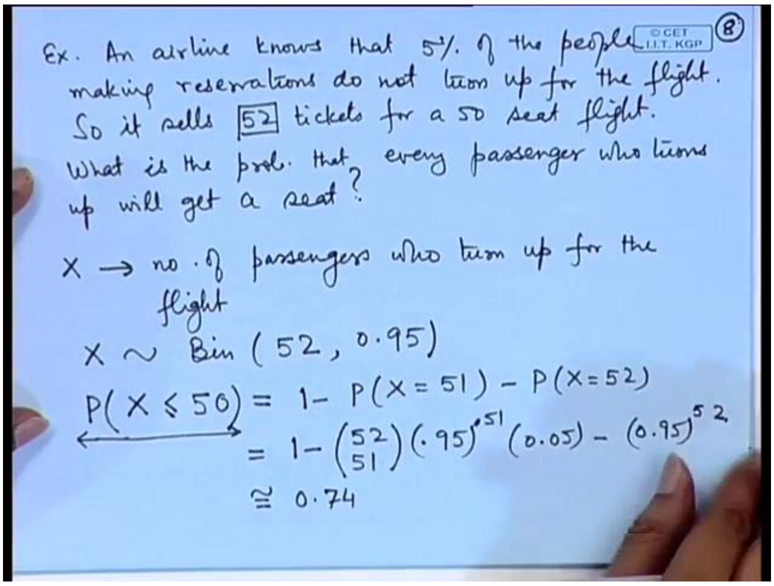
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 20
Special Discrete Distributions – II

Let us take 1 application of the binomial distribution.

(Refer Slide Time: 00:27)



An airline knows that 5 percent of the people making reservations do not turn up for the flight. So, it sells 52 tickets for a 50 seat flight, what is the probability that every passenger who turns up will get a seat?

So, let us look at this. So, if a passenger makes a booking, he may not turn up for the flight with probability 0.05 and may turn up for the flight with probability 0.95. So, if I say X is the number of passengers, who turn up for the flight. Then this can be considered as a collection of Bernoullian trials. That means, a passenger purchasing a ticket he may turn up or he may not turn up. So, turning up maybe considered as success and may not turn up as a failure or vice versa.

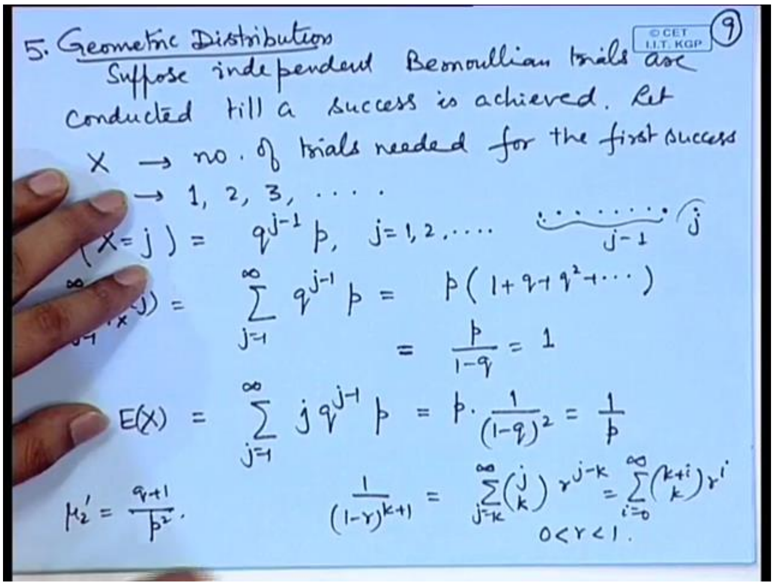
So, if we look at X as the number of passengers who turn up for the flight, then the distribution of X will be binomial 52, because there are total 52 tickets that have been sold 0.95. So, a passenger will get a seat provided the number of people who turn up is less than or equal to 50. So, we may consider it as 1 minus probability x is equal to 51 and probability x I's equal to 52. So, using the form of the binomial distribution these can be easily calculated, this is 52 c 51, 0.95 to the power 0.5, to the power 51, 0.05 minus. So, this will become simply 0.95 to the power 52. So, these values can be easily evaluated and this probability turns out to be approximately 0.74.

Now, it may sound to be somewhat reasonable that if 5 percent of the people do not turn up of the flight, then the airlines sells 52 tickets for a 50 seat flight and it may consider that we are quite safe that there will be hardly any complaints, but you see the probability that less than or equal to 50 passengers turn up for the flight is 0.74, you can say roughly three-forth of the passengers or 75 percent of the passengers, so what does it mean?

That means 25 percent of the times there will be a case, where a passenger turns up for the flight and he will not get a seat. So, which is not a very good advertisement for the flight he as; that means, they must not sell more tickets basically, maybe they will sell only 51 tickets that maybe all right. Now if we consider this Bernoullian trial in a slightly different way, here what we did is that we looked at the total number of trials, how many successes are there; but we may also have a situation where we are looking at a specified number of successes or a specified number of failures.

For example, you consider the trials for a particular drug for a certain disease. So, generally in all medical trials, the trials are done repeatedly over many subjects, and if a specified number of successes are obtained, then only the medicine is approved to be applied to the general population. So, we may find out that a particular medicine is given and there is a target group, and if say 15 people are successfully treated then we may say that the medicine is alright, suppose 100 people are successfully treated then we say that the medicine is successful.

(Refer Slide Time: 05:54)



So, here the trials are conducted till we get a specified number of successes. So, this way of looking at it, suppose independent Bernoullian trials are conducted till a success is achieved. Let X denotes the number of trials needed for the first success. So, what are the possible values of the X? You may get a success in the first trial; you may get the success in the second trial and so on. So, unlike the binomial distribution this is an infinite, but countably many valued random variable.

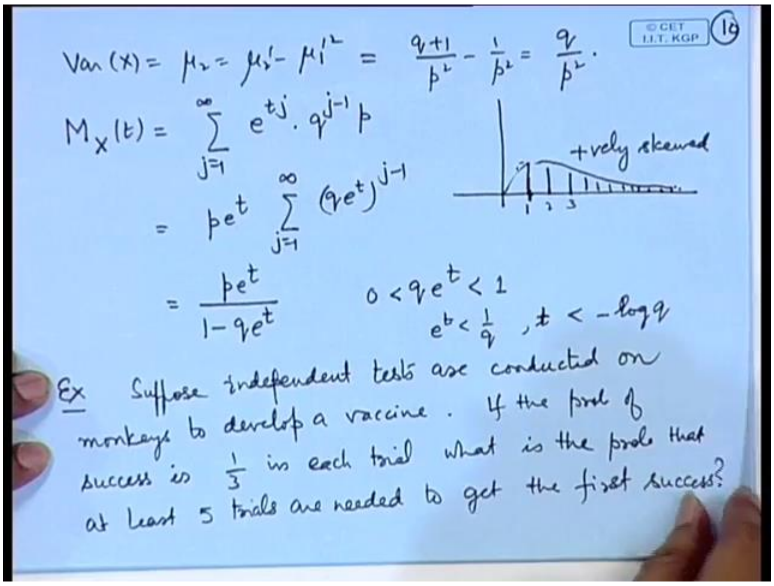
You will have probability X is equal to j is equal to; now we are saying that in the jth trial the success is observed; that means, before that in the j minus 1 trials all of them are failures and all the trials are fail, because first is a failure second is a failure up to j minus 1 th it is a failure, the jth 1 is a success. So, the probability of this becomes q to the power j minus 1 into P. So, if we look at say sigma P x j, j is equal to 1 to infinity that is equal to sigma q to the power j minus 1 into p, j is equal to 1 to infinity, that is equal to P and then you have 1 plus q plus q square an infinite geometric series this is equal to P divided by 1 minus q that is equal to 1.

Since here the terms are involved from a geometric series, so that is why this is known as a geometric distribution. So, we may look at it is various characteristics such as mean, say mu 1 prime expectation X that is equal to sigma j q to the power j minus 1 P. Now from the proof that this was a proper probability mass function, we express the sum as a infinite geometric series; that means, it is an expansion of a term of the type 1 minus q to the power minus 1; therefore, if you want to calculate higher order moments, we have to make use of the similar term. So, this 1 we can write as expansion of 1 by 1 minus q square. So, it becomes 1 by p; that means, the average number of trials needed for the first success is the inverse of the probability of success in a single trial.

So, suppose we consider tossing of a coin where the probability of head may be 1 by 3; then we require on the average 3 throws of the coin to get a head. Suppose we are having a die, where the probability of success is say 1 by 6, you will need on the average 6 trials to get the first 6.

Now, if you want to calculate say mu 2 prime and so on, then you will need to consider expansions of 1 minus q to the power minus 3 and so on. So, in general we will be making use of the formula 1 by 1 minus r to the power k plus 1 is equal to sigma j c k, r to the power j minus k, j is equal to k to infinity or we can also write it as k plus i c k, i is equal to 0 to infinity r to the power i, where r is a this is j minus k, here r is a number between. In fact, this is a mathematical formula and we need to have r between minus 1 to 1, but since here q is a probability, so it has to be between 0 and 1.

(Refer Slide Time: 10:55)



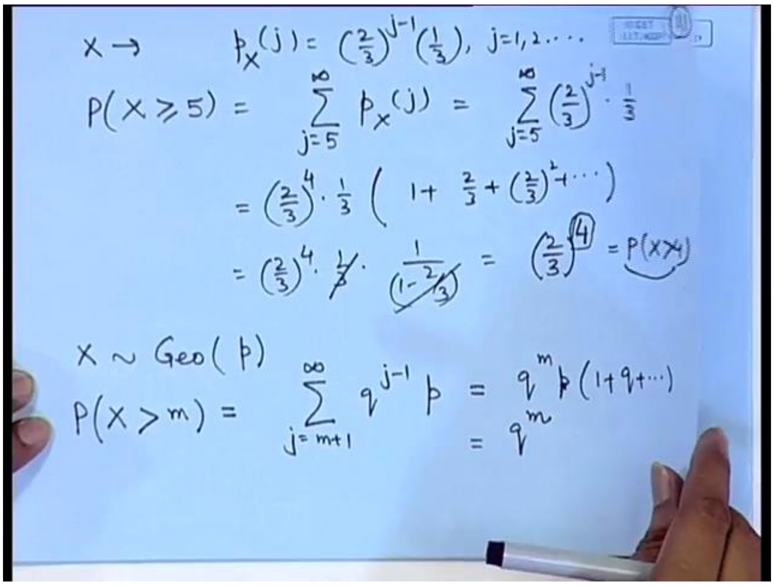
So, using this we may write other terms that is mu prime turns out to be q plus 1 by p square, and therefore we may calculate the variance of x that is mu 2, that is mu 2 prime minus mu 1 prime square and it will be equal to q plus 1 by p square minus 1 by p square, that is equal to q by p square and we may write down further higher order moments say mu 3, mu 4 etcetera to look at the shape of the distribution. However, that can be looked at from the direct description of the probabilities also.

Here the probability that x is equal to 1 is p, probability that x is equal to 2 is say p q, now since q will be between 0 and 1 therefore, this will be less than this, the probability that x equal to 3 will be q square p. So, naturally you can see that there is a decreasing trend there. So, it is a positively skewed distribution. We may also look at the moment generating function of this distribution, expectation e to the power t j. So, it is sigma e to the power t j, q to the power j minus 1 p, j is equal to 1 to infinity.

So obviously, we can see here that this can be expressed as p e to the power t, q e to the power t to the power j minus 1, j is equal to 1 to infinity, which is nothing, but the infinite geometric sum as an expansion of 1 by 1 minus q e to the power t. Now this expansion will be valid provided q e to the power t is less than 1 of course, this is positive; that means, e to the power t is less than 1 by q or t is less than minus log of q; since q is a number between 0 and 1, log of q is negative and therefore, minus log q is positive. So, in a neighborhood of t is equal to 0, this moment generated function exists. So, this can be utilized to calculate all the moments of this distribution.

Let me give one example, here suppose independent tests are conducted on monkeys to develop a vaccine, if the probability of success is 1 by 3 in each trial, what is the probability that at least 5 trials are needed to get the first success? So, this is perfectly a case of geometric distribution.

(Refer Slide Time: 14:37)



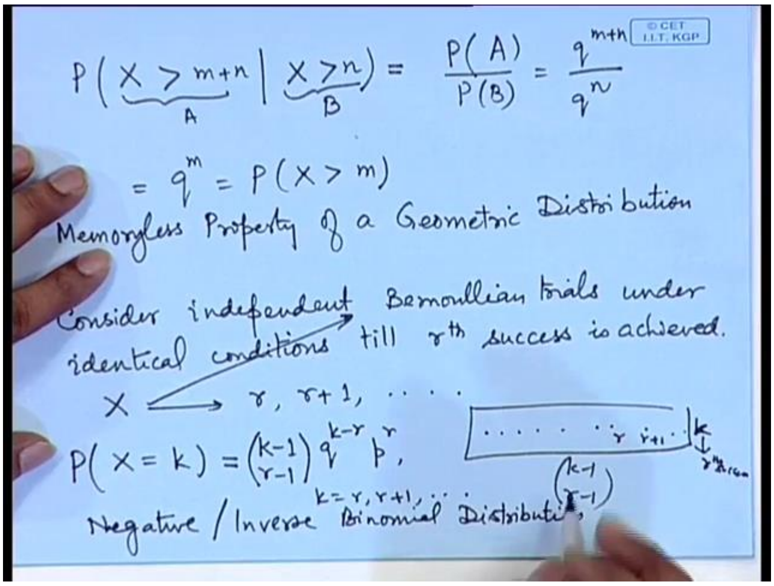
If I am considering X as the number of trials needed, then you will have P x j is equal to 2 by 3 to the power j minus 1 into 1 by 3, for j is equal to 1 2 and so on. That means, in the first trial you will get a success with probability 1 by 3, in the second trial you will get a success with probability 2 by 3 into 1 by 3 and so on.

So, what is required here is probability X greater than or equal to 5; that means, sigma P x j, j is equal to 5 to infinity, which is equal to sigma 2 by 3 to the power j minus 1, 1 by 3 j is equal to 5 to infinity. So, this is equal to. So, if we look at the first term of this series, this is 2 by 3 to the power 4 into 1 by 3. So, we can keep it out and thereafter it will become 1 plus 2 by 3 plus 2 by 3 square and so on. So, this becomes 2 by 3 to the power 4, 1 by 3, 1 divided by 1 minus 2 by 3. So, this cancels out and the probability is 2 by 3 to the power 4.

Since this is a discrete distribution, we can also consider it as probability X greater than 4; now notice here probability X greater than 4 is 2 by 3 to the power 4. So, from here we can write down somewhat more general expressions, if we are considering so x is a geometric distribution with probability p of success in each trial, if we look at what is the probability say X greater than m, then this is equal to sigma q to the power j minus 1 p, j is equal to m plus 1 to infinity.

So, this is q to the power m that is the first term here we correspond to j is equal to m plus 1, that will give us q to the power m into p and then we will have infinite geometric series 1 plus q plus q square etcetera. So, this is 1 by 1 minus q, which is cancelling with q with p, so you will get q to the power m.

(Refer Slide Time: 17:04)



Let us consider probability X greater than m plus n, given that X is greater than n. Now suppose this as event A this as event B. So, this is conditional probability of A given B, it is equal to probability of A intersection B, now here if you look at the sets A and B then clearly a is A subset of B. So, this becomes probability of A divided by probability of B. Now by the formula that we have developed just now, probability of X greater than m is of the form q the power m therefore, these probabilities will be q to the power m plus n divided by q to the power n, and this is equal to q to the power m which is probability X greater than m.

Let us look at the say statement carefully, probability X greater than m denotes, that the first success is observed after m trials, because we are saying X is the number or trials needed for the first success and I am saying X is greater than m; that means, that there is no success observed till mth trial.

Now, if you look at the left hand side, this is denoting the conditional probability that no success is achieved till a m plus nth trial, given that no success is achieved till nth trial. You look at the difference here n to m plus n; that means, mth more trials are needed, and here we started from 0 and we say that m trials are needed for the first success; that means, in m trials there is no success. So, this means that a starting point is immaterial in a geometric distribution, this is called memory less property of a geometric distribution.

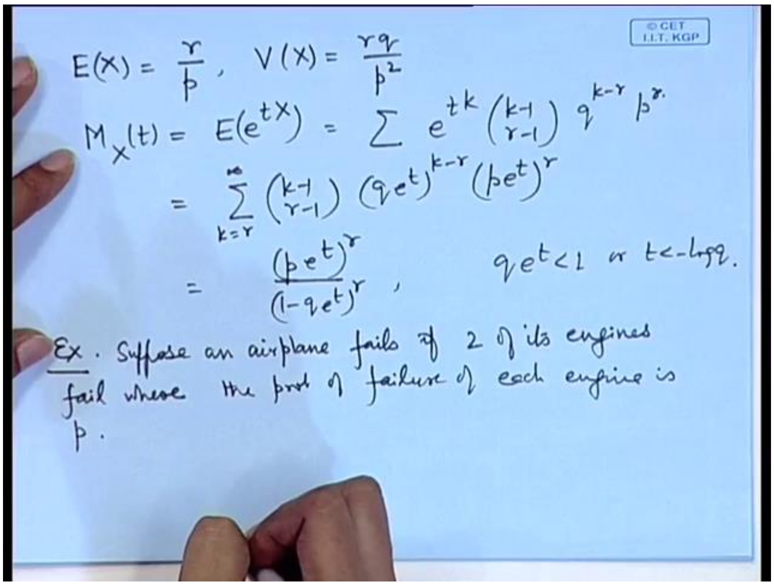
So, this is quite useful when the independent trials are conducted. So, the starting point does not matter if we are interested in to look at how many more trials are needed for a particular kind of either success or failure, because the distribution of success and failure can be interchanged both will be same. Now a further generalization of this distribution can be considered; in the geometric distribution we consider the number of trials needed for the first success, now in place of first success we say that we need 3rd success, 7th success or a particular rth success.

So, if we consider independent Bernoullian trials under identical conditions till rth success is achieved; this is applicable in various industrial applications etcetera for example, a particular mechanical system is there which is having several identical components, and each component may fail or may keep on working. So, a suppose it fails with probability p and the entire component may fail if say 3 of the components fail or entire system fails if 5 components fail; suppose you have an aero plane, there are 4 engines, the plane can fly if at least 2 of them are working. So, if both are at least 2 of them are failing, then the aero plane crashes. So, we may are interested in such kind of events.

So, if we consider X is the number of trials needed for the first time rth success is achieved, then what are the possible values of X? X can take values r, r plus 1 and so on. So, if we write down probability of X is equal to say k, then what this will be? So 1, 2, 3, 4 etcetera say r, r plus 1 k. So, if we are having k trials required for the rth success; that means, the last 1 has to be a success this is the rth success. So, the probability for that is p and before that in k minus 1 trial you should have r minus 1 success, and remaining k minus r are the failures.

So, the probability of this event will become k minus 1, c r minus 1, p to the power r, q to the power k minus r for k is equal to r, r plus 1 etcetera this is known as inverse binomial or negative binomial distribution. The name inverse binomial etcetera is apt here, because in the binomial distribution what we do we fix the total number of trials and we see how many successes are observed; here we fix the number of successes and then we see how many trials are required to obtain that many number of successes. So, it is a something called. So, earlier one if you call binomial sampling, then this is called inverse binomial sampling, this is known as negative binomial distribution. Now in order to see that it is a valid probability distribution you have to sum it, and we if we want to calculate the moments then the calculations which were shown for the geometric distribution the similar type of formula will be applicable, because it will require the expansions of negative powers of 1 minus q.

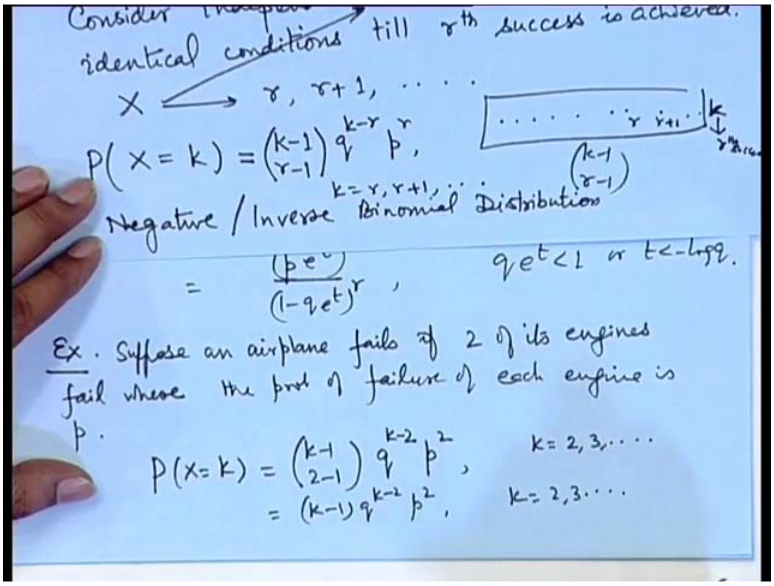
(Refer Slide Time: 23:48)



So, if we make use of that formula, we will get expectation of X equal to r by p, if you look at say variance of X we will get r q by p square, we may look at the moment generating function that will be expectation of e to the power t x that is equal to sigma e to the power t k, k minus 1, c r minus 1, q to the power k minus r, p to the power r that is equal to k minus 1, c r minus 1, q e to the power t to the power k minus r, p e to the power t to the power r k is equal to r to infinity, which is equal to p e to the power t to the power r divided by 1 minus q e to the power t to the power r, it is valid for q e to the power t less than 1 or t less than minus log of q. So, if we are looking at r is equal to 1 this we are reducing to the geometric distribution.

Let us look at one application of this negative binomial distribution; suppose an airplane fails if 2 of it is engines fail, where the probability of failure of each engine is say p. In how many of the flights we will require like in one flight one engine, may fail in second flight the first engine may fail and so on. So, how many flights will be required for the second engine to fail? That is the first time second engine fails.

(Refer Slide Time: 26:04)



So, if we consider probability of x equal to k here, then that will be equal to k minus 1 c r minus 1, that is 2 minus 1 and p is the probability of failure here. So, we need 2, so p square and q to the power k minus r that is k minus 2; for k equal to 2, 3 and so on; that is equal to k minus 1 q to the power k minus 2 p square. Today we have mainly spent time on the one particular kind of trials which are known as Bernoullian trails, and in these trials we looked at from different angles. So, we looked at number of success the distribution of the number of successes or the number of trials needed for the particular specified number of success. So, these are all called Bernoullian trials and we looked at various phenomena related to that.

Now, here you can consider it as the approximation of the real life situation, because in reality it may not happen that each trial will be independent and under identical conditions, but the real probability distributions are approximations of the real life situations and we may have to make certain assumptions to make them applicable. We can also consider finite sampling situations, so we will be taking up hyper geometric distribution and also other kind of distributions where identical situation may not be there to describe other kind of distributions. So, in the next lectures I will be taking up these issues.

Thank you.