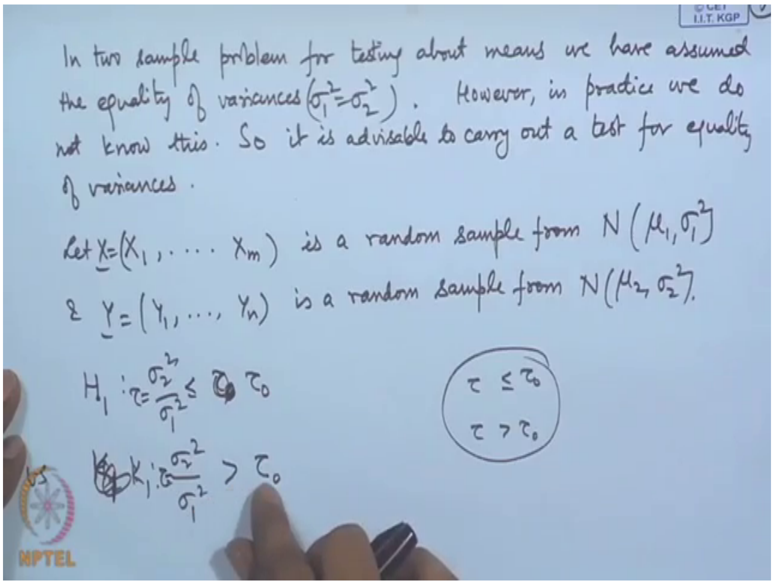
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 57
Likelihood Ratio Tests – VII

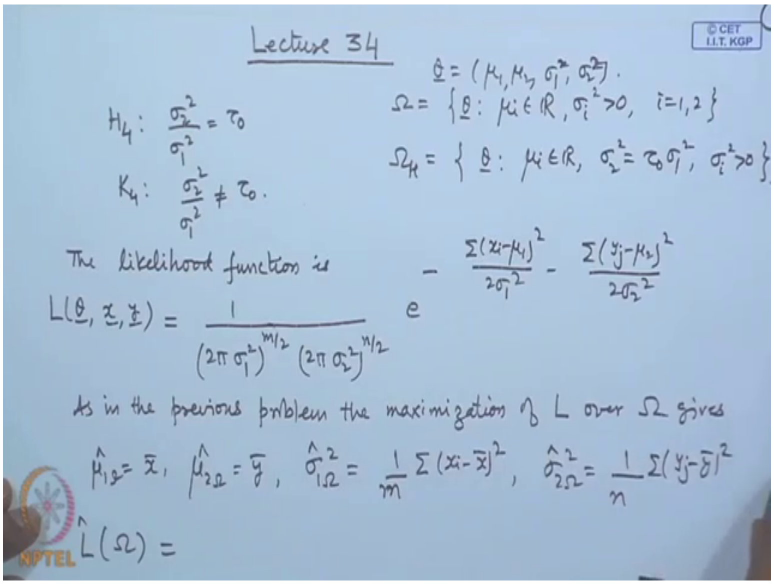
In the last lecture, I introduce the problem of testing or you can say comparing the variances of two normal populations.

(Refer Slide Time: 00:33)



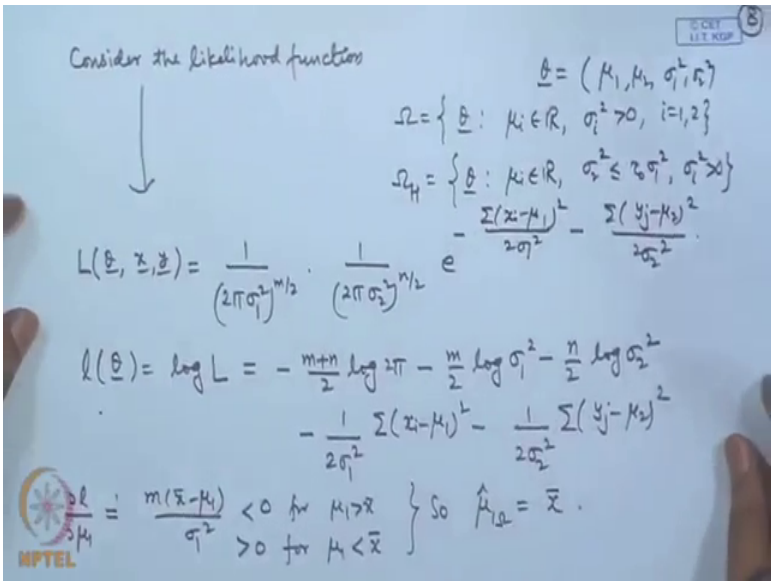
And we had considered one particular type of hypothesis problem that is sigma 2 square by sigma 1 square less than or equal to tau naught against sigma 2 square by sigma 1 square greater than tau naught. And the likelihood ratio test for this was derived.

(Refer Slide Time: 00:52)



Now, we will consider for the same model the hypothesis H 4 that is sigma 2 square by sigma 1 square is equal to tau naught against say sigma 2 square by sigma 1 square is not equal to tau naught ok.

(Refer Slide Time: 01:13)



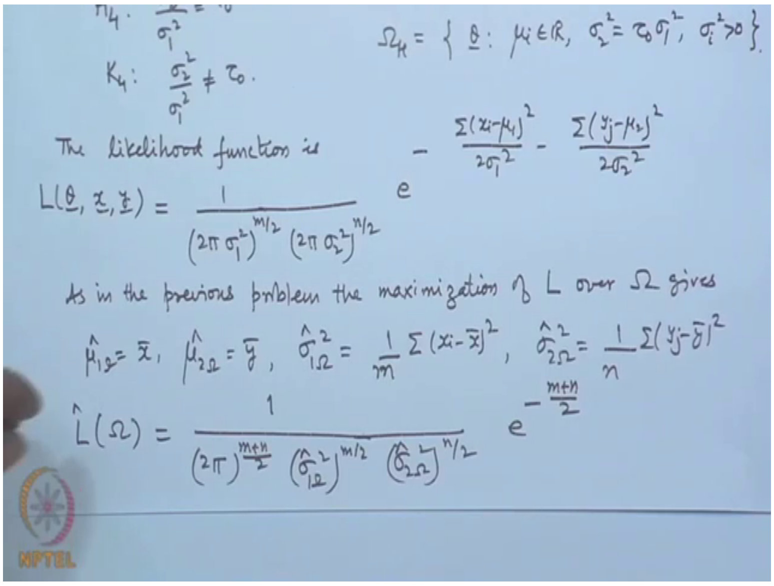
So, we need not write the expressions once again. The likelihood function is given by 1 by 2 pi sigma 1 square to the power m by 2 1 by 2 pi sigma 2 square to the power n by 2 e to the power minus sigma x i minus mu 1 square by 2 sigma 1 square minus sigma y j minus mu 2 square by 2 sigma 2 square. In the previous problem, I considered the maximization of this likelihood function over omega and omega H, now in this case omega H was sigma 2 square less than or equal to tau naught sigma 1 square.

Now, in this new case, so let me repeat it here the log likelihood function the likelihood function and the log likelihood function, which we need to write. So, L theta, x, y that is equal to 1 by 2 pi sigma 1 square to the power m by 2 2 pi sigma 2 square to the power n by 2, e to the power minus sigma x i minus mu 1 square by 2 sigma 1 square minus sigma y j minus mu 2 square by 2 sigma 2 square.

And we are having the omega is equal to mu i belonging to R and sigma i square is greater than 0 for i is equal to 1, 2; and omega H is now theta mu i belonging to R and sigma 2 square is equal to tau naught sigma 1 square and of course, both are positive. Here our theta is mu 1, mu 2, sigma 1 square, and sigma 2 square.

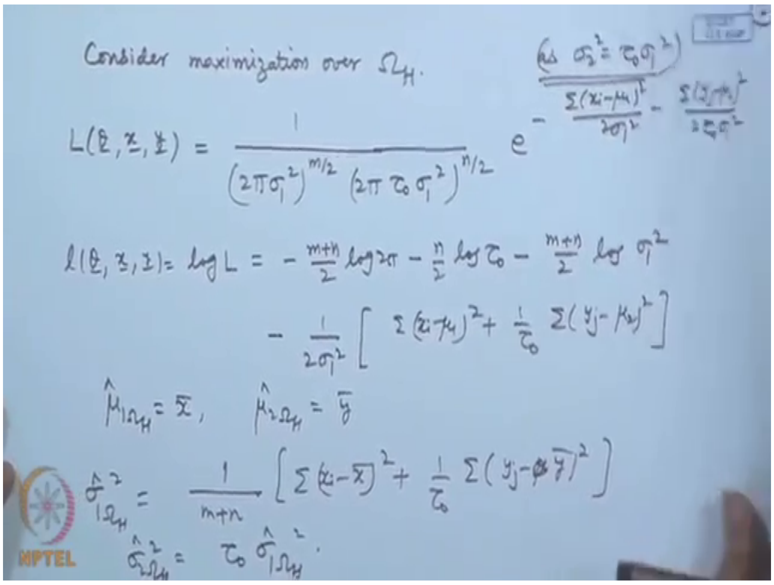
Now, as in the previous problem, the maximization of L over omega gives mu 1 hat omega is equal to x bar, mu 2 omega hat is equal to y bar, sigma 1 omega hat square that is equal to 1 by n, sigma x i minus x bar whole square, this is m here, and sigma 2 omega hat square that is equal to 1 by n sigma y j minus y bar square.

(Refer Slide Time: 04:17)



And as a consequence L hat omega that is given by L hat omega is given by 1 by 2 pi to the power m plus n by 2 sigma 1 omega hat square to the power m by 2 sigma 2 omega hat square to the power n by 2 e to the power minus m plus n by 2.

(Refer Slide Time: 04:59)

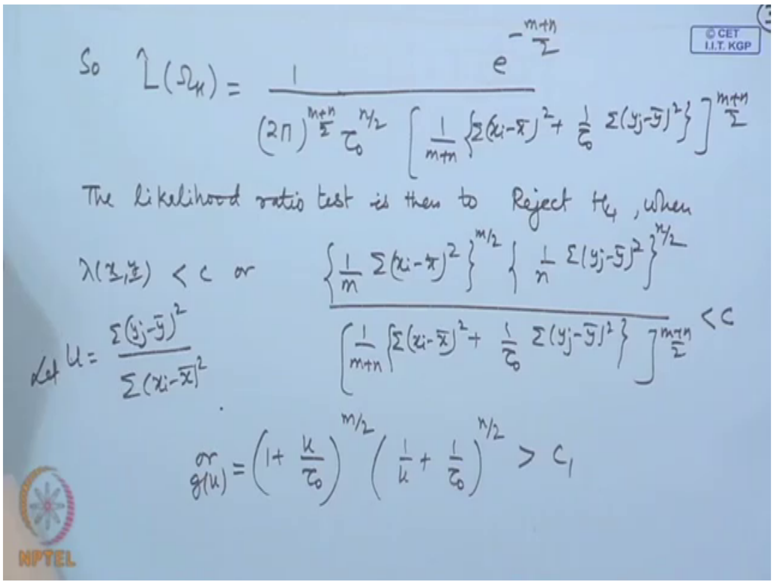


Now, consider maximization over omega H. Now, when we consider over omega H, then the likelihood function gets little bit modified, over omega H L is equal to 1 by 2 pi sigma 1 square to the power m by 2, 2 pi tau naught sigma 1 square to the power n by 2, because we are having sigma 2 square is equal to tau naught sigma 1 square, when we are considering omega H. And then you have e to the power minus sigma by 2 sigma 1 square minus sigma y j minus mu 2 square by 2 tau naught sigma 1 square.

So, of course, if we considered the log likelihood here log of L that is equal to minus m plus n by 2 log of 2 pi minus n by 2 log of tau naught plus minus m plus n by 2 log of sigma 1 square minus 1 by 2 sigma 1 square sigma x i minus mu 1 square plus 1 by tau naught sigma y j minus mu 2 square.

So, if I consider mu 1 hat omega h that will be x bar and mu 2 omega h hat that will be y bar. However, if I consider maximization with respect to sigma 1 hat square, I will get 1 by m plus n sigma x i minus x bar square plus 1 by tau naught sigma y j minus y bar square. So, sigma 2 omega H square that will be equal to tau naught sigma 1 omega H hat square.

(Refer Slide Time: 07:42)



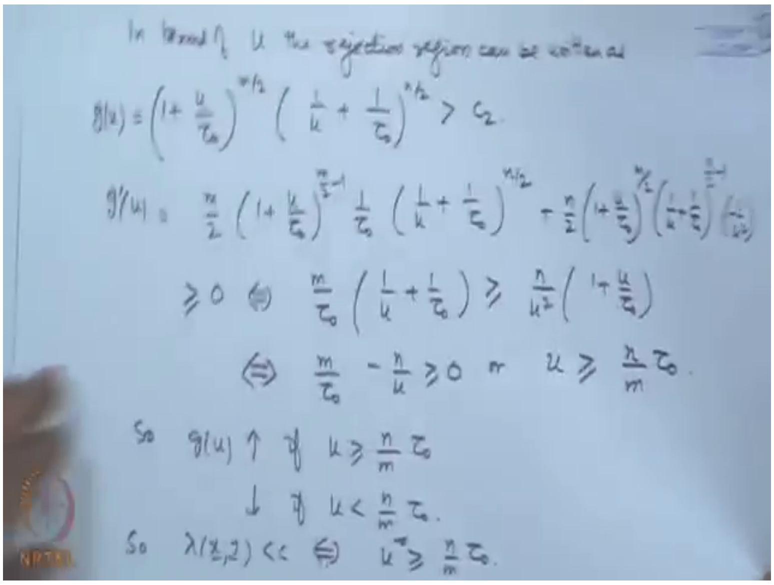
As a consequence if I substitute these values in the log likelihood function, so we get L hat omega H s, I will get this value as 1 by 2 pi to the power m plus n by 2 tau naught to the power n by 2 and then 1 by m plus n sigma x i minus x bar square plus 1 by tau naught sigma y j minus y bar square to the power m plus n by 2 e to the power minus m plus n by 2. So, this is the term that we will be getting.

Now, let us look at both of these terms here. We are having L hat omega as this term and L hat omega H s this term. So, the likelihood ratio test can then be written as the likelihood ratio test is then to reject H 4 when lambda x is less than c, which is equivalent to, now we have already derived the expression.

So, I will just substitute here L hat omega H, I will get this in the denominator, and then divided by L hat omega. So, this will go in the numerator, so that gives me sigma 1 by m sigma x i minus x bar whole square to the power m by 2 1 by n sigma y j minus y bar whole square to the power n by 2 divided by 1 by m plus n sigma x i minus x bar square plus 1 by tau naught sigma y j minus y bar square to the power m plus n by 2 less than c.

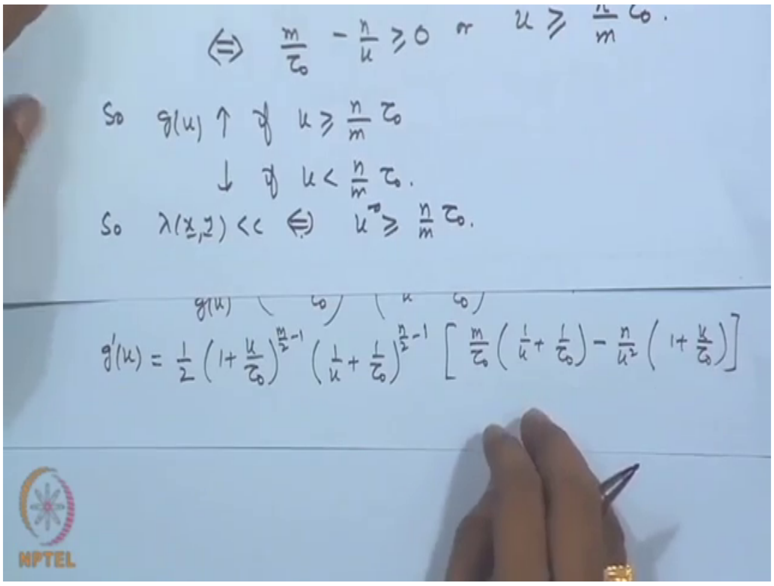
Now, we take reciprocal of this and we use the notation say u is equal to sigma y j minus y bar square divided by sigma x i minus x bar square. If we assume this, then we can rewrite this condition as 1 plus u by tau naught to the power m by 2 1 by u plus 1 by tau naught to the power n by 2 greater than say c 1 ok. Let us assume this to be g u. Now, this function, I considered in the yesterdays derivation of the test function.

(Refer Slide Time: 11:08)



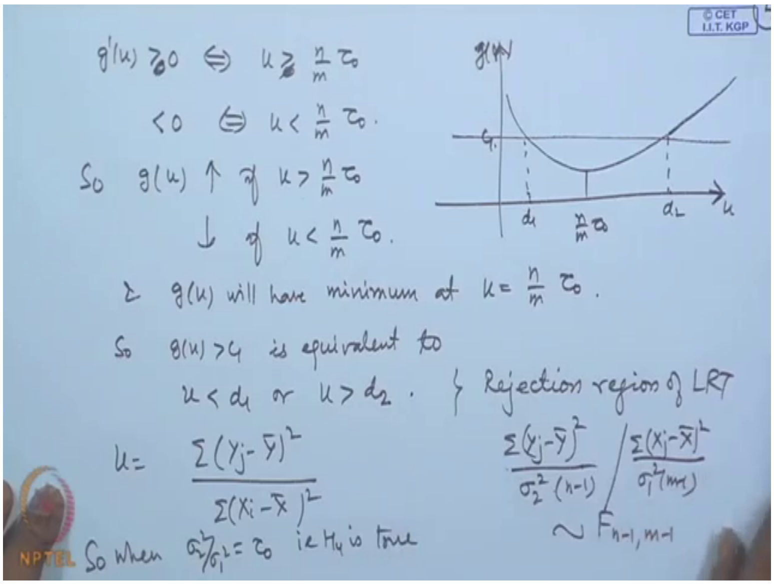
I will show the behavior of it here, I had assumed this term to be g u. Now, this term is same as this term here, and g prime u we had calculated.

(Refer Slide Time: 11:24)



So, this g prime u function then turns out to be so I straight forwardly take the same expression, we obtain g prime u is equal to some term, which is equal to half of 1 plus u by tau naught to the power m by 2 minus 1 1 by u plus 1 by tau naught to the power n by 2 minus 1 m by tau naught into 1 by u plus 1 by tau naught minus n by u square 1 plus u by tau naught.

(Refer Slide Time: 12:13)

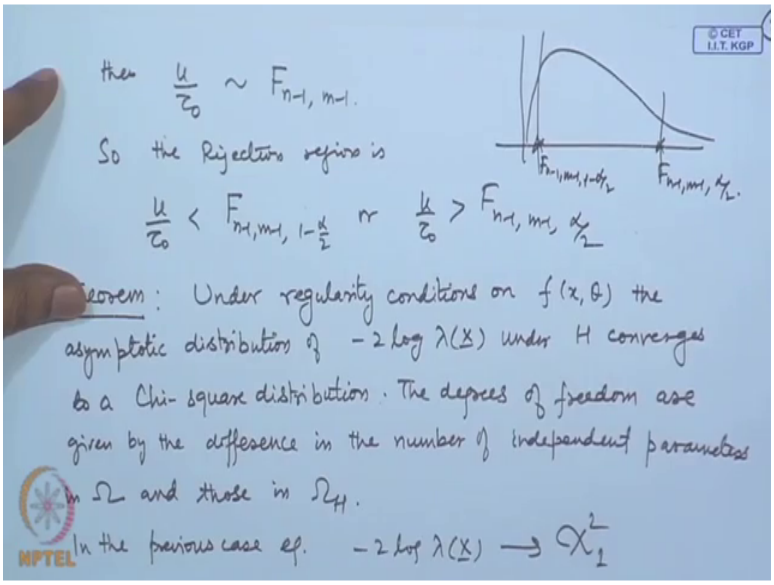


And this we can see as g prime u, it will be greater than or equal to 0 if and only if u is greater than or equal to n by m tau naught; and it will be less than 0 if u is less than n by m tau naught. So, in fact, I can say greater here, greater here. So, g u is increasing if u is greater than n by m tau naught; it is decreasing if u is less than n by m tau naught. And, and g u will have minimum at u is equal to n by m tau naught that means, the nature of this g u function will be something like this. There is a minimum at n by m tau naught. If I am plotting g u function, then it will have something like this.

So, if I say g u is greater than c 1. So, suppose this point is c 1 then this is equivalent to saying that u is either less than certain number or it is bigger than a certain number. So, let me call this number say d 1 and d 2. So, g u greater than c 1 is equivalent to u less than d 1 or u is greater than d 2. So, this is the rejection region of likelihood ratio test. Now, this quantity u that is sigma y j minus y bar square divided by sigma x i minus x bar square.

So, if we considered this, we are having sigma y j minus y bar square divided by sigma 2 square and n minus 1 divided by sigma x j minus x bar square divided by sigma y square m minus 1. This follows f distribution on n minus 1 m minus 1 degrees of freedom. So, when so when sigma 2 square by sigma 1 square is equal to tau naught that is H 4 is true.

(Refer Slide Time: 15:09)



Then u by tau naught that follows f distribution on n minus 1 m minus 1. So, the rejection region is u by tau naught less than F n minus 1 m minus 1 1 minus alpha by 2 r u by tau naught greater than F n minus 1 m minus 1 alpha by 2 of course, this is not necessary I have taken the 2 points symmetric, but F distribution is not symmetric. So, the points may be at different places. So, this is F m minus n minus 1 m minus 1 1 minus alpha by 2 point, and this point will be F n minus 1 m minus 1 alpha by 2. So, this is the likelihood ratio test for the equality of the variances. And you can note here that this test is the same as the u and p and bias test, which I derived for this situation.

So, what we have shown here is that the likelihood ratio tests are applicable for the parameters of the normal distributions. Also they can be applied to some other distributions, where the distribution may not be of the normal type, it could be exponential, gamma, double exponential etcetera. The form of likelihood ratio test is such that it is general.

So, only condition is that you should be able to derive the maximum likelihood estimator for the full parameter is space as well as under the parameter space which is restricted because of the null hypothesis. Once we have that the likelihood ratio test can be written. Now, it is a different matter that whether we can be able to derive the distribution of that are not.

Now, in many cases we may not be able to derive the exact distribution; however, there is a nice asymptotic property of the likelihood ratio test whichever I would like to state here. If you remember for the maximum likelihood estimators, we stated certain conditions which we called regularity conditions. For example, the density should be differentiated differentiable, the parameter space should be a subset of the an open set in the Euclidean space. We had assume that the density function are any expectation of a measurable integrable function should be differentiable under the integral sign. So, these were the regulatory conditions.

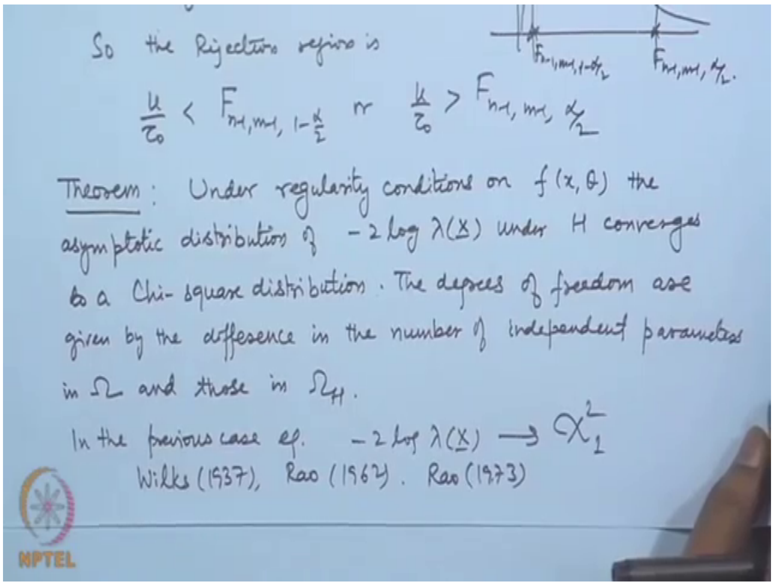
Under those conditions, we had shown that the maximum likelihood estimator exists with probability one. And it is also consistent. And the asymptotic distribution of the maximum likelihood estimator was shown to be normal. Now, under the same regularity conditions, because here we are dealing with the likelihood function, so under the similarly likely regularity conditions the asymptotic distribution of the likelihood ratio test is statistic can also be found.

So, I will state the result without the proof here. Under regularity conditions on f x theta, the asymptotic distribution of minus 2 log of lambda x under H converges to a chi square distribution. The degrees of freedom of the chi square, they are given by the difference in the number of independent parameters in omega and those in omega H. That means, when we are assuming the parameter is space, so we are getting for example, in the previous problem omega had 4 dimension mu 1, mu 2 sigma 1, square sigma 2 square.

Under omega H we had 3 independent parameters, because mu 1, mu 2 are independent, but sigma 1 square and sigma 2 square where related. So, there were three independent parameters, so 4 minus 3, it will become 1. The number of degrees of freedom for chi square of (Refer Time: 20:22). So, basically what will get in the previous case for example, minus 2 log lambda x will be converging to chi square 1 distribution.

Now, this is a very use full thing, because once the convergence is there, then one need not look for the exact distribution all the time. In certain cases, of course, like in the normal distribution cases exact distributions we are able to derive, but many times like in the discrete distributions either we have to deal with the distributions like binomial, Poisson or we have to look at the special tables like we had mentioned about the 2 by 2 contingency tables testing. So, in those cases, we can actually consider this approximation, and usually it is considered to be good.

(Refer Slide Time: 21:19)



More results about these things they are approved by Wilks 1937, Rao he has looked at the asymptotic convergence the rate of convergence of this, and there are many other authors who have actually considered the limiting distributions the book of Rao in 1973 discuss is in detail that Linear Statistical Inference book, they are discuss in detail the asymptotic properties of the likelihood function.

Now, I will consider two advance applications of this likelihood ratio test. So far whether I was discussing the ump test, ump unbiased test most powerful test I have considered, you can say simplistic situations, generally I was dealing with the univariate populations are bivariate populations, but in practice there can be even more complex situations. For example, I may have k-different populations, and I may like to check whether they are means or whether their variances are equal.

Now, in this case certainly the ump tests are ump unbiased tests are very difficult to derive. In fact, we cannot write down the form of the joint density function in a in the form of multi parameter exponential family, so that this parameter, which is to be considered occurs there. And therefore, the likelihood ratio test seems to be a good option only thing is we should be able to derive the maximum likelihood estimators under omega as well as under omega H. So, I will give a couple of examples for applications when we are dealing with multi parameter situations and the number of populations maybe more than two also.