Advanced Engineering Mathematics

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

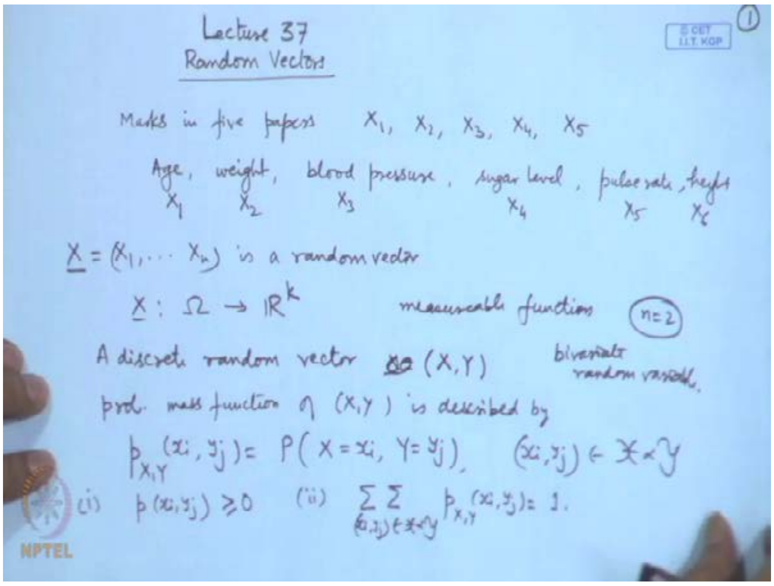
Lecture No. # 38

Topic

Joint Distributions and Sampling Distributions

Here, I have introduced the concept of random variables. So, the concept of random variable is defined in the way that it is a real valued function on the sample space; that means, we are interested in a single characteristic which is a reflected from the sample space single probability or one value we want to take, but many times from the random experiment we want to extract more values. Suppose we are considering certain examination.

(Refer Slide Time: 01:04)



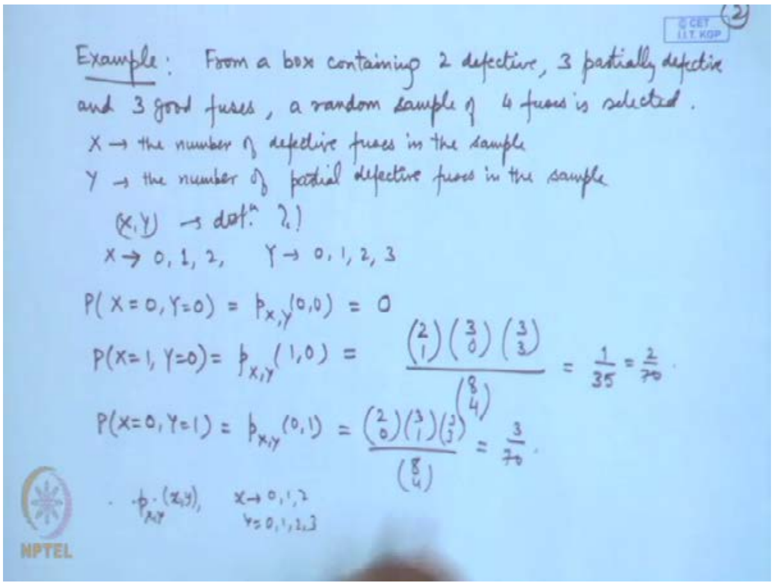
And we are looking at marks of a student. Then we are we may be looking at the marks of the student in say five different subjects. So, marks in 5 papers. So, it could be like X1, X2, X3, X4 and X5. If a patient goes to a doctor and doctor take certain measurements on him. For example, the doctor may ask his age, the doctor will take his weight; the doctor may record called his say blood pressure. So, we may right say for example, in X 1 as the weight. The age X 2 has as the weight; X 3 has as the blood pressure or we may also look at says sugar level his pulse rate may be his height also. In that case from the single random experiment we are extracting 6 dimensional vectors.

So, in general we say that X1, X2,…Xn is a random vector let me denoted by X random vector. Then X is actually a function from omega into R^k and we keep the condition of measurable that is a measurable function. Now like in the case of random variable we may have a discrete or continuous random vector. We may also have in a various type of cases for example, when we are according age. Age may be recorded in the rounded of years. In that case it will be a discrete random variable. Weight may be recorded as a continuous variable blood pressure, sugar level etcetera etc. may be recorded as the say continuous variables.

So, it could be that some components of the random vector are discrete, some components are continuous or all of the components are discrete and all the components are continuous. So, in that case the distribution of the random vector is described differently. Let us take two it's a special case one is when all the components are discrete and second one when all the components are continuous. So, we consider a discrete random vector. So, X is equal to now I will restrict my attention to two dimensions n is equal to 2. let us consider. So, (X, Y). So, this is also called as bivariate bivariate random variable.

So, if we are saying, it is a discrete we can consider the probability mass function probability mass function of X,Y is described by p X,Y(Xi,Yj) that is equal to probability. X is equal to Xi and Y is equal to Y j for X i, Y j values belonging to summer some space of values of in X cross Y. we have two conditions. one condition is that p (Xi,Yj) is greater than or equal to 0. And second condition is that when we some over all the possible values of Xi, Yj in this a space then this is equal to 1.

(Refer Slide Time: 05:07)

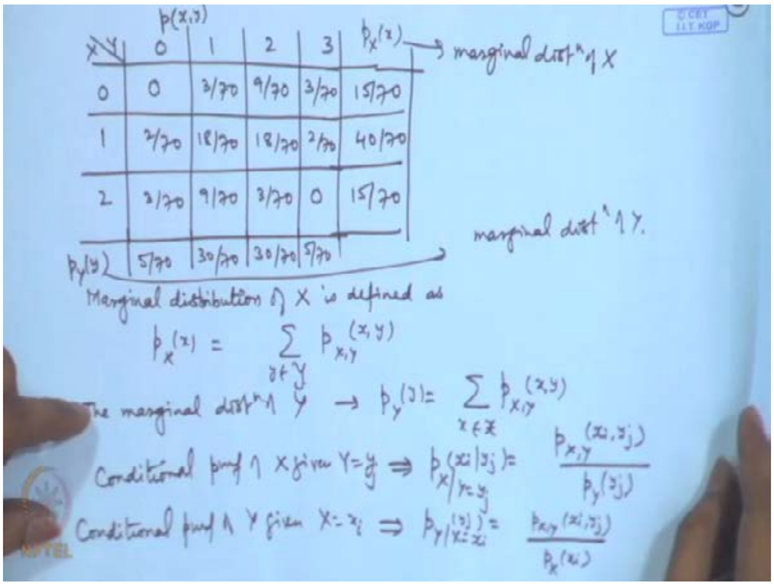


Let we explain through one example here, From a box containing two defective, three partially defective and three good fuses, a random sample of 4 fuses is selected. Now let us consider X is the number of in defective fuses in the sample and Y is the number of partially defective fuses in the sample. We want the distribution of X, Y what is the distribution of X, Y. So, we assume that all the selections are equally likely in that case we can evaluate this probability distribution in the following fashion.

First of all what are possible values are that X and Y can take? Since there are maximum 2 defectives. X can take values 0, 1 and 2. And Y can take values zero 1, 2, 3. And let us consider, what is the probability that X is equal to say 0 and Y is equal to 0. Now so, this is actually p X, Y (0 0). Since the maximum number of good fuses is three 3 and we are selecting 4 four fuses therefore, at least one will be either defective or partially defective it is not possible that all of them are good. Because X is equal to 0, Y is equal to 0 corresponds to the case when all the fuses which have been selected in the sample are good, but that is not possible.

So, this probability is, Let us consider another one probability that X equal to 1, Y is equal to 0. In notational form will write us p X, Y (1, 0). Now these mean that out of our selection from two defectives. one has been selected and from partially defective none or selected and all other good ones are selected. So, basically it is like 2 c 1, 3 c 0 and 3 c 3 divided by 8 c 4. So, we can simplify this turns out to be 1 by 35 or we may write it as 2 by seventeen 70. If we consider say probability X is equal to 0, Y is equal to 1 that is p X,Y (0, 1) then this turns out to be 2 c 0, 3 c 1 and 3 c 3 divided by 8 c 4 and this value can be evaluated as 3 by 70.

(Refer Slide Time: 09:17)



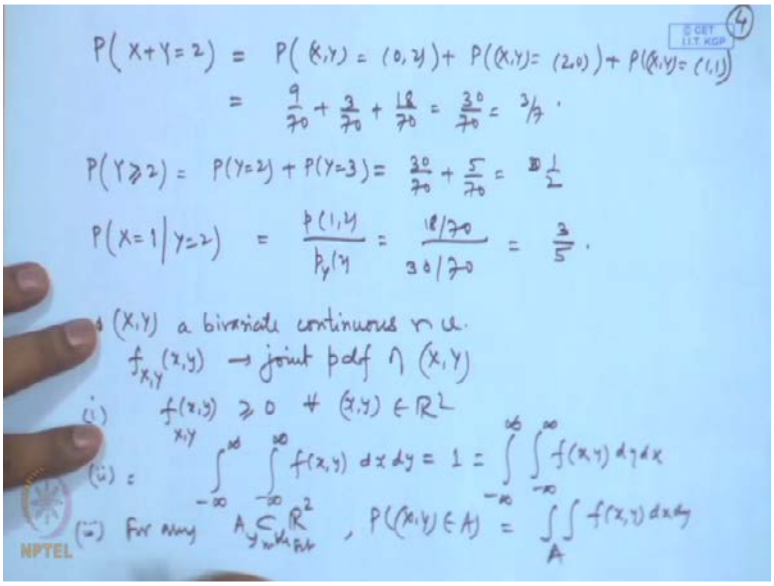
Now likewise we can calculate all other terms that is p X, Y (X, Y) for X taking value 0, 1, 2. Y taking value 0, 1, 2, 3. This p X, Y can be calculated. We can represent it in a tabular form like; on this side we can show the values of X, on this side value of Y 0, 1, 2 and 0, 1, 2, 3. So, this table will represent the probability distribution of X, Y. So, this is zero 0. p zero (0,1) we have calculated it is 3 by 70, p (1,0) we have calculated that is equal to 2 by seventy 70 and in a similar way other values can be calculated. I am substituting this values 9 by seventy70, 3 by 70, 18 by 70, 18 by 70, 2 by 70 and then you have 3 by 70 and 9 by seventy 70 and 3 by 70 and 0.

Because X equal to 0 Y is equal to X is equal to 2 and Y is equal to 3 is not possible. Because total number of selections is only 4. Now we also introduce a concept of marginal and conditional distributions. So, marginal distribution of X is defined as p X(x) as the summation over all the values of Y. similarly the marginal distribution of Y that is defined as p Y(y) if we some over all the values of X. We can also define. So, in this particular case for example, the marginal distribution of X p X(x) it is obtained by summing the probability distribution of X, Y over Y. So, p probability X is equal to 0.

That is obtain by summing over Y is equal to zero 0, Y is equal to 1, and Y is equal to 2, and Y is equal to three 3. So, that is equal to 15 fifteen by seventy 70, forty 40 by seventy 70 and 3 plus 9 plus 3 that is equal to 15 by 70. So, this right most column gives the marginal distribution of X. Similarly if we add with respect to X we get the marginal distribution of Y. for example, if we yet the first column we get the probability of Y is equal to zero 0. So, this is equal to 5 by 70. If we add the values in second column we get 30 thirty by seventy 70.

If we add in the column corresponding Y is equal to 2 we get 30 by 70 and if we add in the column corresponding to Y is equal to 3 we get 5 by 70. So, this is denoting the marginal distribution of Y. We can also talk about the conditional distributions. Conditional probability mass function of X given Y is equal to Y. That is defined by p X given Y is equal to say Y j, Xi given Y j that is equal to the joint distribution of Xi, Y j X and Y divided by the marginal distribution of Y add the point Y j. Similarly we can define conditional distribution of Y given X is equal to X i say that is p Y given X is equal to X i, Y j that is equal to p X,Y(xi, yj) divided by PX(xi). So, once we are having the conditional and the marginal distributions we can answers any questions regarding the probabilities of X,Y or X given Y are the joint distributions of X and Y.

(Refer Slide Time: 14:12)



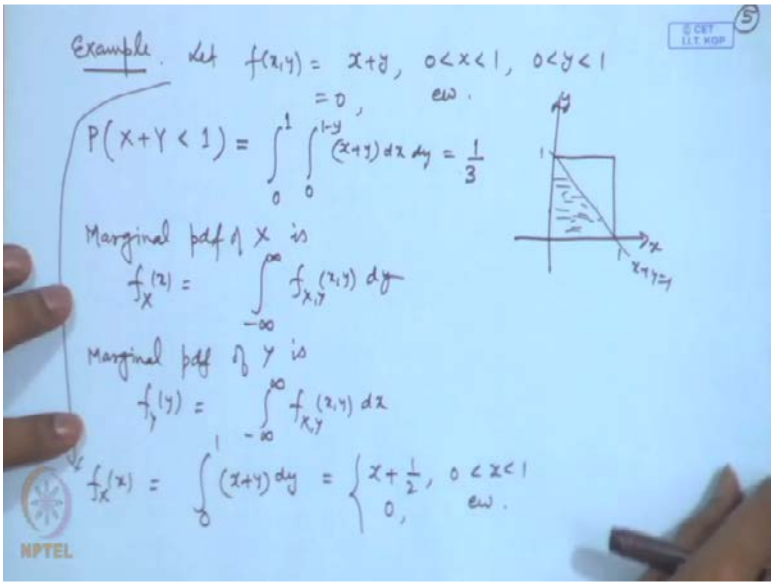
For example, what is the probability here; say we want to find what is the probability that X plus Y is equal to 2 then this is equal to probability of X,Y is equal to say (0,2), to plus probability of X,Y is equal to (2,0), plus probability X,Y is equal to (1,1 one). So, that is equal to 0 to probability is given by 9 by 70, Probability of X equal to 2 and I Y is equal to 0 that is given by 3 by seventy 70 and probability of X equal to 1 and Y equal to 1 is given by 18 by 70 that is equal to 30 by 70 are or 3 by 7.

We can answer any question regarding say Y what is the probability say that Y is greater than or equal to 2 then its equal to probability of Y is equal to 2 plus probability of Y is equal to 3 that is equal to thirty 30 by seventy 70 plus 5 by seventy that is equal to half. Similarly suppose I want to answer, what is the question regarding X is equal to 1 given Y is equal to two 2. So, these is equal to probability (1, 2) divided by probability of Y (2) is equal to 2 now p (1, 2) that is equal to 18 by 70 divided by probability of Y (2) is equal to 2 30 by 70. So, that is equal to 3 by 5.

We can answer probability statement. So, regarding the joint marginal and conditional distributions. Now let me also define the probability density function for a continuous random variable bivariate continuous random variable. So, let us consider X, Y a bivariate continuous random variable say and the probability density function it is called the joint probability density function of X, Y. So, this will satisfy the properties that f (X, Y) has to be greater than or equal to zero 0. For all (X, Y) the integral over the whole region over R 2 of f (X, Y) must be 1.

Of course, here the order of integration is not important. Whether we do dx dy or dy dx both should give the same answer. And for any set say A in R 2 of course, this should be a measurable set.

(Refer Slide Time: 17:32)

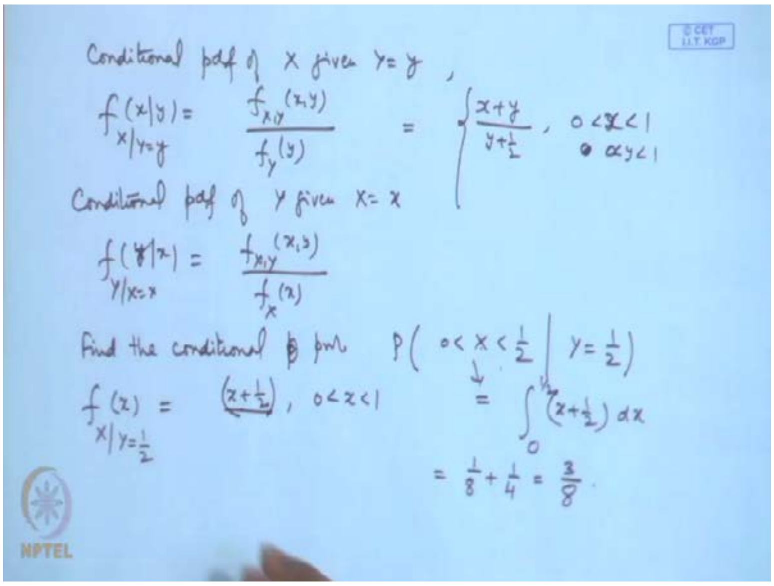


Probability X, Y Belongs to A is given by the integral of the joint density over the region A. Let me explain through one example. Let us consider say f (x, y) is equal to x plus y for x between zero 0 to 1, y between 0 to 1. Suppose I want to calculate what is the probability of x plus y less than 1. Then let us determine the region in the 2 two dimensionally space here x and y both are line between zero 0 to 1 and when we are saying a x plus y is less than one 1. Then this region is so, the region of integration than becomes for x it is from 0 to 1 minus y, this line is x plus y is equal to 1 and for y it will be from 0 to 1. So, 1 one can easily evaluate this integral this values are some to be 1 by 3.

Like in the case of discrete random variable here also we can define the marginal and conditional distributions. For example, marginal probability density function of x is given by integrating the joint distribution with respect to y over the appropriate region. Similarly the marginal distribution of y that is given by f Y (y) is equal to integral of the join density with respect to x over the given region.

For example, in this particular case f X(x) in this case this will be equal to integral of x plus y dy from 0 to 1 that as give as x plus half. Similarly if we consider say here f Y(y) that will be y plus half for 0 less than y less than 1 and 0 elsewhere.

(Refer Slide Time: 20:04)

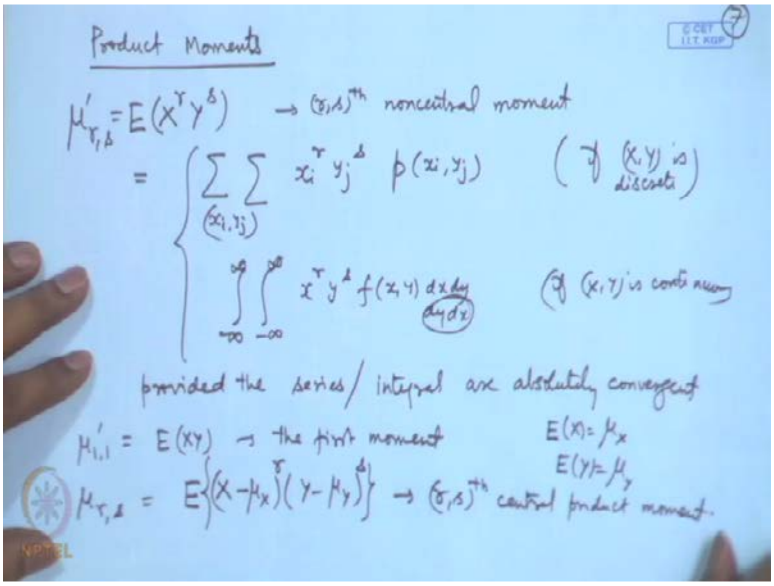


We can also talk about the conditional densities conditional probability density function of X given Y is equal to y. So, that is defined by f x given y is equal to y. It is equal to the join distribution of f X,Y(x, y) divided by the marginal distribution of fY(y). In a similar way we can talk about the conditional probability density function of y given x that is equal to the join distribution of f X, Y(x, y) give divided by the marginal distribution of f X(x).

For example, in this case this value will turn out to be the join distribution is equal to x plus y divided by the marginal of y that is y plus half for 0 less than x less than 1. And here Y is any value fix between 0 and 1. Suppose I say find the conditional probability say probability that say X lies between 0 to1/2. Given that Y is equal to half in that case i need the conditional probability distribution of X given Y is equal to half. So, from here we can substitute the value of y is equal to h i will get X plus half divided by Y plus half.

So, now y is equal to half so, this will become one 1. So, that is t simply X plus half to 0 less than X less than one 1. So, now if I want the conditional probability here of 0 less than X less than half given y is equal to half then I will be integrating this density that is x plus half from 0 to half that is equal to x square by 2 that will be 1 by 8 plus 1 by 4 that is equal to 3 by 8. So, if we have the joined distribution of X, Y. We can find out the marginal distributions we can find out the conditional distributions and we can answer any probability statement regarding the join probability, the marginal probabilities of x or y or the conditional probabilities of x given y or y given x. Now in the case of univariate random variable we have introduced the concept of moments. Now in a similar way in the case of vicariate bivariate random variable also we can talk about moments, we will talk about a slightly original concept we call it product moments.

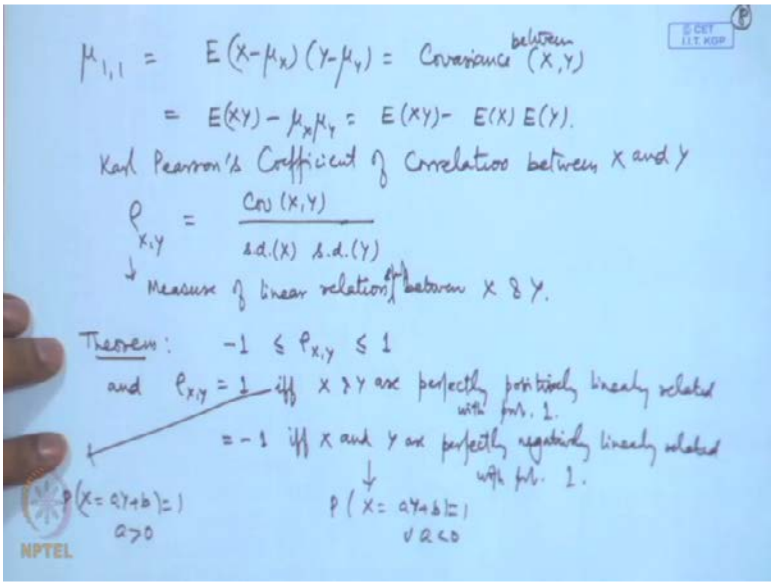
(Refer Slide Time: 23:08)



So, in general we define expectation of X to the power r, Y to the power s. I will use the notation mew crime r,s this is called r s the (r,s)th no central moment. And it is evaluated by xi to the power r, yj to the power s over all x i, yj into the joint probability mass function if f(x, y) is discrete provided these double summation is absolutely convergent.

Similarly if we have continuous then we can write it as x to the power r, y to the power s f (x,y) d x d y and of course, it could be d y d x also. If f (x, y) is continuous provided this vicariate integral is absolutely convergent. Provided the series or the integrals or absolutely convergent. In particular mew 1, 1 one prime that is equal to expectation of x into y, that is called the first product moment. We can also define the central product moments. So, mew r, s that is defined as expectation of x minus. So, let me use a notation expectation of X is equal to say mew x.

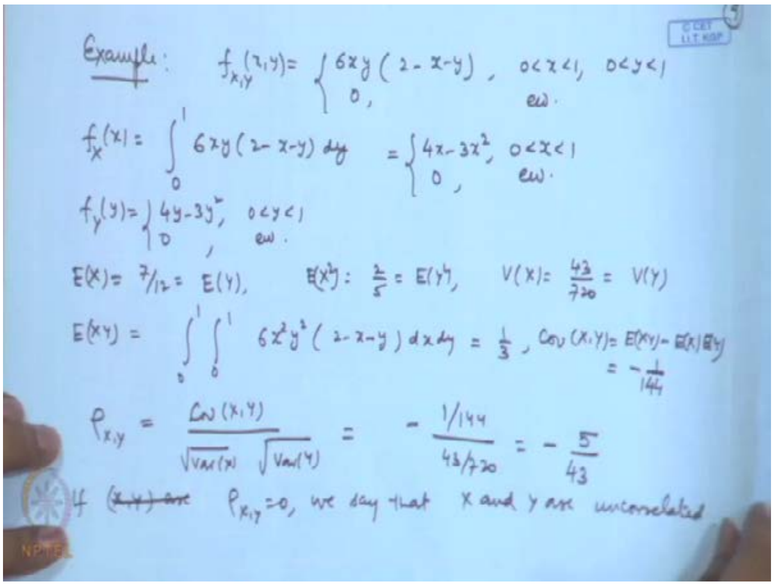
(Refer Slide Time: 25:53)



And expectation of Y is equal to mew y that is the mean of X and mean of Y inters in terms of that we define the expectation of x minus mew x into y minus mew y to the power r and this is to the power s this is called r s the central product moment. In particular if i take r is equal to 1, s is equal to 1 that is called expectation of X minus mew x into Y minus mew y that is defined as covariance between x and y covariance between x and y. Infect this as a simplified version also we can write it as expectation of x y minus mew x mew y or expectation of x y minus expectation of x into expectation of y. Using this we define Karl Pearson's coefficient of correlation between x and y is a rho x,y that is defined to be covariance between x,y divided by a standard deviation of x into a standard deviation of y.

Now this correlation coefficient is actually a measure of linear relation this is a measure of linear relation between X and Y a linear relationship. Infect one 1 can prove that minus 1 less than or equal to rho X, Y is less than or equal to 1 and rho X, Y is equal to 1 If a non leaf x and y are perfectly positivel, linearly related with probability 1 and equal to minus 1. If a non leaf X and Y are perfectly negatively linearly related with probability 1. Actually this positive relationship means that X is equal to aY plus b where a is positive. And negative relationship will mean X is equal to aY plus b, a is equal to 1 where a is negative.

(Refer Slide Time: 29:06)

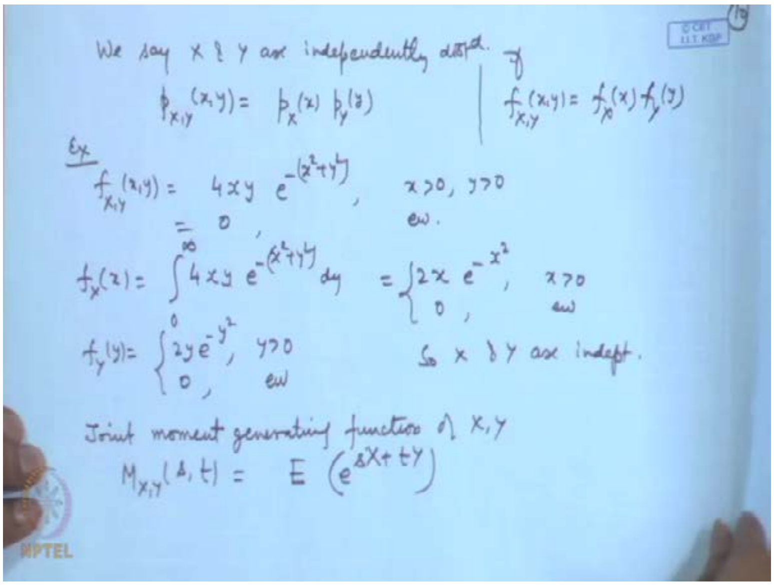


I will explain this concept through calculation in one case, let us consider say f (x, y) is equal to say 6xy of 2 minus x minus y for x between 0 to 1 and y between 0 to 1. Now for this distribution let us can calculate the covariance and correlation etc. So, if you look at the marginal distributions f X(x) that is equal to integral of 6xy into 2 minus x minus y dy from o to one 1. Then this is simplifying to 4x minus 3x square for 0 less than 1 and 0 elsewhere.

And i, i look at f Y(y) that is also of the same form because of the symmetric. So, we can calculate expectation x expectation y etcetera etc. So, expectation x terms out to be 7 by twelve 12 that is same as expectation y.

If we calculate expectation of x square that is equal to 2 by 5 that is equal to expectation of Y square. And if, we calculate variance of x that terms out to be forty 43 by 720 that is variance of y. We also calculate expectation of x into y that is equal to double integral 6x square y square of 2 minus x minus y dx dy 0 to 1 0 to 1. So, 1 one can evaluate this is turns out to be 1 by three 3. So, covariance between X, Y that is equal to expectation of x y minus expectation x into expectation y terms out to be minus 1 by 144 and therefore, the coefficient of correlation that terms out to be that is covariance of x, y divided by a square root of variance of x into a square root of variance of y the terms out to be minus 1 by 144 divided by forty 43 by 720 that is equal to minus 5 by forty 43 which is a very low negative value. So, if x, y are if rho x,y is equal to 0 we say that x and y are uncorrelated.

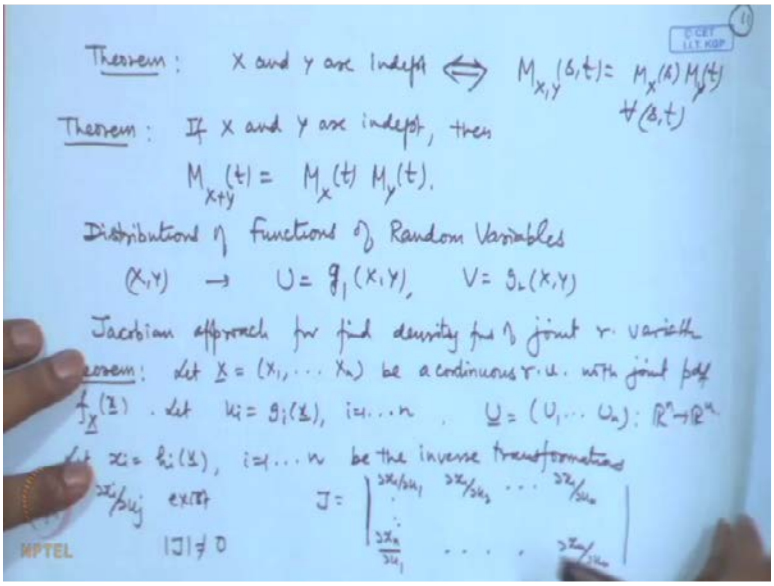
(Refer Slide Time: 31:55)



Now there is a related concept that is of independence. We say x and y are independently distributed random variables. If the joint probability mass function is equal to the product of the marginal probability mass functions. And similarly when we are considering the continuous case it should be the joined probability density function is equal to the product of the marginal probability density functions. For example, let us take f (x, y) is equal to say 4xy e to the power minus x square plus y square.

x is positive y is positive it is equal to 0 elsewhere. Let us look at say f (x, y) f x. So, here I will integrate with respect to y from 0 to infinity. Now if you look at the term y e to the power minus y square if will have integral e to the power minus y square by 2. So, at infinity it will become 0 at 0 it will become one 1. So, you will get twice 2x e to the power minus x square for x greater than 0. Similarly phi look at f y(y) then it will turn out to be 2y e to the power of minus Y square. So, note here that if i multiply f X (x) and f Y(y). I will get f (x, y) therefore, x and y are independent here. Now like in the case of univariate one can talk about joint moment generating function of x, y. As M x, y (s, t) that is equal to expectation of e to the power of s t s x plus t y. Notice here that if i put t is equal to 0, then I will get the moment generating function of s. And if i put to s is equal to 0, then I will get the moment generating function of Y.

(Refer Slide Time: 34:34)



We have powerful results connecting the independence to the moment generating function. We have the following theorem x X and y Y are independent. This implies and implied by the Joint mgf is equal to the product of marginal mgf's. We have another consequence also that if the random variables are independent. If x and y are independent then the moment generating function of the sum is equal to the product of the individual moment generating functions.

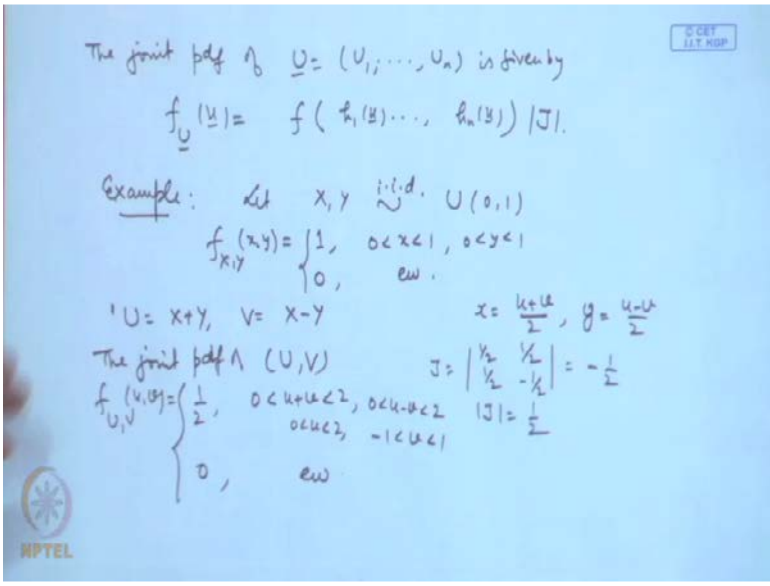
Infect these is one which is use to prove the additive properties of various distributions. Like I give example of the binomial distribution additive geometric distribution adds up to them. Negative binomial distribution, exponential distribution adds up to the gamma distribution, the linearity property of the normal distribution etcetera etc. All of those things where proved using the moment generating functions approach only. Now using the moment generating functions if the random variables are independent one can find out the distributions of the sum, but many times we may be interested in the distribution of functions of random variables.

Distributions of functions of random variables For example, I have (x, y) and I define u is equal to say g1 of x, y and v is equal to g2 of x, y. Say if g1, g2 is a measurable function then u, v is also a random vector. And 1 one can find out the distribution of u and v. I am not going to discuss in detailed the distribution here. Like in the case of 1 one variable when we had the x as a discrete random variable and g (x) also for a discrete there was direct way of obtaining the distribution of g; however, if the distributions where continuous in that case we had differential formula; that means, we are f (x) return at g inverse y and then we multiplied by dg inverse y by dy or you can say dx by dy when you have dealing with more than 1 one variable then the d term is represented or placed by a Jacobin term.

I will state in the form of the following theorem. So, Jacobin approach for finding density functions of joint random variables. Let x X is equal to x X1, x X2,…x Xn be a continuous random vector with joined joint probability density function say f X(x) here x is denoting the vector here xX1, xX 2,…xX n. And let us define ui I is equal to gi I of x bar, for I i is equal to 1 to n 1…n and this u is equal to uU1,U2,…uUn it say 1 to 1 mapping from rRn to r Rn.

And we also have the inverse transformation say X x i is equal to hi I of u bar, for I i is equal to 1 ton 1…n. Let these be the inverse Transformations and if the partial derivates Del xi I over Del u j exist and we define the Jacobin of the transformation as Del x1 by Del u1, Del x1 by Del u2 and Sso on, Del x1 by Del Un and Sso on. Del x n by Del u1 and so on, Del xn by Un.

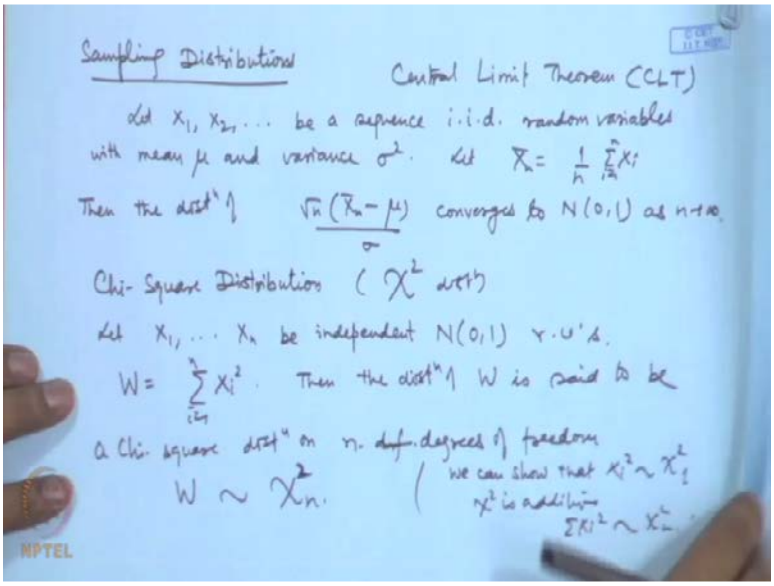
(Refer Slide Time: 39:31)



And we assume that this is non-zero over the range of the transformation. In that case the joint pdf of u U is equal to u U1, u U2,… uU n is given by f (u)that is equal to f of now xi's I are replaced by H h1 of u and Sso on. hn of u multiplied by the Jacobin the absolute value of the Jacobin. For example, let us consider say let x and y be independent uniform (0, 1) random variables; that means, the joined distribution of this is given by 1. 0 less than x less than 1, zero 0 less than Y less than 1 and 0 elsewhere. Let us define say uU is equal to xX plus yY and vV is equal to xX minus Yy then if you see this you can find out the inverse transformation.

X x is equal to u plus v by 2, y is equal to u minus v by two2. So, if I look at the Jacobin the derivatives of Del x by Del u is the half, dDel x by dDel v is half, dDel y by dDel u is half, dDel v by dDel y by dDel v is minus half. So, this is equal to minus half. So, modulus of Jacobin is equal to half. So, the join distribution of uv then is equal to half, 0 less than u plus v less than 2. 0 less than u minus v is less than 2 and of course, if you write down the absolute. Regions of uU and Vv are from 0 to 2 and v is from minus 1 to 1 and 0 elsewhere. So, this is the joint probability density function of UV.

(Refer Slide Time: 41:34)

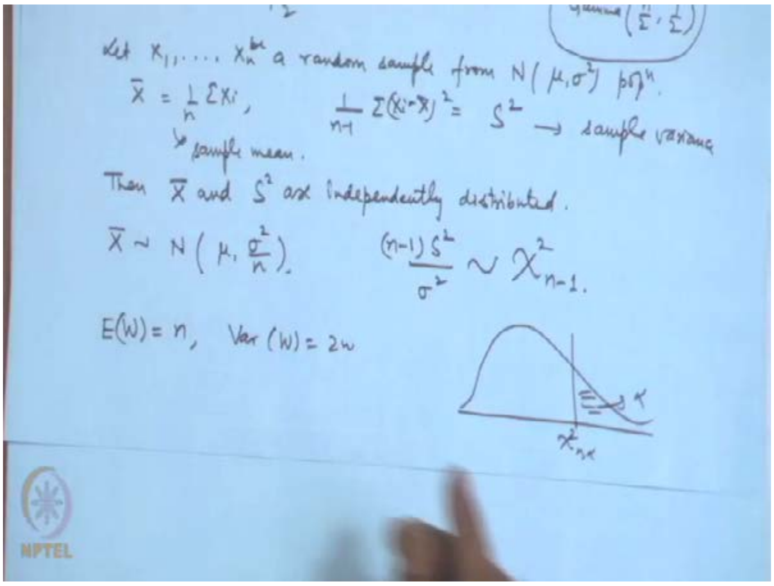


I introduce the concept of sampling distributions. Now so firstly, let us consider say x1, x2 and Sso on be a sequence of independent and identically distributed random variables with mean mew and variance sigma square. Let us use a notation x Xn bar is equal to 1 by n sigma x I Xi since. There is a sequence if I i consider the first n of these observations i take the mean as Xn bar. Then the distributions of root n (Xn bar minus mew) by sigma converges to normal (0, 1) as an n tense to infinity. Now this is the very powerful statement.

I am considering a side to be sequence of any independent and identically distributed random variables, but mean is mew and variance sigma x square is given to us. Then the distribution of the sample means that is xXn bar is approximately normal. So, this famous result is known as a Central limit theorem. This result as further generalizations for example, we may have independent, but none identically distributed random variables are or we may not even have independent random variables. So, then under certain conditions the distribution of the sample mean or the sample some still can be approximated by a standard normal distribution.

So, this is 1 one of the first results in the case of sampling distribution. What you mean by sampling distribution, if we are considering several observations for the same population or with the same distribution then if I i consider any characteristic of that. For example, mean variance etc. the distribution of that is known as the sampling distribution. So, we can say normal distribution itself is a sampling distribution we define some more sampling distributions 1 one is the well-known chi square distributions or chi square distributions. Let xX1, x X2…xXn be independent normal (0, 1) random variables. And let us define w is equal to sigma x I Xi square I i is equal to 1 to n 1…n. Then the distribution of w is said to be a chi square distribution on n degrees of freedom. n degrees of freedom that terminates used here and we user notation w follows chi square n. So, 1 one can actually derive this distribution because I i can derive the distribution of X1 one Square from normal (0, 1). xX2 is square x Xn is square and then we use an identity property to prove this we can show that each of the x I Xi square follows by the chi square 1. And chi square is additive therefore, sigma x I Xi square will follow chi square n.

(Refer Slide Time: 45:41)



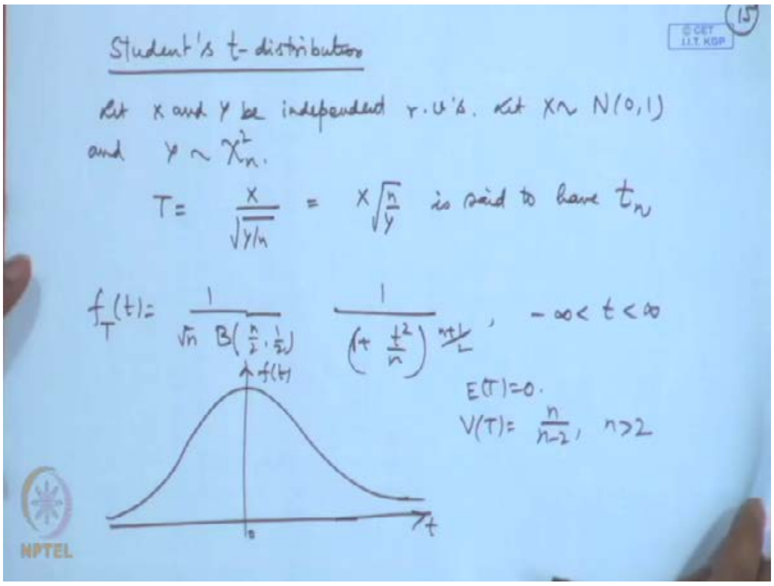
And actually the density is nothing but a gamma density. The distribution of w is 1 by 2 by to the power n by 2, gamma n by 2 e to the power minus w by 2, w to the power n by 2 minus 1. Which is actually a gamma distribution with parameter n by 2 and half, but this is specific form is known as a sampling distribution. Because it is a rising in a sampling from a normal distribution.

In fact we have a general result that if I i consider X1, X2…Xn a random sample from normal mew sigma square population. And let us consider xX bar as the mean and 1 by n minus 1 sigma x I Xi minus xX bar whole square as the sample variance this is called the sample variance. And this is known as the sample mean.

Then the result is that X bar and S square are independently distributed further the distribution of x X bar of codes we have seen earlier it will be normal mew sigma square by n and the distribution of n minus 1, s S square by sigma s S square that follows chi square on n minus 1 degrees of freedom. Therefore, when we have doing sampling from a normal distribution we can answer probability statements, regarding the sample mean or the sample variance we have some properties of W. For example, expectation of w W is equal to n variance of w W is equal to 2n etcetera etc. The since gamma distribution is positively skew distribution. Chi square distributions will also the positively skew distribution. The tables of chi square distribution are given if this probability is alpha this point chi square n alpha is tabulated in the tables of a chi square random variable. I will show you here.

The tables of a chi square distribution they are given in this particular fashion. So, here you see if this point is l chi square alpha on n degrees of freedom. Then the probability beyond this on the curve of probable, chi square probability density functions. This probability will be equal to alpha. So, for different values of alpha like .995, .999, .95 point zero .05, point zero.005 etcetera zero .25 etcetera etc. And for different values of n the point chi square n alpha are tabulated here we talk about some further sampling.

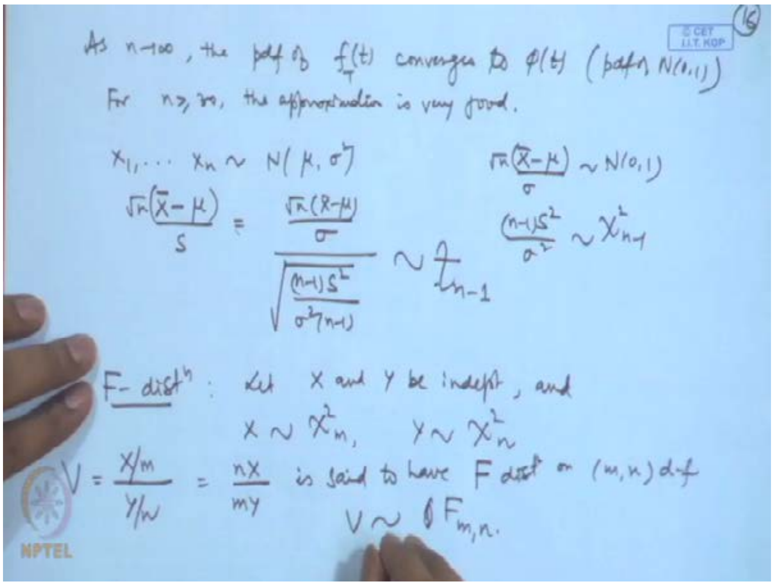
(Refer Slide Time: 49:08)



Distribution another 1 one is called the famous student's t-distribution. Let X and Y by are independent random variables. Let X follow normal (0, 1) and Y follow chi square n and let us define t T is equal to X divided by root Y by n that we can write as X into root n by Y. Then this is set to have t distribution on n N degrees of freedom. The probability density function of t is obtained as 1 by root n beat beta n by 2 half. 1 by 1 plus t square by n to the power n plus 1 by 2.

Where t lies between minus infinity to infinity as you can see this is a symmetric distribution around t is equal to zero 0. The mean of this will be zero 0 and the variance of this is n by n minus 2 which are valid for n greater than 2. As n tense to infinity as n tense to infinity the probability density function of t converges to phi (t) that is the probability density function of normal zero (0,1) random variable therefore, for large values of n for n greater than or equal to thirty 30.

(Refer Slide Time: 51:04)

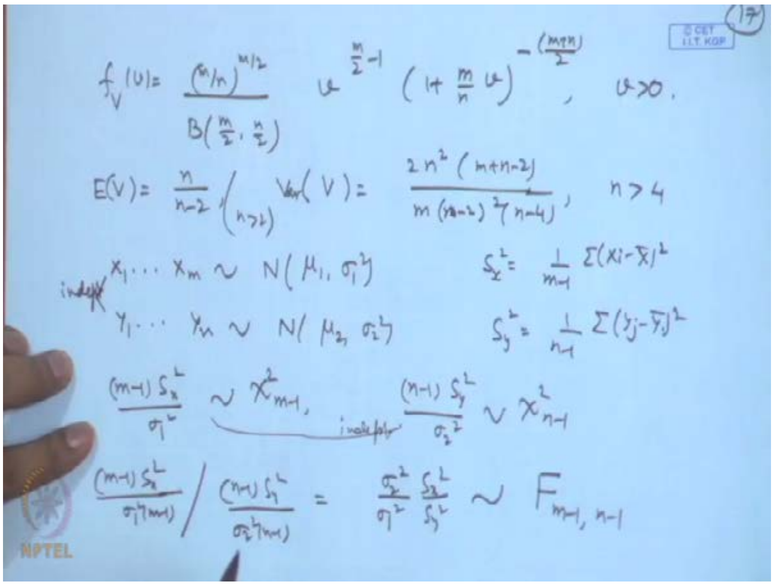


The approximation is very good. Once again on the curve of t distribution suppose I consider this probability to be equal to alpha then values of tn alpha that is the point beyond which the probability is alpha. Then these values are tabulated in the tables of t distribution in this particular fashion. So, this point is t n alpha the probability beyond this is alpha. So, for different values of alpha like .005, .01, .025, .05 etcetera etc. And for different values of the n degrees of freedom the point's t n alpha are tabulated here.

If I consider say x X1, x X2… xXn from normal mew sigma x X square and we define x X bar. So, x X bar minus mew by s S then this we can write as this square root and this. So, this we can write as because we have seen that x X bar minus mew by sigma into root n that will follow normal (0, 1) and we are having n minus 1 s S square by sigma Y square following chi square. On n minus 1 degrees of freedom. So, if I ride this as root n X bar minus mew by sigma divided by a square root of n minus 1 s S square by sigma s S square and 2 n minus 1 then this is same this will follow t distribution on n minus 1 degrees of freedom.

Therefore, this is also a sampling distribution then we also consider f distribution. Let X and Y are independent and X follows say chi square on m degrees of freedom and Y follows chi square on degrees of freedom. Then X by m divided by Y by n that is equal to n X by m Y this is set to follow f distribution on mn degrees of freedom. That is we write this let me use a notation say v follows f on m n degrees of freedom the probability density function of f can also be derived

(Refer Slide Time: 54:10)

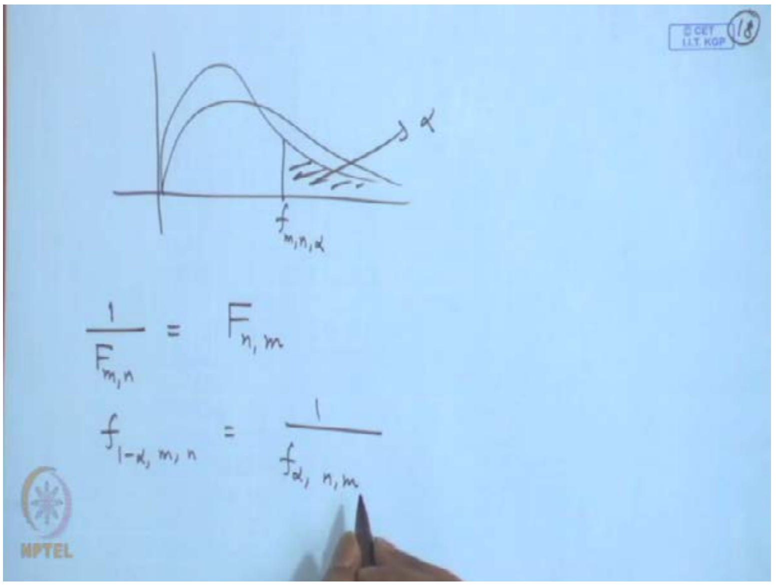


That is equal to m by n to the power m by 2 divided by beta m by 2, n by 2. v to the power m by 2 minus 1. One 1 plus m by n v to the power minus m plus n by Two 2. where we as positive the v of this is equal to n by n minus 2 the variance is equal to twice 2n square into m plus n minus 2 divided by m into n minus 2 a square into n minus 4. This of course, valid for n greater than 4 this is valid for and n greater than 2. We can actually see that it is a sampling distribution if we consider say a random sample X 1, X2… X m from say normal mew 1, sigma 1 X square and say Y1, Y2…Y n is another random sample.

From normal mew to sigma 2x Square and suppose these samples are considered independent. Let us define says X square as 1 by m minus 1 sigma Xi minus X bar square and Sy square is a 1 by n minus 1 sigma Yj minus Y bar whole square then m minus 1 Sx square by sigma 1 square this follows chi square distribution on m minus 1 degrees of freedom and n minus 1 s Y Sy square by sigma 2 square follows chi square on n minus 1.

Degrees of freedom. So, if I take the ratio if the samples are independent then these 2 variables are going to be independent these 2 are independent. So, if I consider the ratio m minus 1 s X Sx square by sigma 1 square into m minus 1 divided by n minus 1. s Y Sy square by sigma 2 square into n minus 1 that is equal to sigma 2 square by sigma 1 square s X by s Y square, that will follow f distribution on m minus 1 n minus 1 degrees of freedom. So, f distribution is a sampling distribution as you can see this is a positively is q.

(Refer Slide Time: 56:37)



Distribution of course, the lay will of scenes of will depend upon the parameters m and n here and if I consider F distribution and this probability is alpha then this point is called F m,n,alpha. That tables of F m,n,alpha are tabulated for selected values of alpha and selected values of m and n.

For example, you can see here this corresponding to alpha is equal to point 1. On this side m is varying on this side n is varying and for these values F m n values are f alpha m n values are tabulated these for alpha is equal to point 0.1 and similarly for alpha is equal to 0.5 also these tables are given. So, most of the statistical tables contains the tables of F variable. There is a relation of F in the terms of reciprocal because if I i take reciprocal of F variable then also we get an F distribution.

So, actually if I i say f F m n 1 by that. That then this is equal to F n m and therefore, if we consider the point f of 1 minus alpha m n that is equal to 1 by F alpha n m. So, these are some sampling distributions and they asked used for inference problems particularly to find out the confidence intervals and testing of hypothesis problems. In the next lecture I will be introducing the problem of inference in particular the problem of point estimation.