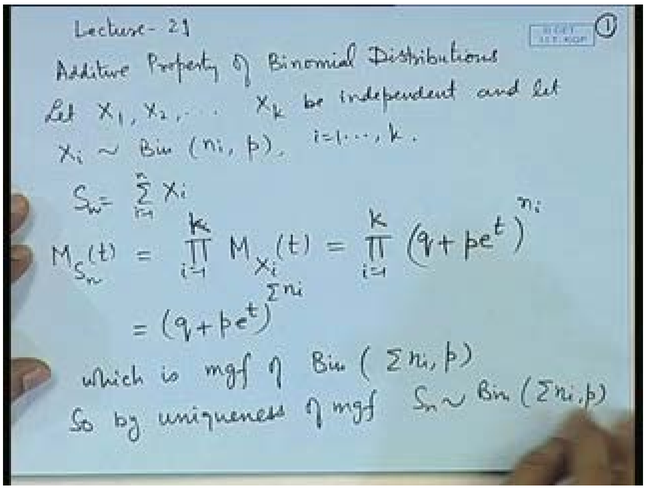
Probability and Statistics
Prof. Dr. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Module No. #01
Lecture No. #21
Joint Distributions – IV

In the last lecture, we haveconsidered jointly distributed random vectors in general K dimensional or Ndimensional random vector.We in particular, defined joint moment generating function and we proved an important property that, if the random variables are independent, then the moment generating function of the sum of the random variables can be expressed as product of the moment 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 illustrate by proving additive properties of certain distributions.

(Refer Slide Time: 01:07)

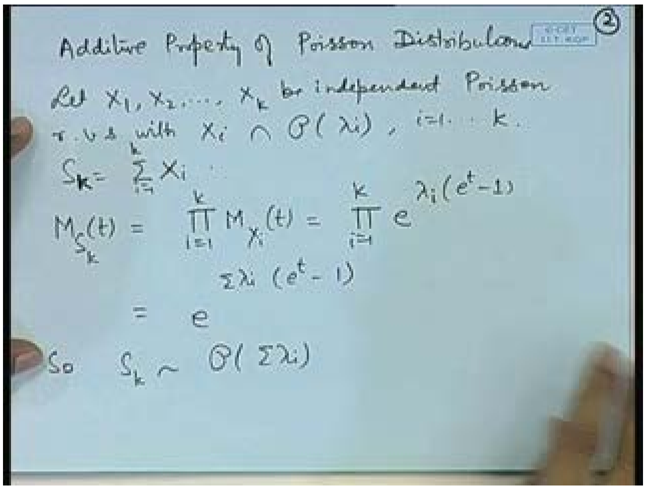


So,firstly, let us prove say Additive Property of Binomial Distributions. So, let us consider sayX1,X2,X k be independently distributed random variables and let Xi follow binomial n i,p distribution for i is equal to 1 to k

I am interested in the distribution of S, that is, sigma X i is equal to 1 to n or rather we can call it S n.If we use themgf here, distribution of the mgf of the sum is equal to product of the mgf of Xi's, i is equal to 1 to 10.Notice here, that mgf of Xi, that is q plus pe 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 pe to the power y to the power sigma n i, which is the mgf of binomial sigma n i, p distribution. So, by uniqueness property of the mgf, S n must follow binomial sigma n i, p distribution

(Refer Slide Time: 04:27)



This additive property of binomial distribution can be expressed physically also.Here, you can see that X 1denotes the number of successes in a sequence of n 1 independent and identically conducted Bernoullian trials, where the probability of success is p.

X 2 denotes the number of successes in n 2 independent and identicallyconducted Bernoullian trials with the probability of success p andso on.Therefore, sigma x i can be considered as a total number of successes inn 1 plus n 2 plus n k independent and identically conducted Bernoulliantrials, where the probability of success is p

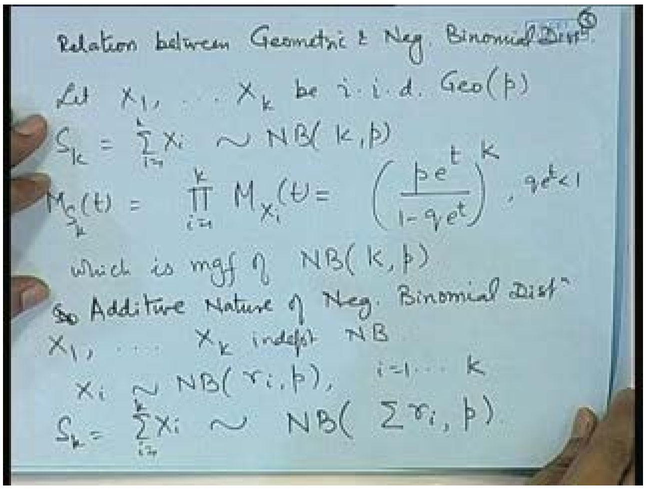
So, this physical fact is confirmed by this additive property which we are able to prove here using the momentgenerating functions.Let us prove a similar property for Poisson distributions,so additive property of Poisson distributions. So, letX 1, X 2,X k be independent Poisson random variableswith x i having a Poisson lambda i distribution.

Once again, we are interested in the distribution of sigma Xi equal to 1 to k.I think, I made a mistake here.This should be S k, that is, sigma Xi 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 genetic 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 Xi, this becomes e to the power lambda i, product i is equal to 1 to kwhich 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 Sk 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 physical terms.Here, we are k different Poisson processes,X 1denotes the number of arrivals in the Poisson process with the arrival rate lambda 1,X 2 denotes the number of arrivals in the Poisson process with arrival rate lambda 2 andso on.Therefore, sum of the xi's denotes the total number of arrivals in a Poisson process with the arrival rate sigma lambda i.

(Refer Slide Time: 07:06)



Let us consider, say a relation between geometric and say negative binomial distribution. So, letX 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 Xi.Now, if I am looking at the mgf of S k, now the mgf of a geometric random variable ispe to the power t divided by 1 minus qe to the power t, where q e to the power t is less than 1.

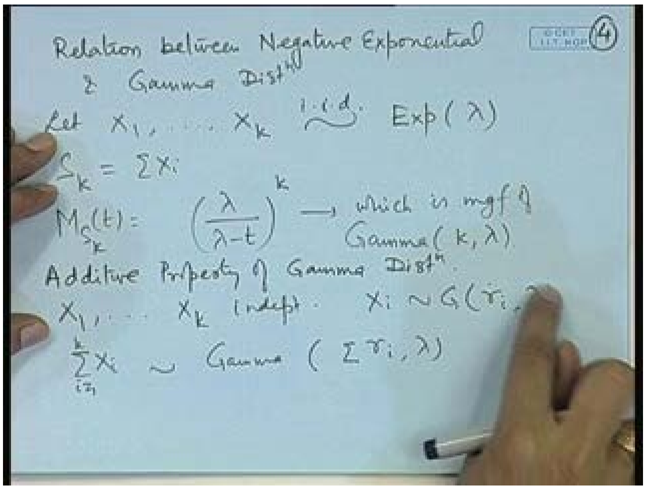
Now, when we are multiplying at k times, this becomes power k which ismgf of negative binomial with parameter k and p. So, this proves that sums of independent geometric variables with the same probability of success is negative binomial k, p.

Once again, we can look at the physical interpretation of this result.X 1denotes the number of trials needed for the first success in a sequence of independently and identically conducted Bernoulliantrials.X 2 denotes the number of trials needed for another successfor the first time in a sequence of independent and identically conducted Bernoulliantrials. Therefore,X 1plusX 2 plus X k denotes the number of trials needed for the first time k success in a sequence of independent and identically conducted Bernoulliantrials and that we know that, it has 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 haveX 1, X 2,X k independent negative binomials and say,Xi 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 Xi is equal to 1 to k, then by this property, when we are multiplying the moment generating functions,I will be multiplying pe 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 success is constant, negative binomial distribution also follows an additive property.

(Refer Slide Time: 10:51)



Let us look at a relationship between, say negative exponentialand gamma distributions. So, letX 1, X 2,X k, they will be independent and are identically distributed exponential variables with parameter lambda.

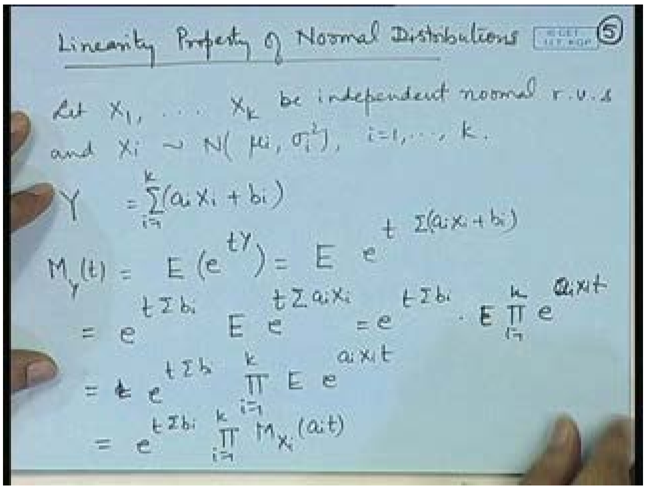
Now, let us consider the distribution of the sum. So, momentgenerating function becomes lambda by lambda minus t to the power k which is mgf of gamma distribution with parameters k and lambda.That means, sums of independent exponential variables is a gamma variable. So, physically if we represent this result,if we are observing a Poisson process with rate lambda,X 1denotes the waiting time for the first occurrence,X 2 denotes the waiting time for first occurrence at another point of time,X k denotes the waiting time for the first occurrence in a kth observation of the process.

So, if we combine these, that isX 1plusX 2 plus X k, we look at, that is whenX 1is observed, we start observing the process once again.X 2 is the time added thereafter.X 3 denotes the time starting from whenX 2 has been, that is the second occurrence has been observed, then we observe. So, thenX 1plusX 2 plus X k denotes the waiting time for the first time kth occurrence in a Poisson process and that we know, that it followsa 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 Xi follows gamma, say r i lambda.Then, sigma X i is equal to 1 to k, that will follow gamma with parameter sigma r i lambda because here, we can consider Xi as the waiting time for the first time r ith occurrence in a Poisson process with rate lambda.

So, when we add these timings, it means that it is a total waiting time for sigmar i occurrences in a Poisson process with rate lambda.Therefore, this gamma distribution also satisfies an additive property, provided the Poisson process parameter remains the same

(Refer Slide Time: 14:24)

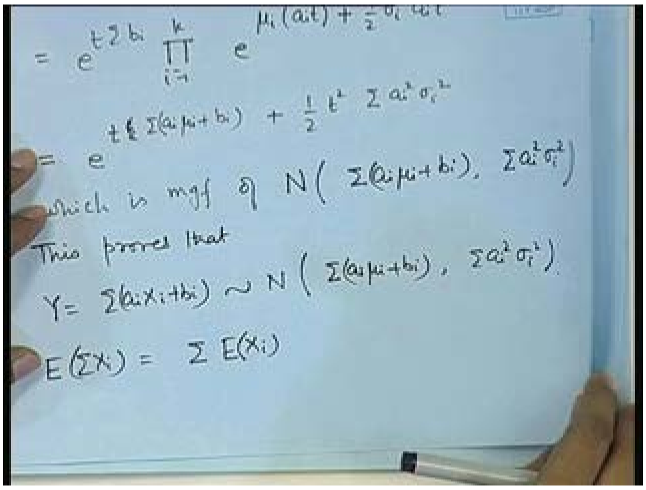


In the case of normal distribution, we have much more general property. In fact, we have a linearity property.Let us consider, sayX 1, X 2,X k independent normal variables and Xi follows, say normalmu i sigma i square for i is equal to 1 to k.

Let us consider a linear function sigma a i x i plus bi,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 ty, that is expectation of e to the power t sigma ai 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 ai x i and t.

Now, this we can express as, e to the power t sigma b i and this term we can split.We can consider it as, e to the power expectation of product e to the power sigma ai xi t.Now, here x i is the 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 ai xi t.

(Refer Slide Time: 17:20)



Now, this is nothing, but the moment generating function of the random variable xi at the point ai. So, e to the power t sigma b i product i is equal to 1 to k momentgenerating function of xi at a i t.Now,xi's follow normal distributions,therefore the moment generating function of x i can be retained.

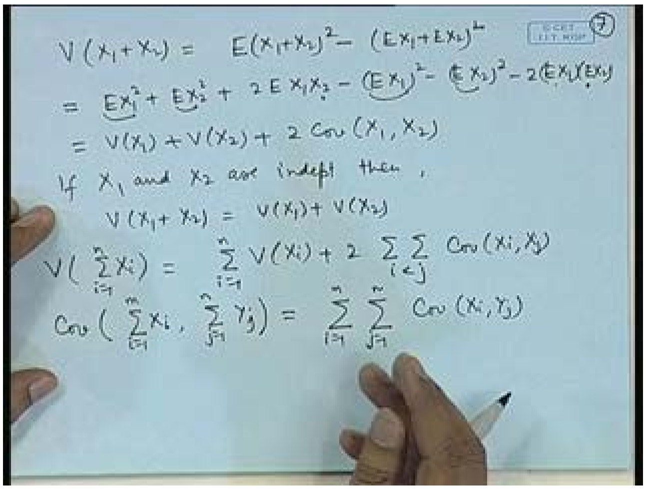
So, we substitute that here, to get e to the power mu ia i t plus half sigma i square a i square t square. So, after adjusting the terms, we get e to the power t sigma a i mu i plus b i plus half t squaresigma a i square sigma i square.

Now, this we can identify as mgf of a normal distribution with mean sigma ai mu i plus b i and variance sigma a i square sigma i square. So, by the uniqueness property of the mgf's, this proves that, that y, that 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 linear combination of the independent normal variables follows a normal distribution.Another important thing to notice here is that, in normal distribution's 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 2 parameters are there when we are considering several independent random variables, we were varying only oneof the parameter andoneof 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.The property is also more general, rather than just talking about thesums; we can talk about any linear function.

There are certainresults which are related to the calculation of the moments of sums variances of the sums etcetera,so these I will state here.For example, if we look at, say expectation of sigma Xi, it is equal to sigma expectation of Xi.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 or the summation signs because here, it is expectation of a summation. So, either you will have a, if the random variables are discrete, we 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 insideand it will prove this property.

(Refer Slide Time: 21:09)



In the case of variance, let us write for 2 of them, that is, variance of sayX 1plusX 2.Now, this is equal to expectation ofX1plusX 2 whole square minus expectation ofX 1plus expectation ofX 2 whole square. So, this is equal to expectation ofX 1square plus expectation ofX 2 square plus twice expectation X 1,X 2 minus expectation of X 1whole square minus expectation of X2 whole square minus twice expectationX 1into expectation of X 2.

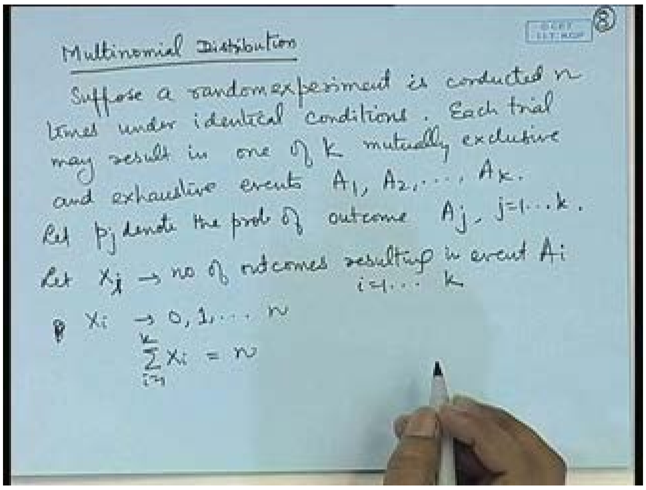
So, these terms, if we combine expectation ofX 1square with expectation of X 1whole square, this is variance ofX 1. In a similar way, expectation of X 2 square can be combined with expectation of X 2 whole square that leads to variance ofX 2.

Now, this cross-border term, that is, expectation of X 1, X 2 minus expectation ofX 1into expectation ofX 2 is nothing, but the co-variance terms variance of a sum is equal to sum of the variances plus twice co-variance of X 1, X 2. So, there is an additional term here.

Now, ifX 1and X 2 are independent, then co-variance will be zero and therefore, variance ofX 1plusX 2 will be equal to variance of X 1plus variance ofX 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 XiX j, where i is less than j.

Obviously, if the random variables X 1, X 2,X n are independent, then these co-variance's 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 co-variance of a sum with co-variance of another sum, say this is i equal to 1 to m, this is j is equal to 1 to n, then this is equal to double summation co-variance of x i y j.That meansco-variance of each term in the first summation is taken with co-variance with the each term in the second.These properties are quite useful in calculation of the moments of the sums of distributions.

(Refer Slide Time: 24:18)



We will consider a fewmultinomial distributions, which are quite commonly used.Oneof them is a generalization of binomial distribution, the so-called multinomial distribution.In the binomial distribution,we are considering a sequence of Bernoullian trials in which each trial of the experiment results in two options.Oneis called success and another is called a failure, that is, two types of outcomes are possible.

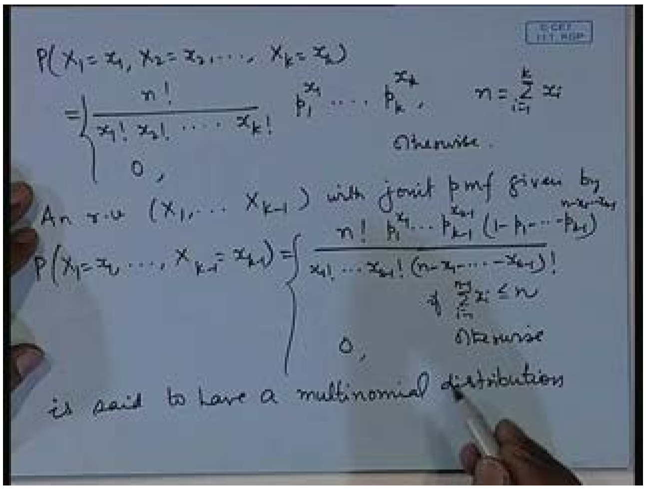
However, there are a variety of trials in which we may be interested in categorizing not only in 2, but in k type of outcomes. So, for example, if you are looking at a tossing of a di, then you have the faces coming up 1 2 3 4 5 6.If you are looking at drawing a card from a pack of cards, then it could be any of the four suits, say heart spade club or diamond or if you are looking at, saywhether what is the number on that, that is 1 2 3 upto 13.

So, there are avariety of experiments where the possible outcomes can be more than one. So, if the probability of ending up anoutcome 1 is, say p 1,ending an outcome 2 is p 2, getting an outcome k is p k and then, if we conduct a certain number of trials, say n,so out of that, say X 1is the number of outcomes resulting in first type.X 2 is the number of trials resulting in the second type of outcome etcetera,what is the distribution of that. So, that is called a multinomial distribution

So, suppose a random experiment is conducted n times under identical conditions.Each trial may result in 1 of k mutually exclusive and exhaustive events.Let us call them,say A 1,A 2,A k.Let p j denote the probability of outcome A j for j is equal to 1 to k.

So, let us consider, sayX 1number of outcomes resulting in event, say Ai.Xi denotes this for i is equal to 1 to k.Then, what are the possible values of X i's?X i's can take value 0, 1 to n subject to the condition that sigma Xi is equal to n because n is the total number of trials.

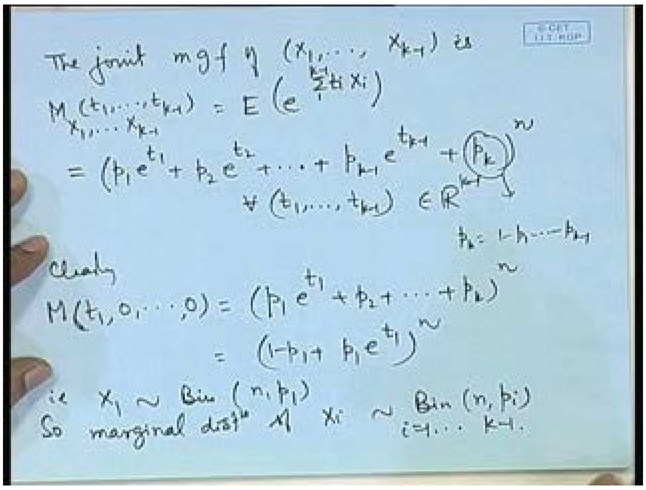
(Refer Slide Time: 28:22)



So, if wewrite down probability of, sayX 1is equal to X 1, X 2 is equal to sayX 2,X k is equal to X k, then this is equal to n factorial divided byX 1factorial,X 2 factorial andso on,X k factorial p 1 to the powerX 1,p 2 to the power X 2, p k to the power X k, where n is equal to sigma Xi, i is equal to 1 to k.It is equal to 0, otherwise.

Now, if you look at this distribution here, ifX 1, X 2,X k minus 1,that is any k minus 1 of these variables are given, the lastonecan be determined in terms of n minus the sum of the remaining ones.

(Refer Slide Time: 31:18)



So, if we consider the joint distribution of a random variable, say X 1, X 2,X k minus 1 with joint probability mass function given by this. So, we consider it as probability ofX 1is equal to X 1, X 2 is equal toX 2,X k minus 1 is equal to X k minus 1,that is n factorial divided by X 1factorial,X 2 factorial,X k minus 1 factorial n minus X 1 minusX 2 minus X k minus 1 factorial, p 1 to the powerX 1andso on, p k to minus 1 to the power X k minus 1 and then, 1 minus p 1 minus p 2 minus p k minus 1 to the power n minus X 1minus X 2 minus X k minus 1.If sigma Xi is equal to 1 to n minus 1 is less than or equal to n 0, otherwise.This is said to have a multinomial distribution.

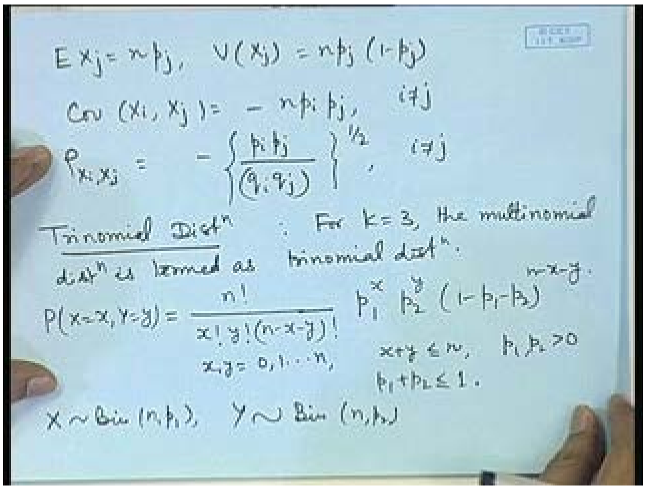
So,in fact, if you look at these 2, they are the same,but formally we define a multinomial distribution to be k minus 1 dimensional because the last value is determined automatically.Like in the binomial distribution, we talk about the distribution of the number of successes; we do not say that distribution of number ofsuccesses and failures.

Now, from a multinomial distribution, we can consider the joint mgf, the joint moment generating function of X 1, X 2, X k minus 1. So, it is evaluated at the point t 1,t 2, t k minus 1,that is expectation of e to the power sigma t i x i from 1 to k minus 1. So, if we look at this distribution here, now the sum of this is a multinomial, that is p 1 plus p 2 plus upto 1 minus p 1 minus p 2 minus p k minus 1 to the power n. So, the sum of this, overall these combinations is actually a multinomial sum.

So, if we want to calculate this term, it will become p 1 e to the power t 1 plus p 2 e to the power t 2 andso on plus p k minus 1 e to the power p k minus 1 plus p k to the power n and this is valid for all t 1,t 2, t k minus 1 belonging to R k minus 1.Now, easily you can see that, suppose I substitute t 2 t 3 upto t k minus 1 is equal to 0, I will get p 1 into the power t 1 plus p k to the power n which will become actually the, sothis p k is actually 1 minus p 1 minus p 2 minus p k minus 1.

So, clearly we can see that, M t 1 0, 0, 0 is equal to p 1 e to the power t 1 plus p 2 andso on plus p k to the power n that we can write as 1 minus p 1 plus p 1 e to the power t 1 to the power n, that isX 1follows binomial n p 1.That means, the marginal distributions of X I’s are binomial n p i for i is equal to 1 to k minus 1.

(Refer Slide Time: 33:52)

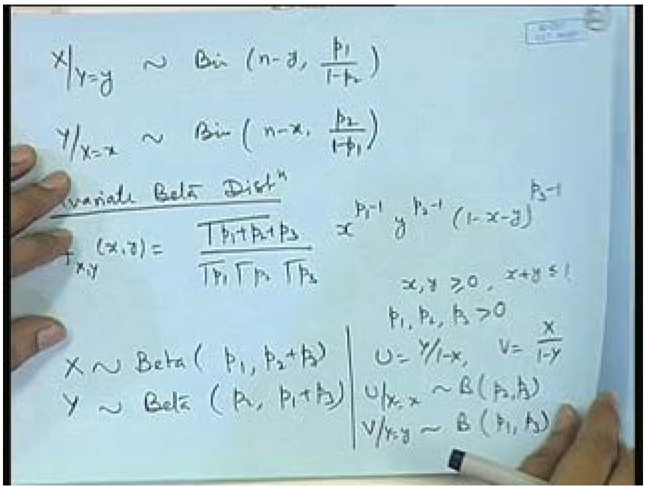


In particular, we can talk about expectations. So, naturally expectation of x j will be n pj variance of x j will be n pj, 1 minus p j.We can also talk about the co-variance terms between x iand x j, that will be minus n pipj for i not equal to j.Therefore, correlation co-efficient between x i and x j can be calculated to be minus p ip j divided by q i q j, where q i q j denotes 1 minus p i and 1 minus p j etcetera to the power half for i not equal to j.So, the correlation co-efficient between 2 of these can also be calculated.

In particular, if in themultinomial distribution, I consider k equal to 3. So, I will have 2 of these variables that areX 1, X 2, that distribution is called a trinomial distribution. So, that is a straightforward generalization of binomial distribution to the casewhen we are having 3 categories as the outcomes. So, it is called Trinomial distribution, that is, for k equal to 3.The multinomial distribution is termed as trinomial distribution. So, if I say trinomial distribution, we can write the probability mass function as n factorial divided by x factorial y factorial n minus x minus y factorial p 1 to the power x, p 2 to the power y, 1 minus p 1 minus p 2 to the power n minus x minus y.

Here, x and ycan take values 0, 1 to n subject to the condition that x plus y will be less than or equal to n and of course, p 1,p 2 are greater than 0 subject to the condition that p 1 plus p 2 is less than or equal to 1.

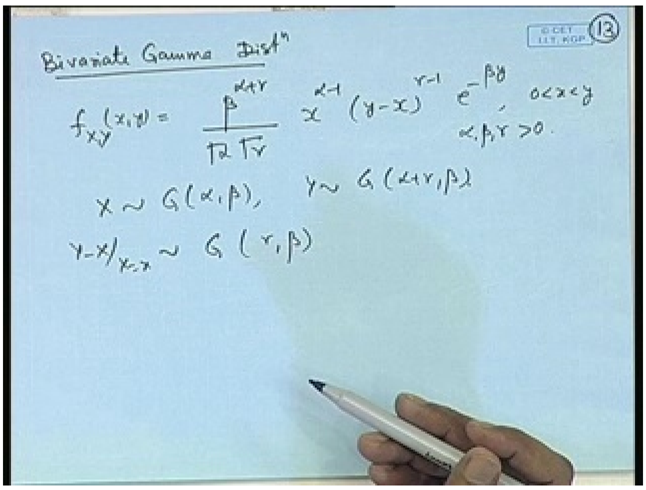
(Refer Slide Time: 36:41)



So, here, if I look at the marginal distribution of x that will be binomial n p 1.If I look at the marginal distribution of y, that will be binomial n p 2.Not only that, if we look at the conditional distributions of x given y and y given x, then conditional distribution of x given y is binomial n minus y, p 1 by 1 minus p 2 and y given x has binomial n minus x p 2 by 1 minus p 1.

Of course, when we write this p 1 by 1 minus p 2 and p 2 by 1 minus p 1,we are assuming that the sums have become, the number is between 0 to 1.

(Refer Slide Time: 39:46)



So, this isoneparticularbinomial distribution which is trinomial distributionand a general multivariate distribution, that is, a multinomial distribution. So, it is a generalization of the univariate binomial distribution.We have a couple of more generalizations,for example, we have done beta distribution. So, a beta distribution can be generalized as a bivariate beta distribution inthe following way fxy as gamma p 1 plus p 2 plus p 3 divided by gamma of p 1,gamma of p 2,gamma of p 3, x to the power p 1 minus 1,y to the power p 2 minus 1,1 minus x minus y to the power p 3 minus 1,where x and y are greater than or equal to 0 and x plus y is less than or equal to 1,p 1 p 2 p 3 must be positive.

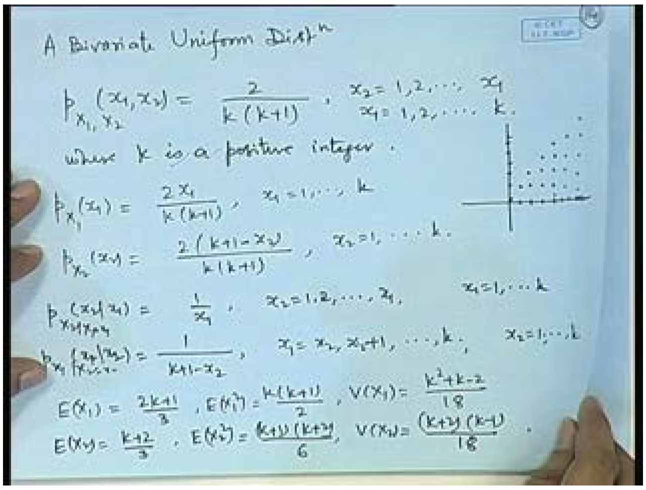
Here, we can see that the marginal distribution of X is beta with parameters p 1 and p 2 plus p 3.In a similar way, marginal distribution of Y can be calculated; it is beta with parameters p 2 and p 1 plus p 3.

The conditional distributions are also beta with a little scaling.For example, if I consider U is equal toY divided by 1 minus X or V is equal to X divided by 1 minus Y, then U given X is equal to X follows beta distribution with parameters p 2 p 3,if we consider V givenY that follows beta distribution with parameters p 1 and p 3.

In no way, this is a unique generalization of a beta distribution to do two-dimension.We can generalize in different ways also.What we are trying to see here is that the marginal distributions also have beta distribution. So, then in that case, we are calling it as a bivariate beta distribution.A very similar thing is done for gamma distribution. So, a bivariate gamma distribution can be defined as f xy beta to the power alphaplus gamma divided by gamma alpha gamma gamma, x to the power alpha minus 1,y minus x to the power gamma minus 1,e to the power minus beta y, where 0 less than x less than y alpha beta gamma greater than 0.

Here, if we see the marginal’s, x follows gamma alpha beta and y follows gamma alpha plus gamma and beta.Also, y minus x given x, this follows gamma,gamma, beta. So, this is another generalization of a univariate gamma distribution to a bivariate gamma distribution.

(Refer Slide Time: 41:00)



We consider a bivariate uniform distribution; it is a discreteuniform distribution. So, consider probability of x 1is equal to x 1, x 2 is equal to small x 2 as 2 by k into k plus 1,where x 2 takes values 1 to x 1and x 1takes values 1 to k.For example, if we consider the points 1 2 3 upto k, then if I take x 1is equal to 1,then x 2 will take value 1.If I take x 1is equal to 2, then x 2 can take values 1 and 2, that is 1 and 2.If I take x 1is equal to 3, thenx 1can takex 2 can take values 1, 2 and 3.

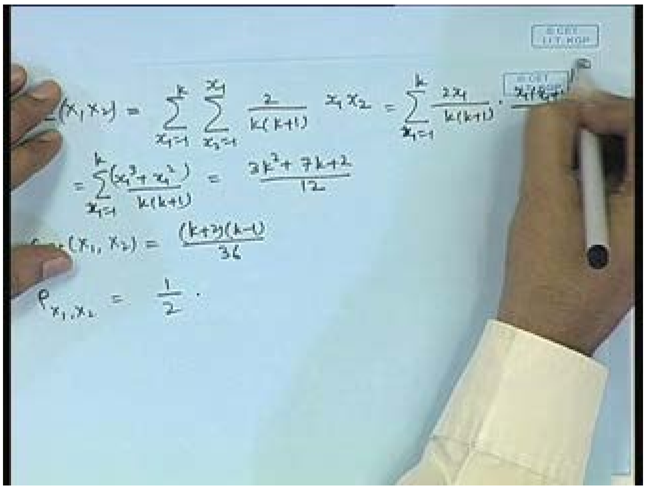
So, you can see that,this distribution is a uniform discrete uniform distribution with probabilities concentrated on these diagonals.This you can say, it as the one half of the square.Actually, we can easily see that the marginal distributions, if we sum over x 2 from 1 to x 1, this will give 2x 1by k into k plus 1 for x 1is equal to 1 to k.

So, the marginal distribution of x 1is obtained like this.In a similar way, if we sum over x 2,sum over x 1from x 2 to k, then we get the marginal distribution of x 2 as 2 into k plus 1 minus x 2 divided by k and 2k plus 1 for x 2 is equal to 1 to k.

Important thing here to notice is, that if I take the conditional distributionsof x 2, given x 1, it is a discrete uniform distribution on 1, 2 upto x 1. Similarly, the conditional distribution of x 1, given x 2 is a discrete uniform distribution univariate 1 by k plus 1 minus x 2 that isranged from x 2 to k.

So, this generalization of a bivariate of a discrete uniform distribution to two-dimensions has the conditional distributions as discrete uniforms, but the marginal are not uniform.We can calculate certain moments about this distribution which can be obtained from the marginal distributions, like expectation of x 1, expectation of x 1square, variance of x 1and similar characteristics for the x 2 variable.

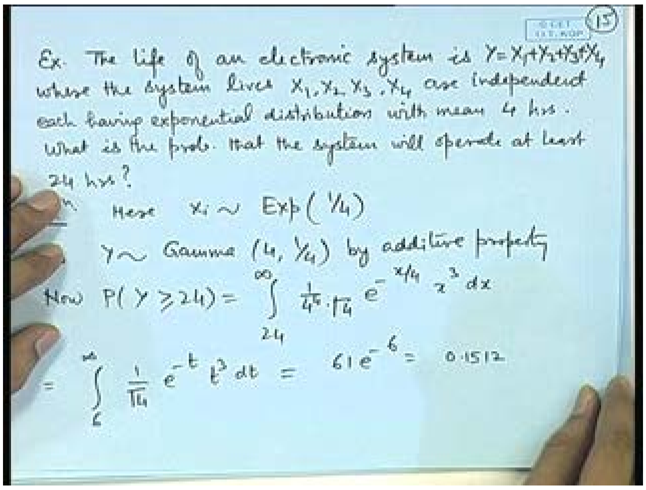
(Refer Slide Time: 44:01)



We can also look at the product moment here,that is expectation of x 1, x 2 which is calculated from the joint distribution of x 1, x 2 and after certain simplification, it turns out to be 3 k square plus 7 k plus 2 by 12.

So, co-variance of x 1, x 2, that is expectation of x 1, x 2 minus expectation of x 1into expectation of x 2. So, after simplification, this quantity turns out to be k plus 2 into k minus 1 by 36. So, now, if we divide co-variance by the product of the square root of the variances, we get simply half because it is k plus 2 into k minus 1 by 18, both of these. So, the value turns out to be half. So, the co-relation co-efficient between the random variables x 1and x 2 is half here.

(Refer Slide Time: 45:07)



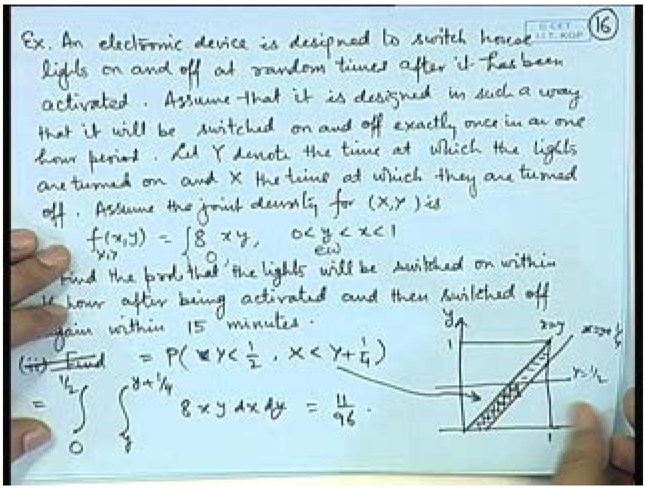
We look atsome application of the additive propertieshere.Let us consider here, suppose the life of an electronic system is described as the sum of 4 independent exponential lives. So, y is x 1plus x 2 plus x 3 plus x 4 and each of the x i's is exponential with mean life 4 hours.

So, what is the probability that the system will operate at least 24 hours?That means, we are interested to find out, what is the probability of y greater than or equal to 24.Now, here, we can use the additive property of the exponential distribution.We have proved that the sums of independent exponentials are following a gamma distribution.

So, here each x i is exponential with parameter lambda is equal to 1 by 4.Here, mean is 4, that is parameter lambda will be 1 by mean because mean is 1 by lambda. So, y will follow gamma distribution with parameters 4 and 1 by 4.Now, the density of a gamma distribution with parameters 4 and 1 by 4 is given by 1 by 4 to the power 4 gamma 4 e to the power minus x by 4xs to the power 4 minus 1.

So, we integrate from 24 to infinity. So, after some simplification, this value turns out to be 61 a to the power minus 6, that is 0.1512, that is there is almost 15 percent of the chance that the system will be operating for at least 1 day.

(Refer Slide Time: 46:47)



Let us look at onemore application of bivariate distributions. So, consider an electronic device. So, it is designed in such a way, that it is switchto switch house lights on or off at random timesafter it has been activated.Assume that, it has been designed in such a way, that it will be switched on and off exactly once in a 1 hour period.

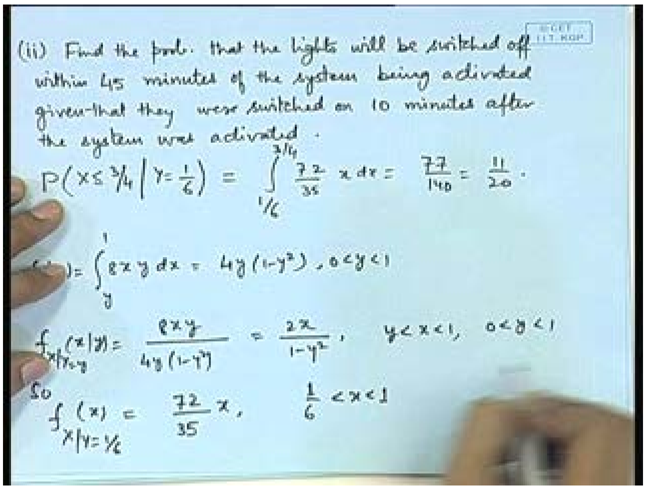
Let y denote the time at which the lights are turned on and x, the time at which they are turned off.That means, firstly, it will be switched on and then they will be switched off and the joint density function for xy is given by 8 xy for y less than x.Of course, since we are considering only 1hour period,so both will lie between 0 to 1and it is 0 elsewhere.

What is the probability that the lights will be switched on within half an hour after being activated and then, switched off again within 15 minutes.That means, what is the probability that y is less than half and x is less than y plus 1 by 4 because within 15 minutes of getting on, it should be switched off. So, xmust be less than y plus 1 by 4.

Now, to determine this probability, we look at the region of integration of the density. So, you see here, the density is defined for in a unit square.The density is defined for y less than x, that is, this region.Now, here y less than half means that we are in the bottom region and x is less than y plus 1 by 4.Now, x is greater than y, that means here, and the line x equal to y is equal to 1 plus 4 is this line.

So, we are basically in this zone. So, we have to integrate the joint density 8 xy over this region. So, the limits of integration are for x y to y plus 1 by 4 and for y, it will be 0 to half. So, this turns out to be 11by 96.

(Refer Slide Time: 49:02)

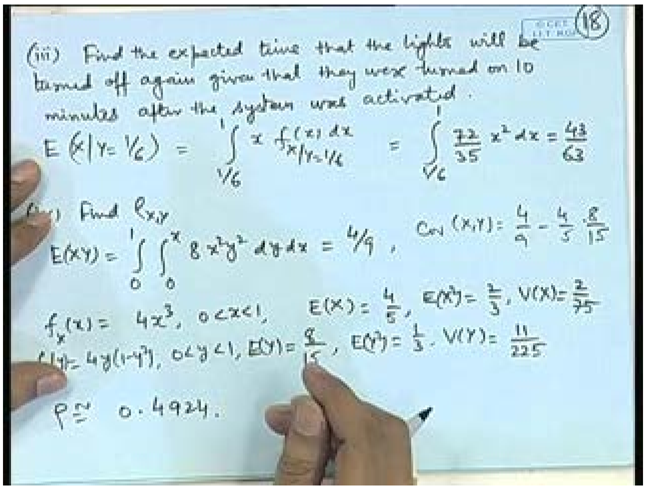


In the same problem, let us look at what is the probability that the lights will be switched off within 55 minutes of the system being activated, given that they were switched on 10 minutes after the system was activated.That means, what is the probability that x is less than or equal to 3 by 4, given that y is equal to 1 by 6.

So, we need the conditional distribution of x given y.So,firstly, we look at the marginal distribution of y. So, here, the joint distribution is 8 xy.We will integrate with respect to x.Now, the region of integration for x is from y to 1. So, after integrating, we get 4 y into 1 minus y square.Therefore, the conditional distribution of x given y is evaluated as the ratio of the joint distribution divided by the marginal distribution of y, which turns out to be 2 xy 1 minus y square for x lying betweeny and 1,where y is a value fixed between 0 to 1.

So, the conditional distribution of x, given y is equal to 1 by 6 is easily evaluated by substituting y is equal to 1 by6 here and we get 72 by 35 x for x lying between 1 by 6 to 1.Therefore, this conditional probability is obtained by evaluating the integral of this density over the region 1 by 6 to 3 by 4 and it is evaluated to be 11 by 20.

(Refer Slide Time: 50:44)



Find the expected time that the lights will be turned off.Again, given that they were turned on 10 minutes after the system was activated.That means, what is the expected value of the distribution that we obtained just now,given y is equal to 1 by 6. So, the density is given here over this region. So, we calculate the expected value as integral of x into the density from 1 by 6 to 1 dx and it is evaluated to be 43 by 63.

Finally, in this problem, what is the correlation co-efficient between the random variables x and y. So, in order to evaluate the co-relation, we need the co-varianceterm and the expectations of x and y and the variances of x and y.In order to evaluate theco-varianceterm, we need the product moments.Here, expectation of xy is x into y into the joint density, that is,8 xy is integrated from 0 to x and x is integrated from 0 to 1,which is 4 by 9.The marginal distribution of x is obtained by integrating with respect to y from 0 to x, which is simply 4 x cube. So, expectation of x turns out to be 4 by 5 variance of x turns out to be 2 by 75.

The marginal distribution of y was evaluated here as 4 y into 1 minus y square. So, we can evaluate the mean and the variance of y also.Therefore, we can find the co-variance between xy as 4 by 9 minus 4 by 5 into 8 by 15 and the co-relation turns out to be, this divided by the square root of the variances of x and y which is .49 approximately, that is nearly half.

So, in this particular problem, the timings of switching on and off of the lights ishaving co-relation nearly half here.

We have in general, discussed the joint distributions,where we initially considered 2 variables.Then, we consideredmultivariate, that is, k dimensional or n dimensional random variables.We have discussed the concept of marginal distributions, conditional distributions, the concept of co-relations and we have also discussed certain additive properties of the distributions. So, now, in the next lecture, we will see that if we consider transformations of the random vectors, how to obtain the distributions of that. So, in the next lectures, we will be covering that topic.Thank you.