Statistical Methods for Scientists and Engineers
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology – Kharagpur

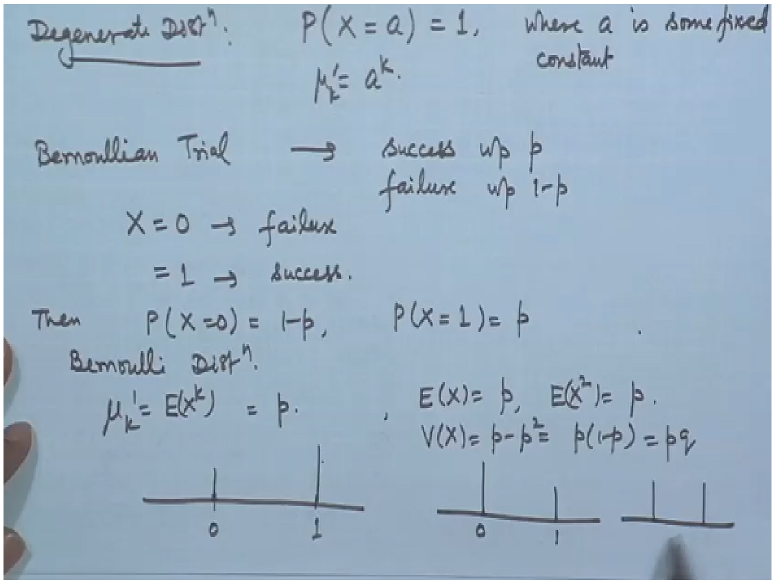
Lecture - 05
Moments and Special Distributions (Contd.)

Friends in the last lecture, I described various characteristics of the distributions such as moments quantiles, moment generating function, the moments in the sense of mean, variance measures of skewness and kurtosis etc. These distributions describe various properties or you can say characteristics about the shape about the center of gravity or measures of location etc., of a distribution.

Now I will look at certain frequently or commonly used distributions and you can say that to each of them has some historical origin that means that distributions have been studied for a long time. Nowadays of course you see that there are so many type of data sets which special distributions and some new distributions are generated but there are certain distributions which have some sort of universal applications they arise anywhere.

And when we discuss them you will see that their importance is in the theory or in the development of the subject itself. Now let us start with say Discrete Distributions.

(Refer Slide Time: 01:47)



So, we talk about say now in the Discrete Distribution we have seen the distributions have positive mass on certain points. Now depending on the number of points itself we can have different distributions one is that what is the probability allotment and another is the number of points itself for example. If there is a surety about something, then what do you say in common sense language we say it is almost sure or we will say it is sure that this event will happen.

Okay now if random variable is allocated to that it will become a degenerated number because we will say Degenerate Distribution that means if random variable X is taking a single value with a probability 1 where a is some fixed constant. For example, the boiling temperature of water under a certain atmospheric pressure. So, it is fixed that it is a 100 degrees Celsius at a certain 760mm per hg atmospheric pressure.

So, if we are reaching 100 degrees Celsius as the boiling point and if X is denoting that then it will be the value will be 100 with probability 1. And similarly for another metals similar thing we can talk about certain chemical compositions and chemical reactions etc that if this is mixed with this then this will be the outcome. So, although this looks a very simplistic distribution but nevertheless it has some uses is in the areas where you have certainty.

And of course all the moments will be opted simply by looking at the value here itself, because this would be simply a to the power k. So, this is a sort of trivial distribution but still one may use it. Now this is a classification based on a number of points so this is based on a single point now another thing could be based on two points. Now what is the two-point distribution or what is the significance of that if we have two-point distribution.

Then you will allocate the probabilities say p1 to one point and p2 to another point such that p1+p2=1 that means p2 is actually=1-p1. Now this simple structure is extremely useful to describe a very large number of physical phenomena or you can say physical realities which are encountered in day to day life. If a person appears in a competitive examination, then the event of interest is whether he qualifies the exam or he does not qualify in the exam.

Although the real outcome could be that how many marks he scored out of the total number of marks. How many students are there above him are below him and so on but ultimately an event of interest would be whether he really qualifies or not for example a clinical trial is conducted to test the efficacy of a medicine. So, there may be various things for example what the side effects of medicine are.

And if the medicine was given and whether the patient was cured or not cured and so on. So, there can be so many things how much time it took but one may be simply interested in knowing that if whether the effect or efficacy of the medicine is more than the efficacy of one previously used medicine or not. Or simply it could be that whether the medicine is effective in treating a certain disease.

For example, it may be used for killing a certain virus or bacteria. So, whether it is successful in doing that. So, although there may be a full-fledged manifestation of an experiment but we may be interested only in a particular type of outcome for example you look at the cricket match in a cricket match you may have for example it is played for 20 overs or 40 overs or for it is a test match which is played over 5 days.

Then there can be so many phenomena which are associated with the number of players scoring runs more than a certain number of people who are taking say wickets. But one may be interested in knowing whether the team A is winning the game or team B is winning the game. So, when we classify the outcome of the random experiment into two possibilities we may associate the term success to one and failure to another one this is termed as Bernoullian trial.

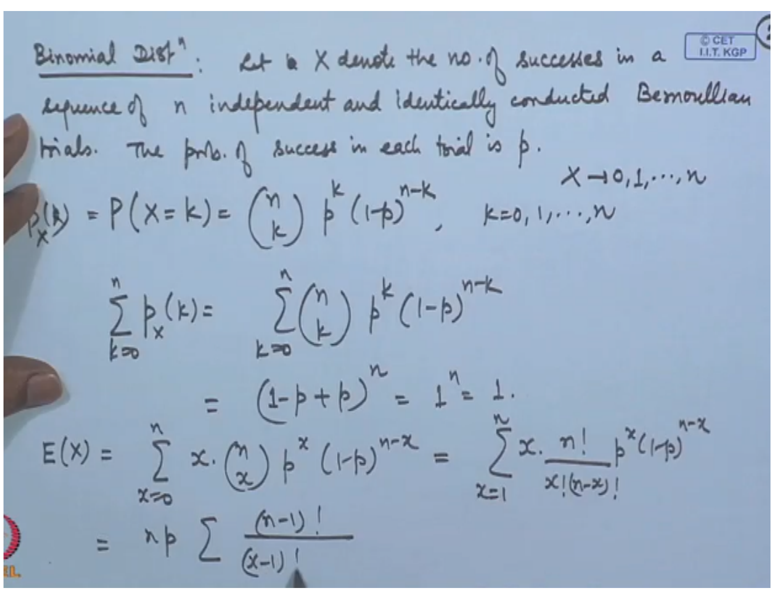
So, in a Bernoullian trial you have success with probability p and a failure with probability 1-p. Let us associate X=0 with a failure and X=1 with a success then we have a probability of X=0=1-p and probability of X=1=p then this is called Bernoulli distribution. One can look at the moment structure for example what is expectation of X to the power k that is mu k prime that is =p in particular what is the mean.

So, mean will be p what is say second moment that is equal to again p. So, what is a variance that will be p-p square that is=p*1-p. One may use a notation say q for 1-p. So, you may write it as also pq and one may write the higher order moments also if we try to plot it looks like this that you have .0 and you have .1 and now depending upon what is p and what is 1-p you may have different shapes.

1-p may be more p may be less here 1-p may be less you may have equality also. So both are equal to 1/2. So, there are various types of representations for this distribution. Now direct generalization of this is to look at several Bernoullian trials in place of one trial for example you are looking at the number of guinea pigs used in a clinical trial and then the event of interest is that how many survive the trial.

So, for example 100 guinea pigs are put in the experiment each may survive the probability p and may not survive the probability 1-p assuming that the effect of the trial is independent on each of them then suppose we say X is the number of survivors what is the distribution of that. So, if we look at a sequence of independent and identically conducted Bernoullian trials and we look at the distribution of the number of successes it is called Binomial Distribution.

(Refer Slide Time: 09:49)



So, this is another important distribution in theory of statistics and it has historical origins as I mentioned this Bernoullian trials are named after Bernoulli the Swiss mathematician which is of course one of the family of the Bernoulli families Binomial Distribution. So, the Binomial Distribution is let X denote the number of successes in a sequence of an independent and identically conducted Bernoullian trails.

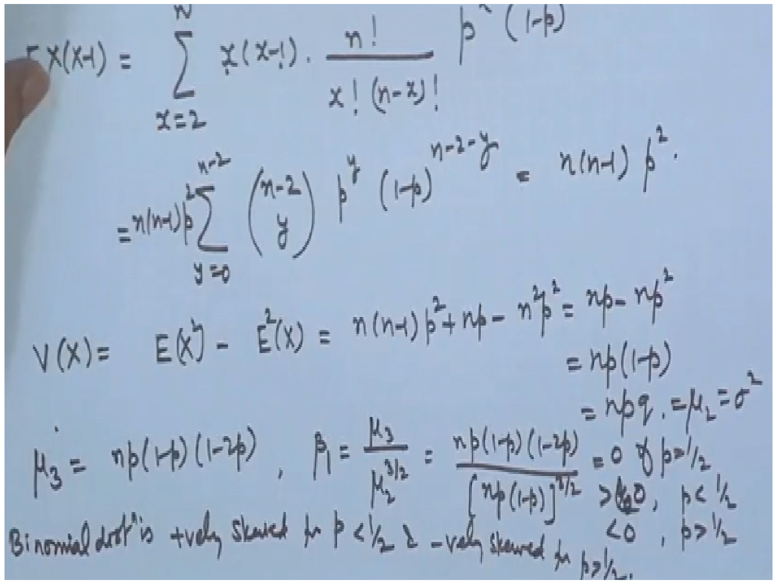
The probability of success in each trail is p. So, what is the probability? What are the possible values of X here? X can take values 0, 1 to n. So, what is the probability of X=k that is =nck p to the power k 1-p to the power n-k where k is taking values 0, 1 to n. The name binomial has come because of the binomial co efficient are appearing and also if you want to check whether it is a valid distribution then you will sum from 0 to n.

So, it is nothing but the binomial expansion of this is actually binomial expansion of 1-p+p to the power n. So, this is actually 1 to the power n and that is =1. So, that is why the name binomial distribution has come. Now, naturally the question arises that what is the characteristics of this distribution for example we may look at expectation X. Now to calculate the expectation of X I have to calculate X=0 to n.

Now easily we can see here we are able to have it as a direct binomial expansion if I multiply by X I cannot have a direct expansion. So, we need little bit of adjustments so we can look at it like this X and then we have n factorial /x factorial, n-x factorial p to the power x*1-p to the power n-x. Now corresponding to x=0 this term will be 0, so we start from 1. Now we can write it as np and here we consider n-1 factorial/x-1 factorial n-1-x-1 factorial.

Then you have p to the power x-1 1-p to the power n-1-x-1. So, this is from x-1=0 to n-1 that means if I put say x-1=y.

(Refer Slide Time: 13:31)



Then the sum is becoming =np sigma y=0 to n-1 n-1 cy p to the power y 1-p to the power n-1-y=np*1-p+p to the power n-1=np. So, the mean of the binomial distribution is actually n times p. So you can easily see if there are no probability of success in one trial is p then the average number of successes in n trials is n*p. Suppose p=1/2 that means roughly half successes you will get that is an average.

So, this is matching with our way of commonsense understanding of the binomial distribution. Now certainly if we are doing the calculation we may look at say expectation x square. What is the variability of this? Now if I look at expectation of x square I will get an x square term here. Now naturally you can see that in this one I had an advantage of cancelling x. But if I look at x square I would not be able to do that.

So, far that we can actually make use of the factorial moments because factorials are involved here. So, I can actually cancel out factorials only so a better way of calculating expectation of x square could be to look at expectation of X*x-1+expectation of X. Now if I apply the same logic which we had here then I will get n-2 here that means I will be taking out n*n-1 p square. Let me just show one calculation and then other calculations can be done in a similar way.

Because for higher order moments we can proceed in the similar way. So, this is x*x-1 n factorial/ x factorial n-x factorial p to the power x*1-p to the power n-x corresponding to x=0 and 1 this term vanishes. So, we can write from 2 to n. So, this is simply becoming n*n-1 p square and you will have n-2cy p to the power y 1-p to the power n-2-y= for y= 0 to n-2 =n*n-1 p square because this becomes again a binomial expansion of 1-p+p to the power n-2.

So, if we use this then we go to expectation of X*x-n*n-1 p square+expectation X will become np=n. So, this is the term here. Now if I want to calculate say variance then variance expectation of X square-expectation of X whole square=n*n-1 p square+np-n square p square=np-np square =np*1-p or we can also say it as npq. One point which we may notice here as in the Bernoulli also that see in the Bernoulli.

We had mean as p and the variant as pq and since p is in between 0 to 1 q is also between 0 to 1 so pq is always less than or =p. In a similar way npq will always be less than or=np that means the variance of a Binomial distribution is always less than or equal to the mean of the distribution This is also one of the observations that we may have about the distributions. Now one may also look at what is say mu 3 prime.

So, for mu 3 prime you need expectation of X*x-1*x-2. So, that will become n*n-1*n-2 pq using that you can calculate expectation of X cube and so on. So, without going into the detail calculations let me just write the expression say for example mu3, mu3=np*1-p*1-2p. So, if I look at say the coefficient of this skewness that is=mu3/2 to the power 3/2, this is actually mu2 that is also called sigma square.

So, it is becoming=np*1-p*1-2p/np*1-p to the power 3/2. Now these terms are positive not negative so we look at this here you can see this is=0 if p =1/2 that means the distribution of course binomial you can see by plotting with these are binomial coefficients and you can see that if p=1/2 then this is exactly symmetric but if p is say greater than 1/2 then this will become negative so this will become I am sorry this is greater than 0 if p is less than 1/2.

And it will be $<$ 0 if p is $>$ 1/2 that means binomial distribution is positively skewed for p less than 1/2. Now if you see carefully this probability is here. If p is less than 1/2 then 1-p is greater than 1/2 then in the beginning, you will have higher probabilities for smaller values of k and for larger values of k you will have smaller probabilities therefore it will be positively skewed.

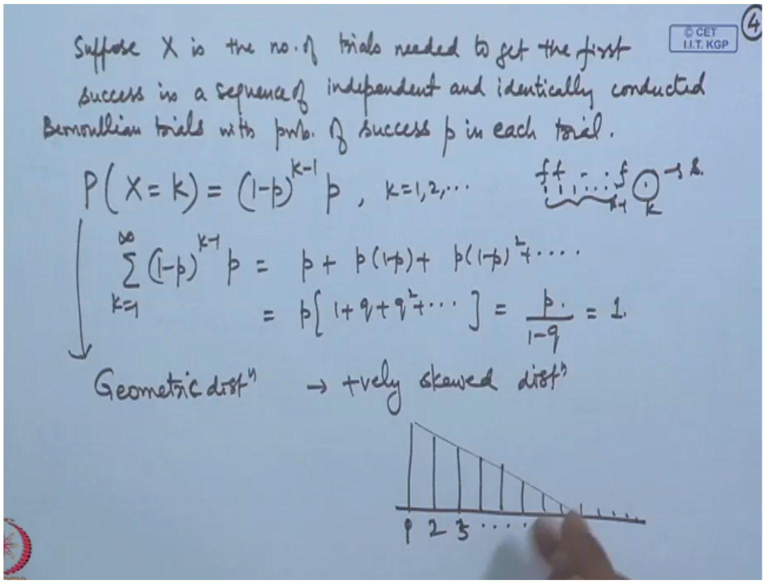
Whereas if I consider say p is greater than 1/2 then for lower values of k you will have a smaller probabilities and for higher values of k you will have higher probability. So, it will become negatively is skewed distribution. Negatively skewed for p would be greater than 1/2 one may also look at the measure of kurtosis but I will not look at this for this particular distribution one may easily calculate mu4 also I am just leaving that here.

In the Bernoullian trails what we have done that we conduct that trial as certain number of times and then we see how many successes are observed that means whatever is favorable event or favorable outcome we want to look at that how many times it has occurred. Now another way of looking at it could be we conduct the Bernoullian trial for example it is a clinical trial and we are testing various medicines ore various chemical substances which one will be successful.

So the first time the success is observed or for example for various kind of diseases one conducts the trials and as soon as a trial is successful than then it is used for constructing the medicine and then it is marketed for treating that particular kind of disease. Now assuming approximately, the Bernoullian structure that means we say their trials are independently and identically conducted. Although in physical reality it may not be so.

But statistical distributions are approximations to the physical reality therefore we may make valid assumptions of this nature. How many trials or how many times you will actually conduct a trial when you get the first success. So this could be another way of looking at the Bernoullian trials.

(Refer Slide Time: 22:01)



So, suppose X is the number of trials needed to get the first success in a sequence of independent and identically conducted Bernoullian trails with probability of success p in each trail. So, what is the probability of X=k. So, it is like this you conduct the first trial you get a failure you conduct a second trail you get a failure and so on up to k-1, up to k-1 you are getting failure this is the k th trial here you have a success before that you have all failures.

Now here the position of the failures and success is fixed unlike the binomial distribution, here we said k success in n trails so that can be done in nck ways whereas here the position of the failures are fixed position of the success is fixed so what is the probability of this if I am using the independent trials then failures have the probability 1-p. So, k-1 failures and the last one is a success.

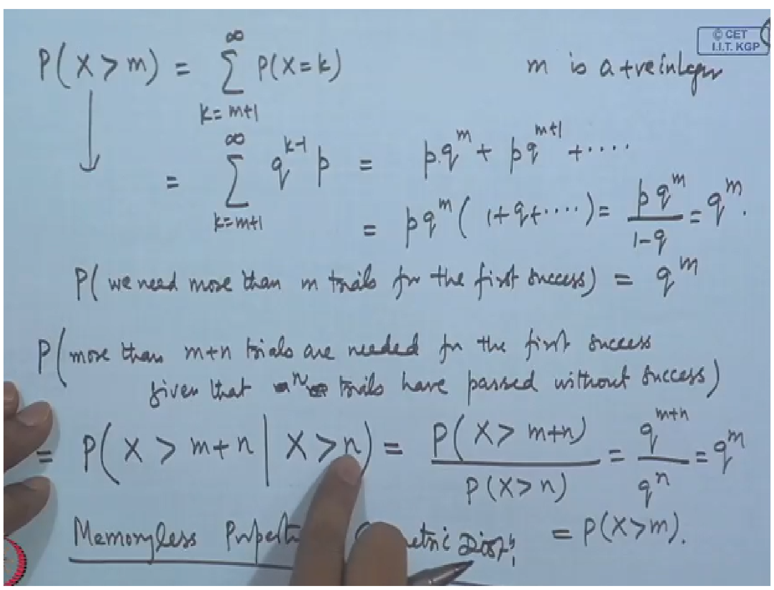
So, we get this distribution if we look at this sum of these probabilities1-p to the power k-1 p for k=1 to infinity then it is actually=p+p*1-p+p*1-p square and so on this is an infinite geometric series so it is =p+1- let me call q here as 1-p so q square and so on. That is =p/1-q so 1-q is p, so it =1 this has become an infinite geometric series with the common ratio lying between 0 and 1. So this is valid.

So, that is why there is a name geometric distribution given to this. Also you look at the probabilities see the first one is p then you have pq and then pq square. So, each successive probability decreases that means if I consider the plotting of this one like binomial distribution here the first one is a p then p*1-p depending upon what is a value the drop may be faster or slower depending upon the value of this.

So, 012 sorry 123 and so on this is the way the distribution will proceed you can easily see that it is a positively skewed distribution. So, this is positively skewed you actually do not need to calculate the third order moments and so on for this because here the form of the distribution is known and you can plot it the measures of skewness, kurtosis etc are helpful when knowledge about the probability mass function or density function may not be very precise.

In that case if you have those things then you may have certain guesswork about that immediately of course it will come that one thing is the decreasing nature of the probabilities another important property that you may notice is.

(Refer Slide Time: 26:07)



Suppose I consider what is the probability of say $X > m$, where m if of course an integer, m is a positive integer then that means I am writing probability of x=k for k=say m+1 to infinity. So, this is=q to the power k-1 p for k=m+1 to infinity. So, this is=p and then the first term here is basically what are the terms p q to the power m+p, q to the power m+1 and so on. So, this is =p q to the power m1+q+ and so on.

That is =p q to the power m and this is again the infinite geometric series. So, this 1-q and p will cancel out and you will get q to the power m. Now a physical interpretation of this is this is the probability that we need more than m trials for the first success and that is =q to the power m. Let us write another what is a probability more than say m+n trails are needed for the first success.

Given that m trials have passed say n trails let me say n trails have passed without success. So, a physical interpretation of this kind of sequence would be that we are actually looking at trials are being conducted already n trials have been conducted and then there is no success. Now suppose the team changes and there is a new person who is conducting the trials now because now he may be interested in knowing that how many trials more will be needed now.

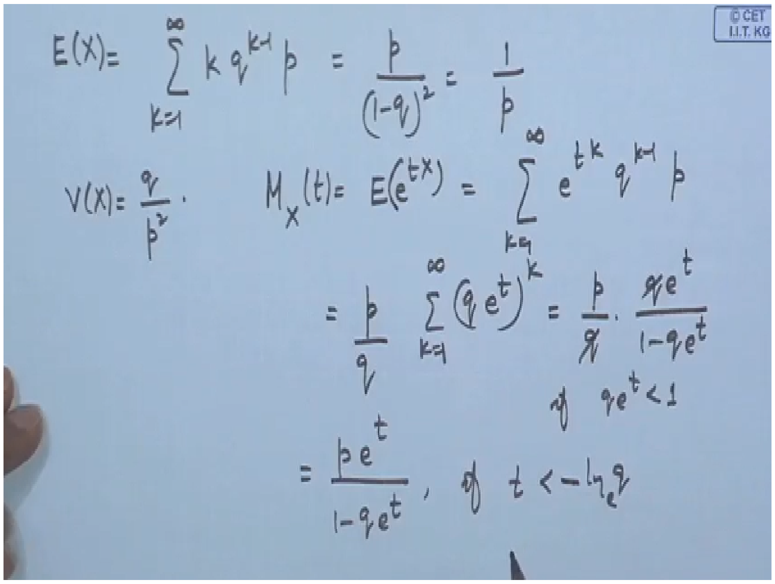
So, that is his interest so we say m+n okay now let us interpret it in terms of the probability. So, more than m+n we may say X greater than m+n given that X is already greater than n. Now this conditional probability you see in the suppose I call this event A and this as event B then this is probability of A intersection B/probability of B. So, now since this event is smaller than this event in the numerator.

I will have that is=q to the power m+n/q to the power n=q to the power m which is nothing but probability of X greater than m. Now look at the physical interpretation of this this means that we need more than m trials for first success now here it says that already n trails have passed now I want to know what is the probability that more than m trials will be needed that means starting from n+1 trail.

What is the probability that more m trials will be needed this probability remains constantly that means a starting point does not matter and that means in the experiments where geometric distribution is applicable? you have something called memory less property. Memory less property of geometric distribution because the starting point does not matter it is as if we started from the scratch.

Now we also look at the mean and variance of this distribution since you have seen that the sum is infinite geometric sum here. So, if I have X here then this will become arithmetic geometric progression, so, you can easily apply the formula for such things I will only write the final value here.

(Refer Slide Time: 30:51)



So, expectation X=sigma k q to the power k-1p, k=1 to infinity=p/1-q square=1/p it is something like this, suppose I am considering a simple coin tossing experiment with probability of head as say 1/3. So 1/1/3 will become 3 that means we need on the average 3 trials to get to the head for the first time. Similarly, you can see say tossing of a dice fair dice and you say probability of 1/6 for each face.

So,1/6 for each phase then what I am expecting a 6 then what is the expected number of trials needed for the first time 6 to be observed then it will become 6 number of trails are needed. So, in that way this is having a very nice physical interpretation variance of x is actually=because you will get 1/p-1/p square so that is becoming q/p square. One may write higher order moments also but I am not interested in that.

Let me simply write down the moment generating function for example what is the moment generating function that is expectation of e to the power tx so that is= sigma e to the power tk q to the power k-1 p. Now this type of a structure is extremely helpful because you can actually adjust this term here k=1 to infinity. So we write it as p to the power say since here we have q we can consider it as q e to the power t to the power k, k=1 to infinity.

So, that means I have adjusted one q here, now you look at this term, the sum this will become=p/q and then q e to the power t/ 1-q e to the power t. if q e to the power t is less than 1 then only this will be a convergence series so this is becoming= p e to the power t/1-q e to the power t provided t is less than–log of q, see log of q will be a negative number because q is between 0 to 1, so–log q is a positive number.

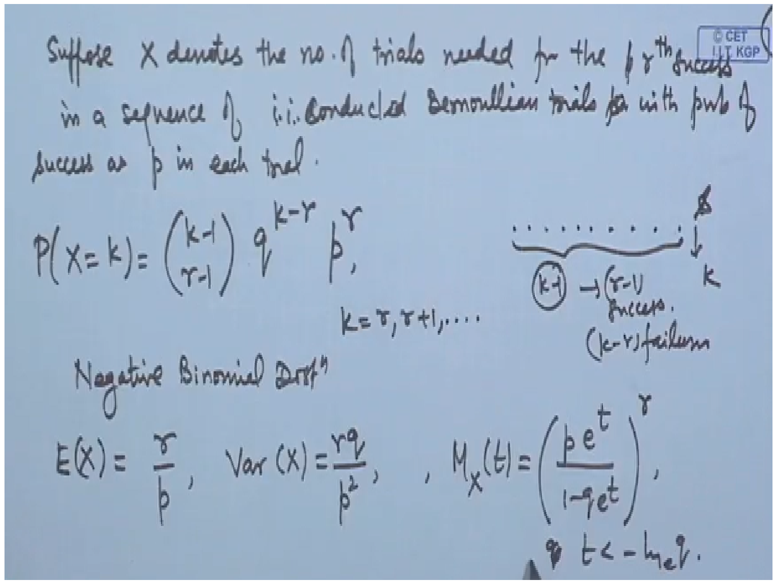
So, in this range this is valid here this expansion is valid here from the moment generating function one may be able to evaluate mu3 prime, mu4 prime and mu3 and mu4 and therefore you can find all the moment measures of is given as kurtosis etc. But of course here you already know the exact structure of the probability distribution now one easy generalization of this geometric distribution could be in the geometric.

We are looking that the number of trails needed for the first success. Now I have given a physical interpretation for example your conduct in the trials to be successful in a clinical trial or somebody is appearing in competitive examinations so many trials you need to get to the success and so on but there are certain other experiments for example you may associate the success with failure just change the name.

So, for example complex mechanical system is there which is consisting of say 100 components and those components are working the machine works as long as say 50% of them are in the working order. So, you are waiting for suppose there are 8 components so you will wait for the first time the 4 th component fails that means 4 number of components fail and then the system will fail.

That means rather than looking for the first occurrence of sequence in a Bernoullian trials you are looking at r th occurrence for certain value of r. This is known as negative binomial distribution. So we consider this simulation suppose let me put it in a new page here.

(Refer Slide Time: 35:49)



Suppose X denotes the number of trials needed for the r th success in a sequence of independent and identically conducted identically conducted, let me use this abbreviation ii independent and identically conducted Bernoullian trails for the with probability of success as p in each trail. So, you look at in this way we have these trials here this last one is the success this is actually the r th success.

Okay here out of so these are total k trail, this is k th trail. Out of k-1 trail you have r-1 success and k-r failures. So, what is the probability that x=k that will become=k-1 c r-1 this one is fixed here. Now you have q to the power k-r that is 1-p and p to the power r where k can vary from r, r+1 and so on, so this is called since this coefficient when we consider this one this is called actually Negative Binomial distribution.

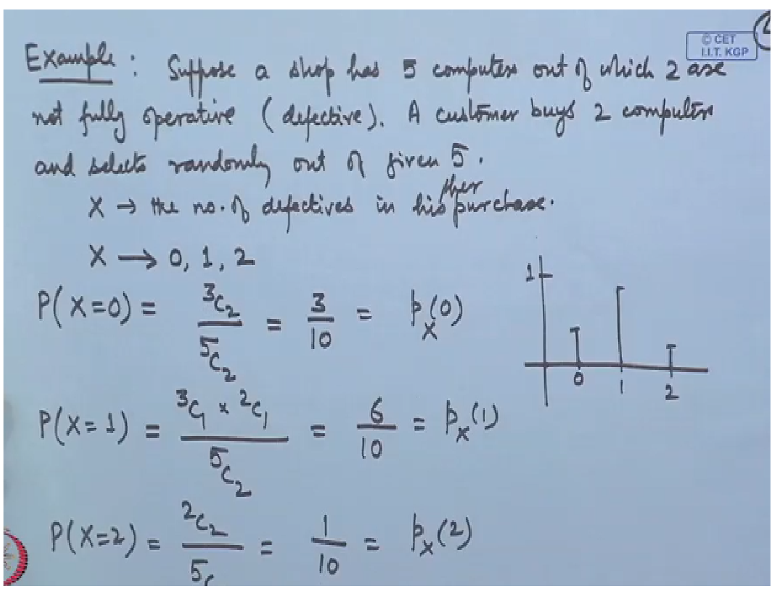
That is also called inverse binomial distribution actually in binomial we are fixing the number of trials and we are looking at how many successes are occurring. In the inverse binomial sampling we are conducting the trials up to a certain number of successes are observed. So, the number of trials is not fixed so there is a difference at looking at the experiment so that is why it is also called Inverse Binomial Distribution or a Negative Binomial distribution.

I am not going to look into the expansions and other things we can actually look at what is expectation of X that will be r/p. The variance of X can be easily calculated to be rq/p square and if we look at the moment generating function then it will become=p e to the power t /1-q e to the power t whole to the power r and of course this is valid for q t < - log of q for this region this will be valid here.

If you look at this trials that I am considering these are approximations to the physical situations as you mean independence of the experiments conducted under identical conditions So, this may be well defined we are looking at a large population size and then the sampling scheme is continued indefinitely like that. But there are other experiments where things are infinite for example you have a small shop.

I gave you example in one of the previous lectures that you have a certain number of say items in a shop where you have a certain amount of defectives. Now if a person is buying what is a probability distribution of the number of that effect is this example a considered in one of the previous lectures let me just show you that.

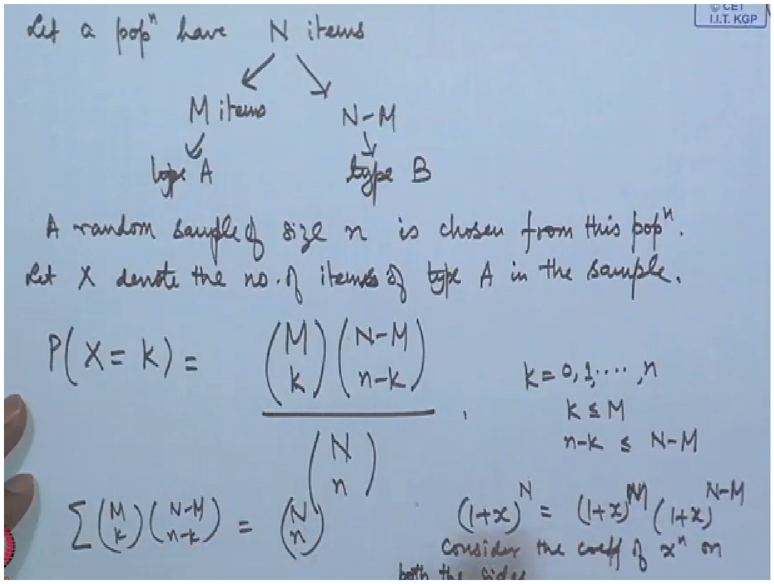
(Refer Slide Time: 40:09)



Let us look at this example so shop has 5 computers out of which 2 are defective now if that is happening and we are considering purchasing from there only then this is a small sample problem. In a small sample problem, the population size also a small the probabilities will change rapidly. So, this carnation of independent and identically conducted Burnollian trails may not hold here.

So, far calculation of the probabilities in this finite situation we need another distribution. So let me come to this one now.

(Refer Slide Time: 40:44)



So, we have let a population has capital N items okay out of which say M items these are of type A and say N-M items these are of type B can this be easily considered like this like you have a class of students. So, you have some boys and there are some girls you may have office staff some of them may be smoking and some of them may not be smoking you may have say people visiting a shopping mall and then you may have people of different origin.

That means how many of them are of a particular ethnic group you may have say an organization where you may have upper income group people, middle income group people and so on. So, here I am considering only two split that means I am considering two categories that means they are complementary. Now a random sample of size say small n is chosen from this population then let x denote the number of items of type A in the sample.

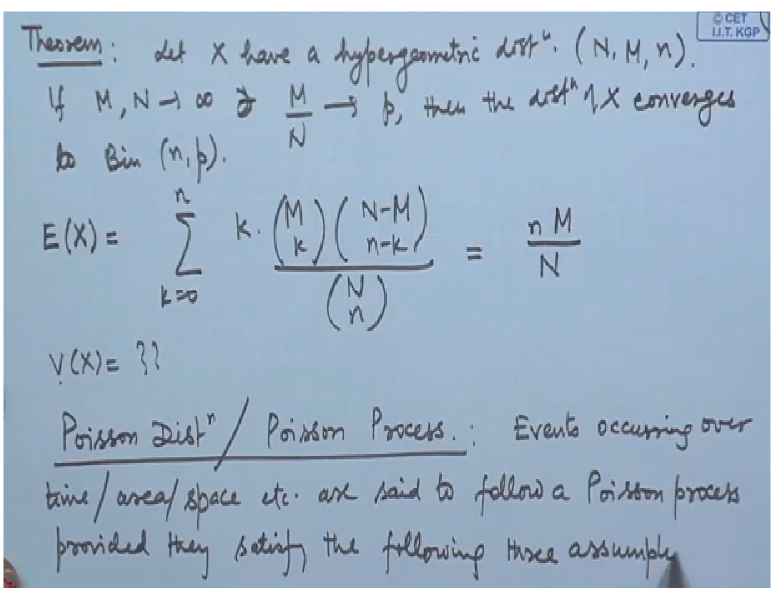
Then what is the probability distribution of X. So certainly you will let me call probability of X=K then this will become=Mck because you have chosen k from here that means n-k items will be from remaining and total number of items are chosen in Ncn and of course you will have this k=0,1 to n. But of course you have some further restrictions such as this k cannot cross capital M. Similarly, n-k cannot be greater than or=N-M.

So, these are some physical restrictions that will come here. Now if you look at the sum of this this is actually=Ncn this is done by considering the expansion of 1/1+x to the power N and you explicit like this 1+x to the power of M *1+x to the power N-M consider the coefficient of x to the power n on both the sides. So, on this side you will have Ncn and here you will have this summation coming here.

Therefore, the sum is=1 this is known as Hyper geometric distribution. So, this is also one of the useful distributions in the theory of probability and this is describing a finite sampling scheme and finite population size here. The Burnollian trails are helpful when we are considering exceedingly large population size that means we can theoretically call it as an infinite population and that is why we are able to maintain the equal probability axiom.

But here small sample that axiom will not be valid and therefore we consider this as a Hypergeometric distribution. Of course see you can consider that suppose here population size becomes actually large this capital N is really large and then name is also large in that case if the proportion is constant then this will converge to a normal distribution.

(Refer Slide Time: 45:07)



So, I can state it as a theorem here. Let X has a hypergeometric distribution with this parametric structure N, M and small n etc. If M and N go to infinity such that M/N goes to p then the distribution of x converges to binomial np distribution. So, that you can prove easily by taking the taking the expansion here and then taking the limits here. I will not be writing the proof I leave it as an exercise here then further you look at the expectations etc.

Let me do one time here another thing you can do yourself. So, if I do the expectation then this is=k times Mck N-M cn-k Ncn. Now as in the case of binomial distribution you have seen that we actually adjust this term k here that means in the divisor we have k factorial. So, you cancel k factorial you will get k-1 factorial. Accordingly, you take common one of the terms here so basically it turns out to be nM/N.

I am not writing the calculations here one can easily check that thing. Another property you can observe that as nM/N goes to p so it is actually NP. So, you look at a physical interpretation of this that if I have M number of type A items in the population then the proportion is N/M in the whole population so what is the expected number of items in the sample of size N. It will be small N times capital M/capital N.

So, it is actually very fine that thing here I am leaving the variance X to the exercise and one can look at of course higher order moments will be a bit complicated and also the moment generating function will have an extremely complicated structure here so we are not looking at that thing. Now you have seen here in each of these problems that I have discussed here it is one can actually describe the experiment in a very proper way.

By physically ascribing the probability of success etc in a precise way. But then there are many other things where it is not really possible for example we are looking at see one is sitting at a railway reservation counter so when you look at a railway station. Now you look at the number of customers arriving at the reservation counter or a ticket counter for example it is a cinema theatre.

So, how many people arrive and between what time to the theater to purchase the ticket see this type of thing is important why one will really study this phenomenon is to consider the appointment of the proper personal or deployment of the people for the service. So, it is something like an arrival and service type of feature and this happens almost everywhere for example how many buses.

You need to deploy to ferry are to carry the people from one place to another place between two cities and between what times so that will need an estimate or a probability distribution of the number of the passengers. How many flights you need at the airport you have the check in counters. How many people you knew could apply how many counters should be open between what time to what time.

Similar type of thing is also appearing for example you are making a traffic analysis so you look at the number of accidents the number of vehicles passing through a particular crossing between certain timing. So, you need to make that road should be so much wide or how many your traffic police would be deployed to avoid any unnecessary circumstances. Here you are looking at the area. In the earlier examples you are looking at the time.

Similarly, you can consider say natural disasters so the disasters for example a flood floods are happening over a period of time floods also happen over geographical region. So, over a particular region how many such events may be observed earthquakes or other kinds of natural disasters you may also look at say space for example one is observing astronomical events. For example, collision of say satellites are passing of satellites through certain portion of the space.

We hear that there is a comet which is supposed to pass very close through the earth and so on and that means we are interested to look at the number of occurrences over a period of time, area or a space and now we want to do the precise distribution unlike the situation which I described in the Burnollian trials are hypothermic distribution. This will have a slightly vague type of distribution because how to say that how many people are arriving in a particular time interval.

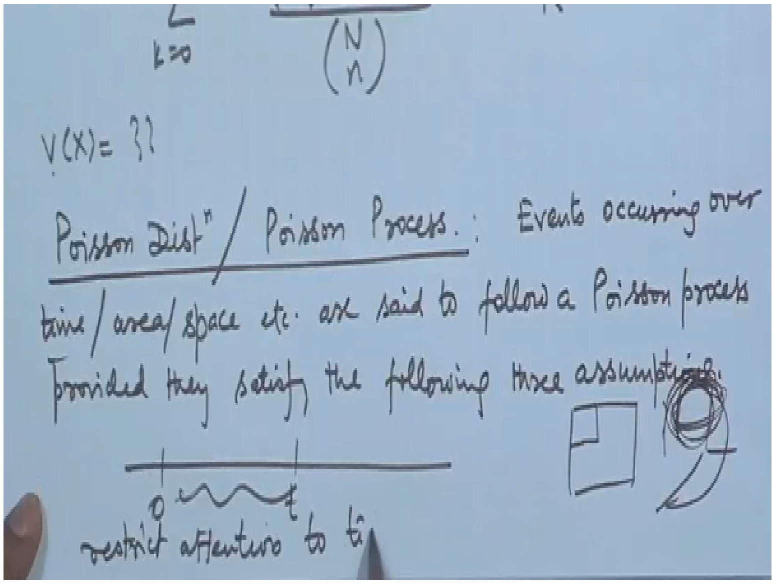
How many accidents are occurring over a stretch of road how many earthquakes are occurring over a period of time here it is not possible to ascribe certain probability like p and 1-p kind of thing that if an earth quake occurs it is occurring with probability p or it is not occurring with probably 1-p. Or if a customer is arriving at a shopping mall, he is arriving with probability p or 1-p we cannot do like this.

So, in that case we need a little bit of different thinking to describe the phenomena we call such events as events occurring in a Poisson process and for that we put forward certain assumptions that if there is despite certain assumptions then we call it as Poisson process. In my next lecture I will be actually obtaining the distribution of occurrences in a Poisson process. And then we will see there.

How it is helpful in describing various phenomena which I have described and then also leading to some other distributions for example distribution of the times for the event for the occurrences and so on. So, I will just give the assumption of the Poisson process in todays class and in the tomorrows lecture in the following lecture I will be describing the distribution and so on. So, we consider Poisson distribution or Poisson process.

So, events occurring over time area, space etc are set to follow a Poisson process provided they satisfy the following three assumptions.

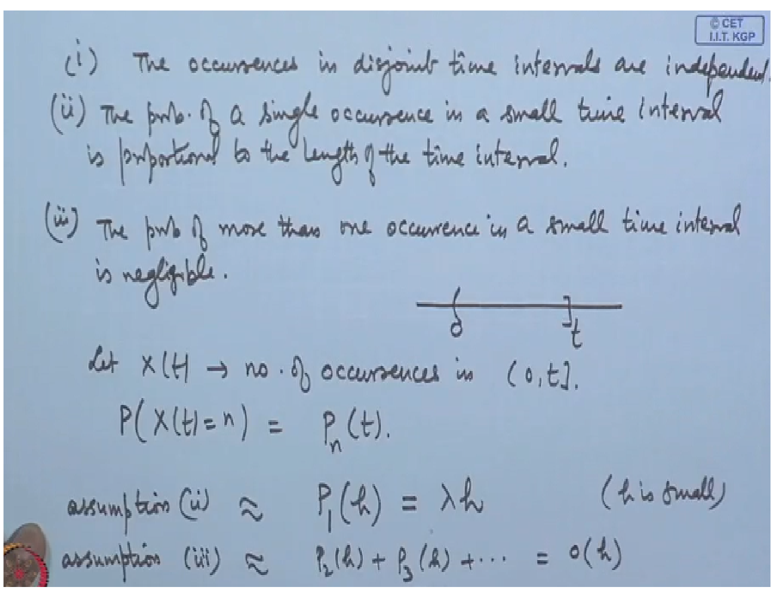
(Refer Slide Time: 53:37)



First assumptions is now another thing is that I will restrict attention to time because see we can consider area, space and so on but for deriving my thing it will be convenient if I restrict attention to time. So, time scale is like this so you start from the t0 and you go up to a time t. When you look at how many occurrences are there. So, same thing will happen if I can see that then what a particular area how many events are there.

I am in a three dimensional space then in a particular stretch of a space how many things are there so that one may look at what I am restricting and enter into their restrict attention to time okay. So, now the first assumption is that.

(Refer Slide Time: 54:23)



We assume that the occurrences in disjoint time intervals are independent. Second assumption is that the probability of a single occurrence in a small time interval is proportional to the length of the time interval. The third assumption the probability of more than one occurrence in a small time interval is negligible. Let me introduce some notation here see we may consider like 0 to t kind of thing.

So, we may keep open on one side and closed on another side which is of course not very strict when we put it in another way also you may put close on this side and open on this side but this is for convenience as you will observe that thing. So, let us consider, say let x t be the number of occurrences in the interval 0 to t. Then we have probability of xt=n. Let me denote it by Pnt. So, this assumption 2 it is equivalent to that P1h=sum lambda h.h usually means h small.

That means probability of a single occurrence in an interval of length h is proportional to the length of the interval and similarly this assumption 3 this is equivalent to p2 h+p3h and so on that is negligible. Now, for negligibility we use a mathematical notation is small o of h. So, basically this is meaning 1-p0h-p1h=oh which is also equivalent to saying that your p0h=1-lambda h-oh-oh or+oh does not matter you may put like this.

And here also you can put small oh it will take care of the approximations here. Under these assumptions the distribution of the number of occurrences will follow it Poisson distribution so in the next lecture I will be showing you how this is following. I will show you the radiation here and then we will also see that how it is leading to describing the distributions of various other phenomena in the Poisson process.

For example, exponential distribution gamma distribution and so on so that I will be doing in the following lectures.