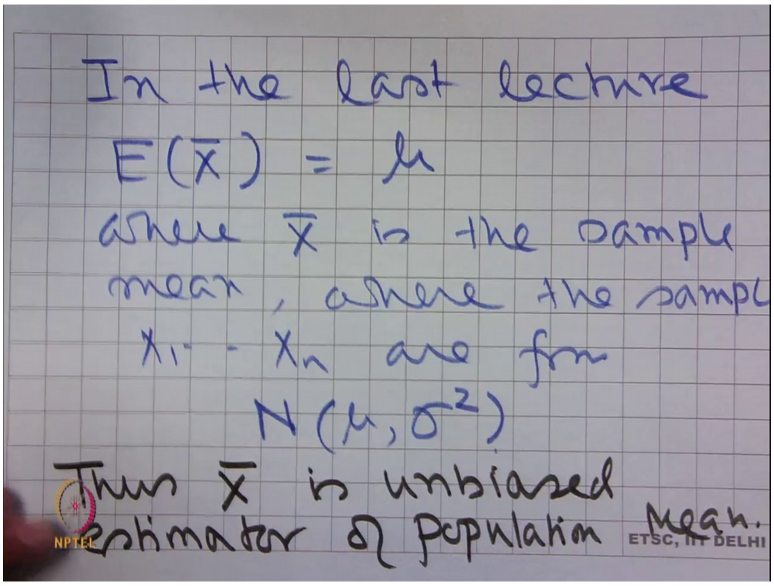
Statistical Inference

Prof. Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture – 08
Statistical Inference

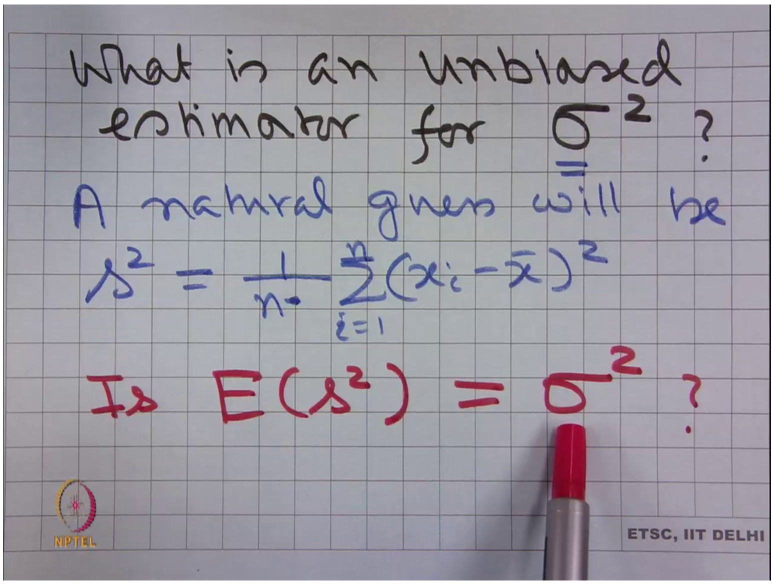
Welcome students to lecture number 8 on the MOOC's course on Statistical Inference.

(Refer Slide Time: 00:32)



In the last lecture, we found that expectation of X bar is equal to mu where x bar is the sample mean, where the sample X 1 up to X n are from normal mu sigma square. Thus, sample mean is unbiased estimator of population mean.

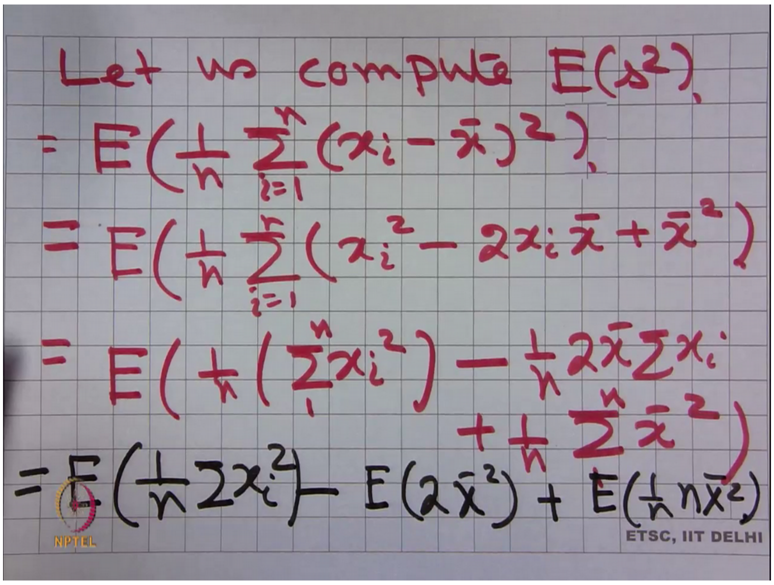
(Refer Slide Time: 01:52)



This we have seen; obviously, the next question is what is an unbiased estimator for sigma square?

That is, we have taken sample x 1 x 2 x n and based on that we want to find an unbiased estimate for the population variance sigma square. A natural guess will be sample variance. So, x 1 x 2 x n are my sample, x bar is the sample mean, therefore, sigma x i minus x bar whole square upon n that is the sample variance. The question is E is the expectation of small square is equal to sigma square or in other words whether the sample variance is an unbiased estimator for population variance.

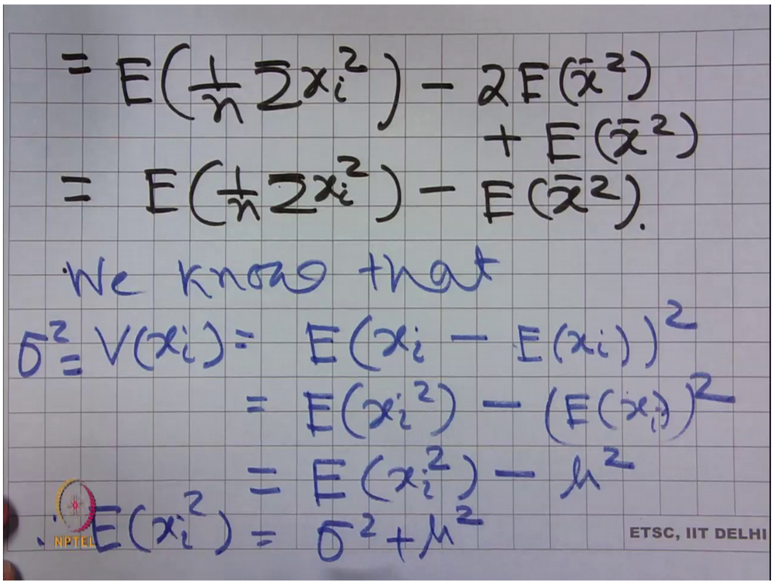
(Refer Slide Time: 03:45)



In order to find the answer, we have to compute the expected value of small square. The expected value of small square is equal to expected value of 1 by n sigma i is equal to 1 to n x i minus x bar whole square; is equal to expected value of 1 by n summation i is equal to 1 to n x i square minus 2 x i x bar plus x bar square, is equal to expected value of 1 by n into sigma x i square 1 to n minus 1 by n 2 x bar sigma x i plus 1 by n sigma x bar square 1 to n. This is equal to therefore, expected value of this whole thing.

So, this is equal to expected value of 1 by n sigma x i square minus expected value of 1 by n sigma x i is equal to x bar. Therefore, it is 2 x bar square plus 1 by n plus expected value of 1 by n into n x bar square.

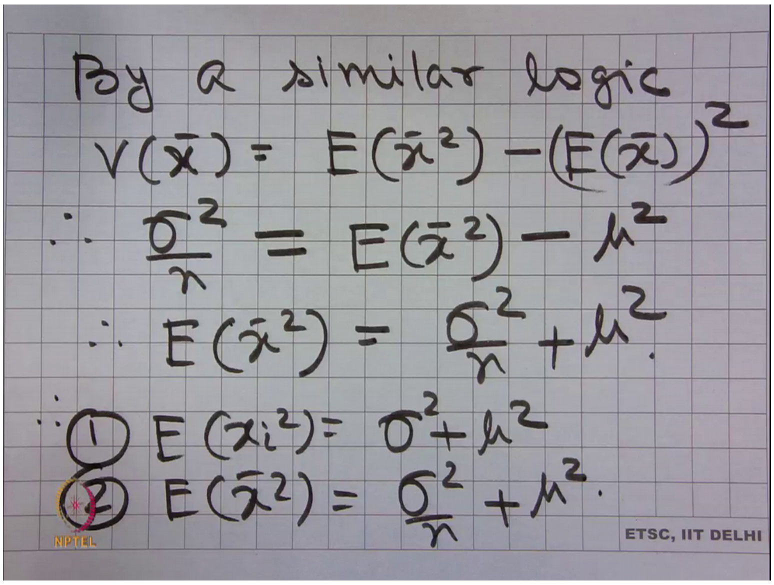
(Refer Slide Time: 06:13)



Is equal to expected value of 1 by n sigma x i square minus 2 times expected value of x bar square plus expected value of x bar square; is equal to expected value of 1 by n sigma x i square minus expected value of x bar square.

Now, we know that variance of x i is equal to expected value of x i minus expected value of x i whole square. So, variants of each x i is equal to expected value of x i square minus expected value of x i whole square, is equal to expected value of x i square minus mu squared. And variance of x i is equal to sigma square. Therefore, the expected value of x i square is equal to sigma square plus mu square.

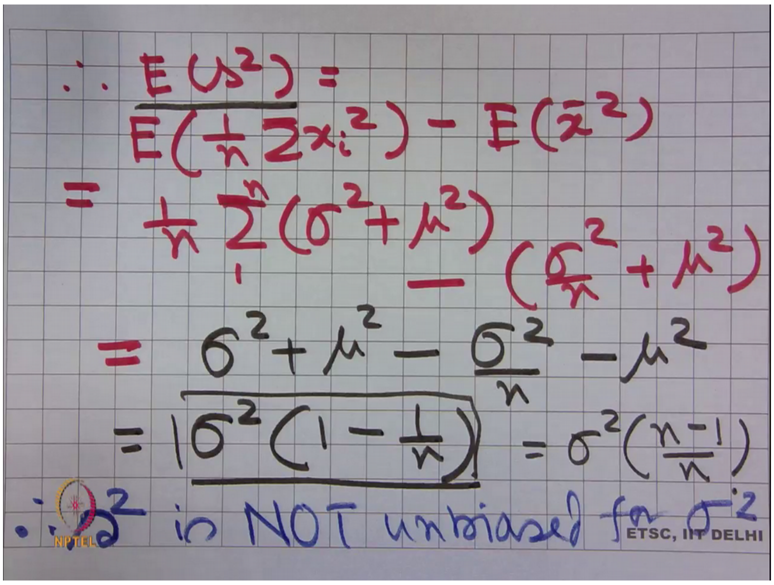
(Refer Slide Time: 08:24)



By a similar logic variance of x bar is equal to expected value of x bar square minus expected value of x bar whole square. And we know that variance of x bar is equal to sigma square by n. Therefore, sigma square by n is equal to expected value of x bar square minus mu square. Therefore, expected value of x bar square is equal to sigma square by n plus mu square. So, we find that one expected value of x i square is equal to sigma square plus mu square 2 the expected value of x bar square is equal to sigma square by n plus mu square.

Now, we have already found that expectation of small square is equal to this.

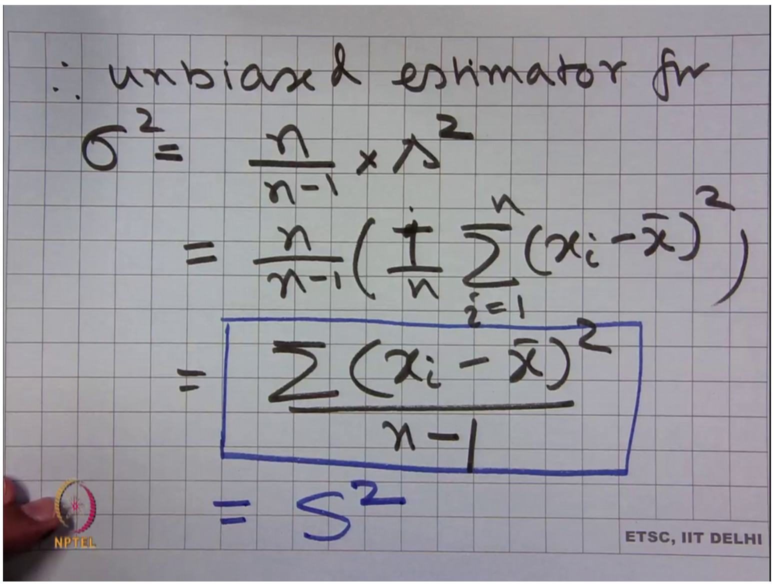
(Refer Slide Time: 10:17)



Therefore by putting these values 1 by n summation over 1 to n sigma square plus mu square minus sigma square by n plus mu square, is equal to sigma square plus mu square minus sigma square by n minus mu square, is equal to sigma square into 1 minus 1 upon n.

Thus we find that expectation of small square is equal to sigma square into 1 minus 1 upon n is equal to sigma square into n minus 1 upon n. Therefore, small square is not unbiased for sigma square.

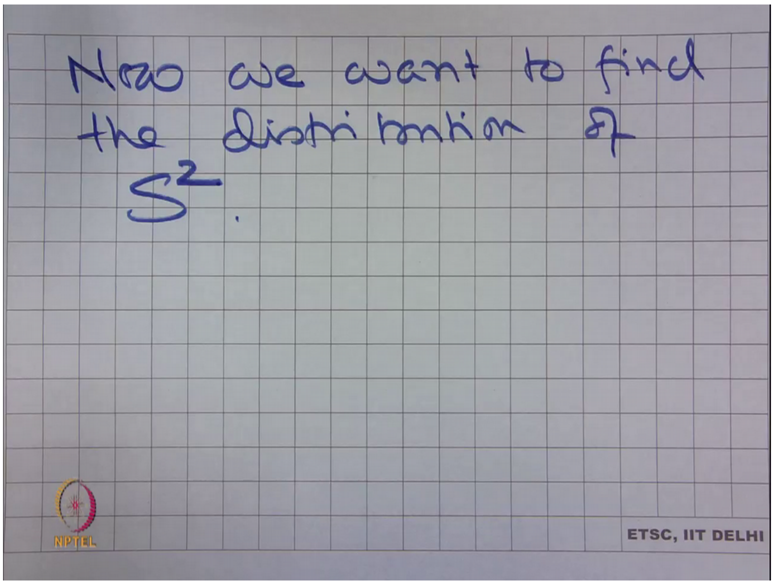
(Refer Slide Time: 12:12)



Therefore, if I ask you what is an unbiased estimator for sigma square, small square is equal to 1 by n into sigma x i minus x bar whole square, i is equal to 1 to n is equal to sigma x i minus x bar whole square divided by n minus 1.

So, that shows that our logical guess that sample variance is going to be an unbiased estimator for sigma square is not correct, the unbiased estimator for sigma square is this quantity which is sigma x i minus x bar whole square upon n minus 1 and we often denoted by s square.

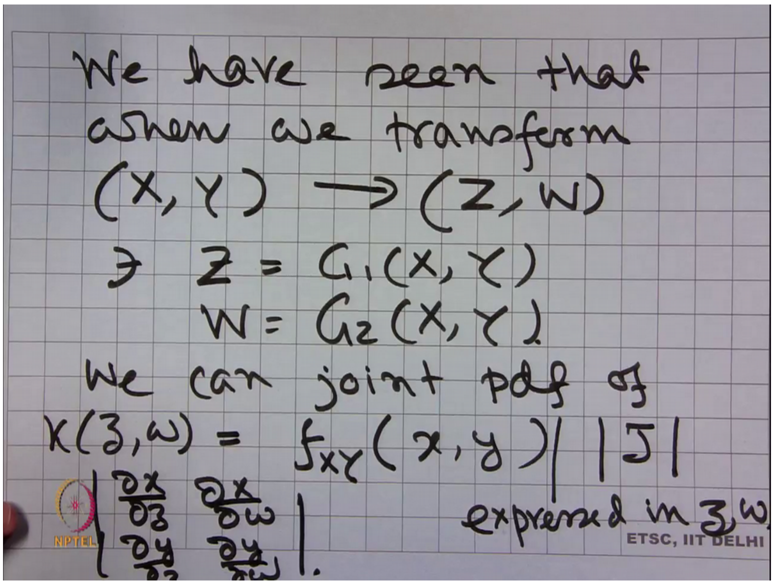
(Refer Slide Time: 13:41)



Now, we want to find the distribution of or we can say we want to find the sampling distribution of capital S square.

In order to do that we first need some mathematical tricks.

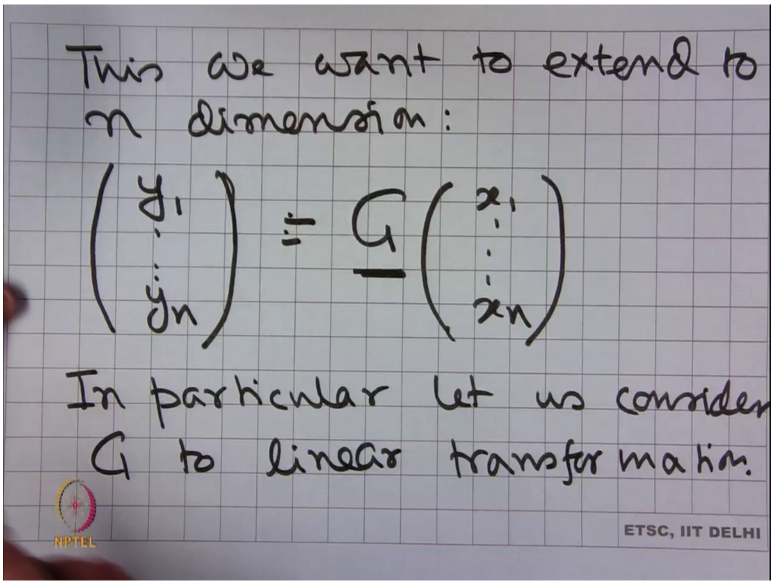
(Refer Slide Time: 14:17)



We have seen that when we transform from the XY plane to the Z W plane such that Z is equal to G 1 of X Y, and W is equal to G 2 of X Y we can get joint pdf of z w as the joint distribution of XY expressed in zw multiplied by the Jacobean where the Jacobean is the determinant of del x, del z, del x, del w, del y, del z and del y del w.

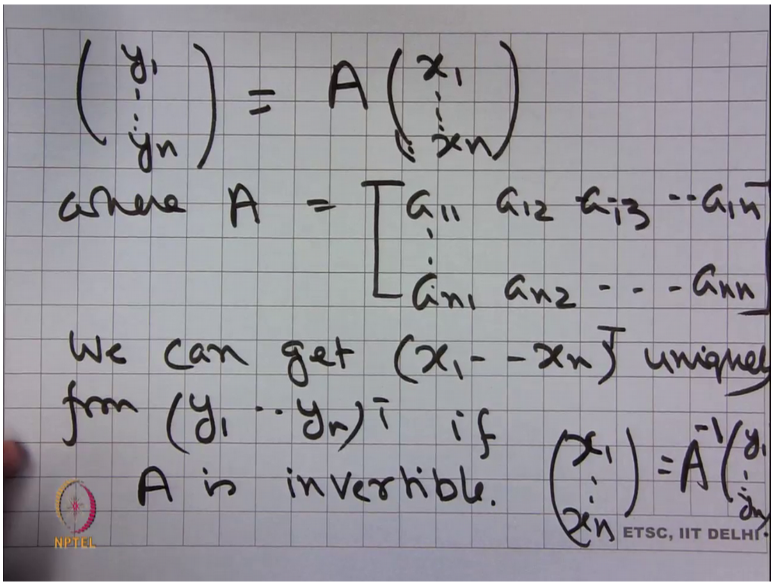
Now, suppose we want to extend it to n dimension. The most important concept was that given z w we can use inverse transformation to get unique solution for x y.

(Refer Slide Time: 16:35)



Now, we want to extend to n dimension. Suppose we write y 1 y 2 y n as a function of x 1 x 2 x n. So, we have n random variables x 1 x 2 x n, we want to make it transformation to new set of variables y 1 y 2 y n using A transformation function g.

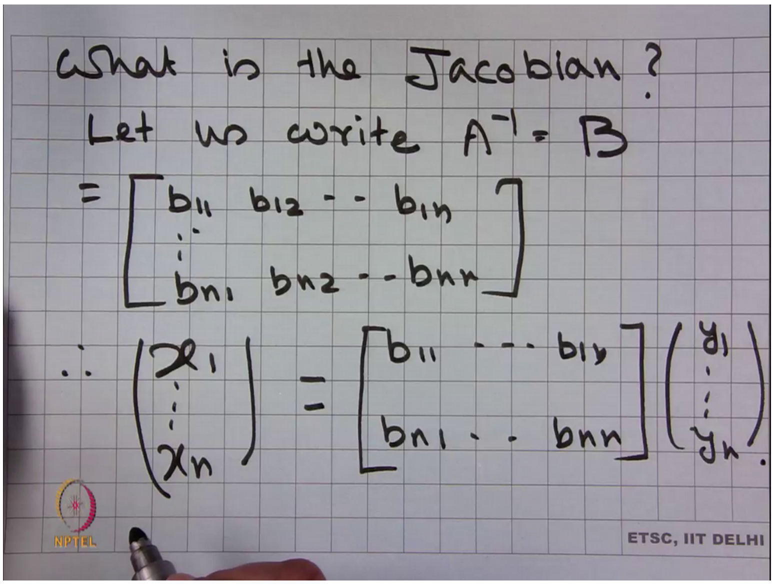
(Refer Slide Time: 17:55)



In particular,, let us consider g to be linear transformation; that is, y 1 y 2 yn is equal to a times x 1 x 2 xn where A is equal to a 1 1, a 1 2, a 1 3, a 1 n, a n 1, a n 2, a n n.

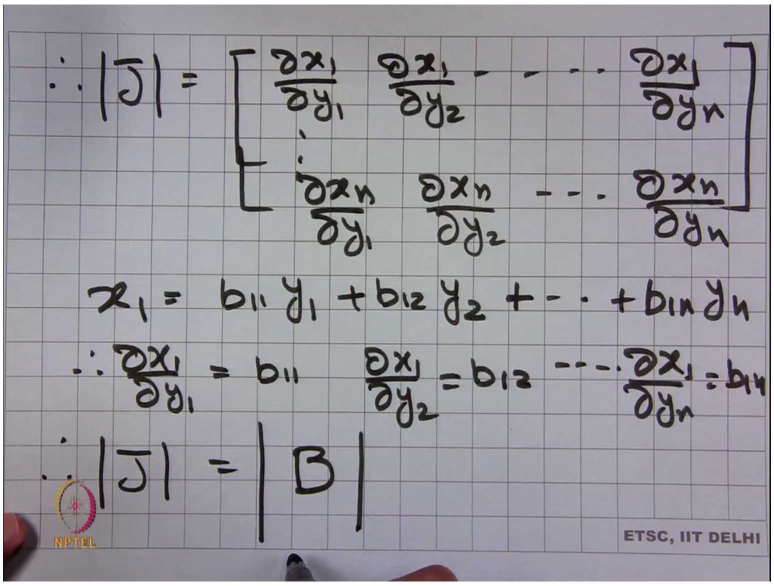
We can get x 1 x 2 xn uniquely from y 1 y 2 yn. Let us use transpose notation for column vectors, if a is invertible. Then we get x 1 x 2 x n is equal to A inverse y 1 y 2 y n. In this case what is going to be the Jacobean?

(Refer Slide Time: 19:25)



Let us write A inverse is equal to B, is equal to b 1 1, b 1 2, b 1 n, b n 1, b n 2, b n n. Therefore, x 1 x 2 up to xn is equal to b 1 1 up to b 1 n b n 1 up to b n n times y 1 y 2 y n.

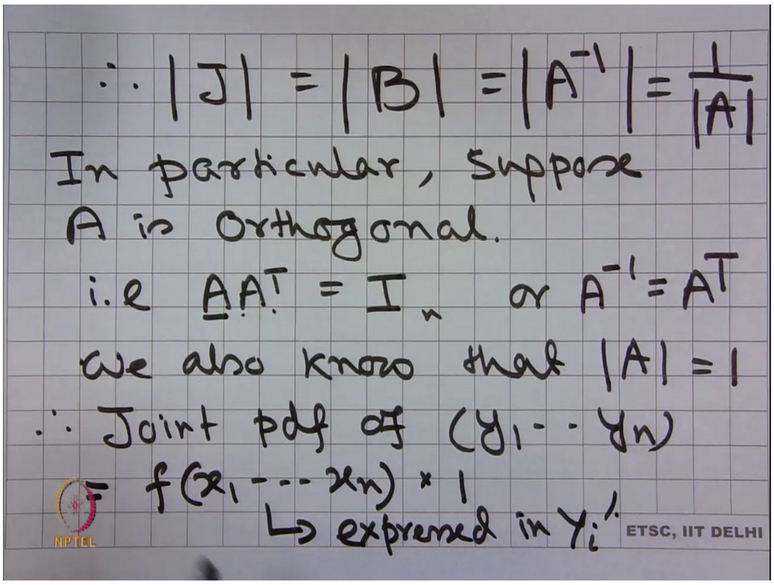
(Refer Slide Time: 20:36)



Therefore, the Jacobean is equal to del x 1, del y 1, del x 1, del y 2, del xn, del x 1, del y n, del xn, del y 1, del xn, del y 2, del xn, del y n.

Now, let us look at one of them. We have x 1 is equal to b 1 1 y 1 plus b 1 2 y 2 plus b 1 n y n. Therefore, del x 1 del y 1 is equal to b 1 1, del x 1 del y 2 is equal to b 1 2, del x 1 del y n is equal to b 1 n. Or in other words, the first row of the Jacobean matrix is going to be the first row of b. In a similar way, we can see that this Jacobean matrix J is nothing but the inverse matrix B.

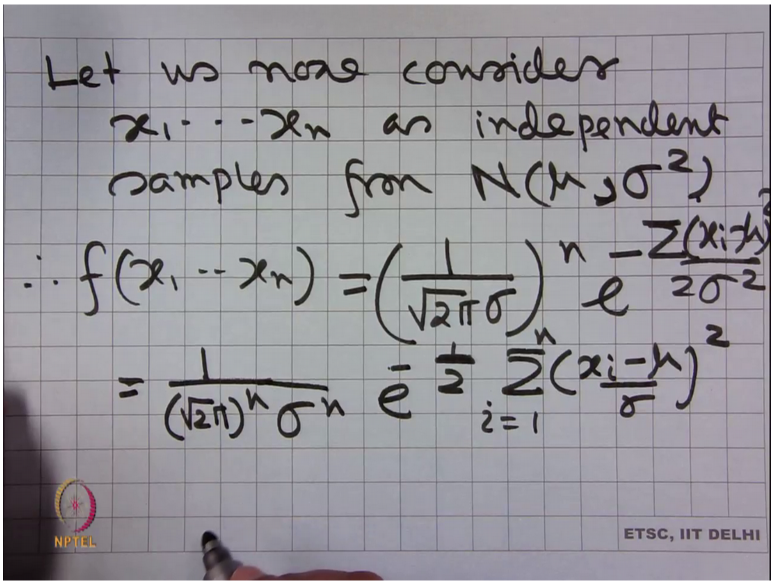
(Refer Slide Time: 22:23)



Therefore the Jacobean of the transformation is equal to determinant of B which is equal to determinant of A inverse, which is is equal to 1 upon determinant of A.

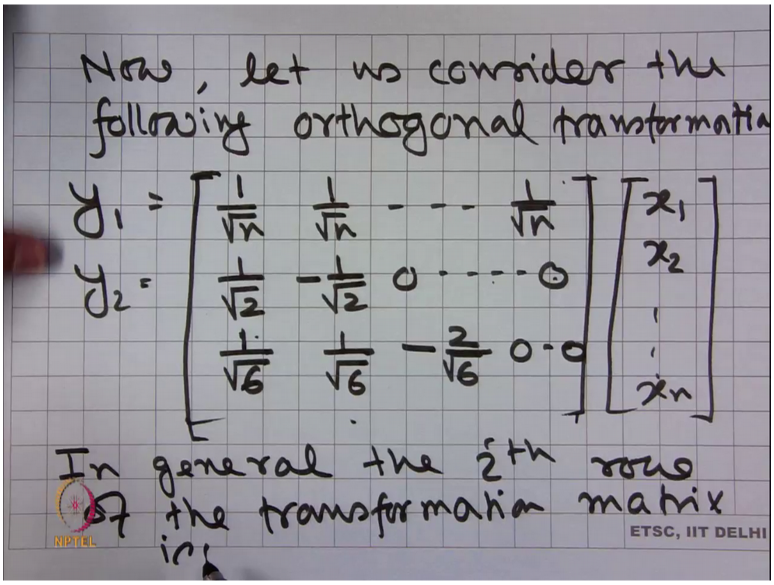
So, this is a very interesting observation. Now in particular suppose A is orthogonal; that is, AA transpose is equal to I. A is an n cross n orthogonal matrix therefore, aa transpose is equal to identity matrix, or A inverse is equal to A transpose. Therefore, therefore, if we consider that transformation matrix to be orthogonal, we get this and we also know that determinant of A is equal to 1 therefore, joint pdf of y 1 y 2 y n is equal to joint pdf of x 1 x 2 xn multiplied by 1 this is expressed in y i's right ?

(Refer Slide Time: 24:21)



Let us now consider x 1 x 2 xn as independent samples from normal mu sigma square. Therefore, f of x 1 x 2 xn which is the joint pdf of x 1 x 2 xn is equal to 1 over root over 2 pi sigma whole to the power n, e to the power minus sigma x i minus mu whole square upon 2 sigma square, is equal to 1 over root over 2 pi to the power n sigma to the power n e to the power minus half sigma x i minus mu upon sigma whole square summation i is equal to 1 to n.

(Refer Slide Time: 25:53)

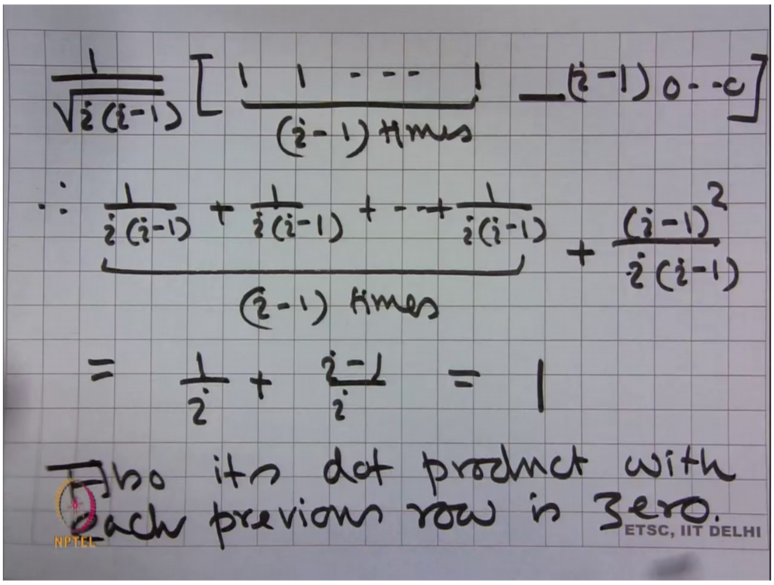


Now, let us consider the following orthogonal transformation. Y 1 is equal to 1 over root n 1 over root n 1 over. So, y 1 is considered to be the dot product of 1 over root n, into x 1 plus 1 over root n into x 2 plus 1 over root n into x n.

Let us consider y 2 to be the second row is 1 upon root 2 minus 1 upon root 2 0 0. Therefore, first thing you note that the sum of squares of these values is equal to 1. Sum of squares of these values is equal to half plus half is equal to 1. Not only that if we look at the dot product of this vector with this we get 0. Therefore, these 2 vectors are mutually orthogonal.

Similarly, we can have the third row is equal to 1 over root 6, 1 over root 6 minus 2 over root 6 rest are 0. Again if you look at it is sum of square is equal to 1 by 6 plus 1 by 6 plus 4 by 6 is equal to 1. And not only that it is orthogonal to this and it is orthogonal to this.

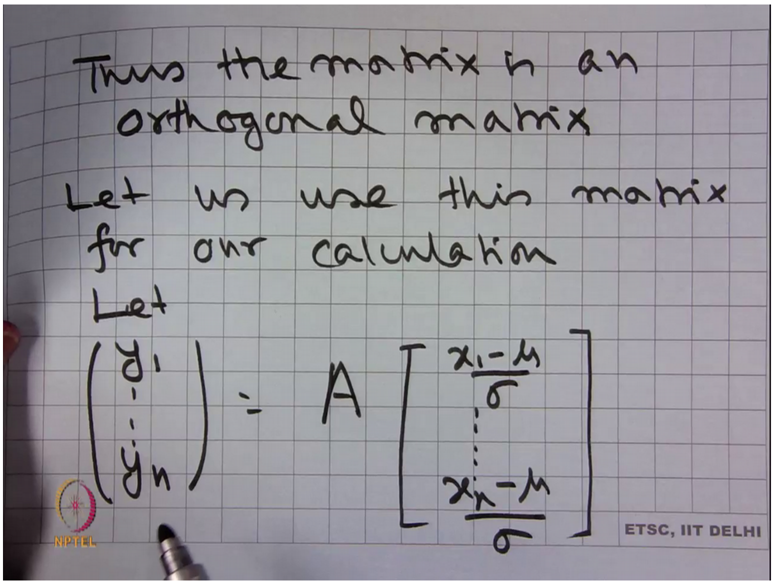
(Refer Slide Time: 28:40)



In general, the ith row of the transformation matrix is 1 over root I into I minus 1 into 1 1 1 i minus 1 times minus I minus 1 and rest are all 0’s.

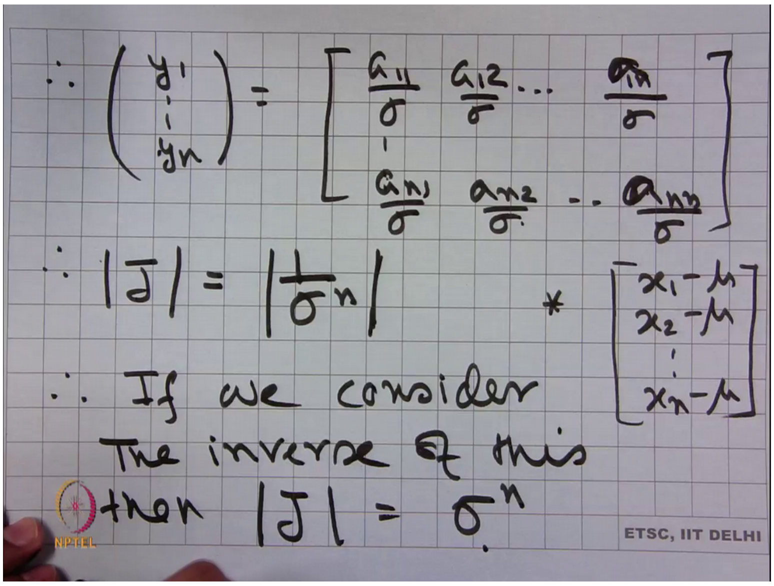
Therefore, the sum square of the element is 1 upon i into i minus 1 plus 1 upon i into i minus 1 plus 1 upon i into i minus 1. This is i minus 1 times plus I minus 1 squared upon i into i minus 1, is equal to 1 upon i plus i minus 1 upon i is equal to 1. Thus the norm the length of each of the rows of the matrix is 1. You can easily verify that it is dot product with all the previous rules going to be 0.

(Refer Slide Time: 30:34)



Thus the matrix is an orthogonal matrix. Let us use this matrix for our calculation. Let y 1 y 2 yn be A into x 1 minus mu upon sigma up to xn minus mu upon sigma. So, this A is the particular A that I have just constructed.

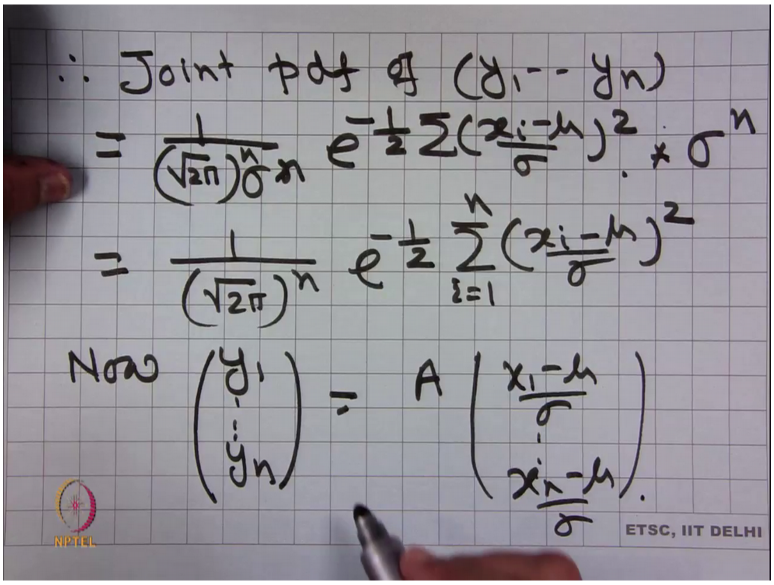
(Refer Slide Time: 31:52)



Therefore, y 1 y 2 y n I can write it as a 1 1 upon sigma, a 1 2 upon sigma, a 1 n up a 1 n upon sigma, a n 1 upon sigma, a n 2 upon sigma, a n n upon sigma, multiplied by x 1 minus mu x 2 minus mu up to xn minus mu. This a 1 is the original, this a ones are the matrix of the A transformation that we have considered.

Therefore, needs Jacobean is going to be 1 upon sigma to the power n, if I take the modulus if I take 1 upon sigma to be out from each of the n columns the determinant is equal to 1. Therefore, that mod of the Jacobean, if I consider this transformation that is going to be 1 upon sigma to the power n. Therefore, if we consider the inverse of it, the Jacobean is going to be sigma power n.

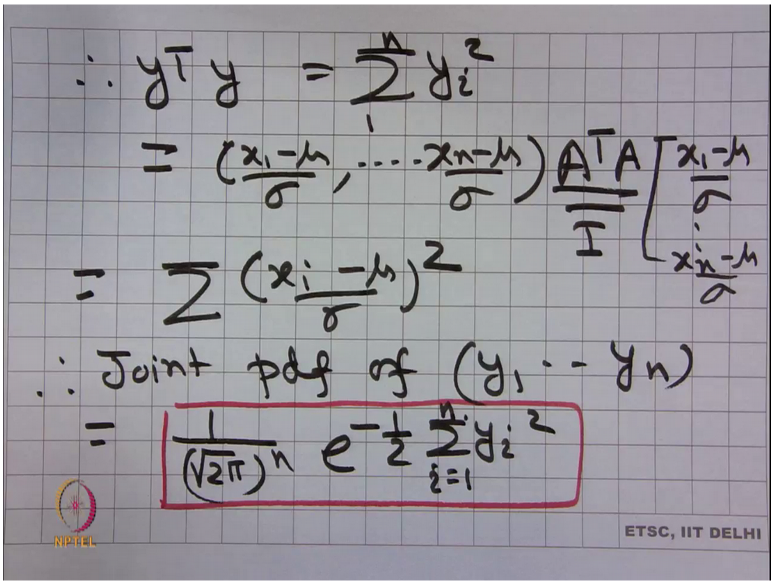
(Refer Slide Time: 34:03)



Therefore joint pdf of y 1 y 2 y n is equal to 1 over root over 2 pi sigma to the power n into e to the power minus half sigma x i minus mu upon sigma whole square multiplied by sigma power n, is equal to 1 over root over 2 pi whole to the power n e to the power minus half and it is sigma x i minus mu upon sigma whole square, i is equal to 1 to n.

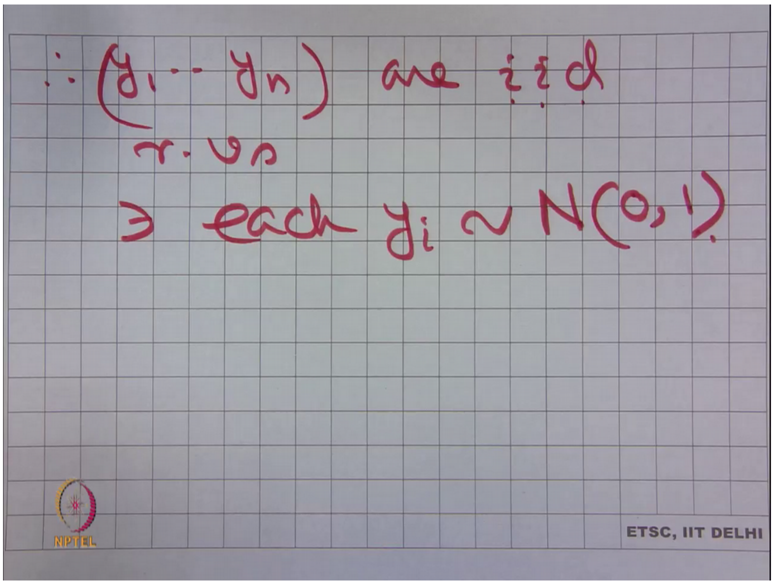
Now, y 1 y 2 yn is equal to A times x 1 minus mu upon sigma up to xn minus mu upon sigma.

(Refer Slide Time: 35:27)



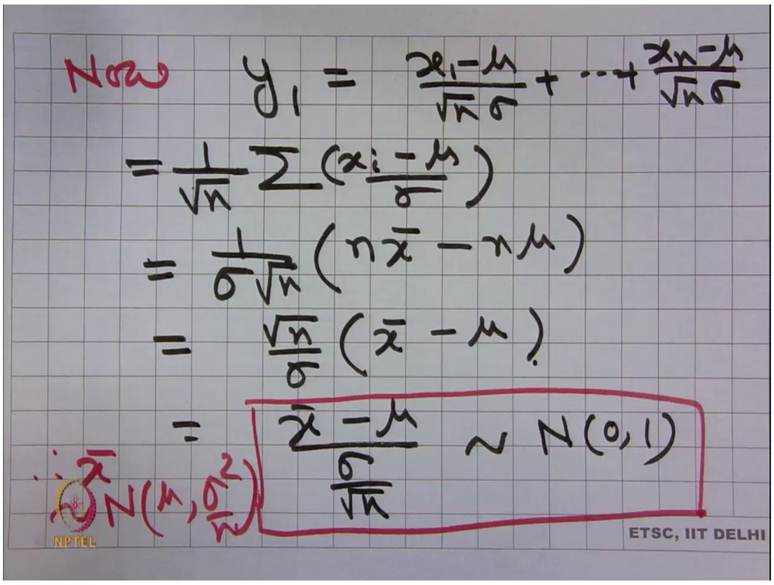
Therefore y transpose y is equal to sigma y i square 1 to n, is equal to x 1 minus mu upon sigma, xn minus mu upon sigma, into A transpose A x 1 minus mu one sigma up to xn minus mu upon sigma, is equal to since this is equal to identity is equal to sigma x y x i minus mu upon sigma whole square. Therefore, joint pdf of y 1 y 2 y n is equal to 1 over root over 2 pi whole to the power n e to the power minus half sigma y i square i is equal to 1 to n. What does it tell us?

(Refer Slide Time: 36:56)



It tells that y 1 1 y 2 y n are iid independent identically distributed random variables; such that each y i follow normal 0 1. So, the advantage of the particular transformation that I have made is that, it converts from x 1 x 2 xn each of which is normal mu sigma square 2 y 1 y 2 y n which are independent, and each y is normal 0 1.

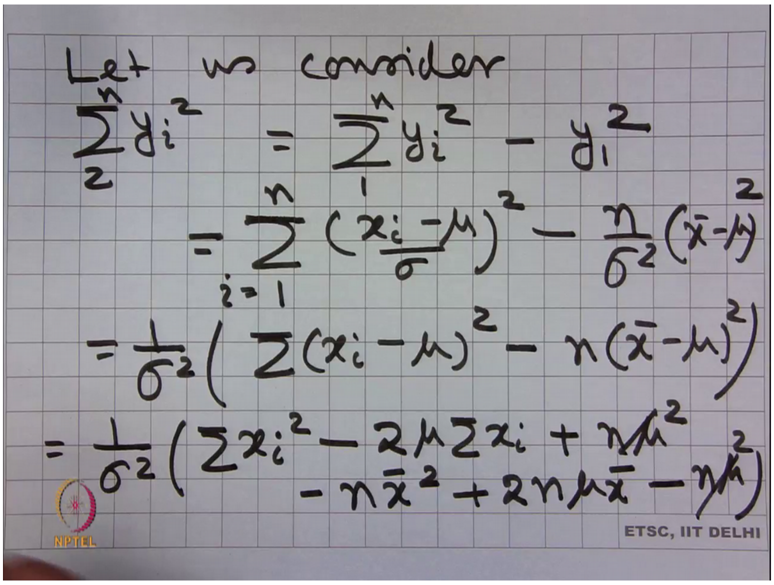
(Refer Slide Time: 37:46)



Now, what is y 1? Y 1 is equal to x 1 minus mu upon root n sigma plus up to xn minus mu upon root n sigma, is equal to 1 over root n sigma x i minus mu upon sigma, is equal to 1 upon sigma root n sigma x i is equal to n times x bar minus n times mu, is equal to root n upon sigma into x bar minus mu, is equal to x bar minus mu upon sigma root n.

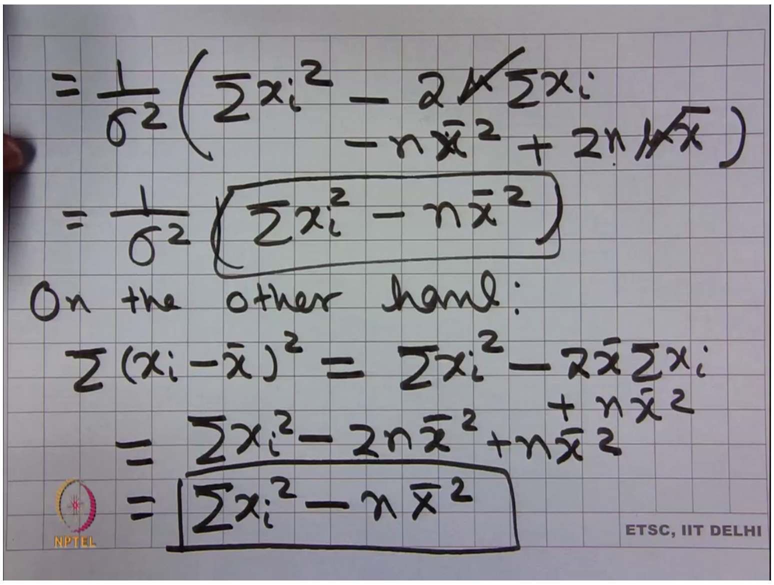
And since y 1 is normal 0 1 this is normal 0 1 as well. So, this also suggests that x bar is normal with mean is equal to mu, and variance is equal to sigma over root n square is equal to sigma square by n. Although this result I have proved before we can find it here also, but we get something extra. What is that?

(Refer Slide Time: 39:50)



Let us consider sigma y i square 2 to n. This is equal to sigma y i square 1 to n minus y 1 square, is equal to sigma x i minus mu by sigma whole square i is equal to 1 to n. Because we have seen some time back that sigma yi square is equal to sigma x i minus mu upon sigma whole square, minus y 1 square, which is equal to n upon sigma square into x bar minus mu whole square, is equal to 1 upon sigma square into sigma x i minus mu whole square minus n into x bar minus mu whole square, is equal to 1 by sigma square sigma x i square minus 2 mu sigma x i plus n mu square, minus n x bar square plus 2 n mu x bar plus minus n mu square. So, it is n times x bar square plus 2 mu x bar minus n times mu square.

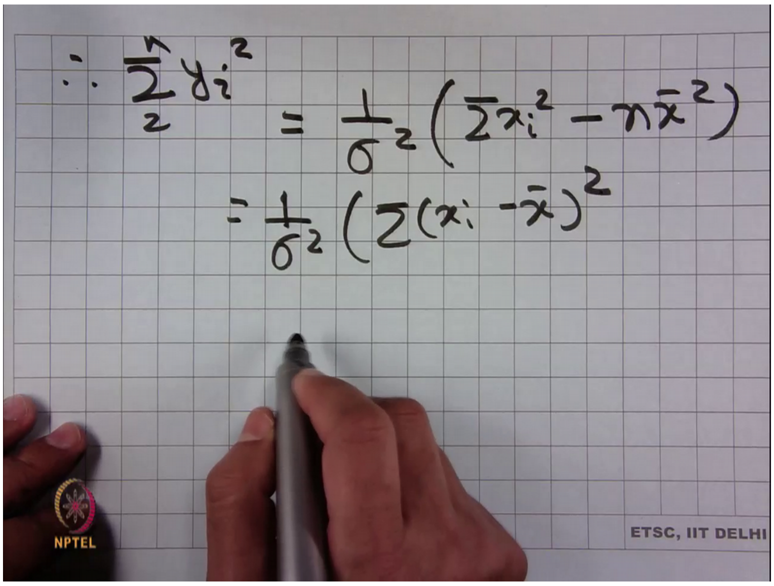
(Refer Slide Time: 42:16)



So, this cancels is equal to 1 by sigma square into sigma x i square minus 2 mu sigma x i minus n x bar square plus 2 n mu x bar. Now sigma x i is equal to n times x bar. So, this also cancels out, is equal to 1 by sigma square into sigma x i square minus n times x bar square.

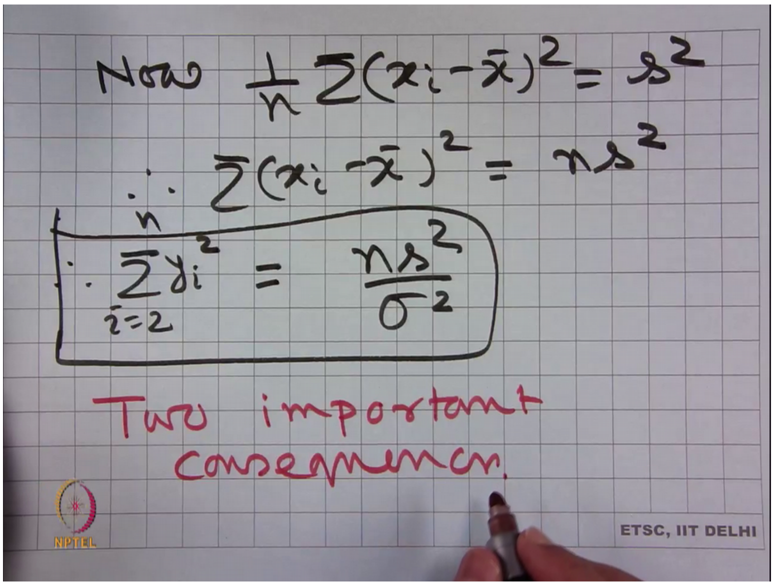
On the other hand,, sigma x i minus x bar whole square is equal to sigma x i square minus 2 x bar sigma x i plus n x bar square, is equal to sigma x i square minus 2 n x bar square plus n x bar square is equal to sigma x i square minus n x bar square.

(Refer Slide Time: 44:16)



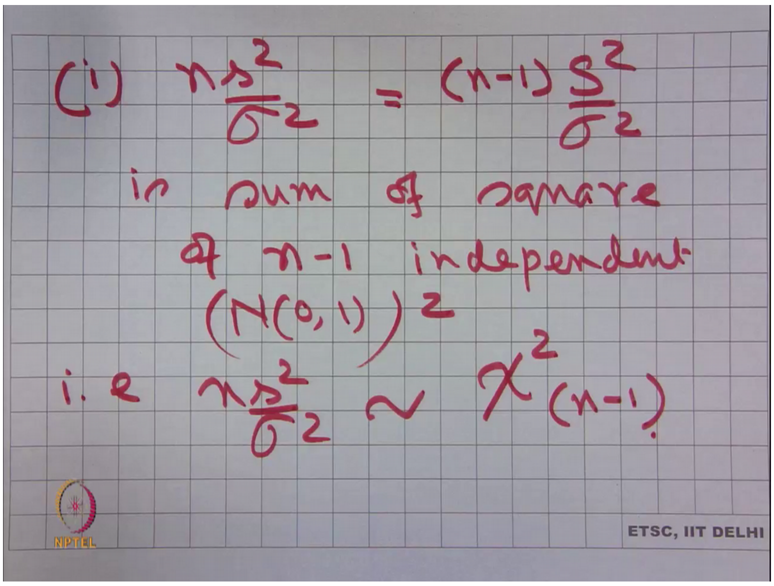
So, these 2 terms are same therefore we can write as sigma y to y i square from 2 to n ,which we came out to be 1 by sigma square into sigma x i square minus n times x bar square, is equal to 1 by sigma square into sigma x i minus x bar whole square.

(Refer Slide Time: 44:44)



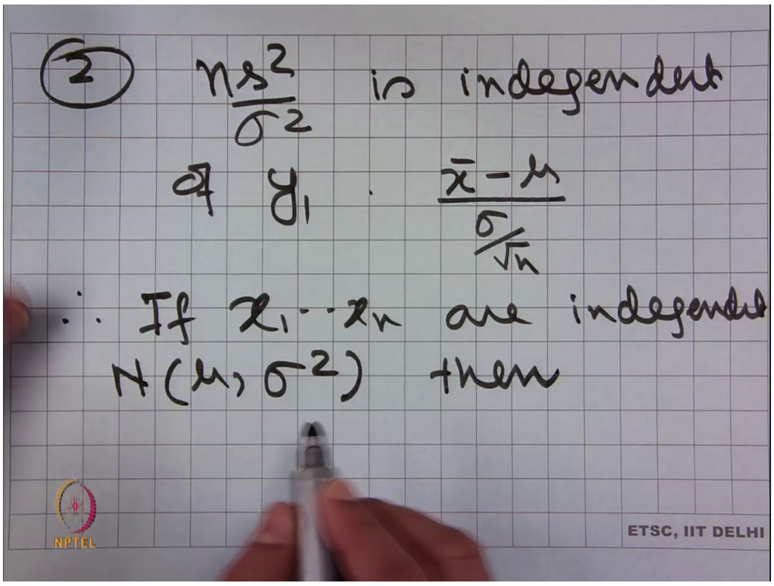
Now, 1 by n sigma x i minus x bar whole square is equal to s square. Therefore, sigma x i minus x bar whole square is equal to n times s square. Therefore, sigma y i square i is equal to 2 to n is equal to n s square upon sigma square. This gives us 2 important consequences.

(Refer Slide Time: 45:50)



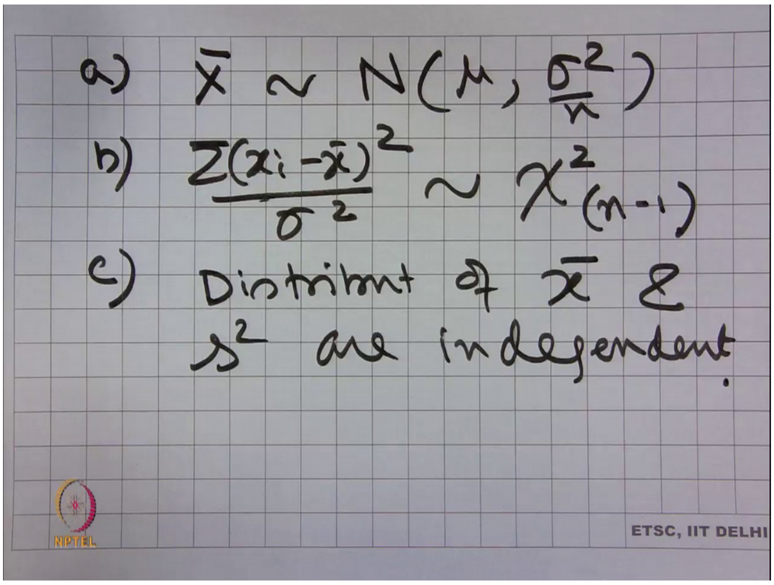
One ns square sigma square which is same as n minus 1 into s square upon sigma square is sum of square of n minus 1 independent normal 0 1 square; that is, n s square upon sigma square is distributed as chi square with n minus 1 degrees of freedom. Because ns square upon sigma square is sum of sigma yi square from 2 to n; that means, n minus 1 of independent normal 0 1 square. Therefore, ns square upon sigma square is distributed at chi square with n minus 1 degrees of freedom.

(Refer Slide Time: 47:10)



And the second point is that n s square upon sigma square is independent of y 1. Because we have found that y 1 y 2 yn are n independent normal 0 1 variate of which n s square sigma square depends only on from y 2 to y n. Therefore, it is independent of y 1 which is nothing but x bar minus mu upon sigma by root n.

(Refer Slide Time: 48:13)



Therefore, if x 1 x 2 xn are independent normal mu sigma square, then a x bar is normal with mean mu variance sigma square by n, b sigma x i minus x bar whole square upon sigma square is distributed as chi square with n minus 1 degrees of freedom. And c distribution of x bar and s square are independent.

So, these are 3 important findings that we get by making a particular orthogonal transformation of x 1 x 2 xn to y 1 y 2 y n. So, these are some important results that we obtain when x 1 x 2 xn are independent samples from normal mu sigma square. With that I stop for today.

Thank you.