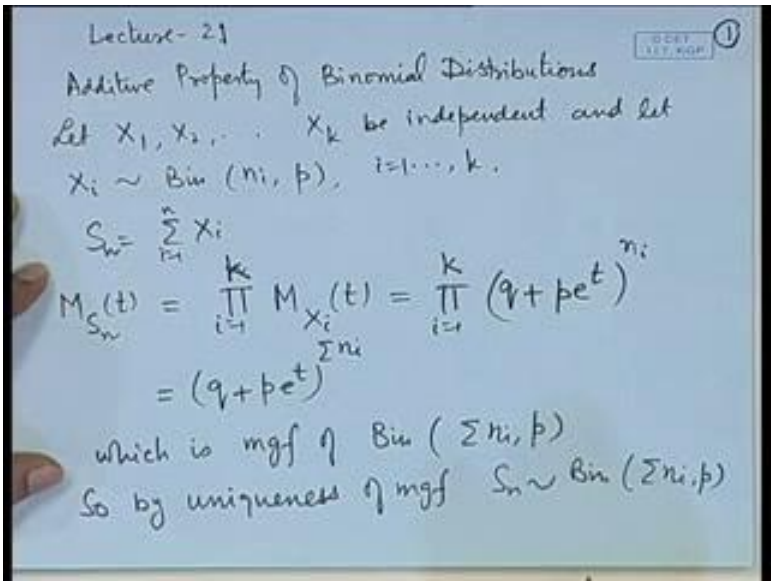
Probability and Statistics

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

Lecture - 41
Additive Properties of Distributions – I

In the last lecture, we have considered jointly distributed random vectors in general a k dimensional or n dimension random vector and we in particular define joint moment generating function and we proved and important property that if the random variables are independent then the movement generating function of the some of the random variables can be expressed as product of the movement generating functions of individual random variables. This last result is extremely useful in determining or deriving the distributions of sums of random variables; let me it by the proving additive properties of certain distributions.

(Refer Slide Time: 01:07)

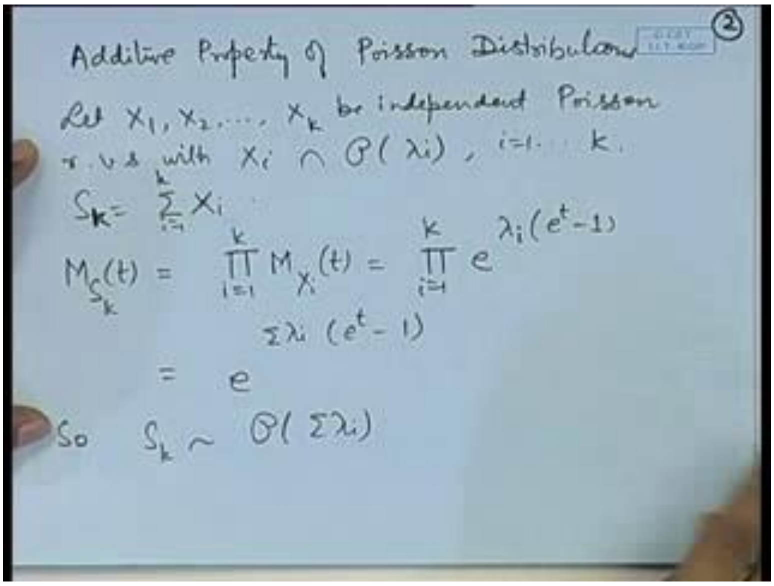


So, firstly let us prove say additive property of binomial distributions; so let us consider say X 1, X 2, X k be independently distributed random variables and let X i follow binomial ni, p distribution for i is equal to 1 to k. I am interested in the distribution of s that is sigma X i; i is equal to 1 to n or rather we can call it S n So, if we use the mgf; here distribution of the mgf of the sum is equal to product of the mgf of X i is i is equal to 1; 2, n. Notice here is that mgf of X i that is q plus p; e to the power t, whole to the power n I; product i is equal to 1 to k, this is also k i is equal to 1 to k.

Now since the term is the same, the powers will be added up and it becomes q plus p e to the power t to the power sigma n I, which is the mgf of a binomial; sigma ni, p distribution. So, by uniqueness property of the mgf; S n must follow binomial sigma ni, p distribution. This additive property of binomial distribution can be expressed physically also, here you can see that X 1 denotes the number of successes in a sequence of n 1 independent and identically conducted Bernoullian trails. Here the probability of success is p; X 2 denotes the number of successes in n 2 independent and identically conducted Bernoullian trails with the probability of success p and so on. Therefore, sigma X i can be considered as the total number of successes in n 1 plus n 2 plus n k independent and identically conducted Bernoullian trails; here the probability of successes p.

So, this physical fact is confirmed by this additive property which we are able to prove here using the moment generating functions. Let us prove a similar property for Poisson distributions.

(Refer Slide Time: 04:28)

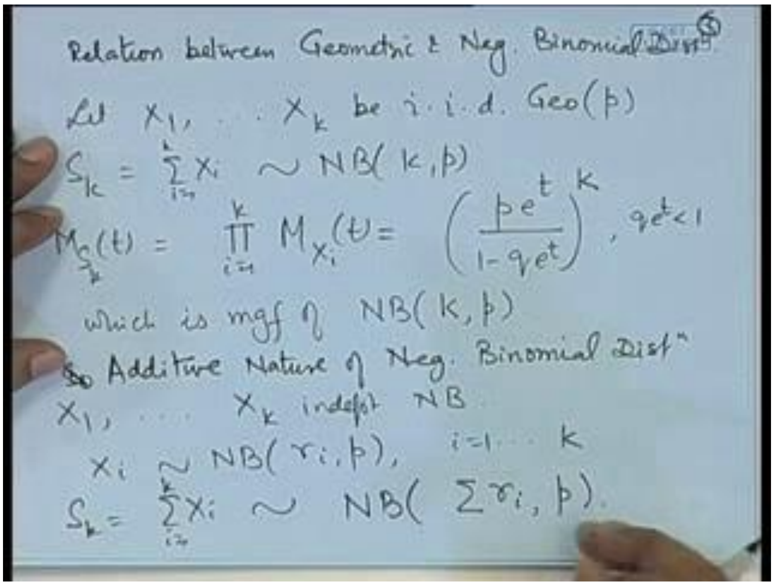


So, additive property of Poisson distributions; so let X 1, X 2, X k be independent Poisson random variables with X i having a Poisson lambda i distribution. Once again we are interested in the distribution of sigma X i; i is equal to 1 to k. I think I made a mistake here; this should be S k that is sigma X i; i is equal to 1 to k, so here also it will be; sum of k variables So, by the independence we can use that the moment generating function of a sum is equal to the product of the moment generating functions.

Now moment generating function of a Poisson distribution with parameter lambda that is given by e to the power lambda, e to the power t minus 1. So, here for X i this becomes e to the power lambda i; product i is equal to 1, 2, k which is becoming e to the power sigma lambda i e to the power t minus 1 So, once again if we use the uniqueness property of the mgf, we conclude that S k follows; Poisson sigma lambda i; that means, sums of the independent Poisson random variables are again having a Poisson distribution.

Once again we can see it in a physical terms, here we are considering k different Poisson processes; X 1 denotes the number of arrivals in the Poisson process with the arrival date lambda 1, X t denotes the number of arrivals in the Poisson process with arrival rate lambda 2 and so on. Therefore, sum of the X i denotes the total number of arrivals in a Poisson process with the arrival rate sigma lambda i.

(Refer Slide Time: 07:04)



Let us consider say a relation between geometric and say negative binomial distribution. So, let X 1, X 2, X k be independent and identically distributed geometric random variables with parameter p; so we are considering S k that is sigma X i. Now, if I am looking at the mgf of S k, now the mgf of a geometric random variable is p e to the power t divided by 1 minus q e to the power t, where q e to the power t is less than 1.

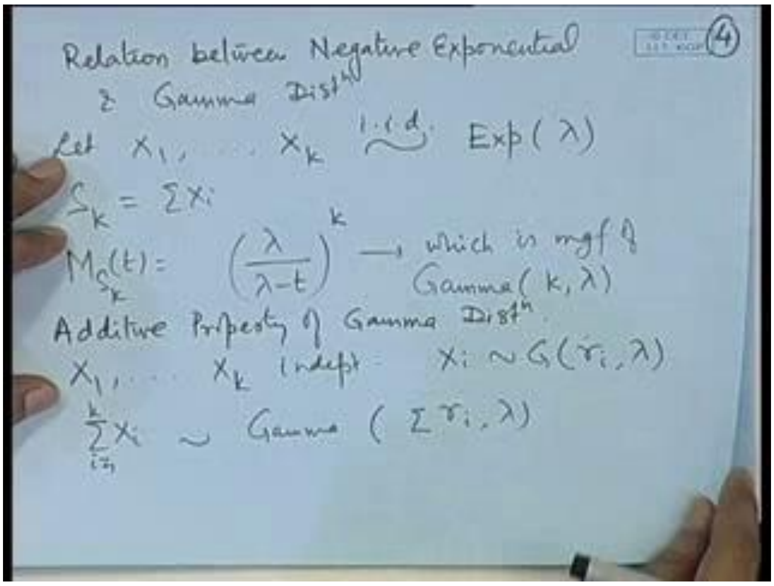
Now when we are multiplying it k times this becomes power k, which is mgf of negative binomial with parameter k and p.

So, this proves that a sum of independent geometric variables with the same probability of success is negative binomial k p. Once again we can look at the physical interpretation of this result; X 1 denotes the number of trails needed for the first success in a sequence of independently and identically conducted Bernoullian trails, X 2 denotes the number of trails needed for the another success for the first time in a sequence of independent and identically conducted Bernoullian trails.

Therefore X 1 plus X 2 plus X k denotes the number of trails needed for the first time k successes in a sequence of independent and identically conducted Bernoullian trails and that we know that it has a negative binomial distributions with parameter k and p. In a similar way, we can prove additive nature of negative binomial distribution also; so if I have X 1, X 2, X k independent; negative binomials and say X i follows negative binomial with parameter say r i and p; i is equal to 1 to k, then if I consider the distribution of S k; that is sigma X i; i is equal to 1 to k. Then by this property, when we are multiplying the moment generating functions; I will be multiplying p e to the power t divided by 1 minus q e to the power t to the power r i for i is equal to 1 to k.

So, the exponent will become sigma r i which will prove that the sum will follow a negative binomial distribution with parameter sigma r i and p. So if the probability of successes constant; negative binomial distribution also follows and additive property.

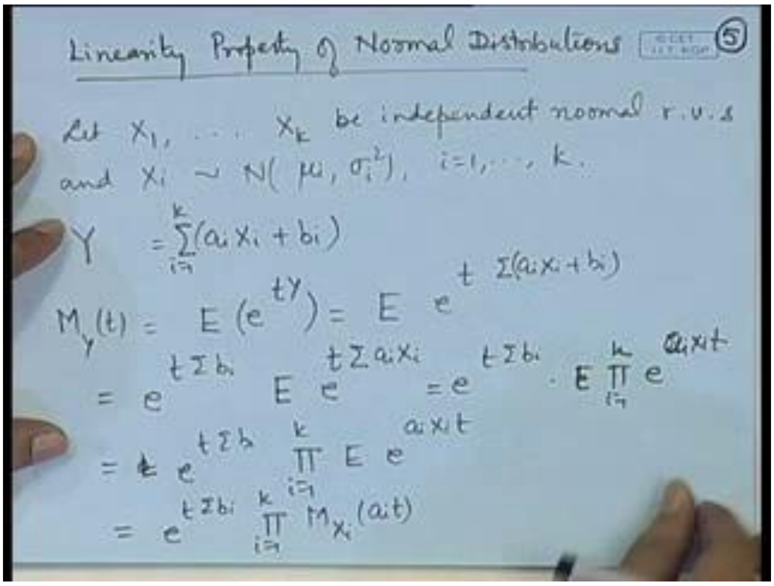
(Refer Slide Time: 10:50)



Let us look at a relationship between say negative exponential and gamma distributions, so let X 1, X 2, X k they will be independent and identically distributed exponential variables with parameter lambda. Now let us consider the distribution of the sum so moment generating function becomes lambda by lambda minus t to the power k, which is mgf of gamma distribution with parameters k and lambda; that means, sums of independent exponential variables is a gamma variable. So, physically if we represent this result, if we are observing a Poisson process with rate lambda, X 1 denotes the weighting time for the first occurrence, X 2 denotes the weighting time for first occurrence at another point of time, X k denotes the weighting time for the first occurrence in a k th observation of the process So, if we combine this that is X 1 plus X 2 plus X k; we look at that is when X 1 is observed we start observing the process once again, X 2 is the time added thereafter, X 3 denotes the time is starting from when X 2 has been; that is the second occurrence has been observed then we observe. So, then X 1 plus X 2 plus X k denotes the weighting time for the first time k th occurrence in a Poisson process and that we know that it follows a gamma distribution with parameters k and lambda.

Likewise we can prove the additive property of gamma distributions also, once again here we can consider say X 1, X 2, X k independent and X i follows gamma say r i lambda then sigma X i; i is equal to 1 to k that will follow gamma with parameter sigma r i lambda because here we can consider X i as the weighting time for the first time r i th occurrence in a Poisson process with rate lambda So when we add these timings, it means that it is the total weighting time for sigma r i occurrences in a Poisson process with rate lambda. Therefore, this gamma distribution also satisfies and additive property provided the Poisson process parameter remains the same. In the case of normal distribution we have much more general property in fact, we have a linearity property.

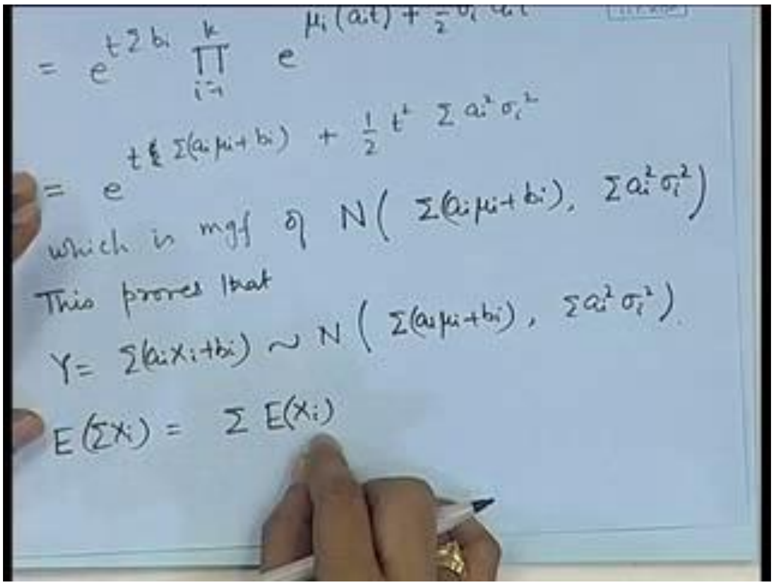
(Refer Slide Time: 14:25)



Let us consider say X 1, X 2, X k independent normal variables and X i follows say normal mu i; sigma i is square for i is equal to 1 to k. Let us consider a linear function sigma a i, X i plus b i; i is equal to 1 to k. Let us obtain the distribution of Y; so M Y; t that is equal to expectation of e to the power t Y that is expectation of e to the power t sigma a i; X i plus b i, here e to the power t into sigma b i can be kept out So, it is e to the power t sigma b i and then we have expectation of e to the power sigma a i; X i and t. Now this we can express as e to the power t sigma b i and this term, we can a split; we can consider it as e to the power expectation of product e to the power sigma a i; x i; t.

Now, here X i's are independent variables therefore, this term is simply e to the power t sigma b i product of i is equal to 1 to k expectation of e to the power a i X i; t. Now this is nothing, but the moment generating function of the random variable X i at the point a i, so e to the power t sigma b i product i is equal to 1 to k; moment generating function of X i at a i t. Now X i is follow normal distributions therefore, the moment generating function of X i can be written.

(Refer Slide Time: 17:18)



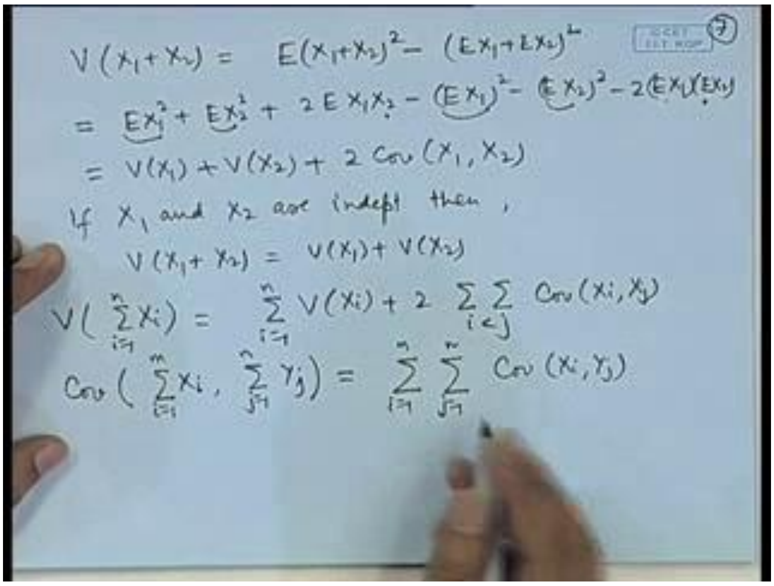
So we substitute that here to get e to the power mu i a i t plus half sigma i square a i square t square So, after adjusting the terms we get e to the power t sigma a i, mu i plus b i plus half t square sigma a i square sigma i square. Now, this we can identify as mgf of a normal distribution with mean sigma a i; mu i, plus b i and variance sigma a i square sigma i square So, by the uniqueness property of the mgf this is proves that Y is equal to sigma, a i; X i plus b i follows a normal distribution with parameter sigma a i; mu i plus b i, sigma a i square, sigma i square.

So, in the case of normal distributions it is not only the sums, but any liner combination of the independent normal variables follows a normal distribution and other important thing to notice here is that in normal distributions case, we can vary both the parameters. Earlier in the additive property of say gamma distribution, additive property of negative binomial distribution or the additive property of binomial distribution where two parameters are there, when we are considering several independent random variables; we were varying only one of the parameter and one of the parameter was kept fixed in order to have the additive property.

But in the case of normal distribution, we can vary both the parameters and the property is also more general, rather than just talking about the sums; we can talk about any linear function. There are certain results which are related to the calculation of the moments of sums, variances of the sums etcetera. So, these I will a state here for example, if we look at say expectation of sigma X i; it is equal to sigma expectation of X i. If we are looking at variance of now the proof of this fact is quite simple, you have to just apply the linearity property of the integral are the summation sins because here it is the expectation of a summation So, either you will have a; if the random variables are discrete, you will have summations or if we have continuous; we will have integral.

So, when we apply the linearity property then the sums can be taken inside and it will prove this property.

(Refer Slide Time: 21:09)



In the case of variance, let us write for two of them that is variance of say X 1 plus X 2. Now this is equal to expectation of X 1 plus X 2 whole square minus expectation of X 1 plus expectation of X 2 whole square. So, this is equal to expectation of X 1 square plus expectation of X 2 square plus twice expectation X 1, X 2 minus expectation of X 1 whole square minus expectation of X 2 whole square, minus twice expectation X 1 into expectation of X 2. So, this terms if you combine expectation of X minus square with expectation of X 1 whole square; this is variance of X 1.

In a similar way expectation of X 2 is square can be combined with expectation of X 2 whole square that leads to variance of X 2. Now this cross product term that is expectation of X 1, X 2 minus expectation; expectation of X 1 into expectation of X 2 is nothing, but the covariance term So, variance of a sum is equal to sum of the variances plus twice covariance of X 1, X 2, so there is an additional term here. Now if X 1 and X

2 are independent, then covariance will be 0 and therefore, variance of X 1 plus X 2 will be equal to variance of X 1 plus variance of X 2.

So, we can generalize this result; variance of a summation is equal to sum of the variances plus twice double summation co variances of X i, X j where i is less then j; obviously, if the random variables X 1, X 2, X n are independent then this co variances will vanish and we will have variances of the sum is equal to, sum of the variances. We can also have a general formula for covariance of a sum with covariance of another sum, so this is i is equal to 1 to M, this is j is equal to 1 to n. Then this is equal to double summation covariance of X i y j. That means, covariance of each term in the first summation is taken with covariance with the each term in the second. Now these properties are quite useful in calculation of the moments of the sums of distributions.