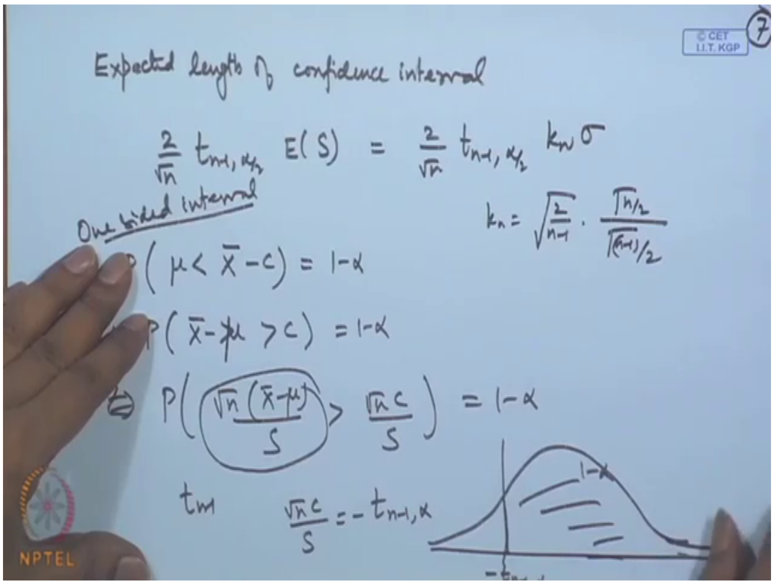
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 62
Interval Estimation – II

(Refer Slide Time: 00:19)

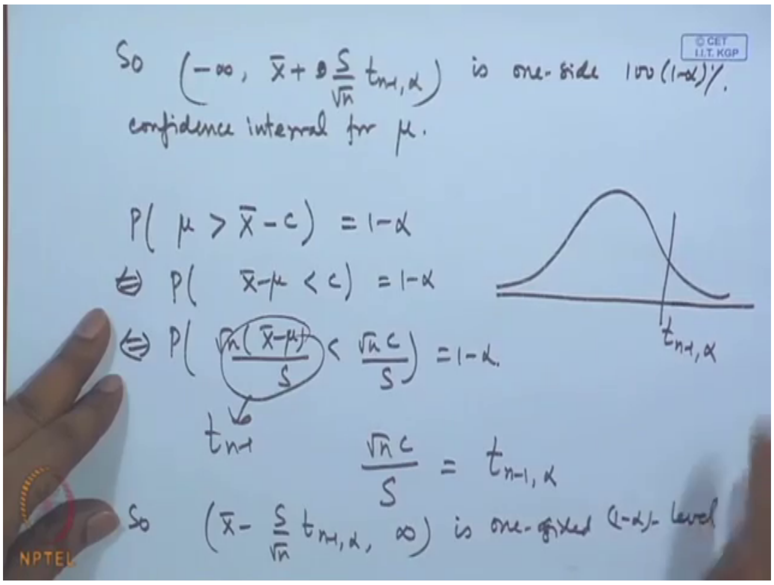


I may also look at expected length. Expected length of confidence interval. That is equal to 2 by root in tn minus 1 alpha by 2 expectation of S. If you remember our lecture on the point estimation when we are considering the minimum variance and bias estimation of sigma and normal distribution then expectation of S was some coefficient times sigma. So, 2 by root n t n minus 1 alpha by 2. Some coefficient I will call it k n times sigma where of course, k n was given by root 2 by n minus 1 gamma n by 2 divided by gamma n minus 1 by 2.

So, this coefficient of course, you can see again it will go to 0 as n tends infinity. So, that is satisfied here. Let me consider here one sided interval also. Because in both the cases I have given 2 sided, but we may sometimes require one sided interval. If we want one sided so, for example, you may say X bar minus c that is equal to 1 minus alpha for one sided interval; that means, what we are saying probability of X bar greater than X bar minus mu greater than C is equal to 1 minus alpha. Or probability of root n X bar minus mu by S greater than root n C by S is equal to 1 minus alpha.

Now, this is having t distribution on n minus 1 degrees of freedom. So, what we are looking at is the point on the t distribution curve beyond which the probabilities 1 minus alpha. So, we may chose root n c by S is equal to minus t n minus 1 basically this is alpha this is minus t n minus 1 alpha beyond which this probability is 1 minus alpha.

(Refer Slide Time: 02:53)

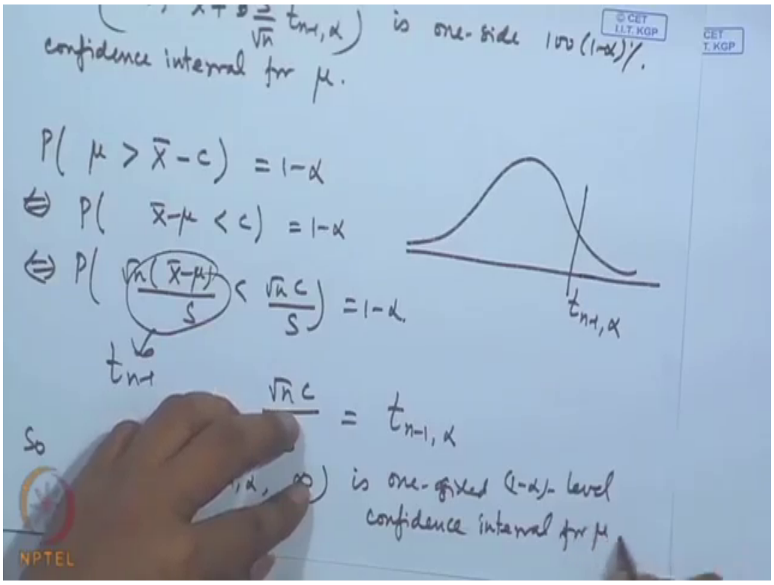


So, the interval is minus infinity to X bar plus S by root n t n minus 1 alpha. This is one sided 100 1 minus alpha percent confidence interval.

Here we have given an upper bound here. We may also take a lower bound. You may say what is the probability of mu greater than say X bar minus C equal to 1 minus alpha. If we see from this point of you then it will become probability of X bar minus mu less than C that is equal to 1 minus alpha are probability of root n X bar minus mu by S less than root n c by S that is equal to 1 minus alpha.

So, this is nothing but a t distribution on n minus 1 degrees of freedom. So, what we are saying is a probability below a certain point is 1 minus alpha. So, this value can be chosen to be t root n C by S can be chosen to be t n minus 1 alpha. So, we are getting X bar minus S by root in t n minus 1 alpha to infinity. This is one sided 1 minus alpha level confidence interval.

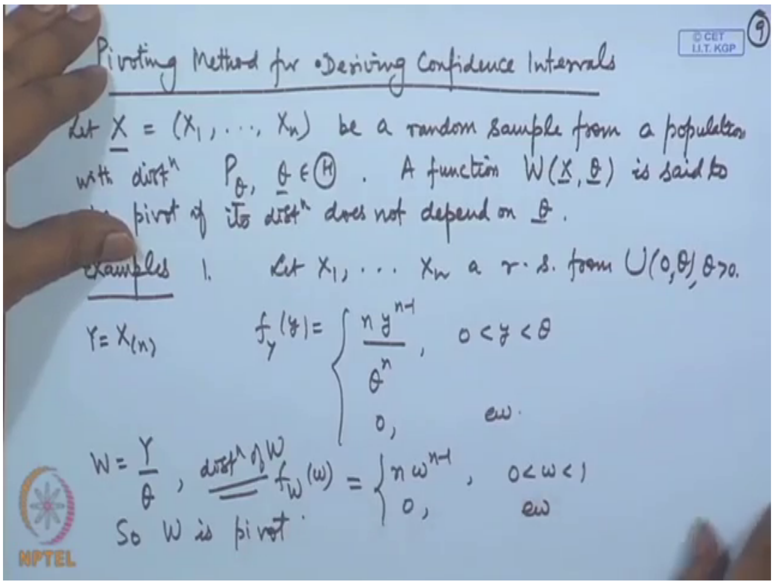
(Refer Slide Time: 04:48)



So, here we have given a lower bound. In this one we are having an upper bound here.

So, this only goes on to show that what we are doing here is that we are considering a quantity based on which it is easy to derive a confidence interval, but that brings us to a question that how to choose this quantity. Like in the case of normal distribution we have chosen X bar, but in other distributions what we will do. So, that is an important concept it is called the method of pivoting.

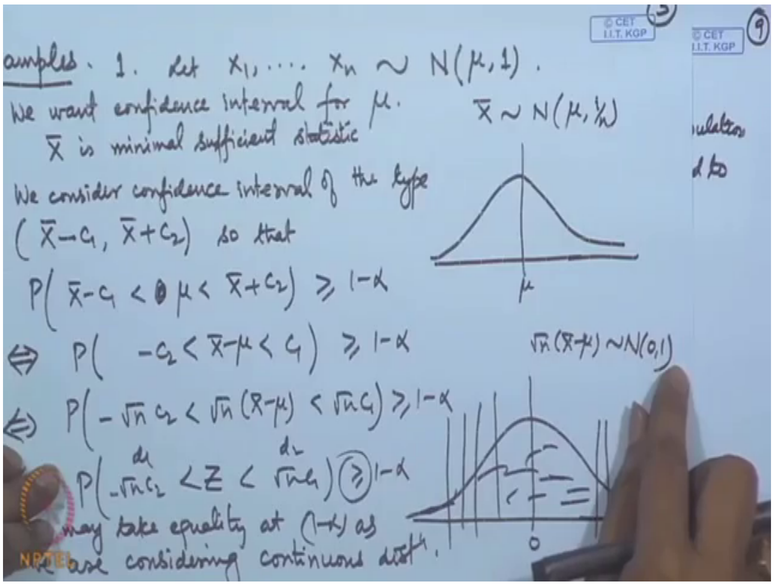
(Refer Slide Time: 05:36)



So, let me discuss that now pivoting method for deriving confidence intervals. So, as before we are having X is equal to X 1 X 2 Xn.

This is a random sample from a population with distribution say P theta in general we may have vector of parameters. Then a function say W of X theta this is said to be pivot if it is distribution does not depend on theta. So, in the previous problem let us consider root n X bar minus mu by S this is having a distribution which is not dependent on mu r sigma square. So, this is a pivot. In this exercise root nx bar minus mu has a distribution which is independent of mu. So, this is a pivot.

(Refer Slide Time: 07:05)

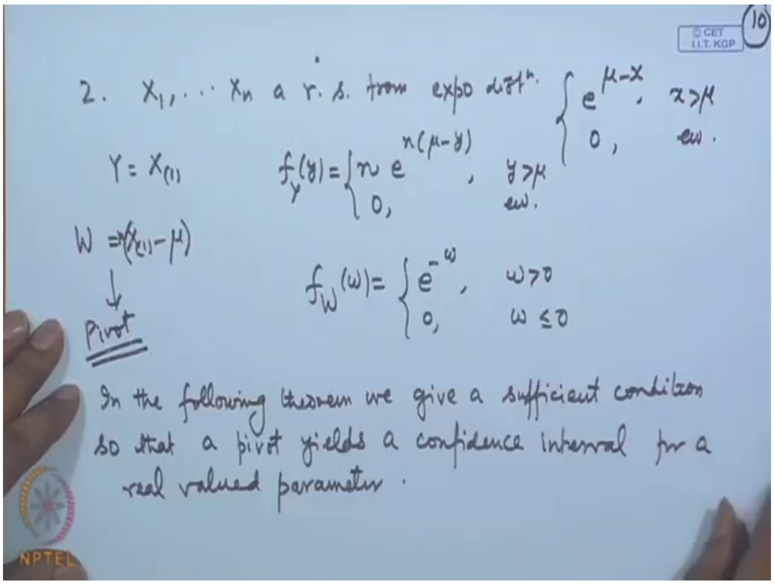


So, in the pivot method we consider a quantity which is having a distribution independent of the parameter. At the same time the pivot must include the random variables as well as the parameter. In such a way that by resolving the interval or you can say by resolving the inequalities we are able to get an interval, which is free from; that means, in between we get the parameter or one side we get the parameter and on the other side we get the quantity which is dependent upon the variable.

So, let me just one or two more examples of the pivot quantity. Let us consider say X 1 X 2 Xn a random sample from uniform 0 theta distribution. Now consider say Xn let me call Y you know the distribution of y that is n y to the power n minus 1 by theta to the power n, it is 0 elsewhere. In this if I define say W is equal to Y divided by theta. Then what is the distribution of W? Distribution of W and this is equal to n W to the power n minus 1 for 0 less than W less than one and it is 0 elsewhere. So, this distribution is free from the parameters. And this W involves the random variable Xn as well as the parameter which is coming here.

So, this is. So, W is pivot in this problem. Let me just give one more example here.

(Refer Slide Time: 09:43)



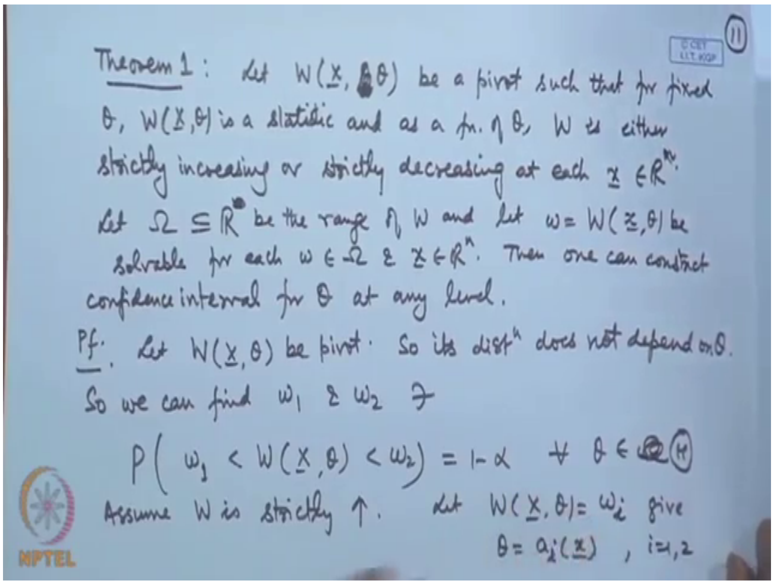
Let us consider say X 1 X 2 Xn a random sample from exponential distribution with density function of the form say e to the power mu minus X let us consider say X 1 what is the distribution of X 1 let me call it Y then the distribution of y that can be calculated to be n e to the power n mu minus y.

So, in that case if I consider X 1 minus mu then what is the distribution of W or I can consider n times that then what is the distribution of W? That is equal to e to the power minus W. That is simply the exponential distribution.

So, this W is a pivot and again the interesting point is that.

This involves a random variable and the parameter which is unknown parameter of the distribution. Now in the following theorem, we give a sufficient condition. So, that a pivot yields a confidence interval for a real valued parameter.

(Refer Slide Time: 12:00)

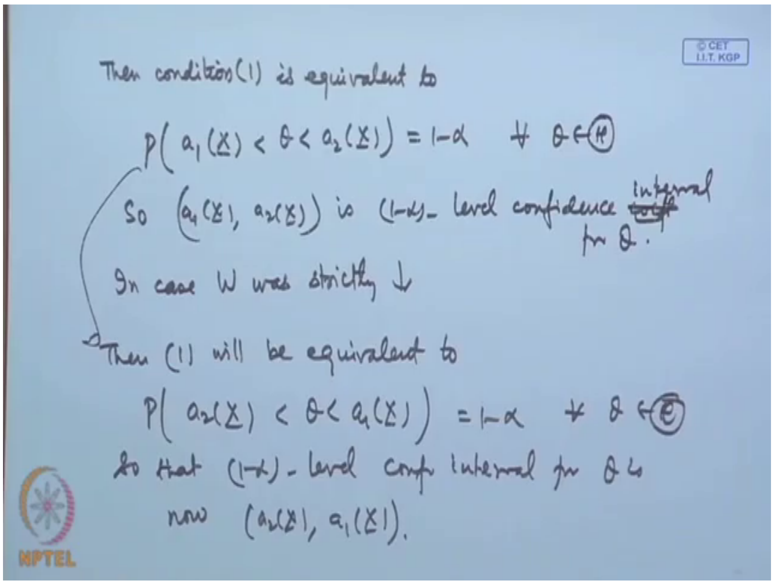


The theorem is as follows. Let me call theorem 1. Let W X theta so, I am taking one dimensional case for parameters while right just theta here, be a pivot such that for fixed theta W X theta is a statistic. And as a function of theta W is either a strictly increasing or a strictly decreasing at each x. Let omega that is a subset of R n be the range of now I am assuming this to be real. So, let this be the range of W and let the equation say small w is equal to capital W of X theta be solvable for each omega and each x then one can construct the confidence interval for theta at any level. Let t W of X theta be pivot. And so, it is distribution does not depend on theta.

If it does not depend upon theta. So, we can find say omega 1 and omega 2 such that probability of omega 1 less than W X theta less than omega 2 that is equal to 1 minus alpha. And this statement now will be true for all theta because the parameter does not depend upon the distribution of W does not depend upon the parameter. See this is I think script theta here the range of the parameter here. Now assume that W is a strictly increasing and let W X theta is equal to omega 1, omega i give theta is equal to say some coefficient a 1 of x for a i of x i is equal to 1 2.

Because we are assuming that this can be solved. If it can be solved then this condition say one then this condition 1.

(Refer Slide Time: 16:28)

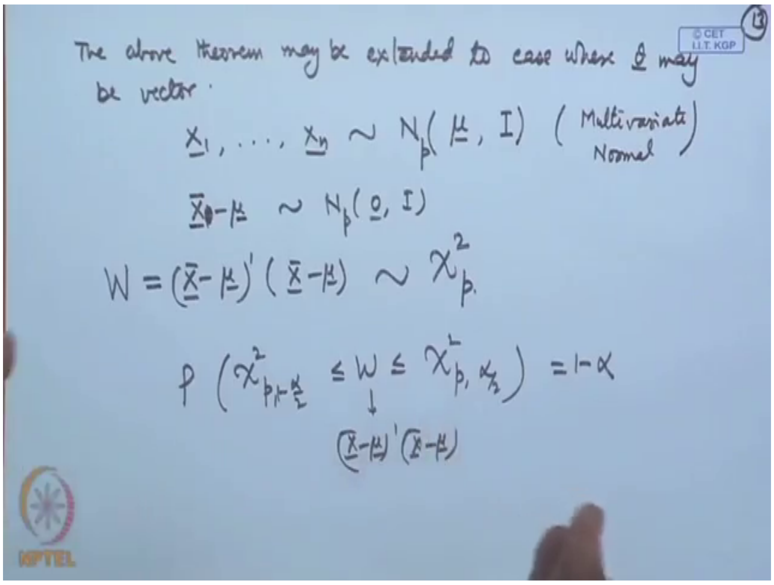


This is equivalent to probability of a 1 X less than theta less than a 2 X is equal to 1 minus alpha. So, a 1 X to a 2 X is 1 minus alpha level confidence; confidence interval for theta. In case W was decreasing suppose it is strictly decreasing then this will change. Then one will be equivalent to probability of a 2 X less than theta less than a 1 X equal to 1 minus alpha.

So, that 1 minus alpha level confidence interval for theta is now a 2 X 2 a 1 X. Now I have considered the simplistic situation where I have put here equality as 1 minus alpha; however, it may happen that the distribution of W is a discrete distribution. In that case we may not be able to achieve 1 minus alpha. We may say it is greater than or equal to 1 minus alpha. So, a value which is slightly higher than 1 minus alpha maybe at achieve, but in that case we will have that as the confidence level. In the case of continuous W any level can be achieved here. Another point when I mentioned this result here, I am considering the solvability and I am considering theta as a scalar.

Suppose theta as a vector. Even in that case it may be possible to write down a statistic which involves the parameter which is free from the which as a distribution free from the parameters. However, in that case if the parameter theta is not a scalar, in that case we do not say confidence interval rather we say a confidence band or a confidence set in general that we are talking about. I will just consider one case when we may deal with multi parameter situations.

(Refer Slide Time: 19:31)

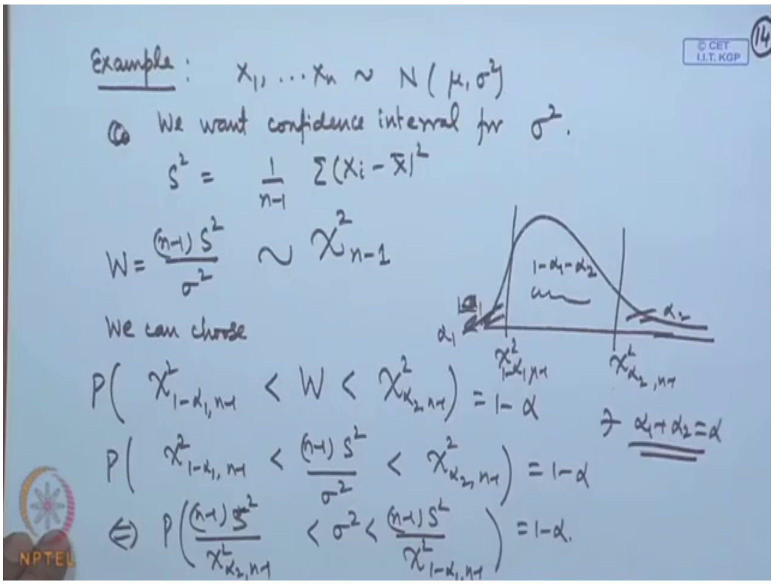


The above theorem may be extended to case where theta maybe vector.

For example, I say X 1 X 2 Xn this follows a multivariate normal with mean vector mu and variance covariance matrix and this is p dimensional. So, this is multivariate normal; however, even this case we may consider thing like xi minus mu. Or we may considered X bar minus mu. That will follow normal 0 i n p and if I consider say X bar minus mu prime into X bar minus mu. That will follow chi square distribution on p degree of freedom this may be treated as W. And in that case what we can write here is W this is chi square.

So, we may write say chi square p 1 minus alpha by 2 less than or equal to W less than or equal to chi square p alpha by 2 probability of this is equal to alpha 1 minus alpha. And then now this W is nothing but X bar minus mu prime X bar minus mu. So, we get a confidence band here. Rather than because if we write down this thing as an equation in mus this is denoting an ellipsoid. So, actually it is concentration ellipsoid that we will be getting here. I will complete this part by giving confidence interval for the variance in a normal distribution by using this method of pivoting.

(Refer Slide Time: 21:53)

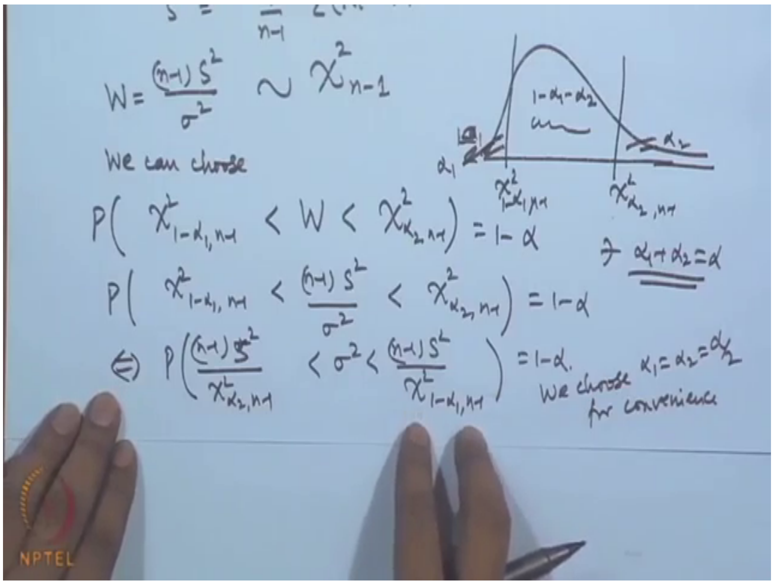


Let us consider say X 1 X 2 Xn following normal mu sigma square. And we want the confidence interval for say sigma square. If I consider this S square as 1 by n minus 1 sigma X i minus X bar whole square, then we may take n minus 1 S square by sigma square as the pivot quantity this as a distribution which does not depend upon the parameter. This is simply chi square distribution on n minus 1 degrees of freedom.

So, this is of course, a skewed distribution. Now we may choose in general 2 points here say chi square alpha 1 chi square alpha 2 on n minus 1 degrees of freedom. So this probability is alpha 2 this probability is 1 minus alpha 1 say in between probability is actually so, we may put 1 minus alpha on say here. So, this probability is alpha 1. So, this middle probabilities actually 1 minus alpha 1 minus alpha 2. So, we can chose probability of chi square 1 minus alpha 1 n minus 1 less than W less than chi square alpha 2 n minus 1 is equal to 1 minus alpha; that means, what we are saying alpha 1 plus alpha 2 is equal to alpha here.

This will give me chi square 1 minus alpha 1 n minus 1 less than n minus 1 S square by sigma square less than chi square alpha 2 n minus 1 it is equal to 1 minus alpha; which is statement we can write as sigma square less than n minus 1 S square by chi square 1 minus alpha 1 n minus 1. And greater than sigma square greater than n minus 1 S square divided by chi square alpha 2 n minus 1 that is equal to 1 minus alpha.

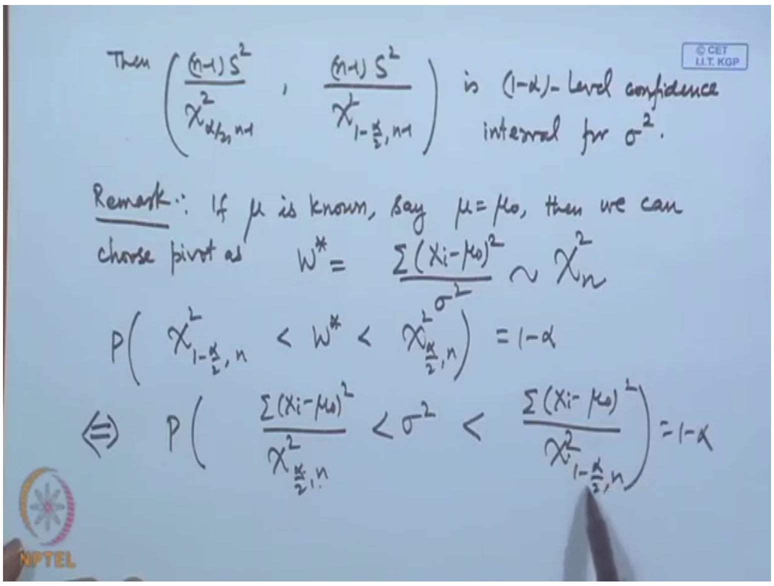
(Refer Slide Time: 24:45)



For convenience we may choose say alpha 1 is equal to alpha 2 is equal to alpha i 2 for convenience. What I am saying here is that we may get actually several intervals which will have the same confidence level 1 minus alpha, but for practical purpose one may fix up so, that we can come to a unique solution.

So, we can choose alpha 1 alpha 2 is equal to alpha by 2.

(Refer Slide Time: 25:17)



In that case then what we are saying is n minus 1 S square by chi square alpha by 2 n minus to n minus 1 S square by chi square 1 minus alpha by 2 n minus 1. This is 1 minus alpha level confidence interval for sigma square. Now see we have considered this pivot quantity when mu is unknown here.

But if mu is known, if mu is known in this problem , say mu is equal to mu naught in that case we can choose pivot as let me call it W star. Sigma X i minus mu naught square by sigma square. Now this is as a chi square distribution on n degrees of freedom. Now if we give the same argument by writing down chi square say 1 minus alpha by 2 n less than W star less than chi square alpha by 2 n that is equal to 1 minus alpha. Then this is equivalent to saying probability of sigma X i minus mu naught square by chi square alpha by 2 n to sigma xi minus mu naught square by chi square 1 minus alpha by 2 n that will be equal to 1 minus alpha.

So, sigma X i minus mu naught square by chi square alpha by 2 n to sigma X I minus mu naught square by chi square alpha by 2 n to sigma xi minus mu naught square by chi square 1 minus alpha this will be the 1 minus alpha level confidence interval for sigma square. So, when mu is known we may use this, but when mu is unknown then suddenly we cannot use this and we make use of this as an interval.

Now, when I was driving this I considered the level of in a confidence level. So, there is a connection between the level of significance in testing problem to the confidence level in a confidence interval. So, I will state a result about this in the next class. We establish a connection between these and derive confidence intervals for various problems. For example, parameters of a normal distribution, parameters of 2 normal distributions binomial proportions etcetera so, that I will be covering in the next lecture.