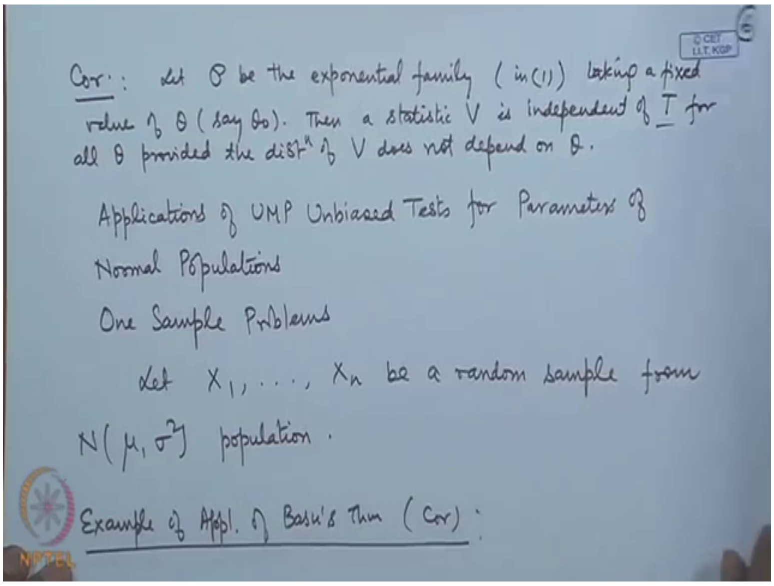
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 48
Unbiased Test for Normal Populations – II

So, I have considered testing for the mean of a normal distribution, testing for the variance in a normal distribution. But the crucial difference was that when I was testing for the mean I had considered variance to be known and accordingly the tests which were either UMP for h 1 and h 2 and 4 h 3 and h 4 it was UMP un biased we were had obtained.

Similarly, for sigma square when I was doing the testing the mu was taken to be known and I had taken without loss of generality to be 0 and once again we had the UMP test for h 1 and h 2 and UMP unbiased test for h 3 and h 4. Now here both the parameters will be unknown, which is actually more practical, in practice when we discuss certain population it is unlikely that one of the parameters may be known.

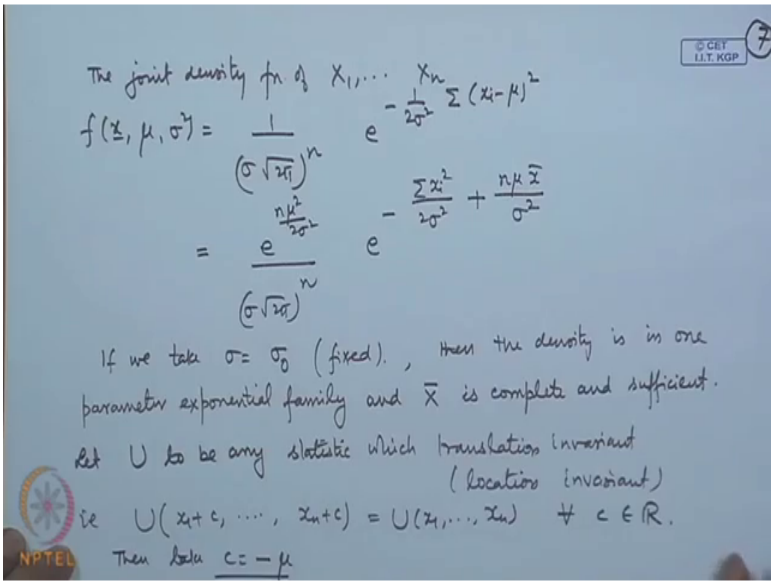
(Refer Slide Time: 01:12)



So, the general situation is both the parameters may be unknown. And in this case we will show that when we test for the mean and when we test for the variance we are in general able to derive UMP unbiased tests for all 4 types of hypothesis that is h 1 h 2 h 3 and h 4. In particular I will describe in detail the test for h 1 and h 4 because they look more natural hypothesis, h 2 and h 3 look more artificial hypothesis of course, you can write down the form of the test functions for each of these cases

So, let us firstly, start with applications of Basu's theorem firstly, example of application of Basu's theorem or this corollary in the normal population case.

(Refer Slide Time: 02:13)

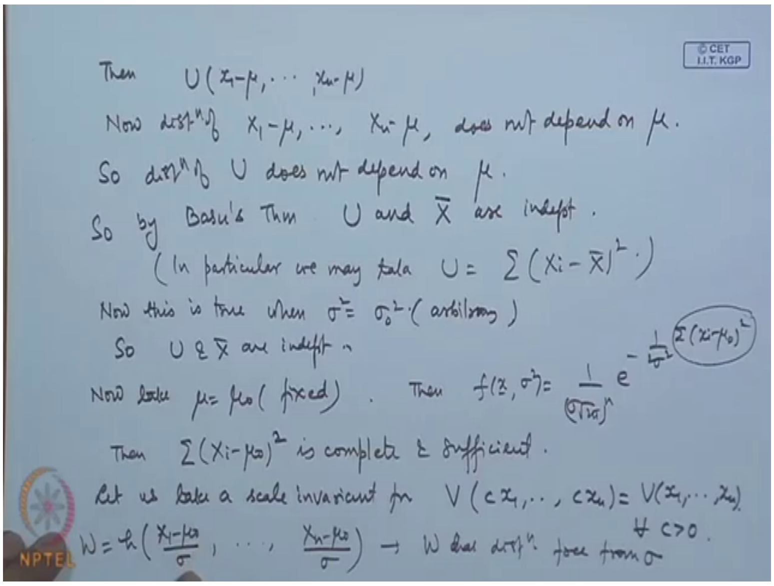


So, let us look at the joint density function of X 1 X 2 X n. So, we are writing it as 1 by sigma root 2 pi to the power n e to the power minus 1 by 2 sigma square sigma x i minus mu square that is equal to e to the power n mu square by 2 sigma square divided by sigma root 2 pi to the power n e to the power minus sigma x i square by 2 sigma square plus n mu x bar divided by sigma square.

Note here this is a 2 parameter exponential family, I am looking at the application of Basu's theorem, so firstly we will look at the independence part here. If we take say sigma is equal to some fixed value sigma naught, if sigma naught is fixed here then this term is going into the hx part of a multi parameter exponential family this will become fixed. So, here you have mu X bar or n mu X bar. So, this is one parameter exponential family then the density is in one parameter exponential family and X bar is complete and sufficient.

Let us take say U to be any statistic which is translation invariant or you can say location invariant; location invariant; that means, I am saying U of x 1 plus c x 2 plus c x n plus c that is equal to U of x 1 x 2 x n for all c on the real line. Now then I can choose c is equal to say minus of mu.

(Refer Slide Time: 05:23)



If I do that then this becomes a function of x 1 minus mu and so on x n minus mu. Now distribution of X 1 minus mu X 2 minus mu X n minus mu does not depend on mu. So, what we are having X bar is complete and sufficient the parameter is mu here and U is a function which is having distribution, so distribution of U does not depend on mu.

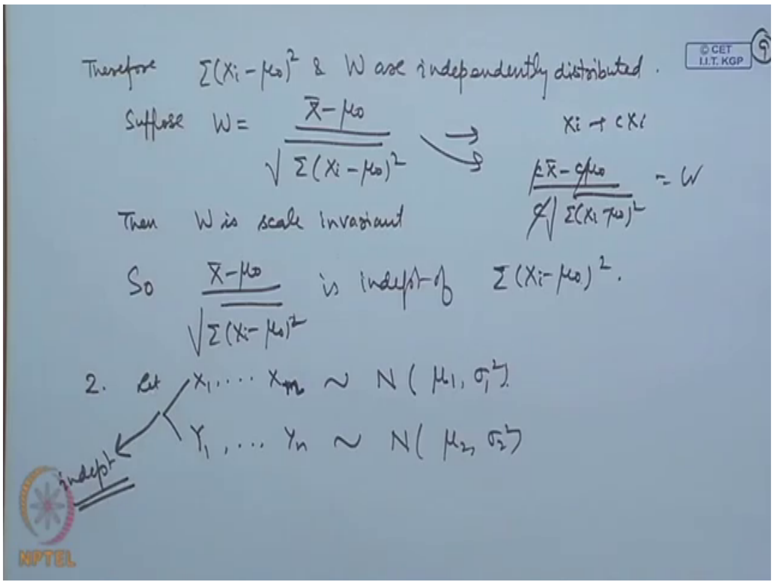
So by Basu's theorem U and X bar are independent, in particular we may take U is equal to say sigma of Xi minus X bar whole square. Of course, in the sampling from normal distributions we know that sample mean and sample variance are independent, but here this is also proved from the Basu's theorem. Now this is true when sigma square is fixed as sigma naught square, but the distribution of a this is arbitrary.

So, if we look at this result here the distribution of sigma Xi minus X bar whole square does not depend upon mu I am sorry this is application of Basu's theorem here V is ancillary here, so this is following from here. So, U and X bar are independent. Now let us take another application here.

Now, take mu is equal to mu naught another fixing of the mean. Then the density is f x sigma square is equal to 1 by sigma root 2 pi to the power n e to the power minus 1 by 2 sigma square sigma x i minus mu naught whole square. So, this is again one parameter exponential family and this becomes sigma Xi minus mu naught square is complete and sufficient.

Let us take a scale invariant function V that is V of cx 1, cx 2, cx n is equal to V of x 1, x 2, x n for all c positive. Then in particular I can consider c this to be a function of say h of X 1 minus mu naught by sigma and so on X n minus mu naught by sigma, then let me call it say W, then W has a distribution free from sigma because the distribution of Xi minus mu naught by sigma is normal 0 1, so this size distribution free from sigma.

(Refer Slide Time: 09:35)

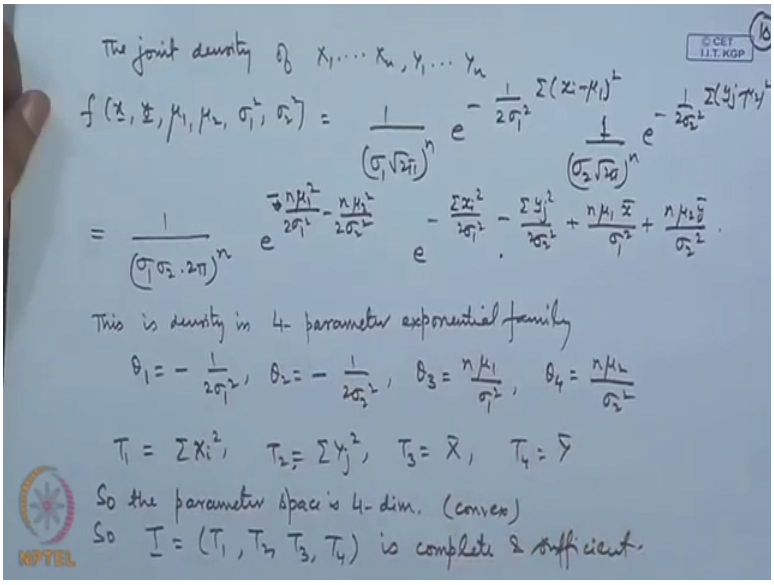


And therefore, sigma Xi minus mu naught square and W are independently distributed. Suppose I take W is equal to X bar minus mu naught divided by sigma Xi minus mu naught square. Then W is a scale invariant because if I change here Xi to c X i then this does not change c X bar minus mu naught divided by c square root Xi minus mu naught square. So, this gets cancelled out, so that is equal to W.

So, this is scale invariant, so X bar minus mu naught divided by square root of sigma Xi minus mu naught square is independent of sigma Xi minus mu naught square. Of course, here mu naught is fixed here unlike the previous application where sigma was not appearing here sigma naught was not appearing. So, independence of X bar and sigma Xi minus X bar whole square was for all sigma whereas, this result depends upon the fixed value of mu naught.

Let me take another example here let X 1 X 2 X n be a random sample let me take say X 1 X 2 X n be a random sample from normal mu 1 sigma 1 square and Y 1, Y 2, Y n be another random sample from normal mu 2 sigma 2 square. I also assume that these two are independent these two samples are taken independently.

(Refer Slide Time: 12:00)



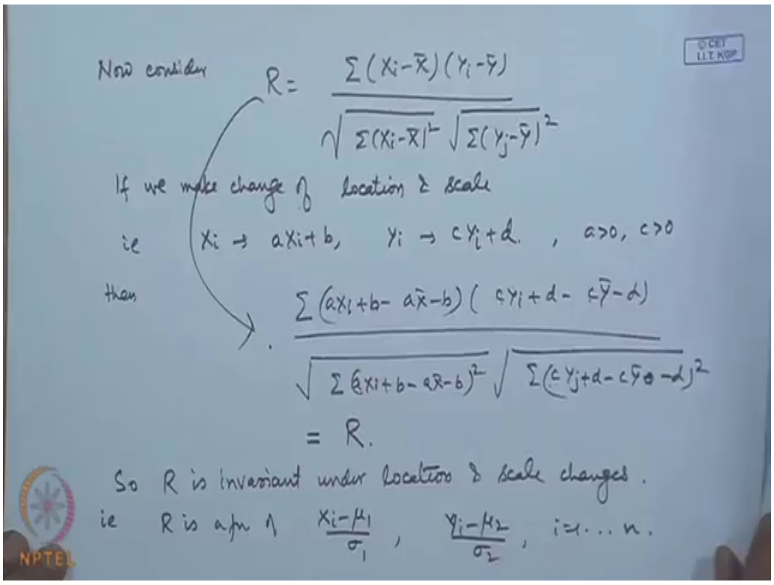
Then let us write down the joint density of X 1, X 2, X n and Y 1 Y 2 Y n, then that will be f x y mu 1 mu 2 sigma 1 square sigma 2 square that is equal to 1 by sigma 1 root 2 pi to the power n e to the power minus 1 by 2 sigma 1 square sigma xi minus mu 1 square 1 by sigma 2 root 2 pi to the power n e to the power minus 1 by 2 sigma 2 square sigma yj minus mu 2 square.

Now, this I will simplify and we can consider this expression is 1 by sigma 1 sigma 2 2 pi to the power n e to the power minus plus n, if I take a square here this will give a minus. So, minus here also it was minus in the earlier case minus n mu 1 square by sigma 1 square minus n mu 2 square by sigma 2 square 2 will come here, e to the power minus sigma xi square by 2 sigma 1 square minus sigma yj square by 2 sigma 2 square plus n mu 1 x bar by sigma 1 square plus n mu 2 y bar by sigma 2 a square.

So, this is a density in 4 parameter exponential family we may write say theta 1 is equal to minus 1 by 2 sigma 1 square, theta 2 is equal to minus 1 by 2 sigma 2 square, theta 3 as say mu 1 by sigma 1 square, theta 4 is equal to n mu 2 by sigma 2 square. We may write T 1 as sigma Xi square, T 2 as sigma Yj square T 3 as X bar, T 4 s Y bar.

So, here the parameter space is natural parameter space, so the parameter space is 4 dimensional; 4 dimensional and also it is convex. So, T is equal to T 1, T 2, T 3, T 4 is complete and sufficient here. Sufficiency follows from the factorization theorem and completeness follows from the theorem which is given in the lemma that if a k parameter exponential family contains a k dimensional rectangle the parameter space of a k dimensional rectangle family then the corresponding statistic which is appearing in the exponent will be complete.

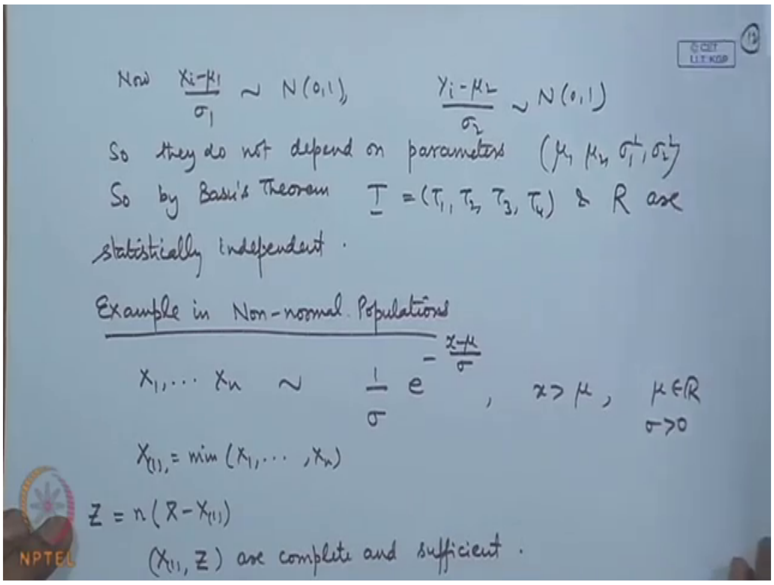
(Refer Slide Time: 15:52)



Now, consider say R that is actually the sample correlation coefficient. If we make change of location and scale, that is if I say Xi goes to say a Xi a plus b and Yi goes to some c Yi plus d. Then what happens here? If you look at this term sigma a Xi plus b minus a X bar minus b c Yi a plus d minus c Y bar minus d divided by similarly here a Xi minus plus b minus a X bar minus b square a square root sigma c Yj plus d minus c Y bar minus d whole square.

So, now this if you see here b will cancel d will cancel c and they will come out and here also a and c will come out. So, we of course, put the condition a greater than 0 and c greater than 0 then this is equal to R again. So, R is invariant under location and scale changes; that means, R is a function of Xi minus mu 1 by sigma 1, Yi minus mu 2 by sigma 2 for i is equal to 1 to n.

(Refer Slide Time: 18:15)



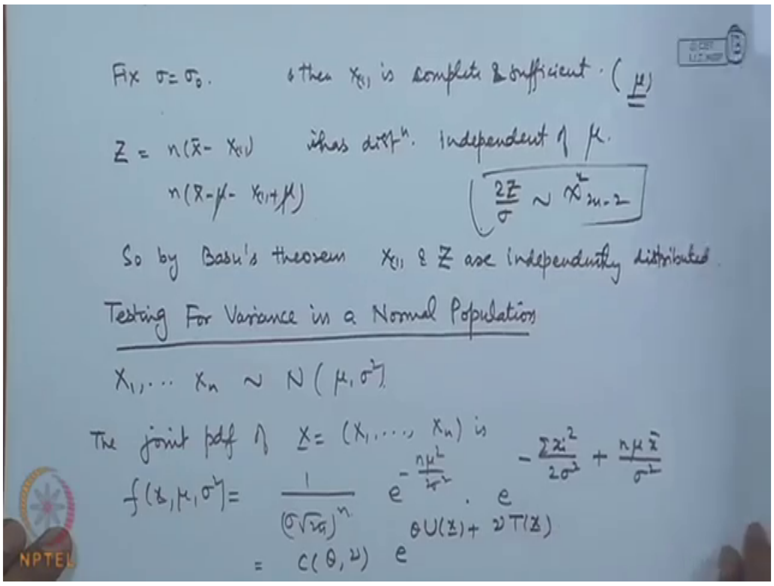
Now, the distributions of Xi minus mu 1 by sigma 1 that is normal 0 1 distribution of Yi minus mu 2 by sigma 2 that is normal 0 1. So, they are they do not depend upon parameters mu 1, mu 2, sigma 1 square, sigma 2 square. So, what we are having now? We have the parameters appearing theta 1, theta 2, theta 3, theta 4 which is a 1 to 1 function of mu 1 mu 2 sigma 1 square and sigma 2 square.

The corresponding complete sufficient statistic is T 1, T 2, T 3, T 4 or which is a 1 to 1 function of X bar, Y bar and sigma Xi square, sigma Yj square. Now what we have demonstrated the sample correlation coefficient has a distribution which does not depend upon the parameters. So, by Basu's theorem then T that is T 1, T 2, T 3, T 4 and R they are independently distributed.

So, this Basu's theorem is very useful result and I have given applications to the normal distributions, but it has also applications in exponential populations, inverse Gaussian populations there are various distributions where this Basu's theorem is extremely useful. Let me give an example which is different from this normal population's example in say non normal; non normal populations.

Let us take say X 1, X 2, X n following 1 by sigma e to the power minus x minus mu by sigma, where mu is any real number sigma is a positive parameter; that means, I am considering a 2 parameter exponential distribution. Now this is not a exponential family; however, we have earlier shown X 1 that is the minimum of X 1, X 2, X n and if I consider say n times X bar minus X 1. Then let me give some name to this let me call it say Z then X 1 and Z they are complete and sufficient.

(Refer Slide Time: 21:46)

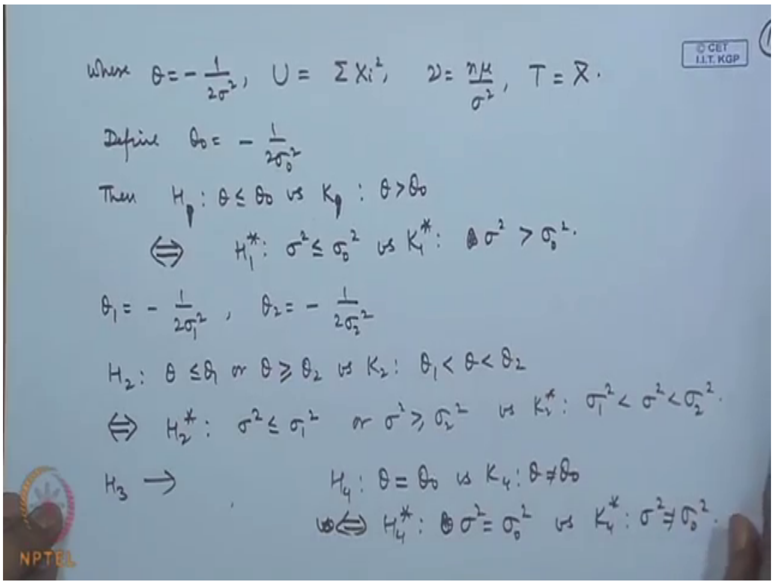


Now if I fix a sigma fix sigma is equal to sigma naught then what happens we will get then X 1 is complete and sufficient X 1 is complete and sufficient. Now Z if I take n X bar minus X 1 this is how. So, now, this is completed and sufficient and the parameter is actually mu now here and if I consider say n X bar minus mu minus X 1 minus mu, so this gets cancelled out.

So, this is this has a distribution independent of mu in fact, we know the distribution 2 Z by sigma actually follows chi square on 2 and minus 2 degrees of freedom. So, by Basu's theorem you will have that X 1 and Z are independently distributed. So, I have given an example of a non normal population now let us pay attention to the testing problems, testing for say variance in a normal population

So, if I am writing down the density function of X 1, X 2, X n f X mu sigma square that is equal to 1 by sigma root 2 pi to the power n e to the power minus n mu square by 2 sigma square e to the power minus sigma xi square by 2 sigma square plus n mu X bar by sigma square. So, here I am considering testing for sigma square, so I will consider here this part as say theta nu e to the power theta U x plus nu T x.

(Refer Slide Time: 24:53)



So, I have defined here theta is equal to minus 1 by 2 sigma square, U is equal to sigma Xi square nu is equal to say n mu by sigma square and T is equal to X bar. If I consider say testing problem, so now, by the theory that I have developed for that UMP bias test I can say that UMP bias test will exist if I test for theta hypothesis say H 1 versus K 1 H 2 versus K 2 x 3 versus K 3 and x 4 versus K 4.

For all the 4 hypothesis I will have UMP bias test for theta. Now what does testing for theta means? Since theta is equal to minus 1 by 2 sigma square the test is exactly reflected and this is an increasing function of sigma square minus 1 by 2 sigma square. So, if I write down a test like this suppose I define theta naught is equal to minus 1 by 2 sigma naught square then H naught theta less than or equal to theta naught versus K sorry H 1 theta better than theta naught. This testing problem is equivalent to let me call it H 1 star sigma square less than or equal to sigma naught square versus K 1 star the sigma square greater than sigma naught square.

Similarly, if I define say theta 1 is equal to minus 1 by 2 sigma 1 square, theta 2 is equal to minus 1 by 2 sigma 2 square, then if I write down the hypothesis problem theta less than or equal to theta 1 or theta greater than or equal to theta 2 versus K 2 theta 1 less than theta less than theta 2. Then this testing problem is equivalent to H 2 star, sigma square less than or equal to sigma 1 square or sigma square greater than or equal to sigma 2 square versus sigma 1 square less than sigma squared less than sigma 2 square.

And in a similar way if I consider H 3 and similarly H 4 that is theta is equal to theta naught that is equivalent to H 4 star theta sigma square is equal to sigma naught square versus K 4 star sigma square is not equal to sigma naught square. Therefore, we have demonstrated here that the theory of UMP un bias test for multi parameter exponential families, where we test for one parameter in the multi parameter exponential family other parameters are treated as constants.

Then that theory is applicable for two parameter normal population when we want to test for the variance I have put it exactly in that framework. Now let me write down the test here by using theorem 2 that I gave yesterday, where the test was conditional of U given T. Now what is U given T here? It will be sigma Xi square given X bar; that means, the test will be conditional test on sigma Xi square given X bar and I mentioned in the beginning of this lecture that this is slightly complicated problem.

What we need to do is to apply the theorem 3 that I gave today; that means, I define suitably a function. So, either in the condition U is greater than or equal to c naught T etcetera. I modify the condition in such a way that the term becomes free from T and similarly if I apply the theorem 3 which I gave today then I need to suitably define a function of this type.

So, in the following lecture I will be doing this work in the detail that is how to define this function W is equal to H of U T, I will consider it for a H 1 versus K 1 problem and that will be applicable for H 2 versus K 2, it will also be applicable to H 3 versus K 3; however, there will be a problem if I consider H 4 versus K 4. So, in H 4 versus K 4 a new function which is a linear function of U that has to be defined. So, we will demonstrate all of this in the following lecture here.