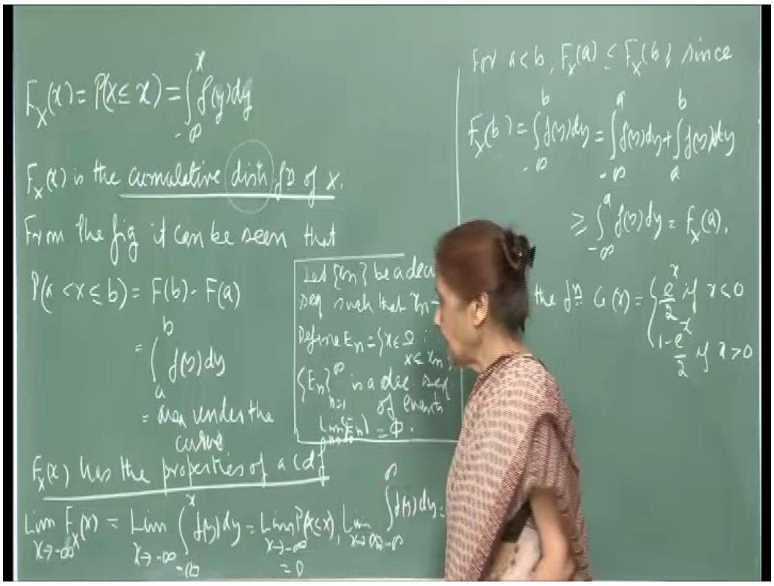


So, let me give you some more idea about continuous random variable. So, you know one can also define a random variable x, as a variable whose distribution function is continuous everywhere. So, actually the name continuous random variable, has come from here, because the distribution function of a continuous random variable is continuous; that is why we call the random variable continuous. So, actually there is nothing about, you know calling a random variable continuous or discrete. We actually say random variable is discrete, because it is cumulative distribution function is discrete; that means, it has jumps, and we say that a random variable is continuous, if its distribution function is, f x is continuous. So, this is everywhere; that is 1 definition.

Another one can be, random variable x is said to be absolutely continuous, if their exist an integrable function f x from r to r. So, that f x is non negative, for all x belonging to r, and its distribution function f x satisfies the equation that f x of x is equal to minus infinity to x integral of f x t d t, x belonging to r for n for n real number. That means, the probability x less than or equal to small x, that probability is obtained as integral of minus infinity to x f x t d t x belonging to r. And then if you see that, since the distribution function has the property that limit f x x as x goes to plus infinity is 1. So, therefore, you see the value of this integral minus infinity to infinity f x t d t will also be equal to 1, and hence these function small f x that we are saying has to be non negative, is actually the p d f, for the function x. So, either way, either you define it through F, and then you say that f, the small f will be the p d f, or sometimes you may define this small p d f, and then small f to be the p d f and then you define the distribution function anyway.

So, the thing is that, most of the time in this course, I will not saying absolutely continuous of the time, but whatever is absolutely continuous I will call it continuous, and then I will distinguish between mixed random variable, so; that means, discrete random variable, mixed random variable, and continuous random variable. So, what I refer to, is continuous random variable, will be is actually by definition absolutely continuous, because the way we have been handling, we are defining the continuous random variables and the p d f's in this course, this definition this is a right one. I mean we are following this definition that capital f x x is equal to minus infinity to x f x t d t x belonging to r. So, whichever way, but I just thought that ,one needs to say little more then what I said in the lecture, about a continuous random variable, and of course, through examples we will come to know, quite bit more about the various kinds of continuous random variables that we come across, and their properties. Now, if b is an interval, then probability x belonging to a, b will be a to b f x d x, and this what I am trying to show you; that if this is, the curve of f x, then this implies that it is actually the area of the curve.

(Refer Slide Time: 38:26)



And of course, I have written it out here, that in terms of your yeah I will come to that. So, then and if a equal to b, then this this reduces to probability x equal to a, and that means, a to this sorry b a. So, the integral will be from a to a f x d x and by again definition of the integral, this is 0. So, therefore, mass at a point for a continuous random variable is 0; that all the probability that a continuous random variable, assumes a fixed value is 0, this is what we are saying by this. right and therefore. So, I will come to this point is that see the interval I have been writing as an open interval, but here I have written as closed. So, it does not matter, because where that you include the point a or the point b or both, the masses at the individual points at the fixed points a and b are 0; that means, no probability attains to fixed values of the random continuous random variable, therefore, it does not matter, whether I write it as this or as a I write it is in open interval.

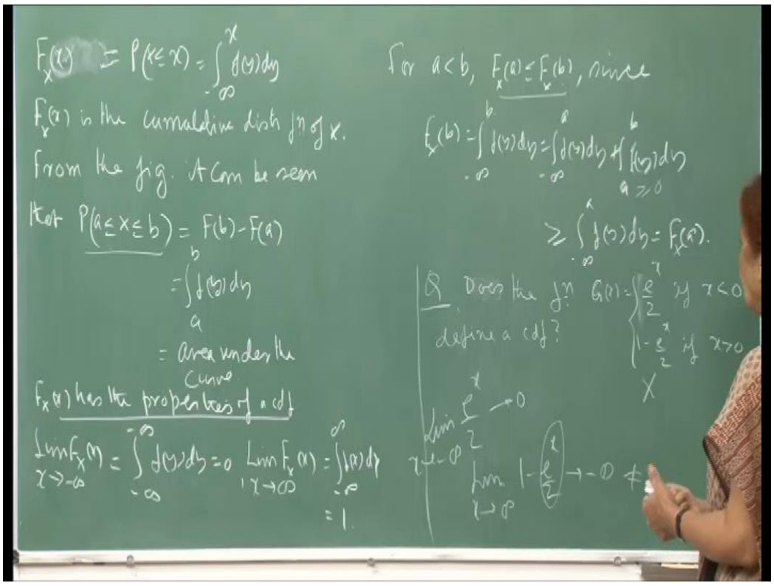
And then, since the probability x less than or equal to x. am sorry this is not correct I want to say here, this is X. So, this is now we are defining the cumulative distribution function, the probability x is less than or equal to x, and this will be in by our definition minus infinity to x of f y d y. So, this is the cumulative distribution function of x, and from the figure you can see that, this is again f b minus f a, so; that means, this will be that area from minus infinity to a under the function f x; that is f a and then this area up to b would be f b. So, you are essentially looking for the area, between in the script inside the script and so this is area into the curve. So, I have shown you that, for the x is a continuous random variable, then the cumulative distribution function, has been defined according to this, and so we can also then, have the concept of a of the probability in an interval a to b and that is the area under the curve.

Now, let me just make emphasize 2 points here, that you see as a post to discrete random variable, where it was important whether in an interval, when you are talking of probability of random variable in an interval, then whether the n points are included in the interval or not. For a discrete random variable it mattered, because every point has some positive probability, I mean are non negative. Now for a continuous random variable you see, it does not really matter, because there is no concept of a probability at a point, the idea that at a point the probability is 0. So, therefore, whether I say a less than or equal to x less than or equal to b or a strictly less than x less than or equal to b has no relevance. So, therefore, it is understood, and see that is why am writing the integral from a to b here, f y d y. So, for continuous random variable, at a point there is no concept of probability, or the mass, whatever or the density, because it is a density.

So, we measure it on an interval. Secondly, for a continuous random variable, we will call it cumulative distribution function; that is the proper notation. And here again, if some places, it may just happen that, without realizing it I may have used the word density, but does not matter. The proper notation is that it should be cumulative distribution function, when your random variable x, or in fact, for this notation, holds for x continuous or discrete random variable. So, the word is cumulative distribution function. We saw that for a discrete random variable, it was summation, and here it will be in the form of an integral, because you are computing the probability of random variable continuous random variable over the an interval. Now we want to check that f x has the properties of a c d f, and so the first thing you want to check is, that this limit of f x x as x goes to minus infinity will be 0.

So, I should have said that this is equal to 0 here, and of course, this is immediate, because the limit as x goes to plus infinity, would be this integral minus infinity to infinity f y d y, and since f is a probability density function by definition minus infinity to infinity, this integral should be equal to 1, so that part is. So, now for this, again the argument may look repetitive, we have already used it elsewhere, but let me just repeat it. So let x n be a decreasing sequence, such that x n is going to minus infinity. Now we define the events E n, which are all points in the sample space for which X is less than or equal to x n. Then you see again E n, this as n goes to from 1 to infinity, is decreasing sequence of events, and limit E n is n goes to infinity will, be empty, because you know as n goes to infinity, you will be talking of event where x is less than or equal to minus infinity. So, they can be no real number, which is less than minus infinity, and therefore, E n n goes to infinity is phi, is the empty set. And again by continuity of the probability function, you will say that probability of you know limit probability E n as n goes to infinity, is same as limit probability E n as n goes to infinity by continuity and therefore, this will be phi, because the probability of empty set is 0.

(Refer Slide Time: 44:40)



So, this properties also satisfied, and will verify the other properties also, and for to show that it is monotonic, for a less than b. This is f x, we want to show that f x is less than or equal to f b. Now since capital f x b can be written as minus infinity b then b being bigger than a, I can break up this integral into minus infinity to a f y d y plus from a to b f now. This is non-negative, f is a non negative function, this is a finite interval, because b is greater than a. So, again this number is something non negative, therefore, your probability or your f x b is bigger than f x a. So, the function F is monotonic, and therefore, it satisfies all the conditions for cumulative density function, and hence we our definition is proper here. Now the question, if you just define a function like this, and you ask whether is it c d f of the random variable x, then what do we all need to verify. We need to verify that, for minus infinity; that means, you need to verify that minus infinity to 0, E raise to x by 2 d x this is equal to what.

See it should from minus infinity to infinity it should be 1, this should, no sorry I am showing this is a c d f. So, this is not p d f am not checking is p d f. So, what is happening is, E rise to x by 2 limit as x goes to minus infinity, because this is x less than 0. So, as x goes to minus infinity this should go to 0. So, that is fine, because x goes to minus infinity E raise to x will go to 0. So, this goes to 0. Now, look at limit 1 minus E raise x by 2 as x goes to plus infinity. So, what is happening here. This portion goes to infinity. So, therefore, this is going to minus infinity, and not equal to 0. So, therefore, this does not define c d f. So, we can continue I mean 1 can try to see if any one of the condition fails to be satisfied here, we will conclude that the function that we are looking at, is not a c d f. And similarly as we have done it for the discrete case also, we check very 5, and we define continuous random variable, that the whether it is a valid p d f or not, probability density function.

Then I will also try to later on give you examples, where the random variable can be the mixed kind; that means, some for some portion of the real line, it may behave like a discrete random variable, and then for some points on the real some portion of the real line it may be continuous random variable. Now another thing that I want to point out is, that since you are defining your c d f as this. So, necessarily by the theory of integral calculus, it turns out that this has to continuous function.

So, that another way of differentiating between; that means, if for a certain part of a real line, your c d f is continuous, then we will assume that the corresponding the part of the random variable is a continuous random variable. And we saw that when it is a discrete random variable, your graph of capital the c d f has jumps, and the jumps are equal to the probability of the random variable at that particular point. So, therefore that means, you can have now c d f, which is for in part step function, and in part it is a continuous function. So, I will try to give you examples of such, and then that case we will say, that the random variable is the mixed kind. So, that means, all kinds of random variables exist, discrete, mixed kind, and continuous.