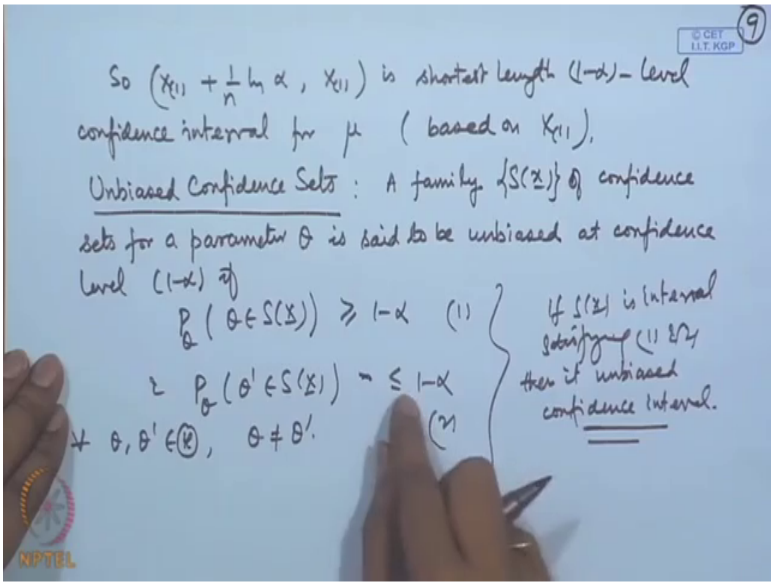
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 64
Interval Estimation – IV

Now I will also consider the situations where in the testing problem we had considered where UMP test does not exist therefore, we were trying to find out UMP unbiased tester also. Now corresponding to that, here we have UMA unbiased confidence intervals. So, what is an unbiased confidence interval let me define that.

(Refer Slide Time: 00:47)

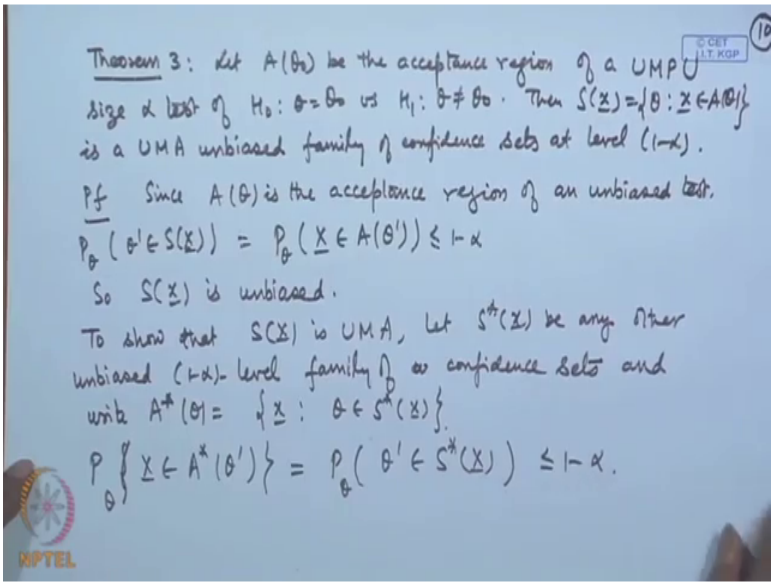


A family S x of confidence sets for a parameter theta is said to be unbiased at confidence level 1 minus alpha if probability of theta belonging to S x is greater than or equal to 1 minus alpha and P theta, theta prime belonging to S x is less than or equal to 1 minus alpha for all theta and theta prime where of course, theta is not equal to theta prime.

So, so this is a 1 minus alpha level confidence set and in case S x is an interval, then this will be called if S x is interval satisfying this conditions 1 and 2, then it is unbiased confidence interval. So, what is the actually interpretation of this? What we are saying here is that the if the true value is theta then the true value is included in the confidence set with probability at least 1 minus alpha and if theta is the true value then theta prime is a false value then that value is included with probability less than or equal to 1 minus alpha; that means, this confidence interval will actually cover the true parameter value with a higher probability than a false parameter value; that means, the value which is true should have a higher chance of getting included in the interval rather than the and the false value that is the wrong value should have less chance of getting included.

So, this is an actual condition of unbiasedness as in the testing problem you remember unbiased test was that beta phi theta is less than or equal to alpha for theta belonging to null parameter is space and for alternative, it was greater than or equal to alpha. So, similarly here it is based on 1 minus alpha. Now there is a direct relation with the UMPU test that is the Uniformly Most Powerful Unbiased test with UMAU that is Uniformly Most Accurate Unbiased test I will state this result here.

(Refer Slide Time: 04:09)



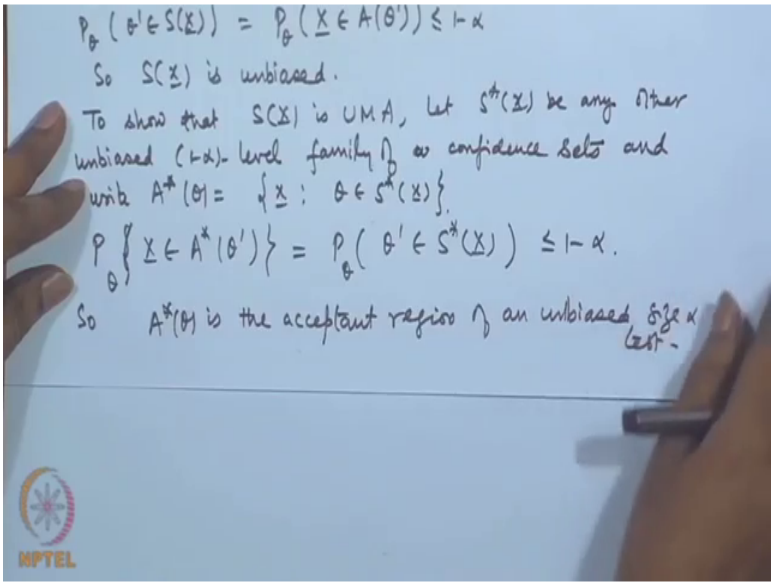
So, these results etcetera they are from Rohatgi and Salehs book they statements of the theorems and also the discussion and proofs of those theorems. Let A theta naught be the acceptance region of a UMPU size alpha test of H naught theta is equal to theta naught against say H 1 theta is not equal to theta naught then S x is equal to theta such that x belongs to A theta this is a UMA unbiased family of confidence sets at level 1 minus alpha.

Let us look at the proof of this. Since A theta is the acceptance region of an unbiased test P theta, theta prime belonging to S x that is equal to P theta x belonging to a theta prime is less than or equal to 1 minus alpha that is the probability of rejecting when h naught is true is the same as this that is equal to less than or equal to 1 minus alpha. So, S x is unbiased.

To show that it is UMA, let us consider S star x be any other unbiased 1 minus alpha level family of confidence sets and we write here A star theta is equal to x such that theta belonging to S star x; that means, the corresponding acceptance regions are denoted by A star theta.

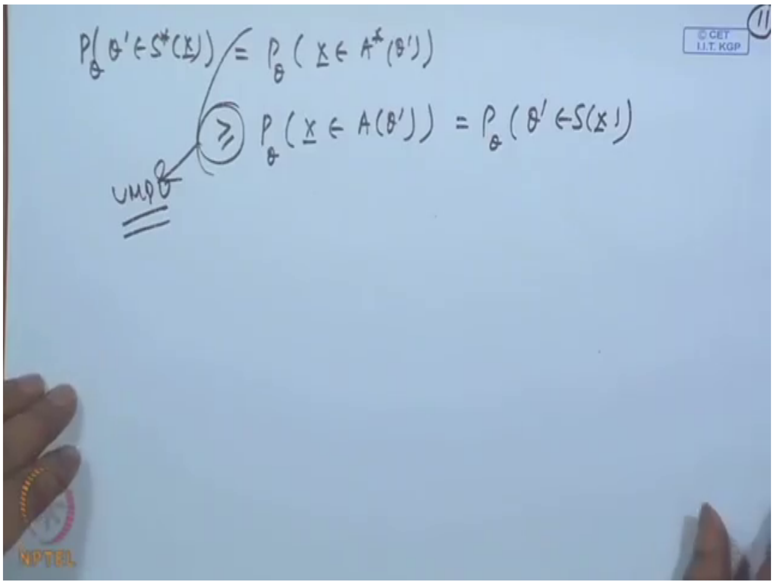
So, probability of x belonging to A star theta prime that will be probability of theta prime belonging to S star x that is less than or equal to 1 minus alpha.

(Refer Slide Time: 07:25)



So, we can say that, A star theta is the acceptance region of an unbiased size alpha test.

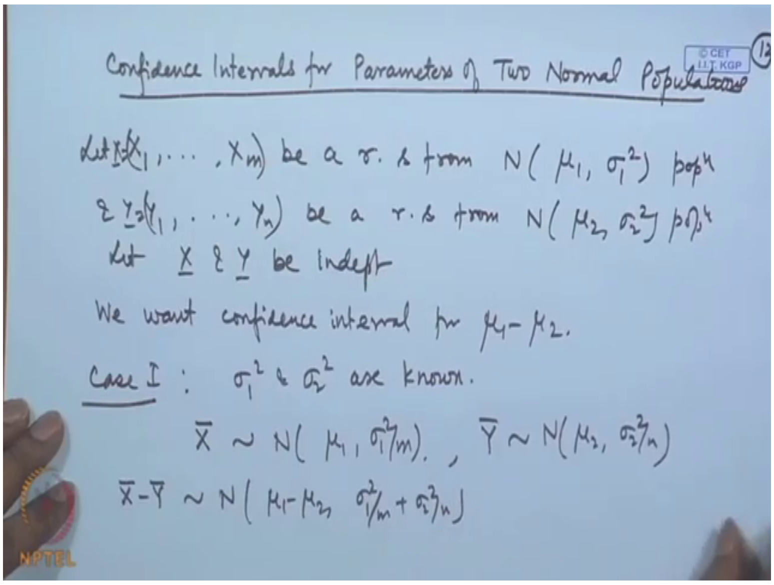
(Refer Slide Time: 07:47)



Therefore we can write the statement that probability of theta prime belonging to S star x is equal to x belonging to A star theta prime that is greater than or equal to P theta x belonging to A theta prime that is equal to P theta, theta prime belonging to S x.

So, this inequality is there because of UMPU test. So, this proves that this is the uniformly most accurate region here. Now I will consider certain confidence intervals for parameters of normal populations. I have already given the confidence intervals for the mean of a normal population both when the variance is known or unknown and also for sigma S square when the mean may be known or unknown.

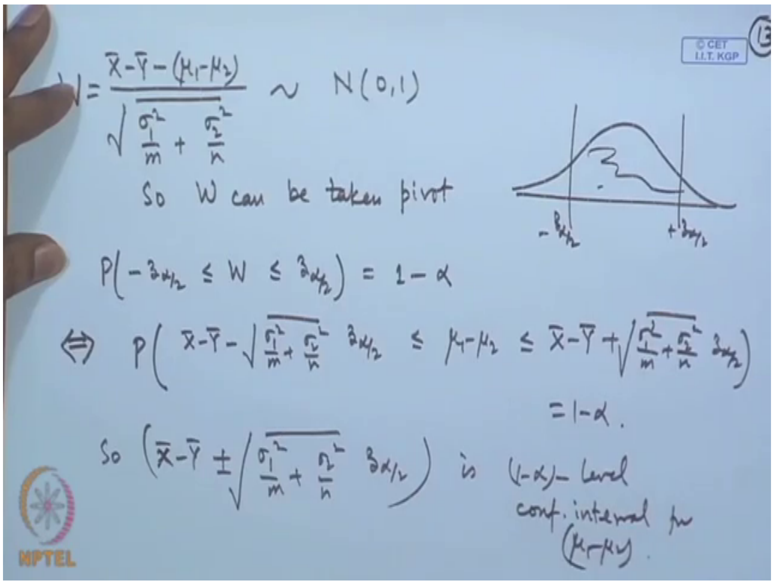
(Refer Slide Time: 08:59)



Now let us consider two sample problems confidence intervals for parameters of two normal populations. Let us consider say X 1, X 2, X m be a random sample from say a normal mu 1 sigma 1 square population and Y 1, Y 2, Y n be a random sample from say normal mu 2 sigma 2 square population and assume that these two samples are independent let me call it X sample and this is Y sample. Let X and Y be independent.

We are interested in finding out confidence interval for say mu 1 minus mu 2. Let us take the first case when sigma 1 square and sigma 2 square are known. If they are known we can formulate nicely the pivot quantity, the distribution of X bar is normal mu 1 sigma 1 square by m, the distribution of Y bar is normal mu 2 sigma 2 square by n.

(Refer Slide Time: 11:17)

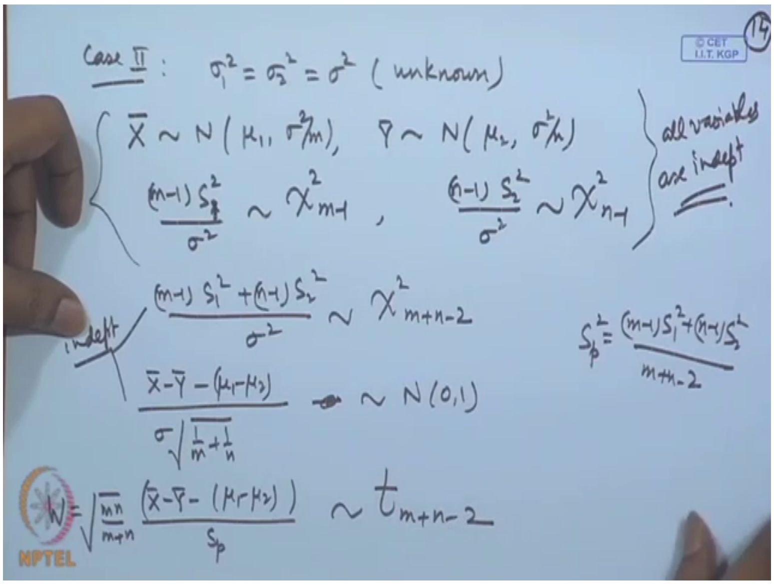


So, the distribution of X bar minus Y bar is normal mu 1 minus mu 2 sigma 1 square by m plus sigma 2 square by n and therefore, X bar minus Y bar minus mu 1 minus mu 2 divided by square root of sigma 1 square by m plus sigma 2 square by n, this has a distribution free from the parameters. So, this can be considered as the pivot quantity. So, W can be taken as pivot because this involves mu 1 minus mu 2 and random variables and the distribution is free from the parameters. So, if I consider the shortest length confidence interval based on this, then I should consider z alpha by 2 and minus z alpha by 2 points.

So, probability of minus z alpha by 2 less than or equal to W less than or equal to plus z alpha by 2 that is equal to 1 minus alpha this is equivalent to the statement X bar minus Y bar minus square root sigma 1 square by m plus sigma 2 square by n z alpha by 2 less than or equal to mu 1 minus mu 2 less than or equal to X bar minus Y bar plus square root sigma 1 square by m plus sigma 2 square by n z alpha by 2 that is equal to 1 minus alpha here.

So, X bar minus Y bar plus minus square root sigma 1 square by m plus sigma 2 square by n z alpha by 2 this is 1 minus alpha level confidence interval for mu 1 minus mu 2, but if sigma 1 square and sigma 2 square are unknown we cannot use this. In that case we consider another pivot quantity sigma 1 square and sigma 2 square is equal to sigma square, but unknown.

(Refer Slide Time: 13:35)

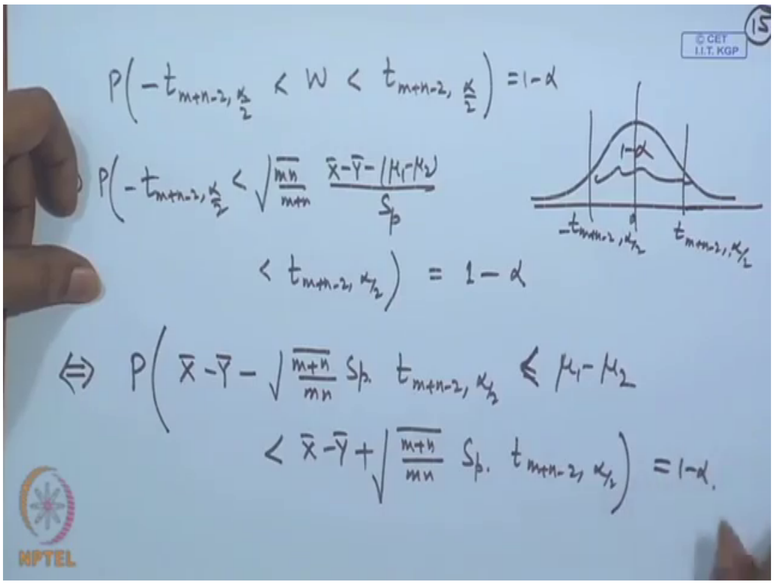


Now this statements are true, that X bar follows normal mu 1 sigma square by m and Y bar follows normal mu 2 sigma square by n, but at the same time let us also consider say m minus 1 S x square or S 1 square by sigma square that follows chi square distribution on m minus 1 degrees of freedom and n minus 1 S 2 square by sigma square this follows chi square distribution on n minus 1 degrees of freedom where S 1 square denotes 1 by m minus 1 sigma x i minus x bar whole square and S 2 square denotes 1 by n minus 1 sigma y j minus y bar whole square that is the two sample variances and these are all independent random variables all these variables are independently distributed.

If they are independently distributed, I can consider say m minus 1 S 1 square plus n minus 1 S 2 square by sigma square that follows chi square distribution on m plus n minus 2 degrees of freedom and therefore, I can write down X bar minus Y bar minus mu 1 minus mu 2 divided by sigma root 1 by m plus 1 by n this follows normal 0 1 and these two are also independent and let me define S p square as m minus 1 S 1 square plus n minus 1 S 2 square divided by m plus n minus 2, then X bar minus Y bar minus mu 1 minus mu 2 divided by S p square root m n by m plus n this will follow t distribution on m plus n minus 2 degrees of freedom.

So, this can be used as a pivot W because this involves certain random variable which is of interest to us because we wanted the confidence interval are mu 1 minus mu 2 random variable is involved here and the distribution is free from the parameters here.

(Refer Slide Time: 16:21)

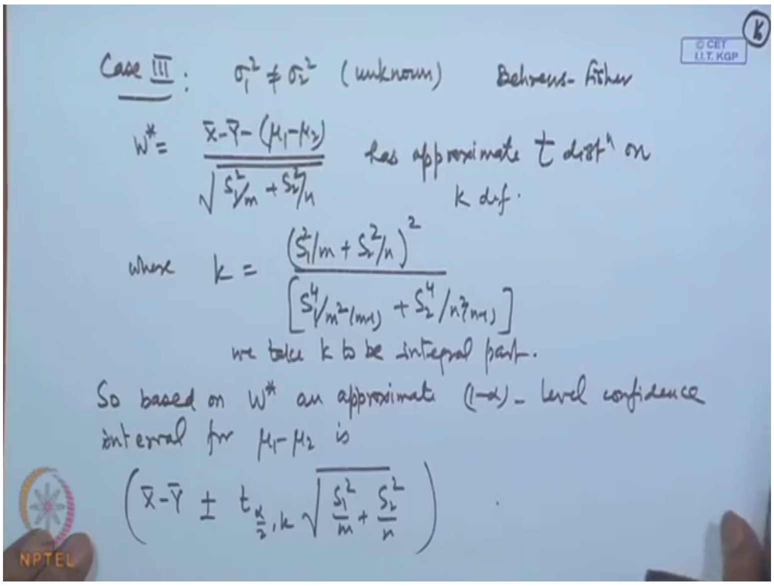


So, this can be used as a pivot quantity based on this I can write down the confidence interval. Probability of minus t m plus n minus 2 alpha by 2 less than W less than t m plus n minus 2 alpha by 2 that is equal to 1 minus alpha. Once again I have chosen symmetric because the t distribution is symmetric. So, shortest length interval will be the 1 which will be symmetric around the origin here. So, this probability is 1 minus alpha.

So, we can simplify this statement here, probability of minus t m plus n minus 2 alpha by 2 less than W so W is root m n by m plus n X bar minus Y bar minus mu 1 minus mu 2 divided by S p less than t m plus n minus 2 alpha by 2 that is equal to 1 minus alpha. So, after certain manipulation, I consider multiplication by S p and divide by this term then I adjust X bar minus Y bar on both the sides that is statement will be then equal to X bar minus Y bar minus root m plus n by m n S p t m plus n minus 2 alpha by 2 less than or equal to mu 1 minus mu 2 less than X bar minus Y bar plus root m plus n by m n S p t m plus n minus 2 alpha by 2, this is equal to 1 minus alpha.

So, we get 101 minus alpha, percent confidence interval as X bar minus Y bar root plus minus root m plus n by m n S p t m plus n minus 2 alpha by 2. Now remember here, we have taken sigma 1 square is equal to sigma 2 square assumption here which could have been done by making a test of hypothesis first that is we test whether sigma 1 is square is equal to sigma 2 square, if it is accepted then we apply this method, but suppose it is not accepted that sigma 1 square by sigma 2 square, the hypothesis sigma 1 square is equal to sigma 2 square is rejected if it is rejected then this test procedure will not be useful.

(Refer Slide Time: 19:15)



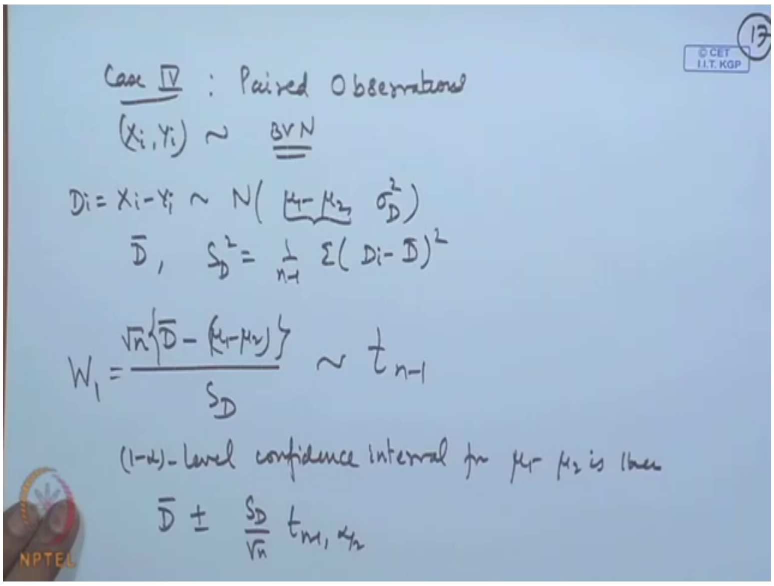
In this case an approximate procedure has been proposed by Welch in nineteen. So, that is let me say case 3, sigma 1 square not equal to sigma 2 square and they are unknown. So, this is known as famous Behrens Fisher situation, even in the testing problem we have not considered this because in this case theory of Neyman Pearsons test fails we cannot find out UMPN bias test in the situation. We can also not derive a likely would ratio test in that situation it does not give a region there.

So, approximate this is famous Behrens Fisher situation and in approximate procedures were proposed and one procedure which is based on let me call it W star that is X bar minus Y bar minus mu 1 minus mu 2 divided by square root of S 1 square by m plus S 2 square by n. So, basically this statistic if you see why it has been considered you considered the case of known sigma 1 square sigma 2 square. In the known sigma 1 square sigma 2 square case this was the statistic that we considered as the pivot quantity.

So, in place of sigma 1 square and sigma 2 square you are placed by S 1 square and S 2 square. So, this has approximate t distribution on say k degrees of freedom where k is actually random quantity. It is a S 1 square by m plus S 2 square by n whole square divided by S 1 to the power 4 divided by m square into m minus 1 plus S 2 to the power 4 divided by n square into n minus 1 and again problem is that this not be an integer. So, we consider we usually take k to be integral part of this. So, based on W star an approximate 1 minus alpha level confidence interval for mu 1 minus mu 2 that will be X bar minus Y bar plus minus t alpha by 2 k square root of S 1 square by m plus S 2 square by n.

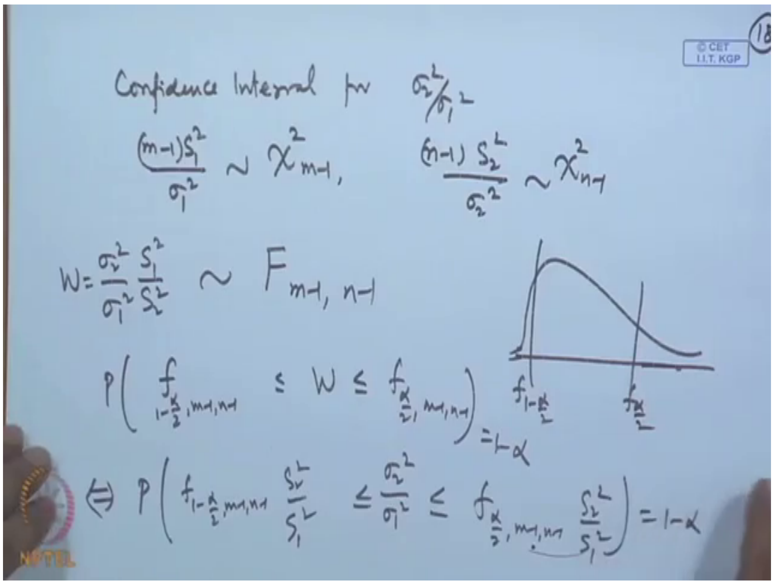
Now another interesting case is that, when the number of observations is a same and there is some sort of pairing if there is a pairing here then like in the paired test paired considered the quantity X bar minus Y bar divided by based on the standard the variance of the differences.

(Refer Slide Time: 22:39)



So, let me consider that situation also, case 4 that is paired observations. So, as before, we are considering X i, Y i this is following bivariate normal. So, in that case, if I am considering D i as equal to X i minus Y i, that follows normal mu 1 minus mu 2 and some variance sigma D square. So, we considered D bar and S D square that is 1 by n minus 1 sigma D i minus D bar whole square. So, root n D bar minus this is mu 1 minus mu 2 by S D this will have t distribution on n minus 1 degrees of freedom. So, this can be considered as a pivot. So, if we use this has a pivot, then the confidence interval 1 minus alpha level confidence interval for mu 1 minus mu 2 is then D bar plus minus S D by root n t n minus 1 alpha by 2.

(Refer Slide Time: 24:13)



I will end up this lecture by confidence interval for variance ratios if I have variance ratio I can consider confidence interval for say sigma 2 square by sigma 1 square, then I can consider this quantity see S 1 square m minus 1 S 1 square by sigma 1 square follows chi square distribution on m minus 1 degrees of freedom n minus 1 S 2 square by sigma 2 square follows chi square n minus 1 degrees of freedom and they are independent. So, if I frame the ratio S 1 is square by a S 2 square sigma 2 square by sigma 1 square let me call it by W this will followed F distribution on m minus 1 n minus 1 degrees of freedom.

So, probability of course, F is a skew distribution, but for convenience we may still consider f 1 minus alpha by 2 and f alpha by 2 points f alpha by 2 and 1 minus alpha by 2 m minus 1 n minus 1 less than or equal to this W less than or equal to f alpha by 2 m minus 1 n minus 1 that is equal to 1 minus alpha.

So, this is equivalent to saying that probability of sigma 2 square by sigma 1 square in the interval f 1 minus alpha by 2 m minus 1 n minus 1 S 2 square by S 1 square to f alpha by 2 m minus 1 n minus 1 S 2 square by S 1 square that is equal to 1 minus alpha. So, we get 1 minus alpha level confidence interval as left and limit as f 1 minus alpha by 2 on m minus on n minus 1 degrees of freedom S 2 square by S 1 square and the right hand limit as f alpha by 2 m minus 1 n minus 1 S 2 square by S 1 square.

I have discussed the general methods of finding out the confidence interval using the pivot method. I have also discussed the a special confidence interval for the parameters of normal distributions, exponential distributions etcetera.

Now in this let me have a summary of this course. We have considered major methods of statistical inference as developed in the classical statistics. The originators of most of this concepts were primarily RA Fisher and who in 1920's and 30's laid the foundations of theoretically statistics and Neyman and Pearson who were credited with the theory of testing of hypothesis, which I have covered in a major way in this particular course and also the theory of confidence intervals was developed by Neman.

I have in this particular course derived the methods; that means, what method should be used using what philosophy that we have done now given a practical problem you will have a data for example, you may have a data. So, you say may treat it as a data from a normal distribution say normal distribution with parameters mu and sigma square. Now you may like to find out a point estimator of mu and sigma square.

So, you may calculate MLEs you may calculate base or minimize estimators you may also consider testing problem so you may use one of the methods, you may also like to consider confidence intervals for the parameters, you may also have say two populations you may have samples from that. So, in this particular course we have derived all the methodologies which one can use as a statistician. I will be putting up certain assignment sheets, that is a problem sets and some of their hints and solutions for those sets which will have actually the practical data sets there. So, from the data sets you actually solve the problem and see which method will be used. So, that will be quite useful.

We have come to the end of this course. So, the topic of inferences actually very very wide I have covered only what is known as classical inference or you can say classical statistics there are many many modern methods which are being used, but I have not been able to cover in this particular course, but these courses will these things will be useful for most of the practical problems that people face. So, with this end we come to an end. I advise all the readers to solve the problem sets that I will be putting.