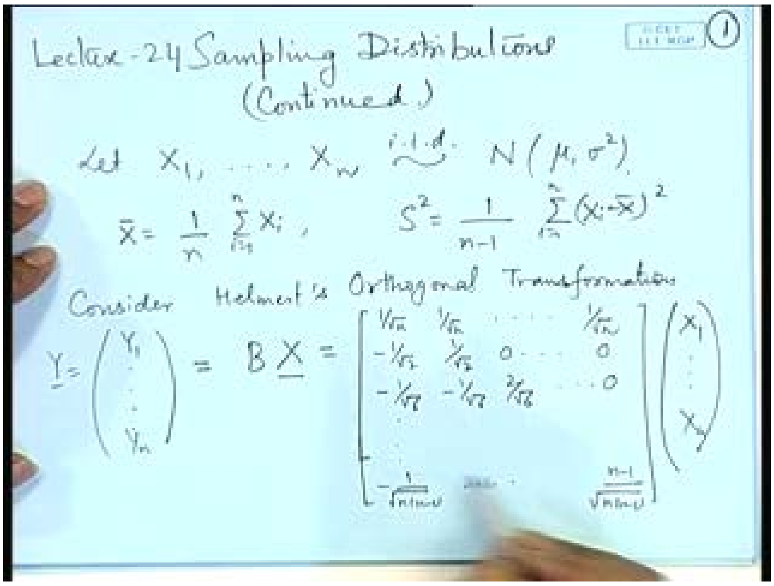
Probability and Statistics
Prof. Dr. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Module No. #01
Lecture No. #24
Sampling Distributions-II

We have introduced a sampling distribution called Chi square distribution and then I showed that if we are doing from sampling from a normal distribution then, the distribution of the sample variance is a Chi square distribution. Therefore, Chi square distribution is a sampling distribution. We used a moment generating function technique to derive the distribution of s square. Firstly, we are proving that sample mean and sample variance are independently distributed then we are sampling from a normal population. Today, I will give an alternative derivation by the method of transformations for the sampling distribution of x bar and s square then, we are sampling from a normal population.

(Refer Slide Time: 01:04)



Let us consider that X 1, X 2, X n, is independently and identically distributed normal mu sigma square random variables. So, we want to derive the distribution of X bar that is sample mean 1 by n sigma X i i is equal to 1 to n and S square that is, 1 by n minus 1 sigma X i minus X bar whole square.

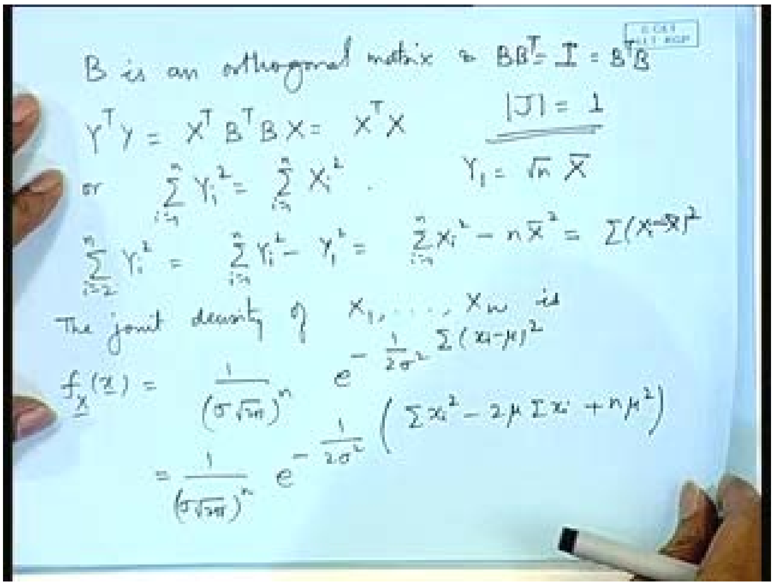
So, the method of proof will be through transformations. So, we will consider a transformation of the set X 1, X 2, X n, in the following fashion. So, let us consider Helmert's orthogonal transformation.

Now, this is a special transformation given in the following fashion that we define y is equal to Y 1, Y 2, Y n, as B X where B is the matrix of coefficients the first row is 1 by root n 1 by root n and. So, on 1 by root n the 2 row it is minus 1 by root 2 and 1 by root 2 and remaining terms are 0 the third one is minus 1 by root 6 minus 1 by root 6 and 2 by root 6 and the remaining terms are 0. Likewise, if we continue in the last row we have minus 1 by square root n into n minus 1 and so on.

Finally, the last term is n minus 1 by root n into n minus 1 multiplied by X 1, X 2, X n,. So, first of all let us observe this B matrix this B matrix is a special matrix which is called Helmert's orthogonal matrix the terms of the first row are same and if you take every next row then it is define in such a way that if I multiply any 2 rows then the product will be 0 that is a scalar product of any two rows is 0 for example, if I take first and second then here it is minus 1 by root 1 here it is plus 1 by root 2. If I multiply the sum will give me 0 suppose I take this with this then again the same thing because these two terms are same; if I multiply here and then this will become 0. Similarly, if I take this and multiply by the first row then minus 1 by root 6 n minus 1 by root 6 n plus 2 by root 6 n. So, again the sum is equal to 0. So, this is special matrix which is constructed for this purpose.

Now, let us see the effect of this. So, what we have done is that we have transform the X 1, X 2, X n, variables to new variables called Y 1, Y 2, Y n, by means of this.

(Refer Slide Time: 04:30)



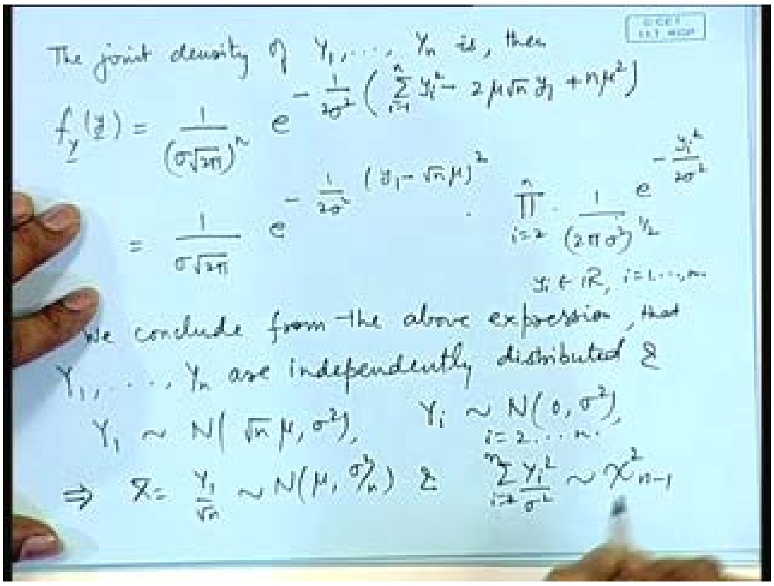
We have B is an orthogonal matrix and we have B B transpose is equal to identity matrix. So, if I consider Y transpose Y then that is equal to X transpose B transpose B X that is equal to X transpose. So, this is also B transpose B that is equal to X transpose X; that means, sigma of Y i square is equal to sigma of X i square i is equal to 1 to n; that means, the original sum of squares is equivalent to the new sum of squares. Also, we can see here by this transformation that Y 1 is equal to root n X bar because, if I consider the multiplication of the this matrix with this vector and look at the first term and the first term will be X 1 by root n plus X 2 by root n plus X n by root n. So, that is root n X bar. So, Y 1 is root n x bar. So, from here we get sigma Y i square i is equal to 2 to n is equal to sigma Y i square minus Y 1 square which is same as now sigma X i square this is i is equal to 1 to n sigma i is equal to 1 to n minus n X bar square that is equal to sigma X i minus X bar whole square which is the term which appears in the S square term that is the sample variance.

Therefore, this new transformation is giving me Y 1 as well as sigma that is which is x bar which is a term of X bar and which is another term which is a term of S square. So, our desired objective was to get the distributions of X bar and S square and this particular transformation helps us in at least representing these two terms in terms of transformed variables.

Now, let us look at the distribution. Firstly, we write down the joint density function of X 1, X 2, X n,. So, each of this X i's are normal new sigma square variables. So, that distribution we write as one by sigma root 2 pi to the power n e to the power minus 1 by 2 sigma square sigma X i minus mu whole square now this we expand we can write as 1 by sigma root 2 pi to the power n e to the power minus 1 by 2 sigma square. And, now we get sigma of X i square minus 2 mu sigma X I plus n mu square.

When we consider the transformation Y is equal to B X and B is an orthogonal matrix then we know that the determinant of an orthogonal matrix is either plus 1 or minus 1. So, jacobian of the transformation that will be having absolute value one. So, this will be used for calculation of the transformed density.

(Refer Slide Time: 08:07)



If I consider the joint density of Y 1, Y 2, Y n, then it is obtained as, in the density of X 1, X 2, X n, let us substitute the transformed values in terms of Yi's. So, sigma X I square will become sigma Y i square and this sigma X i is nothing, but n X bar now X bar is Y 1 by root n. So, this also you can substitute and multiply by the Jacobean of the transformation that is unity. So, we get the transform density as 1 by sigma root 2 pi to the power n e to the power minus 1 by 2 sigma square sigma Y 1 square minus 2 mu root n Y 1 plus n mu square this is i is equal to 1 to n.

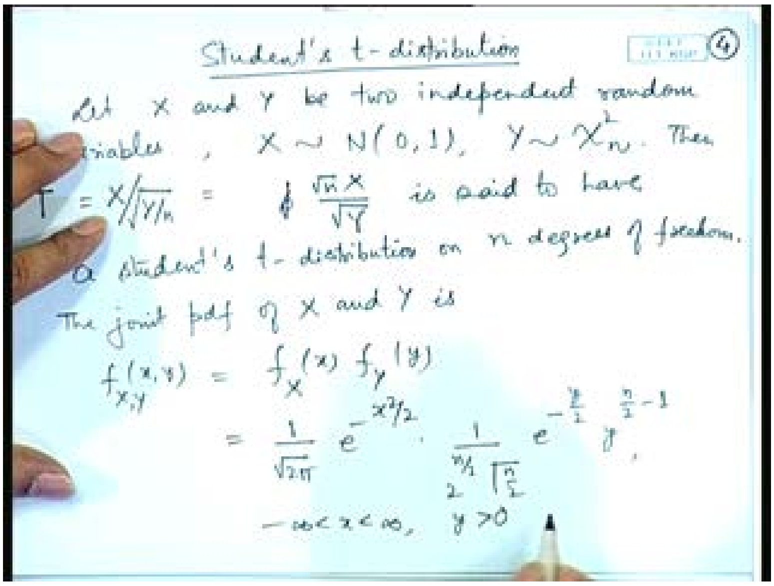
Now, in this particular exponent if I consider this sigma Y i square I take I write this as Y 1 square plus sigma Y i square from 2 to n now this Y 1 square minus 2 mu root n Y 1 plus n mu square becomes a perfect square. So, we can present it as 1 by sigma root 2 pi e to the power minus 1 by 2 sigma square Y 1 minus root n mu square and the other terms we write separately as product i is equal to 2 to n 1 by 2 pi sigma square to the power half e to the power minus Y i square by 2 sigma square.

The range of the transformed variables each of this Y i's are also from minus infinity to plus infinity now you see this representation we are able to express the joint density of Y 1, Y 2, Y n, as product of the certain functions where each function is strictly dependent only on each Y i. So, these are n functions. So, this is a function which is dependent upon Y 1 alone and here we have n minus 1 function, each function is dependent upon Y 2 Y 3 Y n respectively.

So, if we integrate with respect to Y i's we will get individual terms; that means, Y 1, Y 2, Y n, are independent and you are also able to say that these Y i's are independently normally distributed because of the form of the density. So, we conclude that we conclude from the above expression that Y 1, Y 2, Y n, are independently distributed and Y 1 follows normal root n mu and sigma square and remaining Y i's follow normal 0 sigma square for i is equal to 2 to n.

So, this implies that since Y 1 is root n X bar that is X bar that is equal to Y 1 by root n that will follow normal mu and sigma square by n and the sum of the squares of this Y i is divided by sigma square that is sigma Y i square by sigma square i is equal to 2 to n that will follow Chi square on n minus 1 degrees of freedom. But this term is nothing, but n minus 1 S square by sigma square. So, that is following Chi square on n minus 1 degree of freedom and further these two are independent. So, the result which we had proved using moment generating function we have proved using transformations of the variables also this Helmert's orthogonal transformation is quite useful and that actually suggest that a procedure for obtaining the distributions of the sums of random variables and the squares of random variables. So, many times it is taken as quite useful.

(Refer Slide Time: 13:19)



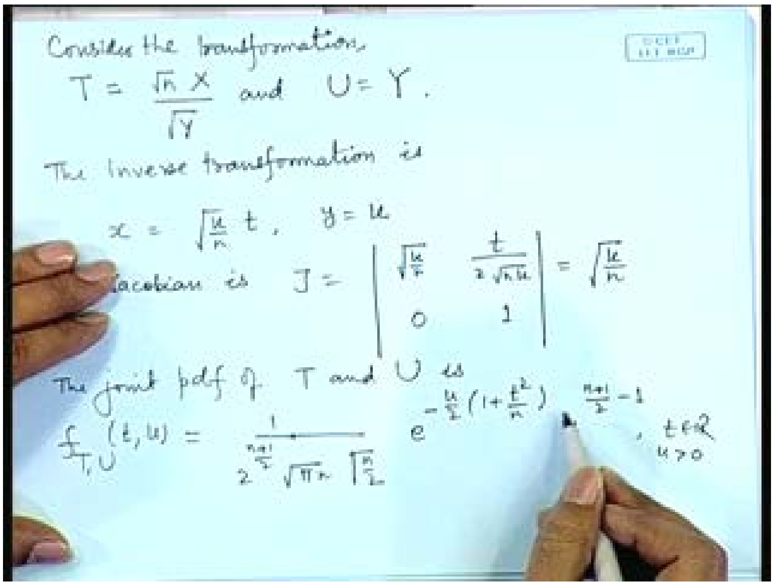
Now, we move over to another sampling distribution called t-distribution. So, we call it students t-distribution there is a story behind that why it is called students t-distribution basically it was discovered by W H Gosset a statistician in England; however, he worked in a brewery and therefore, it was not permitted for him to give his affiliation as working in a brewery. So, he used a (( )) name student and therefore, the distribution became famous as students t-distribution.

So, if I have let X and Y be two independent random variables let us assume that X follows normal 0 1 distribution and Y follows a Chi square distribution on n degrees of freedom if I consider the ratio X divided by Y by n and square root. So, let me write it as square root n X by square root of Y let me call it say T then this T is said to have a student's t-distribution on n degrees of freedom now this degrees of freedom terminology is coming from the Chi square distribution where we express that what is a meaning of the terms degrees of freedom a Chi square distribution on n degrees of freedom was represented as the sum of squares of n independent standard normal random variables.

So, in the definition of t-distribution I am using the degrees of freedom of Chi square and therefore, this t-distribution is said to have said to be on n degrees of freedom now since here X and Y are independently distributed random variables the variation of the density of T is an exercise of deriving distribution of a function of random variables. So, we can write down the joint distribution of X and Y and create a transformation in which one of the variable will be define by T and some other variable and we derive the distribution. So, let us do it in the following way.

Firstly, we look at the joint probability density function of X and Y. So, it is equal to the product of the individual distributions of X and Y now X is normal 0 1. So, the density is 1 by root 2 pi e to the power minus x square by 2 and the density of Y is Chi square n that is 1 by 2 to the power n by 2 gamma n by 2 e to the power minus Y by 2 Y to the power n by 2 minus 1 here the range of the x variable is from minus infinity to infinity and range of Y variable is positive.

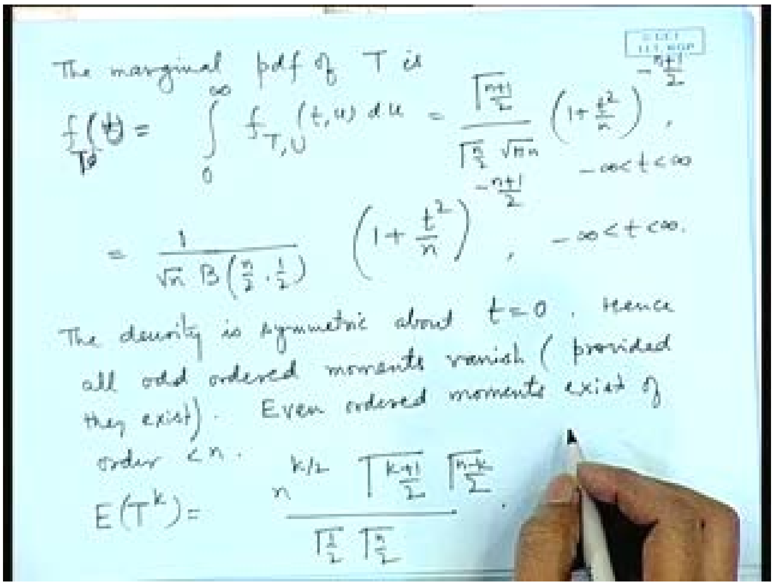
(Refer Slide Time: 17:00)



So, now, we are considering the transformation t is equal to root n X by root Y. So, let us consider this transformation ok. So, the second variable we can consider as say V is equal to or U is equal to say y because we have to consider a 1 to 1 transformation or at least a number of variables should be same. So, that we can find the joint density and then we can integrate out the not desire variable. So, the inverse transformation here will be. So, X is equal to root u by n t and Y is equal to U. So, we consider the jacobian del x by del t that is root u by n del x by del u that is t by 2 root n u del y by del t that is 0 and del y by del u that is 1. So, it is equal to root u by n.

So, if we substitute this in the joint density of X Y and multiply by Jacobean we get the joint probability density function of t and u f t u. So, let us substitute the values here 1 by root 2 pi and all this thing is constant. So, we combine it together it becomes 1 by 2 to the power n plus 1 by 2 root pi n gamma n by 2 e to the power minus u by 2 1 plus t square by n u to the power n plus 1 by 2 minus 1. So, this is after combining the coefficients and another thing you observe here that if you are making this particular transformation the range of t remains from minus infinity to infinity and u is the Chi square variable. So, u is positive. So, t belongs to r and u is positive to get the density of t we integrate this joint density with respect to u. So, if you integrate with respect to u from 0 to infinity you observe this term it is e to the power minus u into something and then u to the power some power which is of the nature of a gamma integral or a gamma function.

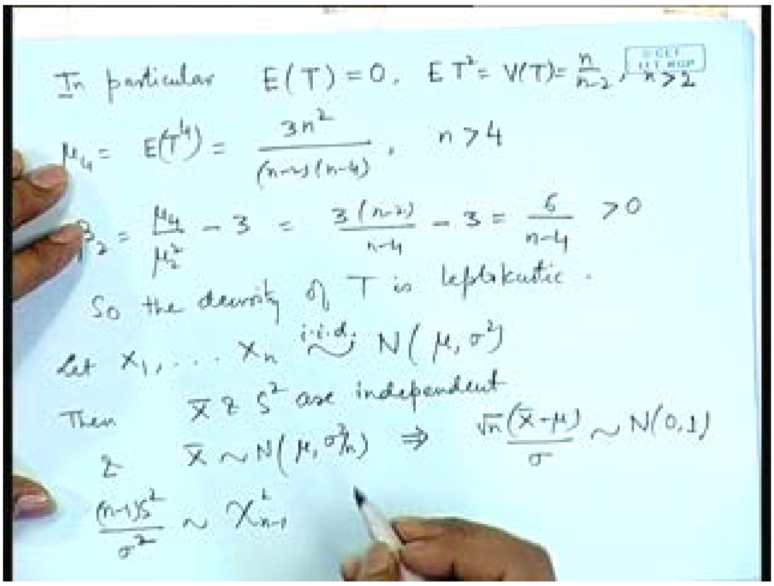
(Refer Slide Time: 20:19)



So, it can be easily evaluated and we get the marginal probability density function of t is let me call it f U that is integral f t u d u from this is f T. So, I have written wrong this is f T from 0 to infinity. So, here the order of the gamma function is n plus 1 by 2. So, in the numerator I will get gamma n plus 1 by 2 divided by u as a multiply half 1 plus t square by n. So, in the denominator I will get half 1 plus t square by n to the power n plus 1 by 2 now there is two to the power n plus 1 by 2 term that will cancel out. So, we are left with these density as gamma n plus 1 by 2 divided by gamma n by 2 root pi n 1 plus t square by n to the power minus n plus 1 by 2 and the range of the variable is from minus infinity to infinity. So, this is the density of the t-distribution on n degrees of freedom this particular coefficient we can write in slightly different way also because root pi if we observe it is gamma half. So, we can utilize the beta function notation and it becomes one by root n beta n by 2 1 by 2 and 1 plus t square by n to the power minus n plus 1 by 2.

Obviously, if you look at this one the density is a symmetric function in t around 0 because if you replace t by minus t you get the same function another thing you observe that as t becomes large this will go towards 0 because the 1 plus t square by n term is in the denominator also you observe that higher the power of n higher the value of n the convergence to 0 will be faster. So, basically that determines the shape of the t-distribution. So, we look at this thing the density is symmetric about t is equal to 0 hence all odd ordered moments vanish provided they exist even ordered moment can be calculated. Now, if you evaluate t to the power 2 k integral of this term then you can reduce it to a gamma function and do the calculation. So, the even ordered moments exist of order less than n. So, we have expectation of t to the power k for the even ordered moment as n to the power k by 2 gamma k plus 1 by 2 gamma n minus k by 2 divided by gamma half gamma n by 2.

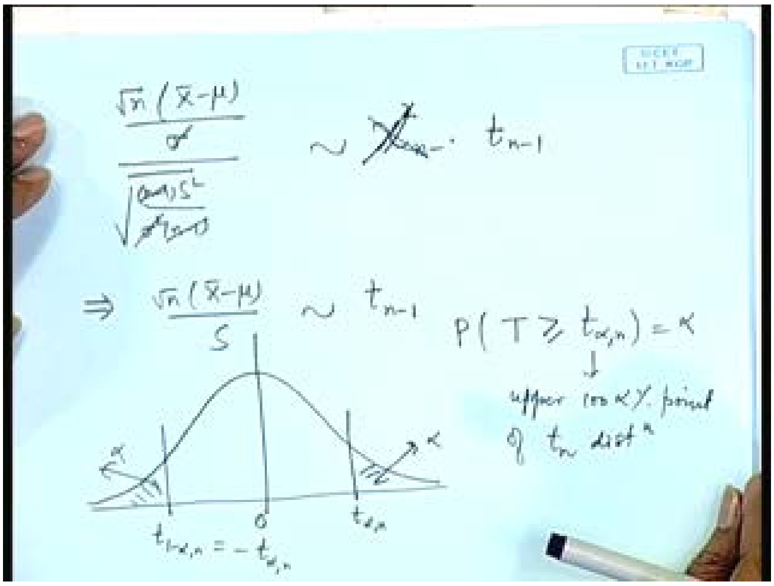
(Refer Slide Time: 24:16)



So, in particular expectation of T is 0; expectation of T square is that is, variance of T that is n by n minus 2 which is existing for n greater than 2 you observe this is somewhat peculiar number n by n minus 2 as n becomes large this becomes close to 1. Mu 4 is expectation of T to the power 4 that is three n square by n minus 2 into n minus 4 which is valued for n greater than 4. So, we can calculated the measure of courtesies that is beta 2 that is mu 4 by mu 2 square minus 3. So, you look at this term we divide by n square. So, this cancels out and we get 3 into n minus 2 by n minus 4 minus 3 which is simply 6 by n minus 4. Obviously, this is positive because we are considering n to be greater than 4 but, you can observe here that if n becomes sufficiently large then, this number becomes small and that means, the courtesies moves towards normality as n becomes large in general it is leptokurtic; density of t is leptokurtic.

Now, let us consider this distribution as a sampling distribution because, right now I have given a distributional theoretic representation of this t variable because we are writing it as only ratio of two variables in a particular form. But, we can make use of the fact that Chi square itself is a sampling distribution. So, whether we can represent t also in the same fashion, let us see. If I consider a random sample from normal mu sigma square distribution then, X bar and S square are independent that we proved and what are the distributions X bar follows normal mu sigma square by n. This means that, I can consider root n X bar minus mu by sigma and this will follow normal 0 1 n minus 1 S square by sigma square follows Chi square on n minus 1 degrees of freedom and this variable and this variable is independent.

(Refer Slide Time: 27:29)



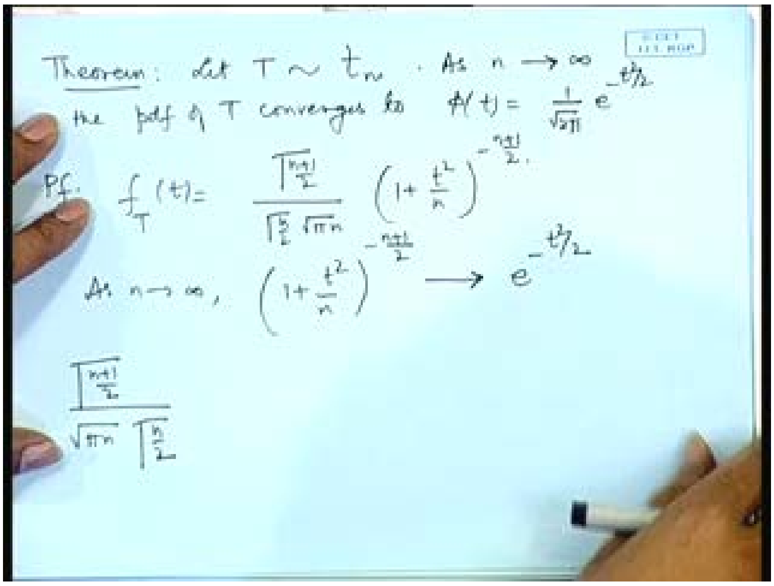
So, if I make use of this definition of the two variables of the t variable then I can write root n X bar minus mu by sigma divided by root n minus 1 S square by sigma square into n minus 1 this must follow Chi square distribution on n minus 1 sorry, this must follow t-distribution on n minus 1 degrees of freedom now if you simply this term here n minus 1 cancels out sigma cancels out. So, we are left with root n X bar minus mu by S this follows t-distribution on n minus 1 degrees of freedom.

Therefore, t-distribution is a sampling distribution another interesting thing is you can observe when I consider root n X bar minus mu by sigma this is standard normal and here root n x bar minus mu by S is there. That means, sigma is replaced by S later on we will see in the inference portion that S is actually an estimate for sigma. So, when sigma is not known we have to work with S and the distribution of that is known. In fact, in the context of this only this distribution was derived.

Now, regarding the probability points of t-distribution. So, the t-distribution is symmetric distribution about zero. So, if this point I call t alpha n then the probability beyond this must be alpha; that means, probability of T greater than or equal to t alpha n is equal to alpha; that means, t alpha n is upper 100 alpha percent point of t-distribution on n degrees of freedom because of the symmetry if you consider t 1 minus alpha n then that will be equal to t minus t alpha n; that means, if this value is alpha then this point is t 1 minus alpha n by this definition, but because of symmetry this will be equal to minus t alpha n.

Now, few things that we observed let us recollect that when I wrote the density I said it is symmetric about 0 the odd ordered moments vanish even ordered moments can be found mean is 0 the variance approaches 1 as n becomes large the peakedness approaches - normal peak as n becomes large. So, these things and another things I said that if you replace sigma by S then you have a t-distribution here you have a normal distribution these shows some sort of close similarity between t-distribution and t-distribution and a standard normal distribution. Actually it is true. In fact, we can prove that as n becomes large the t-distribution can be approximated by a standard normal distribution. So, we prove the following result.

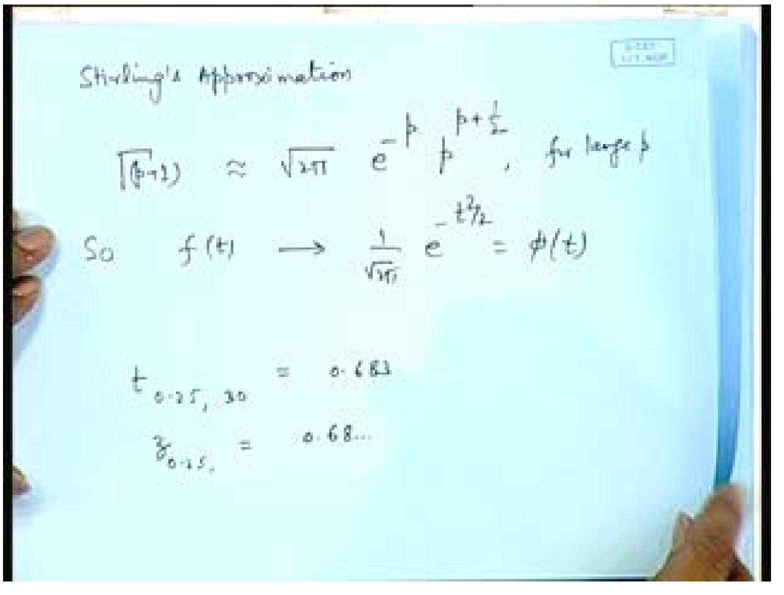
(Refer Slide Time: 30:53)



We consider let T be a t random variable on n degrees of freedom then as n becomes large the pdf of converges to phi t phi t is the probability density function of a standard normal random variable. So, to prove this let us write down the density function of t as derived that is gamma n plus 1 by 2 divided by gamma n by 2 root pi n and 1 plus t square by n to the power minus n plus 1 by 2.

Now, you observe this term as n becomes large or n tends to infinity this converges to e to the power minus T square by 2. So, as n tends to infinity 1 plus t square by n to the power minus n plus 1 by 2 converges to e to the power minus t square by 2. So, let us look at the remaining terms we must actually prove that this remaining term converges to 1 by root 2 pi. So, if we look at this term n plus one by two gamma root pi n gamma n by 2 now there is a formula called Stirling's approximation.

(Refer Slide Time: 32:47)



So, Stirling's approximation is let me write it here that gamma p plus 1 can be approximated by root 2 pi e to the power minus p p to the power p plus half for large p; that means, for large p gamma function can be approximated by an exponential and binomial type of (( )). These are mathematical formula we can use it here. So, if I am saying n is large than I can represent these things as root 2 pi e to the power minus n minus 1 by 2 n minus 1 by 2 to the power n plus 1 by 2 no n minus 1 by 2 by to the power n by 2 and in the denominator you have root pi n root 2 pi e to the power minus n minus 2 by 2 n minus 2 by 2 to the power n minus 1 by 2. So, we can do some simplification this root 2 pi etcetera will cancel out and here you have 1 by 2 to the power n by 2 and 1 by 2 to the power n minus 1 by 2. So, 1 root 2 will come here. So, this is giving rise to 1 by root 2 pi and then we get this e to the power half and here I can take common n in the numerator and denominator.

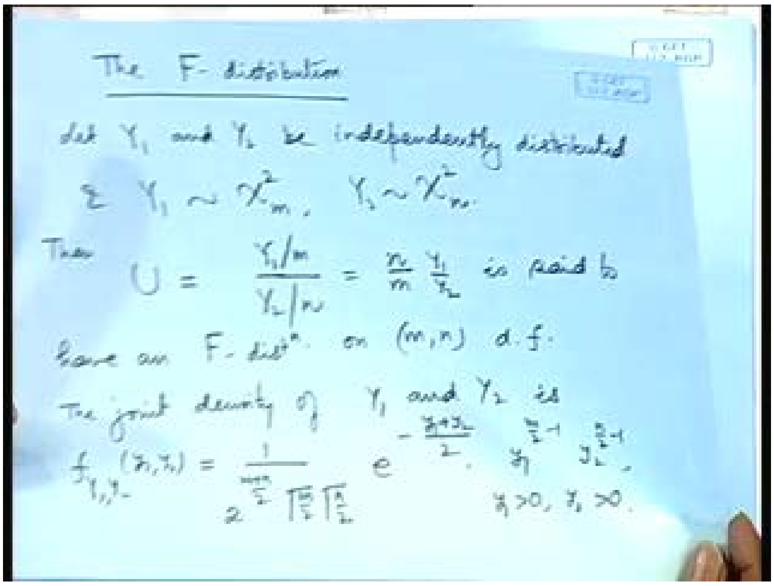
The terms are getting cancelled out that is n to the power n by 2 in the denominator I have n to the power half here and n to the power n minus 1 by 2. So, these all terms get cancelled out and you are left with 1 minus 1 by n to the power n by 2 divided by 1 minus 2 by n to the power n minus 1 by 2.

So, if I take the limit as n tends to infinity this goes to e to the power half and this goes to e. Therefore, the limit is simply 1 by root 2 pi because this e to the power half e to the power half and e they get cancelled out. So, this proves that this f T converges to f T converges to 1 by root 2 pi e to the power minus T square by 2; that is the density function of the standard normal variables.

The question arises that for what sufficiently large value of n is this approximation good the answer is that for n greater than or equal to 30 the approximation is extremely good and the tables of t-distribution most of the times they show. So, if we look at a standard table of t-distribution unfortunately, this cannot be seen here but, I will just write here. The t value say at point 25 and 30 is point 683. If I look at the same value for normal distribution then the point where the probability is point 25 above the given point, it is point 67 that is the if I consider that is z point z point 05 is equal to point 68 and since this tables is given only up to two places. I cannot predict here, but it is pretty close as you can see from here.

In fact, the tables are not given beyond thirty in most of the cases because the approximation is extremely good. In fact, at one twenty the value is almost equal to the next sampling distribution which is important and it is use quite frequently is f-distribution.

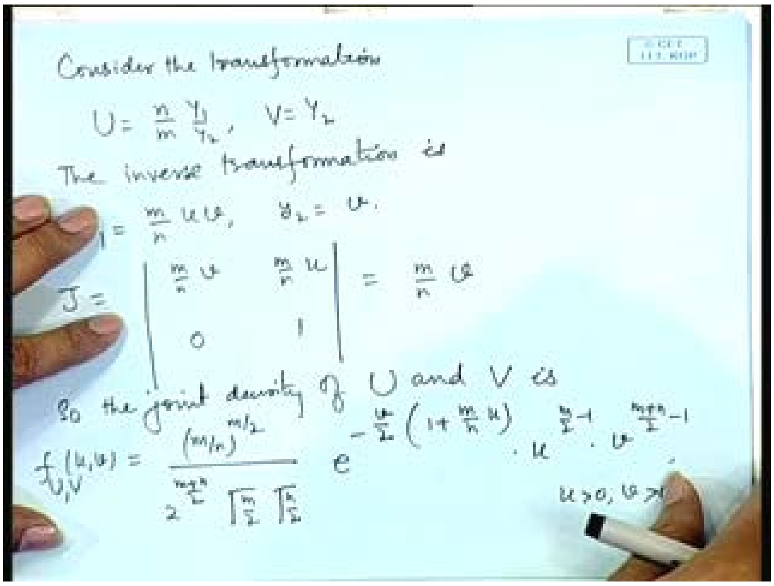
(Refer Slide Time: 37:43)



So, now, I introduce the f-distribution here let Y 1 and Y 2 be independently distributed random variables and Y 1 follows say Chi square distribution on m degrees of freedom and Y 2 follows say a Chi square distribution on n degrees of freedom then if we define a variables called U as the ratio of this Y 1 and Y 2, but divided by their degrees of freedom that is Y 1 by m divided by Y 2 by n that is basically becoming n by m Y 1 by Y 2 then this is said to have an f-distribution on m n degrees of freedom.

Now, here one has to notice that two degrees of freedom terms are coming and therefore, this order is important. If I am having a numerator Chi square variable as m and the denominator Chi square as n then we will write the ordered pair m n. That means, if I write n m it will denote a different f-distribution now by our theory of transformation of random variables u is a function of Y 1 Y 2. Therefore, I can use a new dummy variable V and find out the joint density of U and V to derive the probability density function of U. So, for that propose we write the joint distribution of Y 1 and Y 2 the joint density of Y 1 and Y 2 is f of Y 1 Y 2. Basically, we multiply the individual densities of Y 1 and Y 2 which are basically Chi square densities on m and n degrees of freedom. So, if we combine the coefficients 1 by 2 to the power m plus n by 2 gamma m by two gamma n by 2 e to the power minus Y 1 plus Y 2 by 2 Y 1 to the power m by 2 minus 1 Y 2 to the power n by 2 minus one where both Y 1 and Y 2 are positive.

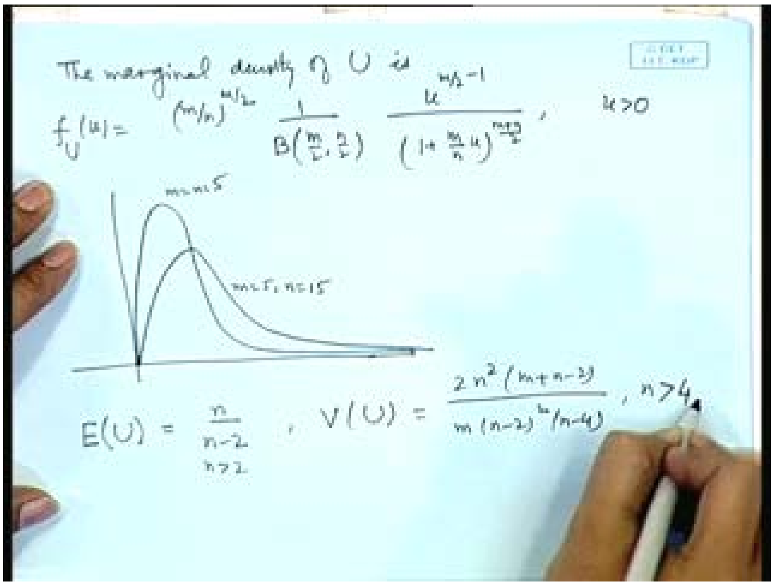
(Refer Slide Time: 40:37)



So, we consider the transformation in which U is this variable. Consider the transformation u is equal to n by m Y 1 by Y 2 and V is equal to say Y 2. So, the inverse transformation is y one is equal to m by n u v Y 2 is equal to V. So, the Jacobean of the transformation can be calculated as m by n v m by n v u 0 1 which is basically m by n v and since the terms are positive. So, absolute value of j will also be the same. So, the joint density of U and V is now you can observe we are having here this constant term and we will be replacing Y 1 by m by n u v and Y 2 by V. Here, V by 2 term will come out and this will give additional powers of V for power of U will be this one only.

So, after adjustment of the terms we can write it as m by n to the power m by 2 divided by 2 to the power m plus n by 2 gamma m by 2 gamma n by 2 e to the power minus v by 2 1 plus m by n u u to the power m by 2 minus 1 v to the power m plus n by 2 minus 1 where u and v are positive variables. So, we can integrate with respect to v from 0 to infinity to get the desire density of f random variable. Again, if we observe the integral of this with respect to u is nothing, but a gamma function where the order is m by 2 and the coefficient is one plus half into one plus m by n u.

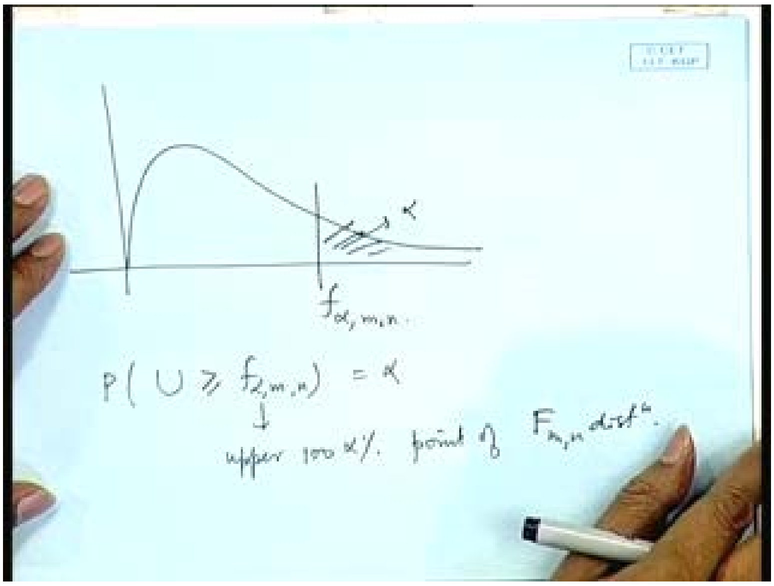
(Refer Slide Time: 43:07)



So, using the standard argument the marginal density will turn out to be the marginal density of U is then f U is equal to m by n to the power m by 2 and after the cancellation of the terms we will get it as 1 by beta m by 2 n by 2 u to the power m by 2 minus 1 divided by 1 plus m by n u m plus n by 2, for u positive. Obviously, this is a distribution of a positive valued random variable and it is positively skewed; however, the shape will vary depending upon the values of m and n.

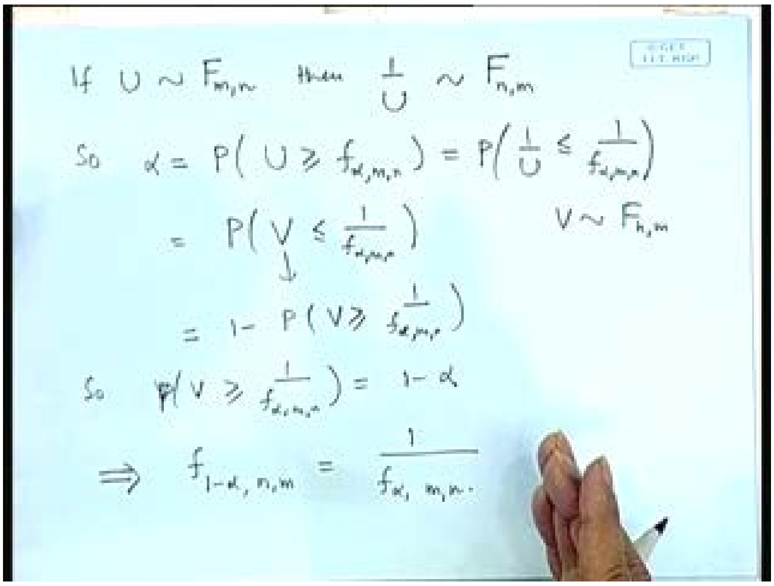
I can just give one example here; if I consider say m and n is equal to 5 then form of the density is somewhat like this if I consider say m is equal to 5 and n is say 15 then the form is something like this. So, likewise for different values of m and n you get different shapes of the curves. Here calculation of the moments will make use of different beta functions; however, I will write mean and variance the mean of this is n by n minus 2 you may be little bit surprised here that it is dependent only upon second variable because m is not appearing here variance of U is equal to here you need n greater than 2 and variance term is twice n square m plus n minus 2 divided by m into n minus 2 square into n minus 4 this is valid for n greater than 4 this positively skewed distribution.

(Refer Slide Time: 45:16)



We concentrate on the points here. So, if I have this probability equal to alpha then this point is termed as f alpha m n that is the probability of u greater than or equal to f alpha m n is equal to alpha that is this is upper hundred alpha percent point of F m n distributions.

(Refer Slide Time: 45:56)

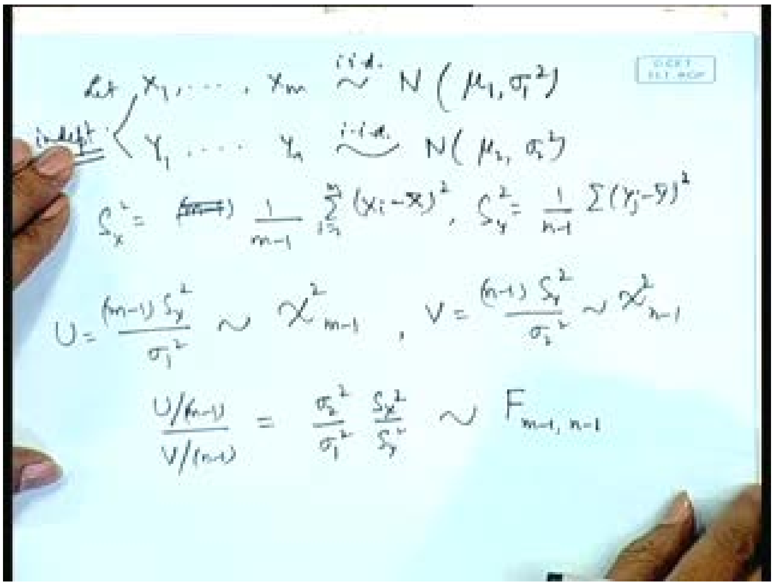


Now, by the definition of this F variable it is clear that if I have U following f m n then one by u will follow F n m because 1 by U remains ratio of Chi square variables divided by their degrees of freedom. However, the numerator degrees of freedom have gone to the denominator and the denominator degrees of freedom have gone to the numerator. So, this becomes f-distribution on n m degrees of freedom.

So, we can derive a formula for the points of f-distribution. So, we define that f alpha m n is the upper hundred alpha percent point of the f-distribution on m n degrees of freedom. So, probability of U greater than or equal to this is equal to alpha. So, if I write it as probability of 1 by U less than or equal to 1 by f alpha m n then this 1 by U is a F n m variable that is this is probability of some V less than or equal to f alpha m n that is this V is F n m variable. So, if I am saying that V greater than 1 by f alpha m n this equality or inequality does not play any role here because of this continuous distributions. So, I am saying probability v greater than or equal than to 1 by f alpha m n is equal to 1 minus alpha, but V is F n m distribution this implies that F 1 minus alpha n m is equal to 1 by f alpha m n this relationship is used for calculasing calculation of the percentage points of the f-distribution and generally the tables are because here it is a two dimensional table you have m and n both varying and therefore, only for selected values of alpha the tables are given.

Now, if they are given for say alpha is equal to point 05 or alpha is equal to point 1 then 1 minus alpha becomes point 95 and point 9 respectively. So, those tables can be automatically derived from the tables of point 05 and point 01 value, etcetera. Now, we look at that how in the sampling it arises.

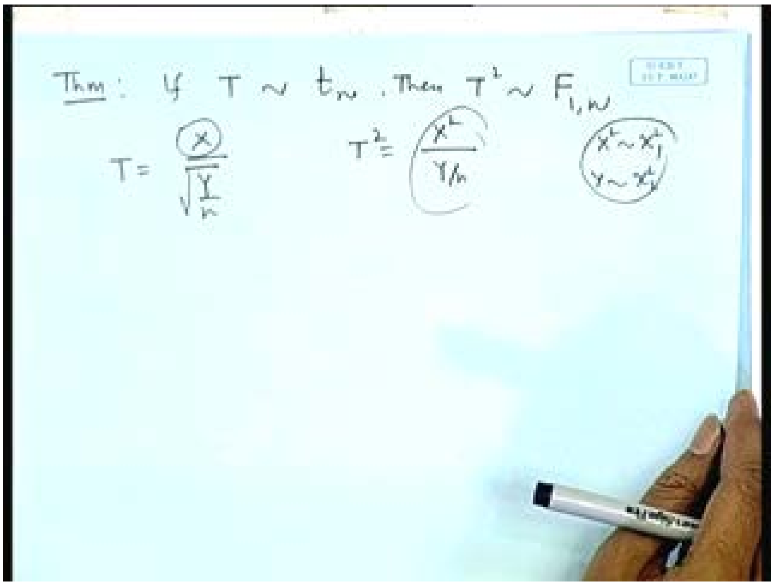
(Refer Slide Time: 48:36)



If we consider say a random sample say x one x two x m following normal mu 1 sigma 1 square and say Y 1, Y 2, Y n, be a random sample from normal say mu 2 sigma 2 square and also I assume that this samples are taken independently. Let us define the quantities say S X square as m minus 1 1 by m minus 1 sigma X i minus X bar square i is equal to 1 to m and say S Y square as 1 by n minus 1 sigma Y j minus Y bar square then by the theory of Chi square m minus 1 S X square by sigma 1 square follows Chi square distribution on m minus one degrees of freedom and n minus 1 S Y square by sigma 2 square follows Chi square distribution on n minus 1 degrees of freedom.

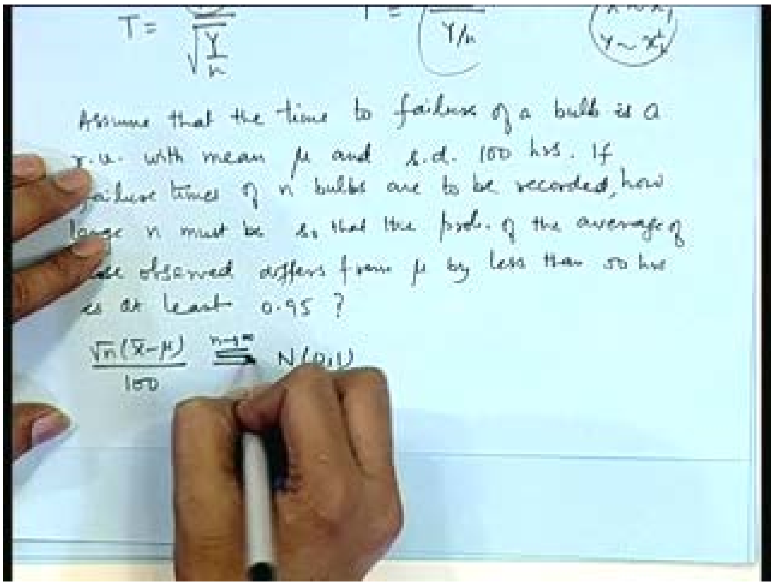
Let me call it say U variable and this as a V variable then if I take the ratio U divided by m minus 1 divided by V divided by n minus 1 then this is nothing, but sigma 2 square by sigma 1 square S X square by S Y square that follows f-distributions on m minus 1 n minus one degrees of freedom. So, these relationships are these result is used quite frequently in drawing inferences on ratios of the variances because ratios of the population variances and ratios of the sample variances is occurring here.

(Refer Slide Time: 50:39)



Another relationship which is coming here is that if say T follows t-distribution on n degrees of freedom then t square follows f-distribution on 1 and n degrees of freedom one can prove it by direct transformation by writing down the density of t and making the transformation u is equal to t square there and write and compare with the forms of the densities; however, one can look at an easy representation see we can write t as x divided by root y by n where X is a standard normal and Y is a Chi square variable. So, if I look at T square that is X square by Y by n now this x square will be Chi square on 1 degrees of freedom and Y be Chi square on n degrees of freedom. So, this is nothing, but the definition f variable on 1 and n degrees of freedom. Sampling distributions are extremely useful in statistical inference and they are used all the time I will give a couple of applications here.

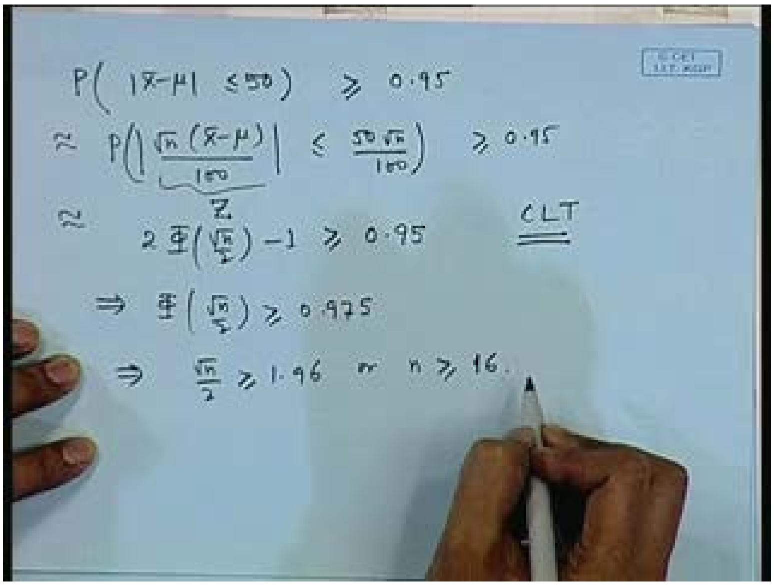
(Refer Slide Time: 52:02)



Assume that the time to failure of a bulb is a random variable with say, mean mu and standard deviation say hundred hours if failure times of n bulbs are to be recorded how large n must be. So, that the probability of the average of those observed differs from mu by less than 50 hours is at least point 95; that means, how much should be my sample size such that the sample average and the population mean should differ by less than 50 percent and this probability should be at least point 95.

So, we can make use of the central limit theorem here because the only information about the distribution that we are having is that it is a particular distribution with certain mean and certain variance. So, the conditions for application of the central limit theory are valid here. So, we will have root n X bar minus mu by sigma that is hundred this will be approximately normal 0 1 as n becomes large.

(Refer Slide Time: 54:14)

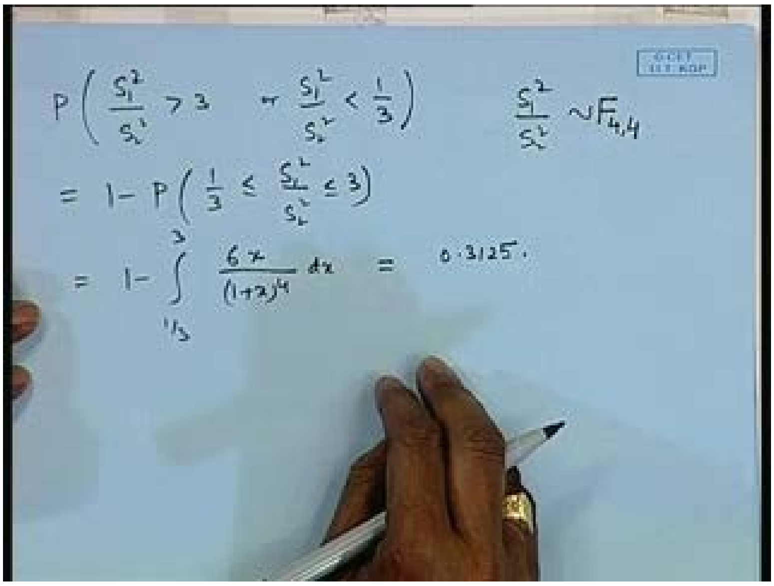


By this statement we have the condition here that probability of modulus X bar minus mu less than or equal to 50 we want to put this to be greater than or equal to point 95. So, we approximate this probability by converting to by making use of the central limit theorem. So, this is less than or equal to 50 root n by 100. So, this is approximately a standard normal random variable then this probability is I can replace this variable by z where z is a standard normal variable. So, this can be written in terms of cdf that is twice phi root n by 2 minus 1 greater than or equal to point 95.

I have made use of the central limit theorem here and this gives us phi of root n by 2 greater than or equal to point 975. So, from the tables of the normal distribution root n by 2 must be greater than or equal to 1 point 96 or n must be greater than or equal to 16. So, we need minimum sample size 16. So, that the sample average and the population average do not differ by more than 50 and the probability of that should be at least point 95 let me give one more problem here. So, we consider independent random samples of size 5 from two normal populations and they have the same variance.

So, what is the probability that the ratio of the larger to the smaller variance exceeds 3? This problem is important in the following sense see we have taken two sample from the same population now we want to check whether the there is too much variability in the sampling process. So, we look at the variances of the two samples and one of them will be naturally larger and one will be smaller. So, we are saying that the ratio of the larger to the smaller exceeds three what is the probability of this event.

(Refer Slide Time: 56:44)



So, if we write in terms of S 1 square and S 2 square basically we are requiring here what is the probability that S 1 square by S 2 square is either greater than 3 or S 1 square by S 2 square is less than 1 by 3 what is a probability of this? Now, if we have taken the samples of size 5 each then by the formula that n minus 1 by m minus 1 S 1 square by this one we get that S 1 square by S 2 square follows f-distribution on 4 and 4 degrees of freedom; that means, for calculation of this we have to look at the either the tables of F 44 or write down the density of F 44 here fortunately the density of F 44 become a quite simple form. So, we can write this as 1 minus probability one by 3 less than or equal to S 1 square by S 2 square less than or equal to 3 and this turns out to be 1 minus 1 by 3 to 3 and the density function of F 44 is 6 x divided by 1 plus X to the power 4. So, this integration can be done easily and the value turns out to be point 3125.

Various problems which relate to the sample means or the sample variances or the comparison of the means or comparison of the variances can be solved using sampling distributions in the portion of point distribution confidence interval estimation and testing of hypothesis we will have frequent uses of these sampling distributions. So, in particular we have consider normal distribution itself as a sampling distribution because, for the large samples any sample mean will be approximately normal - normally distributed under certain condition. Of course, and then when we are sampling from normal distributions then certain functions which are related to the means and the variances they are having Chi square t and f-distribution. So, these are in particular 4 important sampling distributions; there are many more sampling distributions, but they are not as frequently used in practice.