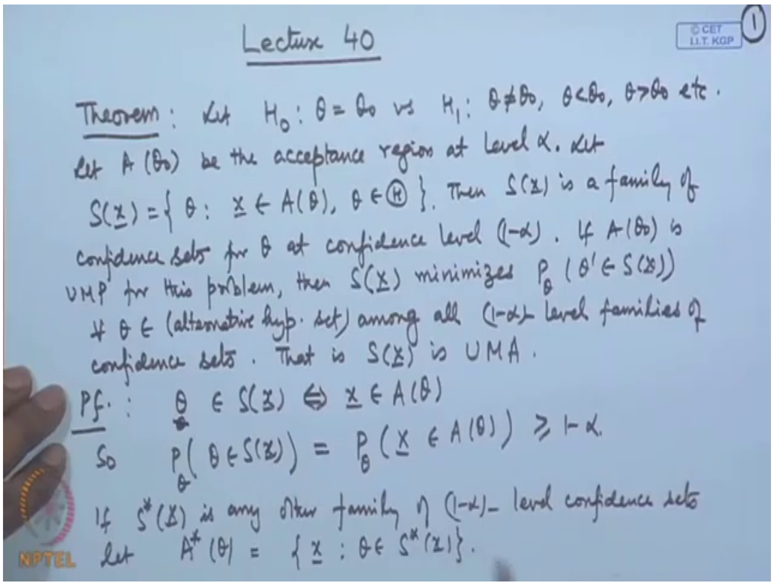
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 63
Interval Estimation – III

In the previous lecture, I have introduced the concept of Interval Estimation confidence intervals with a certain confidence level. We also discussed a method that is the method of pivoting for obtaining confidence intervals for particular confidence levels. Now what I want to mention is that the theory of uniformly most powerful test or the uniformly most powerful unbiased test is intimately related with the theory of uniformly most accurate confidence intervals or uniformly most accurate unbiased intervals.

(Refer Slide Time: 01:08)



So, we restrict attention to the non ms tests and then we have the following result. So, let us consider say hypothesis testing problem say theta is equal to theta naught against some alterative actually that alternative I am not writing it specifically it could be say theta naught is equal to theta naught theta less than theta naught or theta greater than theta naught etcetera and any of the alternatives can be there ok; that means, based on any of the alternatives are rejection region or you can say the alternative hypothesis set will be.

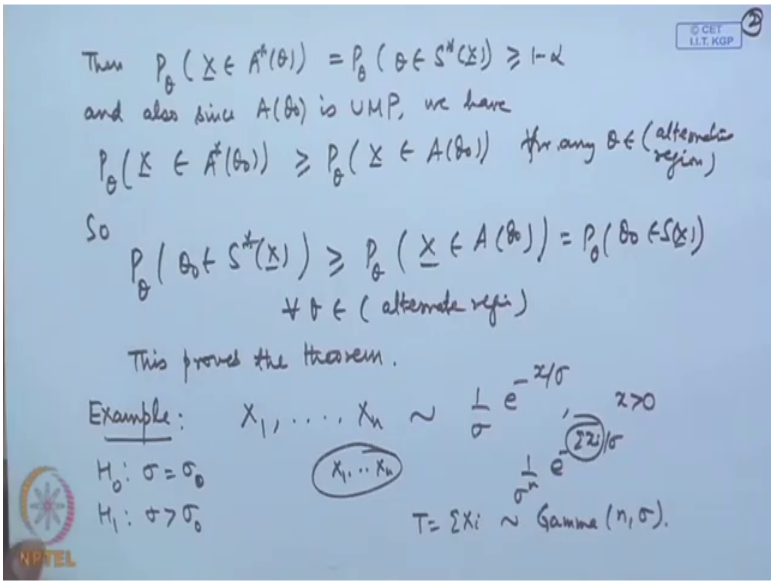
For example, here in the acceptance region you have theta naught in the rejection region you may have all the points except theta naught or you may have all the points below theta naught or you may have all the points above theta naught. Let us consider say A theta naught as the acceptance region; acceptance region at level alpha and let us consider say S x to be the set of parameters theta such that x belongs to A theta.

Then S x is a family of confidence sets for theta at confidence level 1 minus alpha; that means, if I am having critical region or you can say a level alpha test and if I consider the acceptance region for that, then from there I can derive 1 minus alpha level confidence set ok. Now moreover, if A theta naught is uniformly most powerful for this problem then S x minimizes P theta; theta prime belonging to S x for all theta belonging to alternative hypothesis set. As I mentioned here it could be all the thetas other than theta naught it could be thetas less than theta naught or could be theta greater than theta naught etcetera, among all 1 minus alpha level. That is S x is uniformly most accurate confidence region.

Let me look at the proof of this of course, it is very simple because we are just following the definition see if I say theta belongs to S x then this is equivalent to saying that x belongs to A theta. So, the probability of theta belonging to S x that is S x including theta is same as probability of x belonging to A theta greater than or equal to 1 minus alpha.

So, what does it mean? It means that if A theta is acceptance region of a alpha level test then S x will be a family of confidence sets of 1 minus alpha level that is one thing. Now next let us to prove that uniformly most accurate. So, if S star x is any other family of 1 minus alpha level confidence sets, let us define say A star theta that is equal to x such that theta belongs to S star x.

(Refer Slide Time: 05:58)

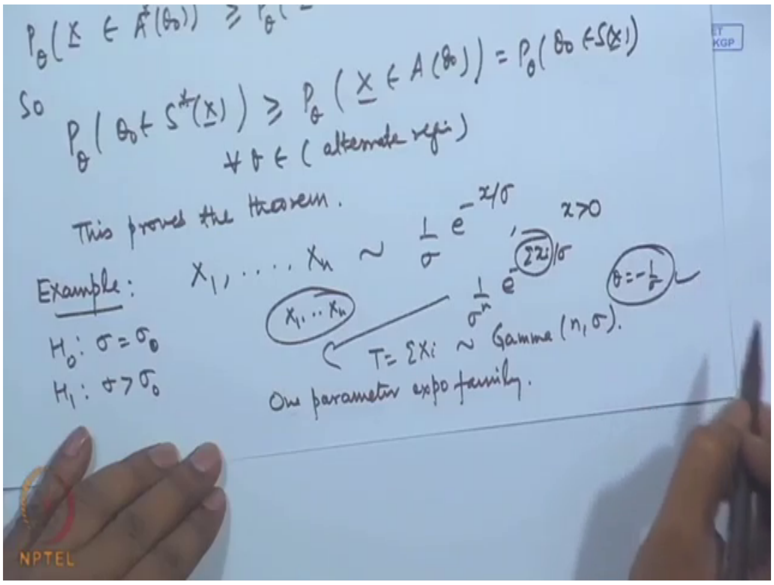


Then P theta x belonging to A star theta that is equal to P theta; theta belonging to S star X that is greater than or equal to 1 minus alpha and also since A theta naught is UMP, we have that P theta X belonging to A star theta naught is greater than or equal to P theta X belonging to A theta naught, for any theta in the alternative region; region of the alternative hypothesis.

So, probability of theta naught belonging to S star X is greater than or equal to probability of X belonging to A theta naught that is P theta theta naught belonging to S X, for all theta belonging to alternate region. So, this proves that the S X family is the family of uniformly most accurate confidence regions or confidence sets.

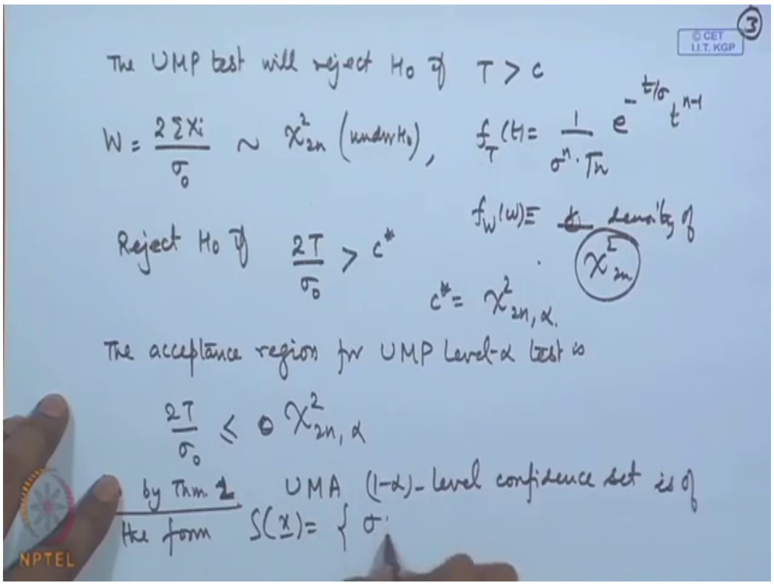
Let me consider one example here say X 1, X 2, X n this follows 1 by sigma e to the power minus x by sigma and let us consider the testing problem sigma is equal to sigma naught against sigma is greater than sigma naught. Now, in this problem we have derived see this is 1 parameter exponential family, if I consider the joint distribution of X 1, X 2, X n, m the join distribution that is 1 by sigma to the power n e to the power minus sigma x i by sigma. So, this is the one parameter exponential family T is equal to sigma X i let us call it T is equal to sigma X i and that follows gamma distribution with parameters n and sigma ok.

(Refer Slide Time: 08:38)



This is one parameter exponential family therefore and also you are having minus 1 by sigma which is it is quickly increasing in sigma. So, theta is equal to minus 1 by sigma therefore, the test will exist, the test will be actually rejecting in favor of the larger value of T.

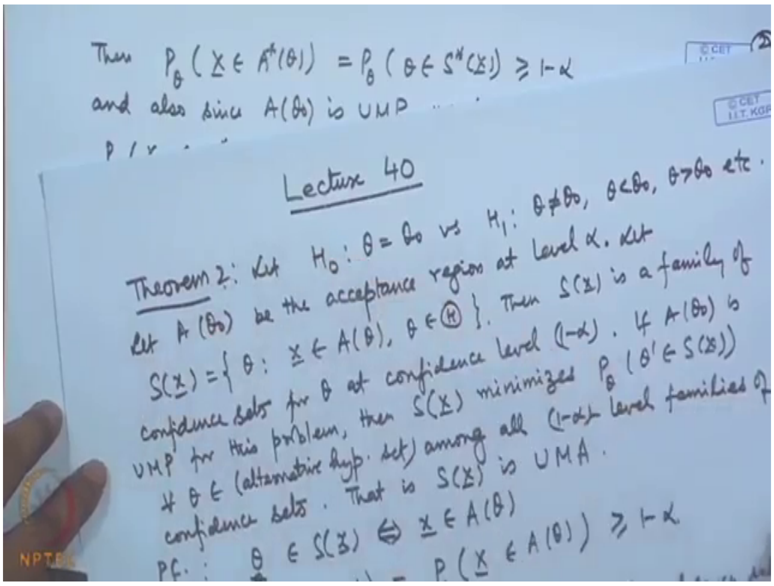
(Refer Slide Time: 09:16)



So, the uniformly most powerful test will reject H naught if T is greater than some c and then we determine this thing actually you look at the distribution of twice sigma X i by sigma naught, let me call it W. So, that will follow a chi square distribution on 2 n degrees of freedom because what is the distribution of T? The distribution of T is 1 by sigma to the power n gamma n e to the power minus t by sigma t to the power n minus 1.

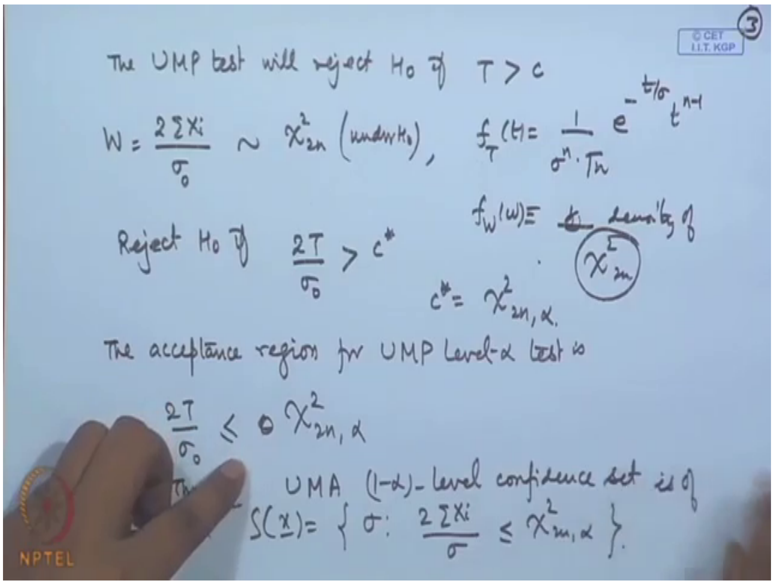
So, if I consider this is under H naught, say if i consider 2 T by sigma naught then the distribution of f w that will come equal to 1 by gamma same thing. So, that will be the density of chi square distribution on 2 n degree of freedom. So, we can write rejection reject H naught if 2 T by sigma naught is greater than say c star where c star is nothing, but chi square 2 n alpha ok. So, the acceptance region for UMP level alpha test is 2 T by sigma naught less than or equal to chi square 2 n alpha.

(Refer Slide Time: 11:36)

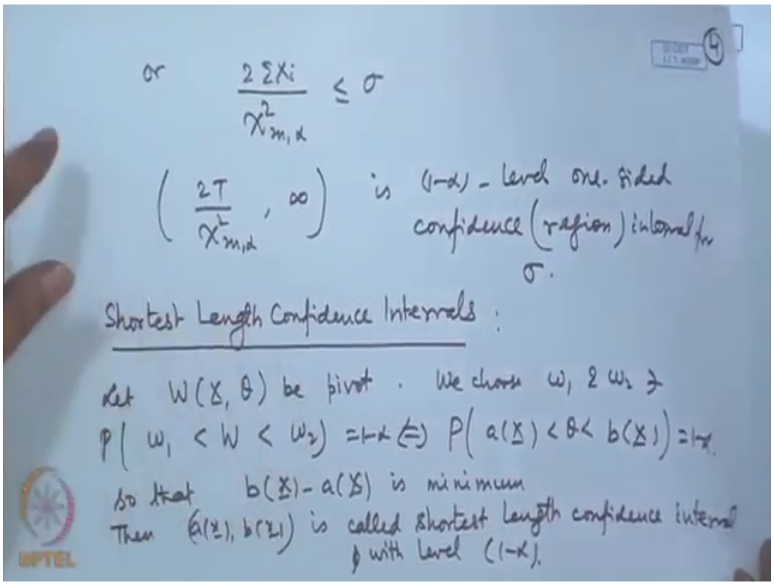


So, by the previous theorem which we call theorem 2 say, we can call this as theorem 2, uniformly more accurate 1 minus alpha level confidence set is of the form sigma such that, twice sigma Xi by sigma is less than or equal to chi square 2 n alpha.

(Refer Slide Time: 12:00)



(Refer Slide Time: 12:12)



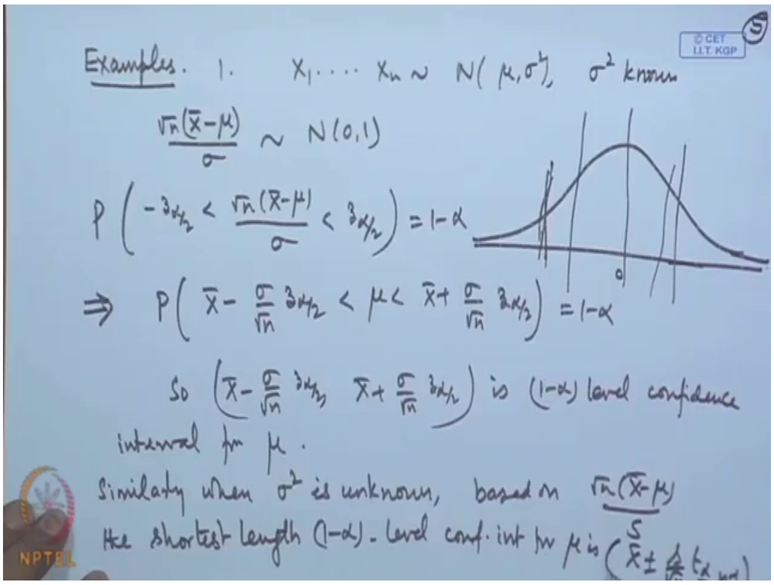
So, you can simplify this condition here twice sigma X i by chi square 2 n alpha less than or equal to sigma. That is we are saying twice T by chi square 2 n alpha to infinity, this is 1 minus alpha level one sided confidence interval actually for sigma. So, the theorem the theory of UMP test is intimately connected with the theory of finding out uniformly most accurate test. So, particularly when we are dealing with the one parameter exponential family etcetera, then straight forwardly we can translate the acceptance region into a confidence interval by replacing that parameter theta naught or theta whatever we are saying by a general parameter theory.

Now, also I talk about the shortest length confidence intervals. So, let W X theta be pivot and we are choosing say w 1 and w 2, such that this is equivalent to some interval. So, that the length of the interval b X minus a X is minimum, so this is our interest. For a fixed confidence level we would like to have the shortest length, as I mentioned if I am giving a 2 statements.

For example, I am looking at the rate of the cure by a certain medicine. So, we say we are 95 percent sure or we can say we are 90 percent confident that the rate of cure by this medicine is at least 75 percent and second statement is we are 95 percent confident that the rate of cure is more than 80 percent. Then which statement will be better? Obviously, the second statement is better because it is giving a shorter interval.

So, the problem of determining shortest length intervals is of interest. So, we say this is the, then a x to b x this is called shortest length confidence interval for confidence interval with level 1 minus alpha of course, it is based on there is W, if we change the W we may get another interval here.

(Refer Slide Time: 16:00)

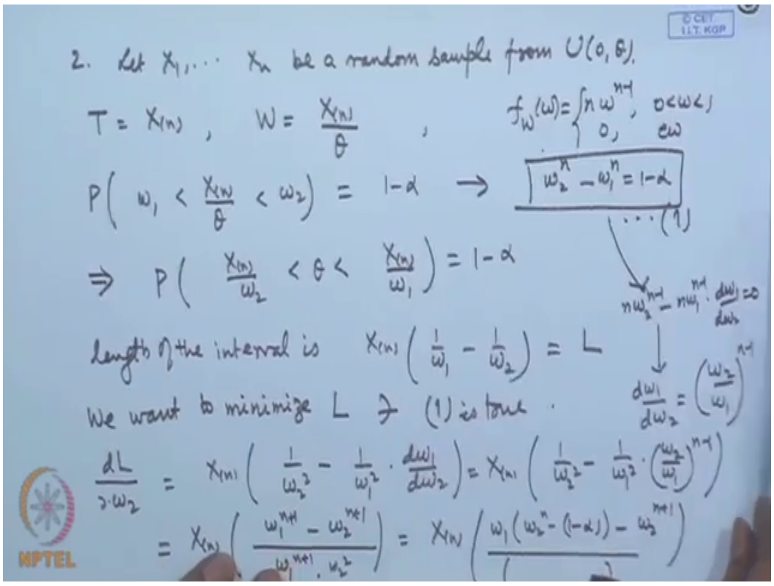


So, let us look at the examples here. So, if we are considering X 1, X 2, X n following normal mu sigma is square; sigma square is known. Then based on X bar minus mu root n by sigma this follows normal 0, 1, since the normal distribution is symmetric if I consider is symmetric interval around 0 this will have the highest probability.

If I consider a non symmetric interval, then the length will increase therefore, the shortest interval will be s symmetric interval about this. So, that if you use this logic then we will get minus z alpha by 2 less than root n X bar minus mu by sigma less than z alpha by 2 that is equal to 1 minus alpha which gives the interval X bar minus sigma by root n z alpha by 2 to X bar plus sigma by root n z alpha by 2, that is X bar minus sigma by root n z alpha by 2 to X bar plus sigma by root n z alpha by 2 as 1 minus alpha level confidence interval for mu.

Similarly when we are considering sigma square is unknown in that case based on square root n X bar minus mu by S, the shortest length 1 minus alpha level confidence interval for mu is X bar plus minus s by root n t alpha by 2 n minus 1.

(Refer Slide Time: 18:43)



And let me solve it as an optimization problem also I will explain this through one example, where let X 1, X 2, X n be a random sample from uniform 0 theta interval. So, we consider say T is equal to X n and here we can use the pivot quantity as X n by theta, we have actually seen the distribution of this is n w to the power n minus 1. So, we may consider an interval of the nature say w 1 less than X n by theta less than w 2 is equal to 1 minus alpha. So, this implies probability of this is equal to 1 minus alpha.

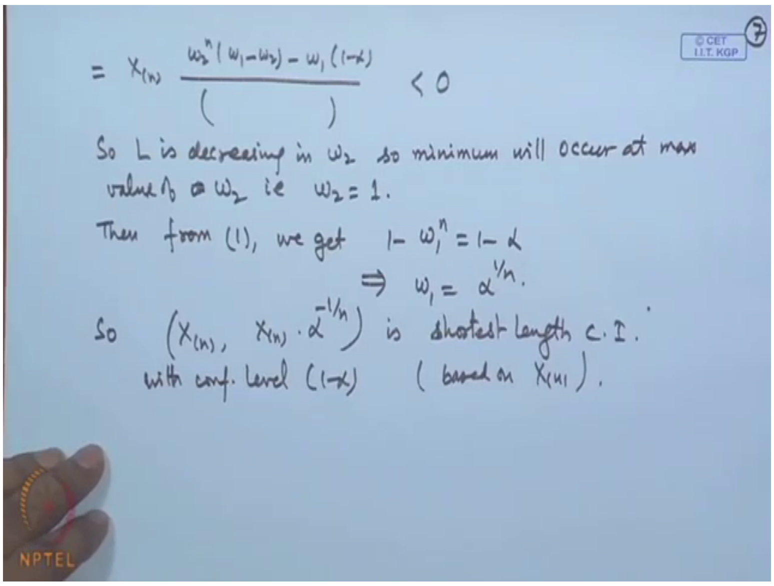
So, what is the length of the interval? Length of the interval is X n into 1 by omega 1 by minus 1 by omega 2, let me write it as equal to say L and at the same time this condition will imply that w 2 to the power n minus w 1 to the power n is equal to 1 minus alpha, let me call this condition as 1. So, we want to minimize L subject to the condition 1 ok.

Now let us consider differentiation of this with respect to say omega 2, then that will give me X n 1 by omega 2 square minus 1 by omega 1 square d omega 1 by d omega 2, but what is d omega 1 by d omega 2? Let us consider also differentiation of this condition. If I differentiate this I get n omega 3 to the power n minus 1 minus n omega 1 to the power n minus 1 d omega 1 by d omega 2 is equal to 0.

But this condition will give me d omega 1 by d omega 2 is equal to omega 2 by omega 1 to the power n minus 1 if i substitute it here I get this as equal to X n 1 by omega 2 square minus 1 by omega 1 square omega 2 by omega 1 to the power n minus 1. So, I can simplify this and I can write it as X n into well omega 1 to the power n plus 1 minus omega to the power n plus 1 divided by omega 1 to the power n plus 1 omega 2 square.

Now, also if we see from this condition I can consider it as X n and of course, this denominator term will come here the numerator term I can substitute here see from here the value of omega 1 to the power n I can put it in terms of omega 2. So, this becomes omega 2 to the power n minus 1 minus alpha minus omega 2 to the power n plus 1.

(Refer Slide Time: 22:49)

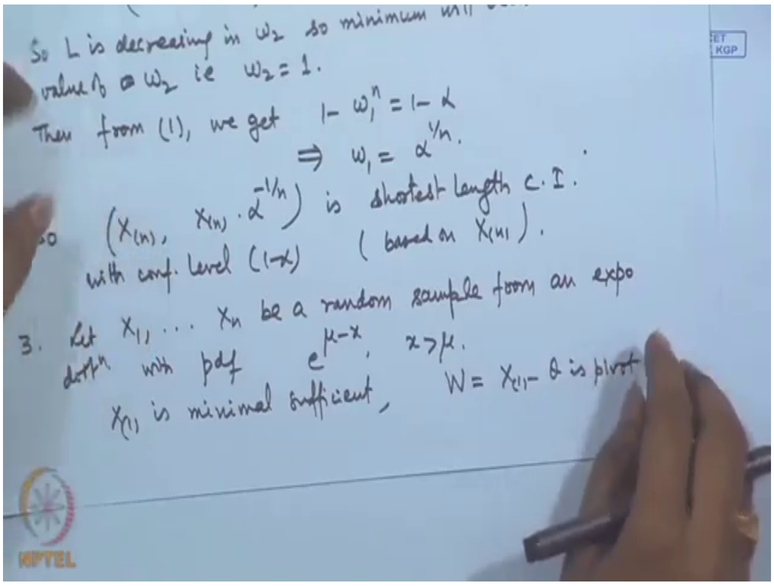


This I can write it as X n and then w omega 2 to the power n omega 1 minus omega 2 minus omega 1 into 1 minus alpha divided by some term this is clearly negative because omega 1 is less than omega 2. So, what it means, so L is actually decreasing in omega 2, so minimum will occur at maximum value of omega 2. Now what are the ranges of omega 1, omega 2? Because here I have taken this and the range of this distribution is 0 to 1. So, the maximum value of omega 2 can be equal to 1 only, so that is omega 2 is equal to 1.

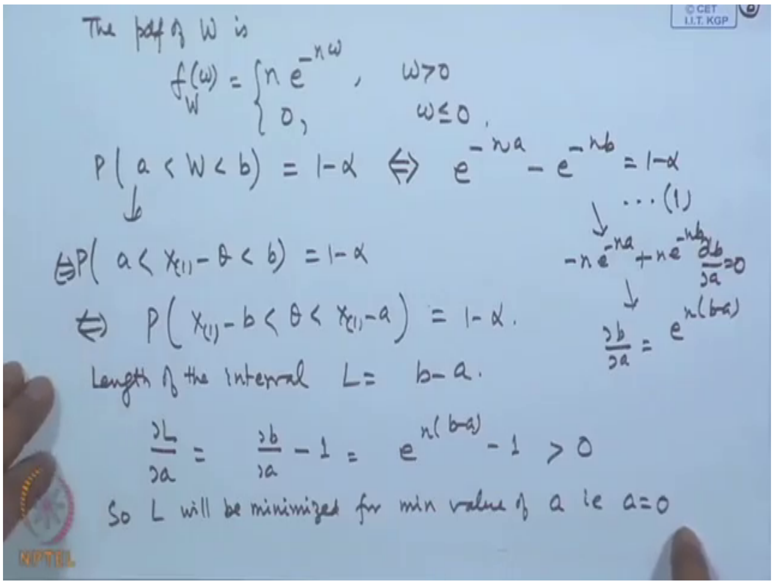
If I have omega 2 is equal to 1, then this will give me from 1 we get 1 minus omega 1 to the power n is equal to 1 minus alpha which gives me omega 1 is equal to alpha to the power 1 by n. So, X n to X n divided by are be multiplied by alpha to the power minus 1 by n, this is shortest length confidence interval with confidence level 1 minus alpha this is based on X n. So, here I have shown through direct optimization that this is the shortest length interval here I have actually minimize the length of the interval here.

Let me give one more example here of an exponential distribution. Let X 1, X 2, X n be a random sample from an exponential distribution with pdf e to the power say mu minus X bar X greater than mu. Now, this example we have considered here if I consider X 1 this is minimal sufficient here and we can consider the pivot quantity W is equal to X 1 minus theta, this can be taken as pivot.

(Refer Slide Time: 26:06)



(Refer Slide Time: 26:13)

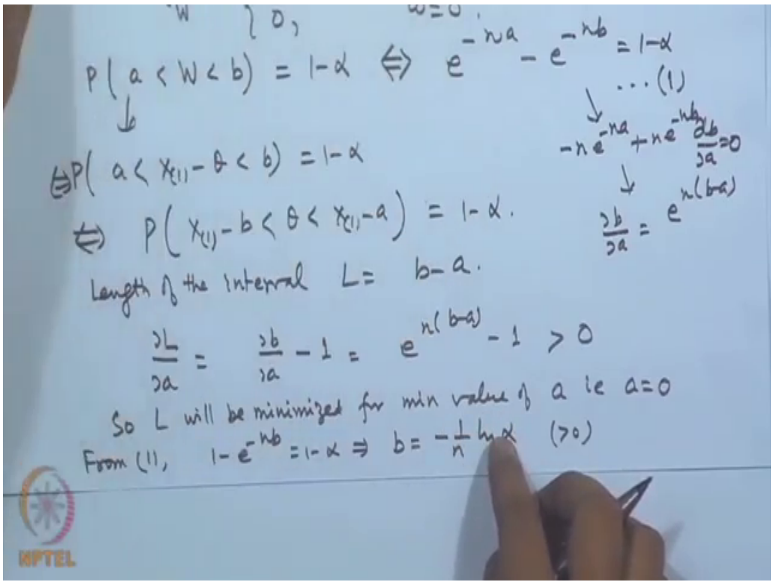


The distribution of w will be the distribution of W is f W is equal to n e to the power minus n W where W is positive it is 0. So, we take probability of say to points a less than W to less than b it is equal to 1 minus alpha. Now in this distribution if I integrate out this condition is equal lent to same e to the power minus n a minus e to the power minus n b is equal to 1 minus alpha this is condition 1.

Now based on this when we are considering, the confidence interval this is becoming X 1 minus theta less than b is equal to 1 minus alpha which is equivalent to probability of theta less than X 1 minus a and it is greater than X 1 minus b that is equal to 1 minus alpha. So, here length of the interval b minus a and we want to minimize this, suppose I want to minimize is respect to a here, then I will get d b minus del a minus 1.

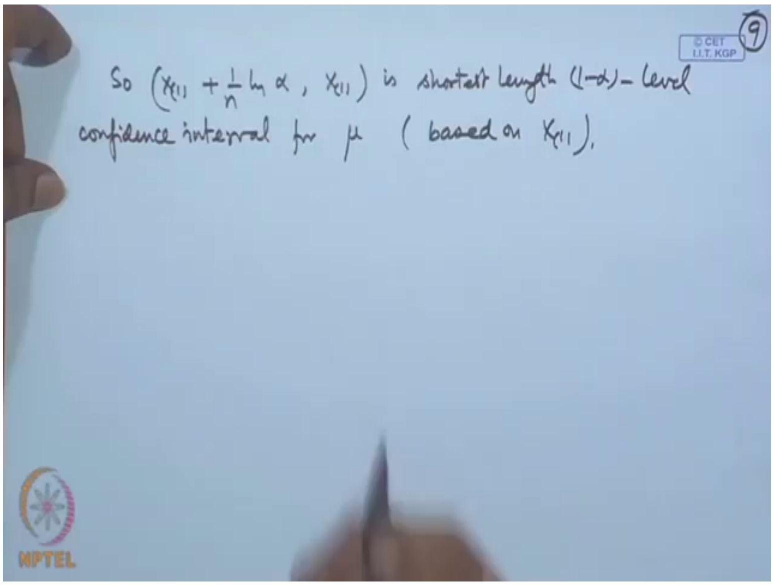
Now if I differentiate this relation I will get minus n times e to the power minus n a plus n times e to the power minus nb del b minus del a equal to 0. Now from here del b by del a that will turn out to be e to the power n times b minus a. So, this will be equal to e to the power n b minus a minus 1; obviously, since b is greater than a this is positive, so this is greater than 1; that means, this is greater than 0. So, L will be minimized for minimum value of a, that is a is equal to 0 because I am considering interval from a to b for W and the range of the W is greater than 0. So, the minimum value of a that can be simply 0 itself.

(Refer Slide Time: 29:00)



If I put a equal to 0 in this one, this will give me 1 minus e to the power minus n b is equal to 1 minus alpha; that means, b is equal to minus 1 by n log of alpha of course, this is positive because alpha is a number between 0 to 1, so this is positive.

(Refer Slide Time: 29:25)



So, the interval that we will get here X 1 plus 1 by n log of alpha to X 1, this is shortest length 1 minus alpha level confidence interval for mu this is based on X 1.

So, here I have shown that directly by considering the minimization of the length of the interval subject to the condition that we are achieving a certain confidence level it is possible to derive the shortest length confidence interval.