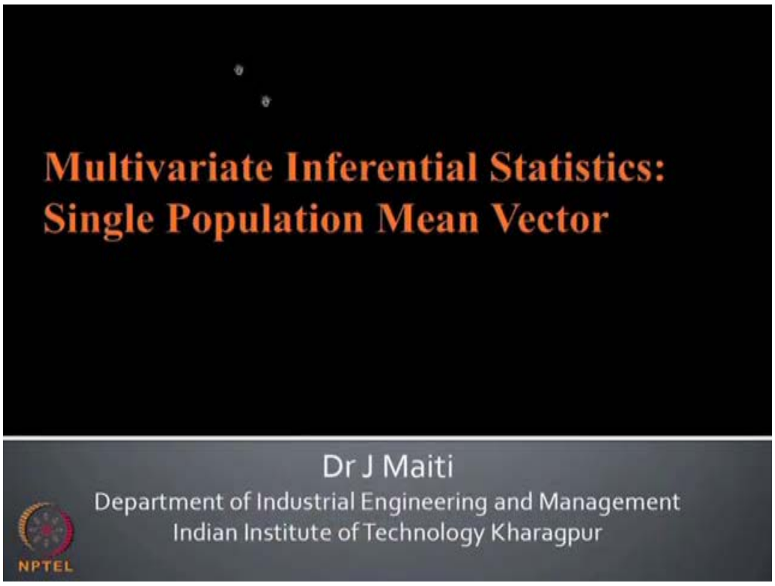
Applied Multivariate Statistical Modeling
Prof. J. Maiti
Department of Industrial Engineering and Management
Indian Institute of Technology, Kharagpur

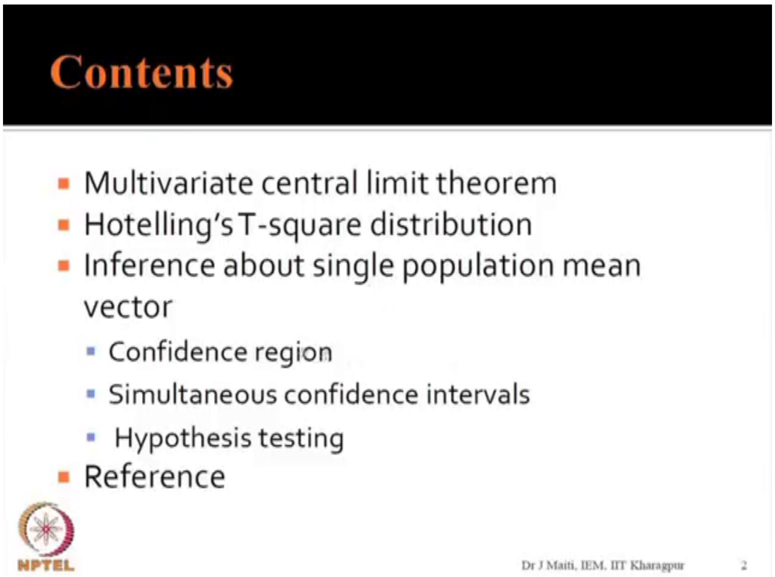
Lecture - 12
Multivariate Inferential Statistics

(Refer Slide Time: 00:24)



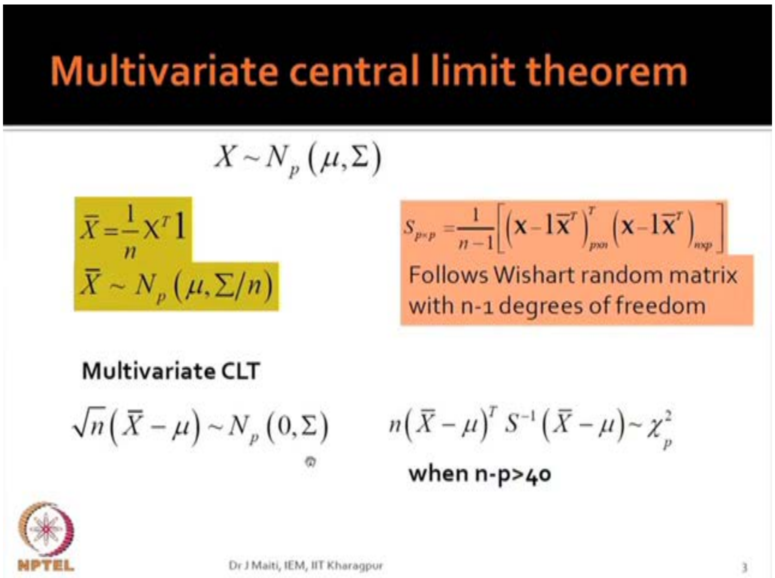
Good evening, today we will discuss Multivariate Inferential Statistics. We will start with single population mean vector, and then we will go for 2 population mean vectors.

(Refer Slide Time: 00:37)



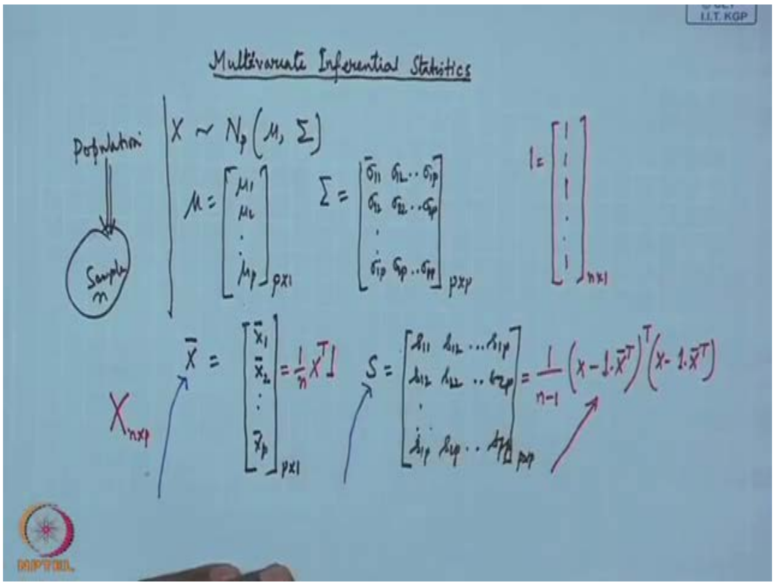
Now, under single population mean vector, these are the items we will be covering today. One is multivariate central limit theorem, hotelling T-square distribution, inference about single population mean vector. Under this confidence region, simultaneous confidence intervals, hypothesis testing followed by references.

(Refer Slide Time: 01:02)



Now, let us think that you are sampling from multivariate normal distribution.

(Refer Slide Time: 01:09)

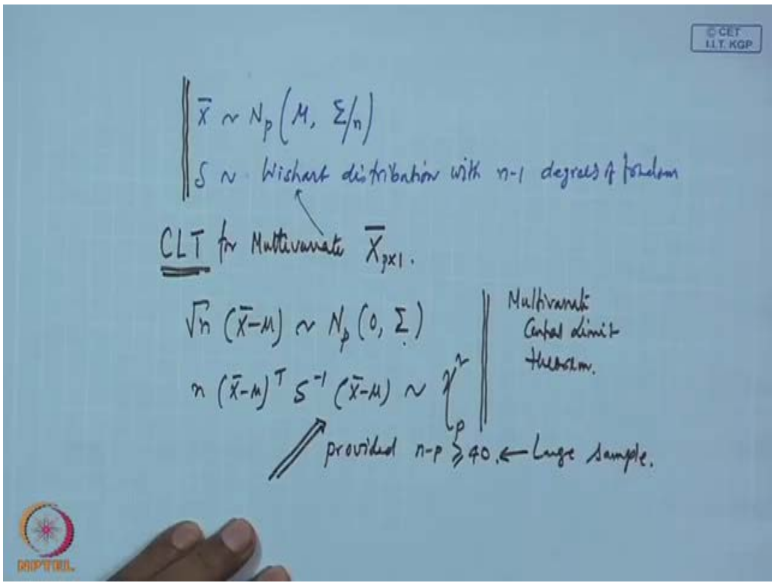


That X is multivariate normal, where the population parameters are mu which is mu 1, mu 2, it is a mu vector p cross 1. And your covariance matrix is p cross p. And you all know that what is the elements that is sigma 1 1, sigma 1 2 like sigma 1 p, sigma 1 2, sigma 2 2, sigma 2 p like this sigma 1 p, sigma 2 p, sigma p p. This is what your population is. When you sample from the population for example, a sample of size n then you compute sample statistics like X bar, which is again a p variable vector x 1 bar, x 2 bar like this x p bar.

And you also compute that covariance matrix, which is s 1 1, s 1 2, s 1 p, s 1 2, s 2 2, like s 2 p and followed by this s 1 p, s 2 p dot dot s p p, p cross p. What formula you have used earlier you have seen already that you have used this formula to calculate this, if your data matrix which X n cross p. And what you have used here, you have used this one is 1 by n then, X transpose. And you are creating the unit vector here, where one is 1 1 1 like this 1 n cross 1.

For sample covariance you have computed like this, that is X minus here it will be p cross p. So, one dot X bar transpose followed by X minus 1 dot X bar transpose, this is the formula you have used fine. Now, when we talk about the inferential statistics, we have seen in univariate case also that we talk about the distribution of the statistic of interest. For example, you may be interested to k. Now, what is the distribution of X bar. Similarly, what is the distribution of S?

(Refer Slide Time: 04:19)

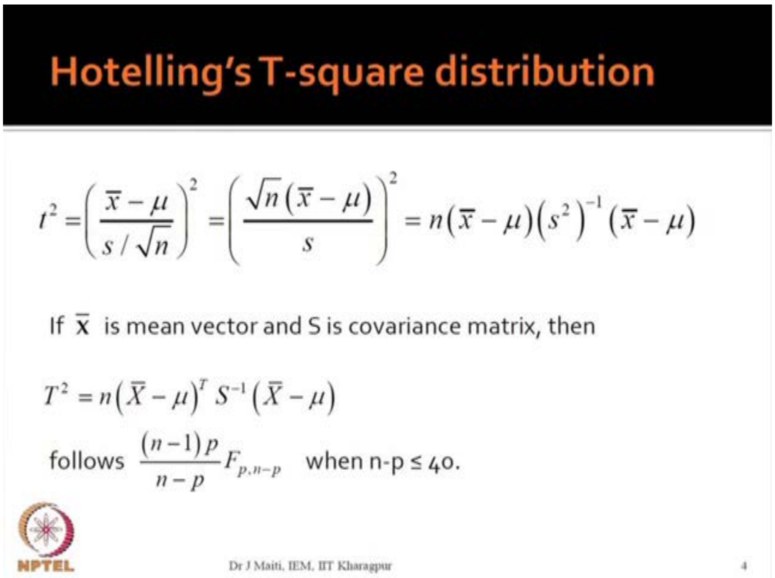


Now, X bar that follows normal distribution. That is multivariate normal with N p variables mu, sigma by n. And S follows Wishart distribution with n minus 1 degrees of freedom. This is a complex distribution Wishart distribution. We will not talk anything about this Wishart distribution input.

Now, then another issue what you have seen in the univariate case also, that is Central Limit theorem. Now, central limit theorem for multivariate data, for multivariate random variable X bar. What will be this? Here you have seen that square root of X bar minus mu that follows multivariate normal with 0, mean vector is 0 and that covariance matrix sigma. Last class you have seen, some of you have seen not all this one.

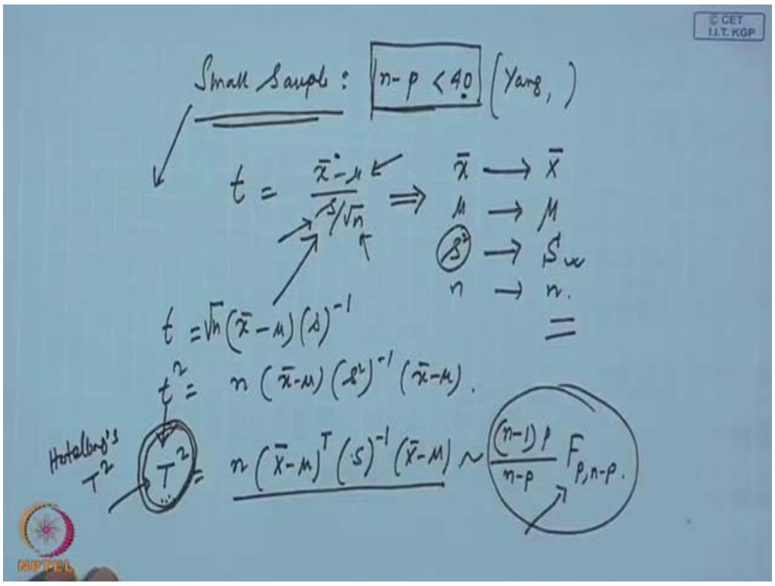
And also we have seen that n X bar minus mu transpose S inverse X bar minus mu. This follows your chi square distribution that is what multivariate CLT, multivariate central limit theorem. This second part, that in X bar minus mu S inverse X bar minus mu that follows chi square distribution provided n minus p greater than equal to 40 that is large sample.

(Refer Slide Time: 07:03)



Now, what will happen when your sample size is small? So, you go for small sample.

(Refer Slide Time: 07:08)



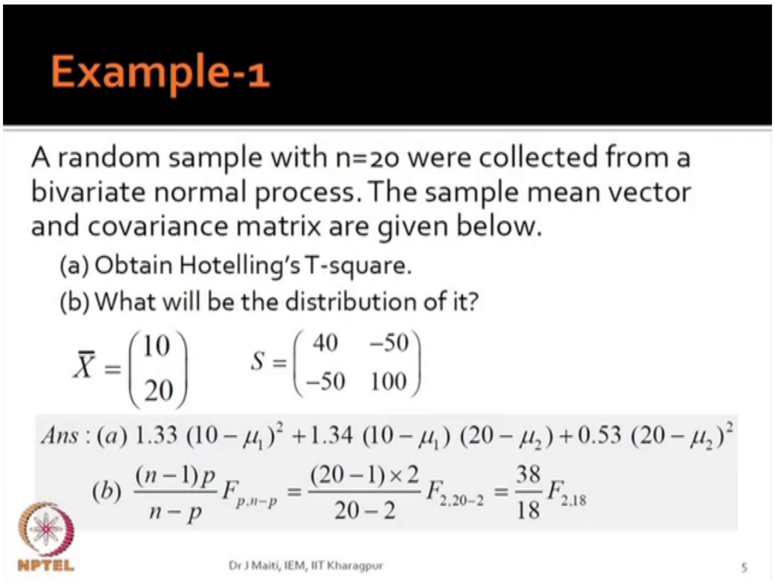
Small sample here we are saying that n minus p less than 40. This I have seen in the book by Yang and one more professor is there, that applied multivariate modeling in quality management. This is what I can say when you say it is small sample less than equal to 40. Now, let us see how we develop this one the distribution for the required statistics for X bar when your sample size is small. What we have seen earlier that in univariate case when your sample size is small, you have used t distribution, t equal to x bar minus mu by s by root n. This x bar is the mean, sample mean and mu is the population mean, s is sample standard deviation, n is the sample size. Now, I want to see what will be its multivariate counterpart, x bar will no longer be your scalar quantity it will be x bar, mu also will be a vector quantity.

And s will be replaced by that this S and n will be remain n. That is the case if you rewrite this, what you will write that t equal to x bar minus mu. Then, here we write square root of n, the S to the power minus 1. Now, if you make t square what you will get n x bar minus mu S square to the power minus 1 x bar minus mu. Now, we like the multivariate normal distribution we developed using univariate normal distribution. Here also we will use the similar formulation. So, instead of this t square we will write capital T square which is known Hotelling T square.

So, this will be written like this n x bar minus mu transpose then, this will become a S. So, what we say here that S that counterpart, we will be taking capital S. Basically, this is

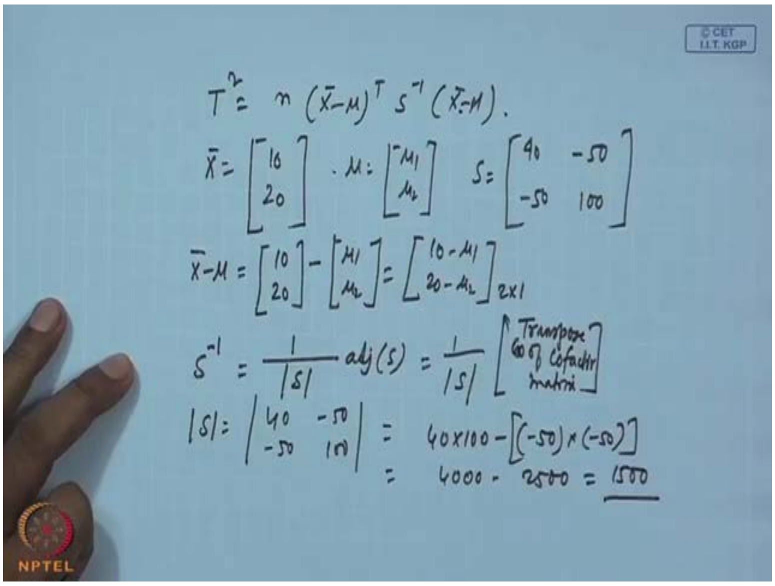
S square in terms of variance. So, S inverse X bar minus mu. And Hotelling has proved that this quantity will follow n minus 1 p by n minus p F p n minus p that is f distribution. It will follow f distribution. So, when you take a small sample size, you will find that this statistics is of importance now. And the derivation of this T square is this, and it will follow this distribution. Now, see the problem what I have solved in another class, again I want to solve it here.

(Refer Slide Time: 11:12)



And it may be repetition for some of you, but for others it is not that. So, random samples with size n equal to 20 were collected from a bivariate normal process. The sample mean vector and covariance matrix are given below. What is this X bar is 10 20 that is the mean vector. And covariance matrix is that is S 40, minus 50, minus 50 and 100. Do you want what will be the value of Hotelling T-square, that statistics value. Answer is this 1.33 10 minus mu 1 to the power square like this, so for some of you who are not gone through that computation.

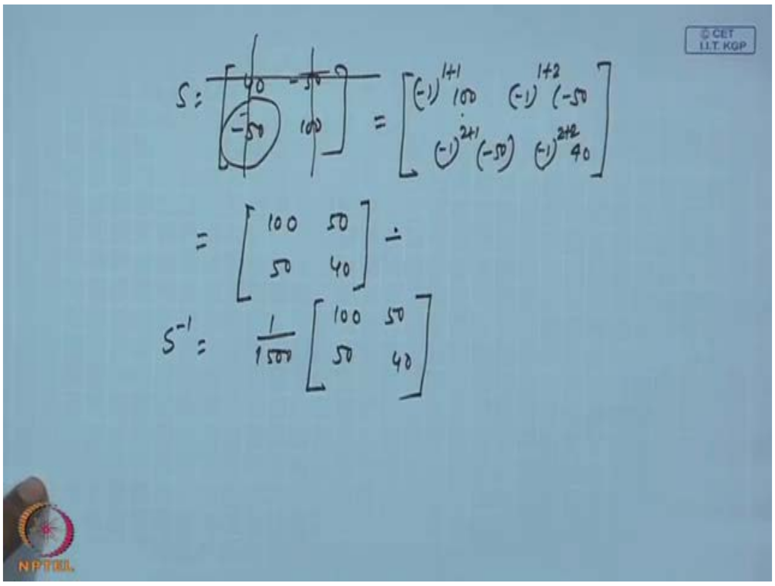
(Refer Slide Time: 12:11)



For you please see what is this T square is n X bar minus mu transpose S inverse X bar minus mu. What is your X bar, X bar given 10 20. What is your mu? mu is mu 1 and mu 2, because it is a bivariate case. What is s given S given 40, minus 50, minus 50, 100. What you require to compute? You require to compute first X bar minus mu which is 10 20 minus mu 1 mu 2, this is nothing but 10 minus mu 1 20 minus mu 2, which will be a 2 cross 1 vector. Then you require to compute S inverse, S inverse is 1 by determinant of S adjoint of S.

Now, you all know that adjoint of S is that I told you in one class. So, this one is that you find out here that cofactor transpose of cofactor matrix, that you want to find out. Anyhow we will see this. So, determinant of S is determinant of 40, minus 50, minus 50, 100. This will be 40 into 100 minus minus 50 into minus 50. So, this will be 4000 minus 2500 which will be 1500.

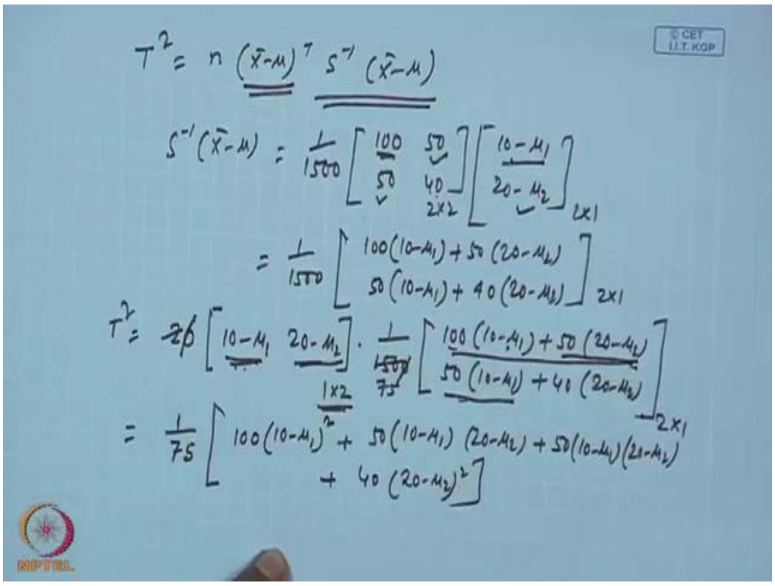
(Refer Slide Time: 14:21)



Now, you see the transpose of cofactor of S. So, S is 40, minus 50, minus 50, and 100. When you find out the cofactors, for this particular element you have to cross that particular row and column. So, ultimately what you get here minus 1 to the power i equal to 1, j equal to 1 then, what is remaining is 100. In the same manner you will go this. And you will find out here minus 1 to the power row 1 column 2 then, what is remaining if I go by this remaining will be this.

And here minus 1 2 plus 1 that remaining portion is this. And here minus 1 2 plus 2 remaining portion will be 40. The resultant quantity will be 100 then 50 then 50 and 40. Now, if you take the transpose you will get the same thing because it is a square symmetry matrix, you will get the same thing. So, S inverse will be 1 by 1500 into this. This will be your S inverse, once you know S inverse what you require to compute now.

(Refer Slide Time: 15:52)

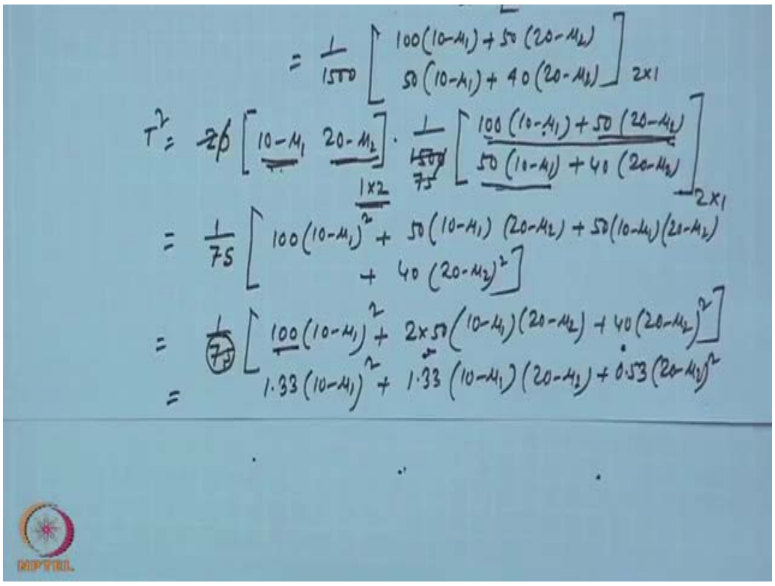


You want to compute now this one. So, I can see this one first. If I write S inverse X bar minus mu then this is nothing but 1 by 1500 into 100 50 50 40. And your X bar minus mu is 10 minus mu 1, 20 minus mu 2. So, 1 by 1500 this is 2 cross 2 matrix, 2 cross 1 matrix so your resultant will be 2 cross 1 matrix, so, 100 into 10 minus mu 1 and this into this plus this into this. So, plus 50 into 20 minus mu 2, then again 50 into this 10 minus mu 1 plus 40 into 20 minus mu 2. This is our 2 cross 1.

Now, you see T square n equal to how much n are 20. And X minus mu transpose will be 10 minus mu 1 and 20 minus mu 2, this will be 1 cross 2 matrix. Then into 1 by 1500 into this total quantity 100 10 minus mu 1 plus 50 20 minus mu 2 by 50 10 minus mu 1 plus 40 20 minus mu 2. So, this cut this one 2 1 50, that is 75. So, if I write like this one this is 1 by 75. Now, you see one is to 2 and 2 cross 1, you will be getting only 1 scalar value. So, multiply this into this and this into this.

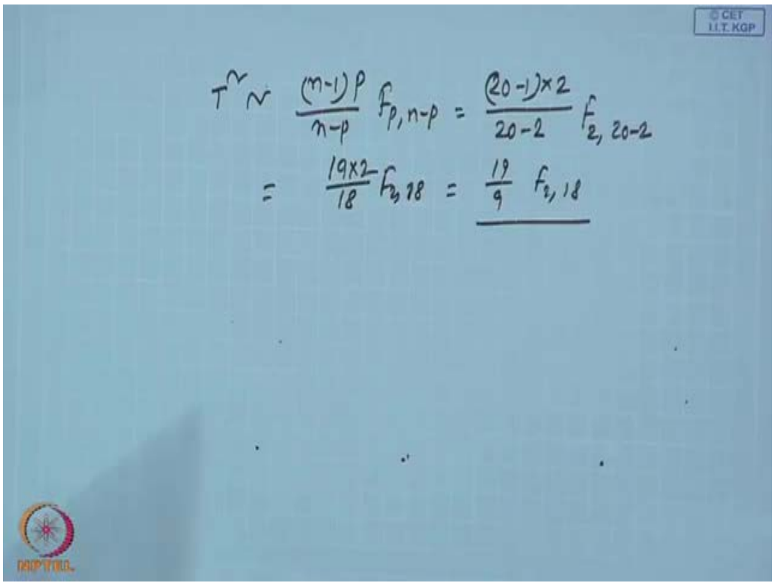
So, your resultant quantity will be 100 into 10 minus mu 1 into 10 minus mu 1 10 minus mu 1 square plus 50 into 10 minus mu 1 into 20 minus mu 2. So, this one multiplied with this totality then you will find out plus 20 minus mu 2 into 50 into this. So, I can write 50 into 10 minus mu 1 10 minus mu 1 into 20 minus mu 2 plus you will get 40 into 20 minus mu 2 square.

(Refer Slide Time: 18:59)



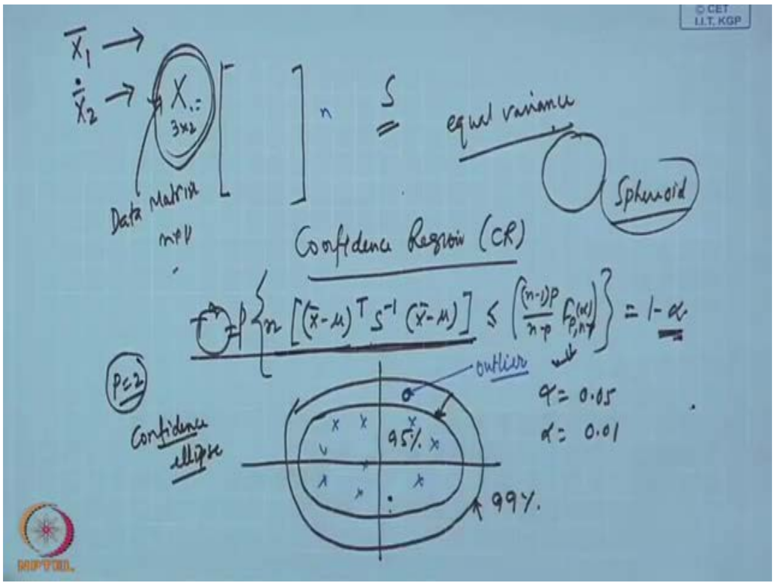
So, the total will be 1 by 75 100 10 minus mu 1 square plus 2 into 50 into 10 minus mu 1 20 minus mu 2 plus 40 20 minus mu 2 square. So, I said this will be 1.33 second one is 1.33 it should be, 1.34 it is written. Then this one is 0.53. So, resultant quantity is if I divide this 1 by 75 that is 1.33 10 minus mu 1 square plus 1.33 10 minus mu 1 into 20 minus mu 2 plus 0.53 20 minus mu 2 square. This type of equation you have seen earlier also. When I ask you to develop the bivariate constant density function, there we have developed this type of equation. It is similar the only difference is this one and X bar is X S will become sigma and that was the issue here. Now, my second question is what will be the distribution of it.

(Refer Slide Time: 20:28)



So, distribution mean T square follows n minus 1 into p divided by n minus p F p n minus p that will be your distribution. So, n is 20 minus 1 into p is 2 divided by 20 minus 2 then F p is 2 and 20 minus 2. So, this will be 19 into 2 by 18 F 2 18. We can write 19 by 9 F 2 18. We will see this slide again. So, our problem is this, and your question is this and your answer is this. And in exam time basically, they are 2 by 2 matrix will be giving, but there is possibility also that you require to compute X given by may be 3 by 2 matrix. This is 3 cross 2, this is the data matrix.

(Refer Slide Time: 21:35)



So, instead of giving you straight way the S what is required to be given because you will be getting this only. Initially when you collect data you will be getting this n cross p and all other things are to be derived. This is what Hotelling T-square is. What Hotelling T-squares give you? Hotelling T-squares gives you a confidence region. Confidence region CR that mean what we are trying to say if we follow this one the T square equal to n X bar minus mu transpose S inverse X bar minus mu n into this then, we are saying that this is an equation of ellipse when P equal to two.

You have seen earlier when you derived something like this. In the equation this one T square equal to this is an equation of ellipse. So, two dimensional case it will be equation of ellipse. And you know, what will be it in the multidimensional situation. What we want to say that this one is like this the T square value is this. This will be less than equal to the required probability. What is the probability distribution here? Your probability distribution is this.

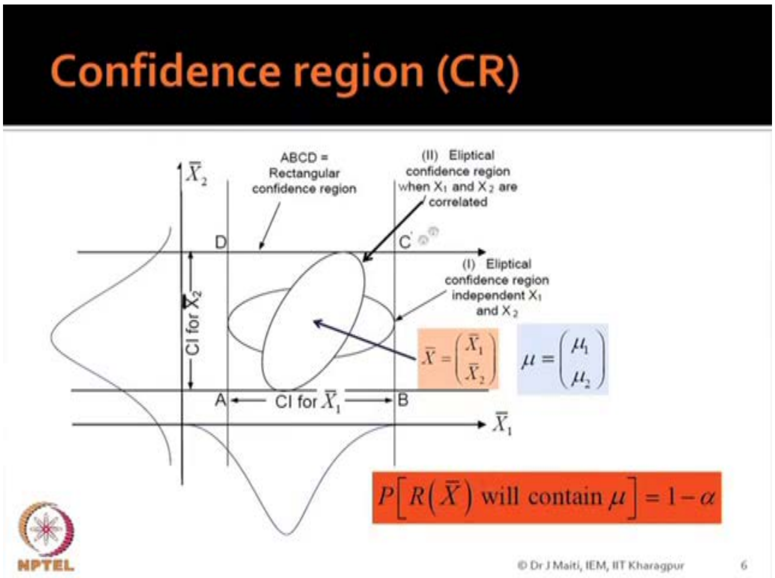
So, if I say this quantity so straight away if I write like this n minus 1 into p by n minus p F p n minus p. The totality I am writing that probability that this quantity will be less than this. What will be the value? If I put here 1 minus alpha in 2 dimensional case. Depending on the alpha, so you will get here a depending on again it all depends on the covariance matrix also. What type of ellipse it will be, but if the variables are independent then, you will be getting ellipse like this. This is what is determined by this quantity. This quantity will determine this.

And by 1 minus alpha if alpha equal to 0.05 that mean 95 percent of the observation will fall on or within the ellipse. If you take alpha equal to 0.01 then, your ellipse will be bigger, reason this contain 99 percent of the observations. This is known as confidence ellipse. Where from this ellipse is coming, that we have already seen that this is the quantity which in equation of ellipse when p equal to 2 and when p equal more than 2, it will be equation of ellipsoid. What will be the equation what will be this equation, when all the variables had equal variance. It will be a circle and for p dimensional scale it will be a sphere.

If all the variables have equal variance ultimately, you get a circle in 2 dimensional and a spheroid I think that will be in the higher dimensional case. Now, we want to understand this confidence ellipse from the original variable point of the like the X bar for the first variable that is the mean for the first variable. And X 2 bar is the mean for the second variable, these are sample mean. We can say fine there are 95 percent observations within the ellipse or there are 99 percent, but how do you interpret this.

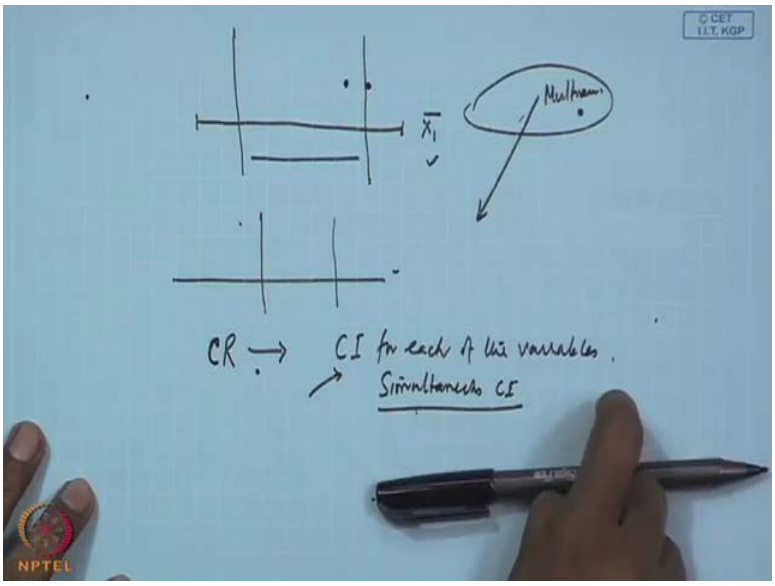
One of the interpretations could be suppose, you find out the ellipse. And then plot, plot in the sense you fix all the points. Suppose n points are there so you may get points are like this. Suppose you get one point here, this is one point. So, if I consider the 95 percent confidence ellipse then, this is a outlier. So, this approach can be used for outlier deduction also.

(Refer Slide Time: 27:20)



Another important issue here is this. You see this figure. This is what we practice in quality control. What we say that that there are 2 dimensions that X 1 bar and X 2 bar. You see that individual confidence interval and the interval collectively is making one rectangular region a b c d. So, that means some observation falling on here. Suppose one observation falling here. Based on individual interval it is within the confidence level, but collectively if you see that this is the collective confidence region then, this point is out of the confidence region. So, what is the interpretation then? Interpretation is if your case is a multivariate case.

(Refer Slide Time: 28:19)



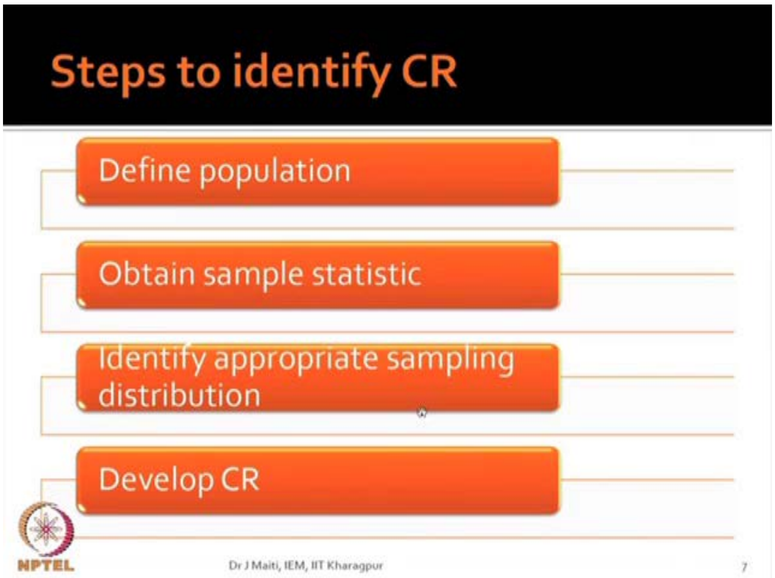
And you go by univariate analysis, and accordingly you develop certain intervals you may find out that the point you are talking about with respect to X 1 bar. This is within the interval even within with reference to X 2 bar, also this will be within the interval. But from their joint density, if we consider this is not within the interval in the sense within the region confidence region. That is very, very important concept that is why what I suggested that, when you infer a multivariate population, you start using the multivariate characteristics of the population. You can go for univariate, but not first. It is very clearly shown here, that if I take this as an quality control example then what I will find out.

I will find out that if I go by univariate that control level UCL and LCL then a point which is lying here, I will accept it as within control, but it is not within control that will be reflected only when you go by the confidence ellipse in 2 dimensional case, in multidimensional case confidence ellipsoid. The problem with this confidence ellipse is manifold. What is this? So, you are getting a region, but you cannot make decision based on this region you only say that it is outlier. For example, it may so happen that the point is here, which is outlier or out of control.

It is for which variable. Is it out of control for variable X 1 or out of control for variable X 2? You cannot make any comment based on this confidence interval, in confidence region that confidence ellipse. Confidence ellipse are ellipsoid will give you idea that yes one point, which is out of the control from quality control point of view. But even other way from simple data analysis point of view, you can say that it is an outlier, but why it is outlier, for which variable it is outlier?

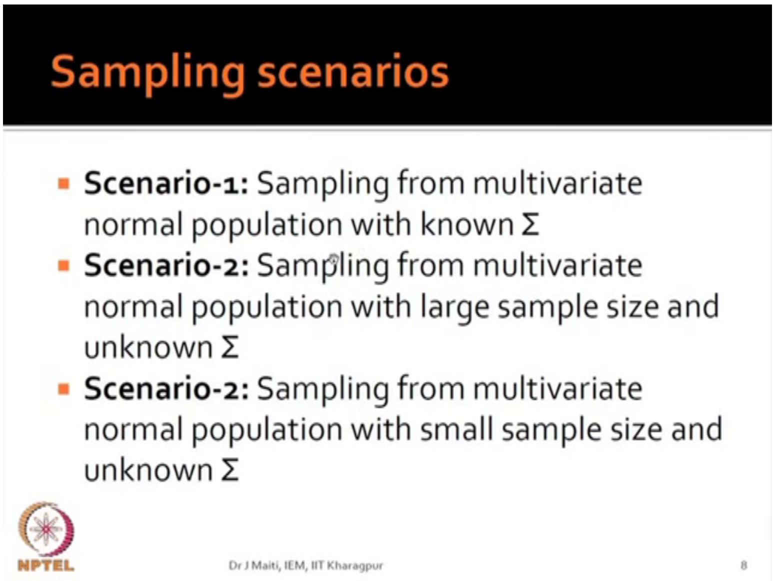
If you want to find out that you have to transform from confidence region to confidence interval. What I am trying to say here now. I am telling you that yes confidence region is the first step you have to develop it. If you do not develop it your inference will be faulty. Now, once you have confidence region. Suppose one point observation is out of control or outlier then, what is required to be known that, which variable is responsible for this? Then you cannot do anything with this region. You have to go back to confidence interval for each of the variables. That is what is known as simultaneous confidence interval. I will discuss this, but before that I want to do little more on confidence region.

(Refer Slide Time: 32:25)



What are the steps you have to follow to identify or define confidence region. First you define the population definitely it is we are considering multivariate normal population our things become easier. Then you find out the sample statistic. Identify the appropriate sampling distribution. So, if you do not know the sample statistic and the appropriate statistical distribution that is the sampling distribution. You cannot go for confidence interval also. You are not fit in the sense you have not enough information to develop confidence interval for univariate case, confidence region for multivariate case.

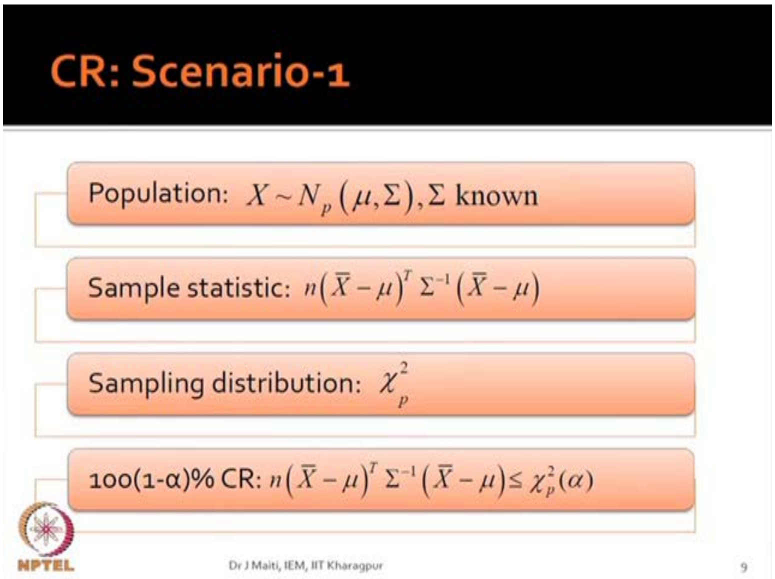
(Refer Slide Time: 33:15)



Now, by define population what do you mean? Define population then obtain sample statistic then, identify the appropriate sampling distribution then, develop CR. Again these are all broad steps, in each broad step there are sub steps also. We will solve one problem for confidence region by define population we are considering three scenarios. Scenario one sampling from multivariate normal population with known sigma, that means you know the covariance matrix for that population. This is basically parameter matrix, covariance parameter matrix this is my first scenario.

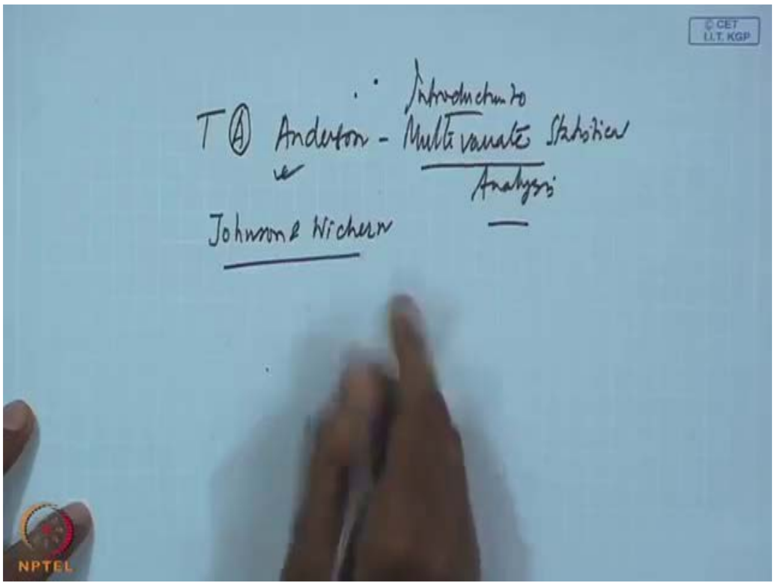
Second scenario is sampling from multivariate normal population with large sample size, but unknown sigma. And third one is sampling from multivariate normal population with small sample size and unknown sigma, there are more cases also. Suppose sampling from non normal population what is the guarantee that your population will be multivariate normal always, it can be non normal also. So, it is possible suppose when you sample from non normal population if your sample size is large because of the benefit obtained from central limit theorem. You can still say that X bar follows multivariate normal distribution and you can proceed we will not discuss that part.

(Refer Slide Time: 35:35)



We are discussing these three scenarios and let us see that what will happen where your scenario 1 is like this. Scenario one is what sampling from multivariate normal population with known sigma. We see the case. Now, if you want to know that how things are derived in the first scenario there is one good book, which I forget to tell you in the beginning that is T A Anderson.

(Refer Slide Time: 36:09)

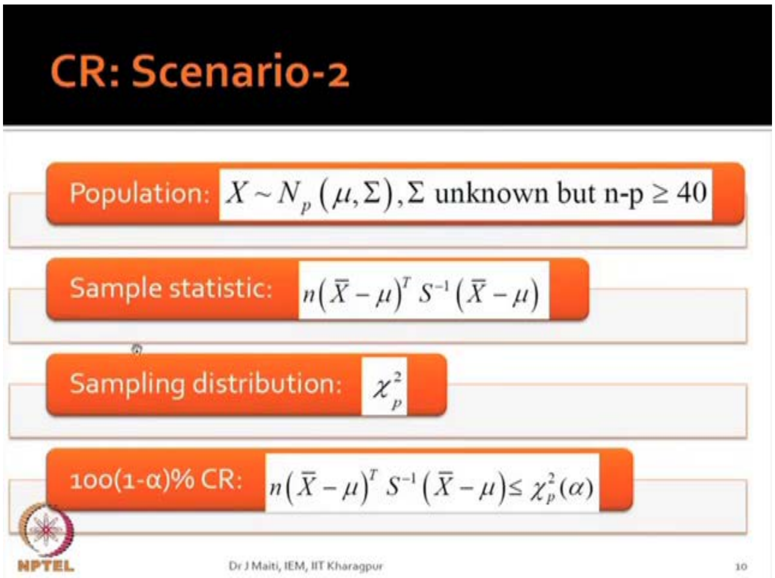


I think T R Anderson or T A, this one you check Anderson this is multivariate data analysis, introduction to multivariate statistical analysis. This is a very good book and in earlier than Johnson and Wichern, but the treatment is little different in the sense that it will be difficult for us to follow.

So, the first part is available in this book, second and third part is available in Johnson and Wichern scenario two and scenario three. So, you write down this one is scenario 3 sampling from multivariate normal population with small sample size, and unknown sigma. Scenario one the sample statistics is this n X bar minus mu transpose sigma inverse X bar minus mu because we know sigma, we straight away use this one. And based on our earlier discussion all the earlier classes, if you remember I do not require writing that the sampling distribution will be chi square because you all know this.

Earlier you have seen that this square term will follow chi square distribution with p degrees of freedom. And what is your 100 into 1 minus alpha percent confidence region. This is less than equal to chi square p alpha, this quantity will be less than equal to chi square p alpha and if I say that probability that this quantity is less than this equal to 1 minus alpha, see define steps. Define population identify the appropriate sample statistics. Find out the sampling distribution then, find out the confidence interval because these are the steps. So, for all cases not necessarily this case for all cases these are the steps.

(Refer Slide Time: 39:09)



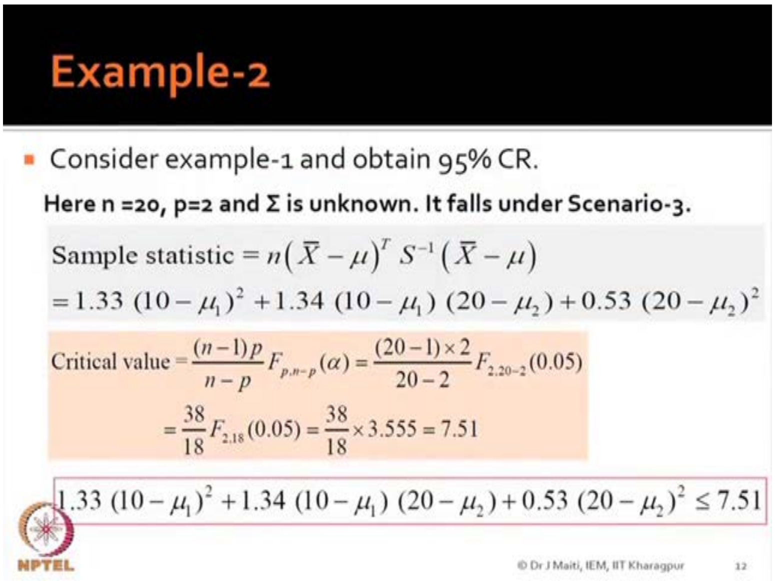
Can I go to the next slide? Suppose scenario 2 you are sampling from multivariate normal population sigma unknown, but your sample size is large. What is the difference here? Difference is simply that if you see this sample statistic, it is almost same to the earlier one except this quantity. A symbol instead of sigma inverse because now sigma is not known, S is the estimate of sigma. So, long your sample size is large, this quantity also follows chi square distribution. And always the degree of freedom will be the number of variable you have considered in defining the multivariate population.

And your probability that n X bar minus mu transpose S inverse X bar minus mu less than equal to chi square p alpha that will be 1 minus alpha. Now, if I ask you what will be the case for the third scenario, what will be the sample statistics for the third scenario, what is our third scenario? Scenario 3 sampling from multivariate normal population with small sample size and unknown sigma.

Everything is T square here this under different situation, but definitely for small sample the T square is known Hotelling T-square. What distribution it will follow, F distribution. Let us see this is scenario 2, this is scenario 3. So, population is, this is the condition sample statistics remains same as earlier for the scenario 2. And sampling distribution is changed to F from chi square distribution, and if you know this three the other thing the confidence region for this very simple.

So, you will get this 1 is less than equal to n minus 1 this quantity. I hope that there will be no problem at least from confidence region point of view. And you will be able to identify the confidence region properly. And you will also require to little bit study is required little bit where understanding that what distribution you require to use.

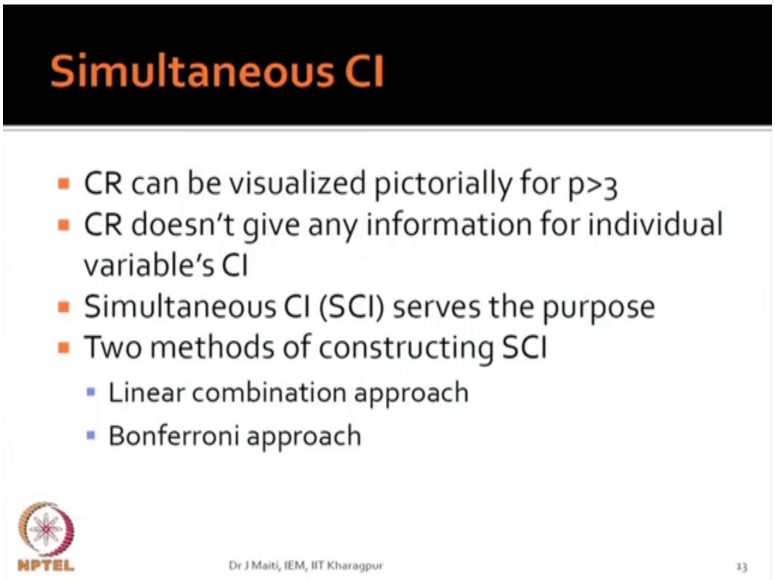
(Refer Slide Time: 42:32)



This is my example now. So, we have done lot of calculation in example one, example one was that what will be the T square. So, I can go to this, this is my T square, I think T is not visible, now it is visible. This is your T square, for the example one you also have found out that your distribution is this. Then what you require to do you require to find out what is this value, F 2 18 this value is 3.555. You see here I checked from the table.

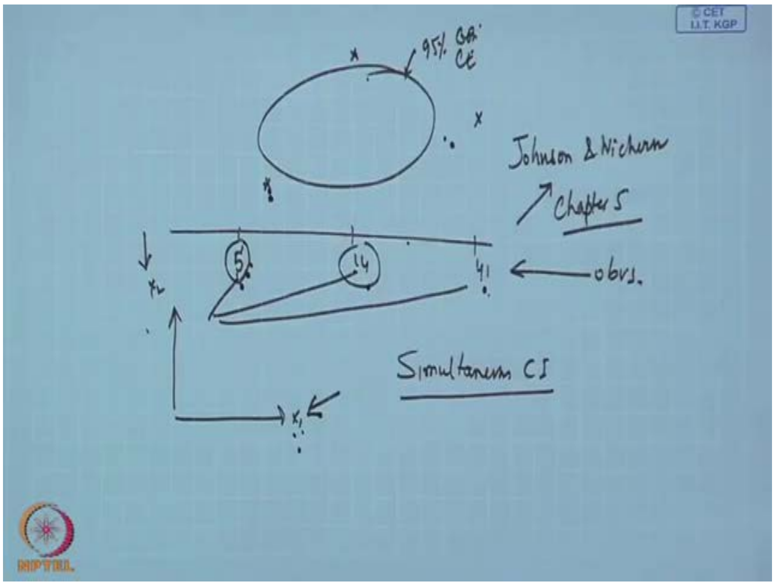
You check from the f distribution table, if there is mistake you please correct it I do not think it is wrong. So, 38 by 18 into 3.555, this is 7.51. Now, this is the T square value it is less than 7.51, this is my confidence region. In this case this is the confidence ellipse. I think you will be able to compute. There is nothing great here it looks very complicated if you see the equations. As if there is something big things, but ultimately the understanding point of view it should not big. But if you say from derivation point of view, this is very big. And we are not statistician, we are basically applied engineers all of us almost or applied either unanimously in the applied sciences we are working. And our aim is to make this complex statistics as simple as possible and then use it that is the issue.

(Refer Slide Time: 44:56)



You check this calculation also by your own. I think I have to show this otherwise they will switch over to me that radio mode. So, this is 1.33 this point already you have computed this one. Now, it is less than 7.51, this is the issue this quantity you will be getting. If you find it difficult please again come to me I will tell you. Now, let us go to the second portion which is also equally important here because we have done one study in quality control, where we have found out that several variables are there. And then we have used this confidence ellipse using 2 variables simultaneously. And we found out that what are the outliers or the out of control observations.

(Refer Slide Time: 45:58)



We found out something like this, the observation number 5, observation 14 may be observation 41. These are out of control observations in quality control point of view. How we found out, if we use the confidence ellipse you will be finding out what are the observation out of control. Now, question my question is 5 is out of control because of X 1 or X 2. 15 is out of control, whether it is both the variable are contributing to this or one variable is contributing to this like this that was the starting point what you have thought of…

Student: ((Refer Time: 46:37))

What five?

Student: When the confidence interval you know ((Refer Time: 46:46)).

Five variables we are not taking here. We are talking about confidence values 2 variables.

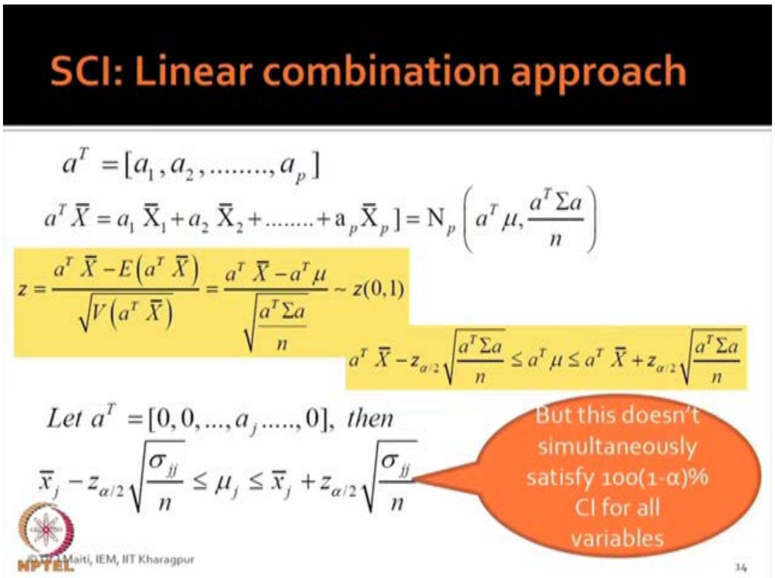
Student: Five what you have shown.

Five I have show somewhere which one. That means you have not listened properly, I said these are observations. Observation number 5, observation number 14, observation number 41, your confidence ellipse is this let it be. And your 5 is falling here suppose 14 is falling here and 41 if I write somewhere here, suppose it is falling here let it be here itself no problem. When these are not falling within my 95 percent confidence ellipse, if I say confidence ellipse what I am saying here 5 is out of control, 14 is out of control and 41 is also out of control.

You want to know whose variable is causing this observation out of control. Is it X 1 alone, is it X 2 alone or X 1 X 2 combiningly that joint if it is also coming here. May be we do not know that is possible only when you will, but you know this only when you go for simultaneous confidence interval. Now, we will discuss simultaneous confidence interval. How you compute simultaneous confidence interval? And you will find its applications also, and you will be happy to apply this in your project work. What is given here, CR can be visualized pictorially for p greater three.

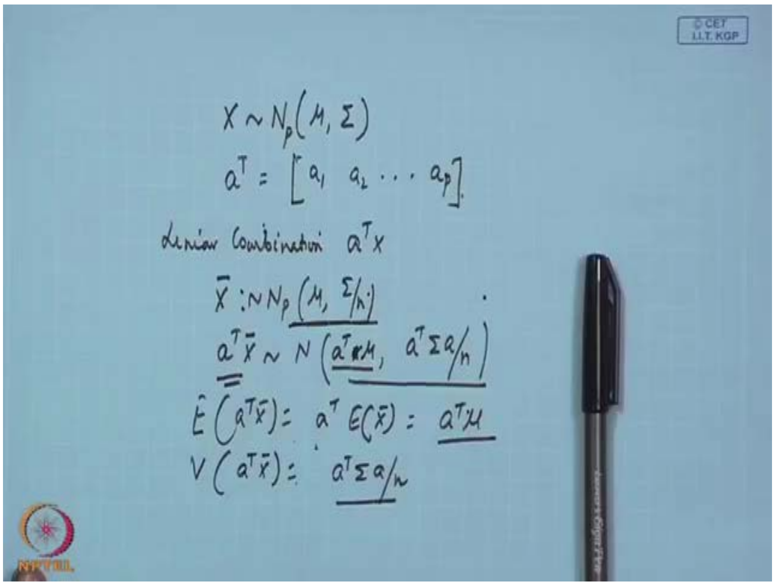
CR does not give any information for individual variables confidence interval. Simultaneous confidence interval serves a purpose two methods of constructing SCI where one is Linear combination approach other one is Bonferroni approach. There are two approaches both the approaches are available in the book of Johnson and Wichern. Please go through chapter 5.

(Refer Slide Time: 49:36)



What is this linear combination approach? Let me give you the basic idea here and then you will be able to compute by your own. If X is multivariate normal and let a transpose is a some constant a 1, a 2 like a p. So, a 1 to a p and you have seen the linear combination in last class.

(Refer Slide Time: 50:16)



I told you the linear combination is a transpose x. Now, X bar is also your multivariate normal. So, you can create linear combination of X bar also. And one of the properties of multivariate normal distribution is the linear combination will be univariate normal. Can you remember I told you four properties of multivariate normal distribution. First one is if X is multivariate normal, individual variable will be univariate normal. Second one is subset of that multiple variables will be multivariate normal. Third one is linear combination of this variables will be univariate normal.

So, our X bar is multivariate normal with mu and sigma by n. These are the parameters of the multivariate normal distribution. Then a transpose X will be univariate normal with a transpose mu that will be the mean value then, a transpose sigma a by n that will be the variance in this case. Why variance because it is univariate, and this quantity will be 1 cross 1 this is the case.

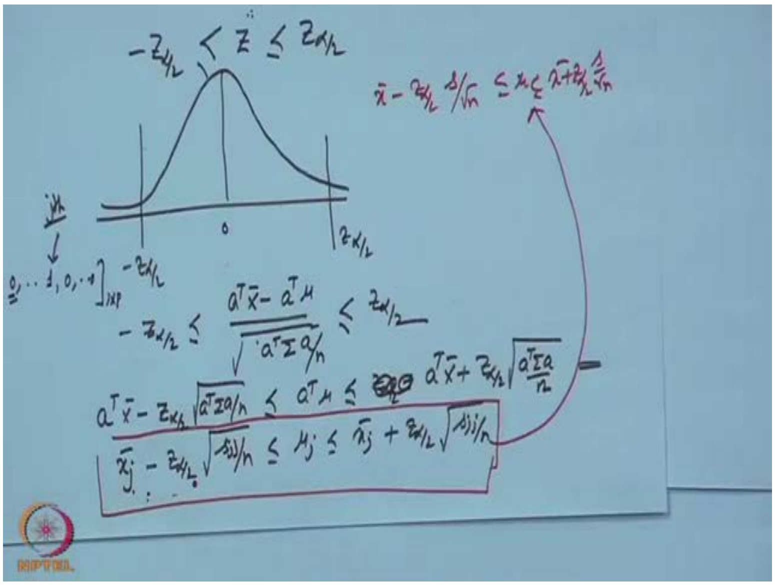
Now, if you know this one can you not create this matrix a statistics here. You see here the statistics what we are writing here. Z is the random variable minus expected value of this random variable divided by square root of variance. I think you all have seen this one earlier when you have developed Z. So, this is the case when this one a transpose X bar minus a transpose mu because expected value of a transpose X bar is a transpose mu. It is a transpose expected value of this.

This one is a transpose mu and also you have seen that variance of a transpose X bar will be a transpose sigma a by n that you have seen earlier. And as we have taken this quantity a transpose X bar minus a transpose mu this, this is your z distribution.

Student: ((Refer Time: 53:31))

A transpose you know this is n not p, this p that is the cotton pest problem. So, Z is because I told you the property says that linear combination will be univariate normal. Now, Z is this quantity, if we just derived in this manner you will be getting this Z distribution which is 0. Now, if you know this quantity Z distributed, can you not create the interval?

(Refer Slide Time: 54:22)

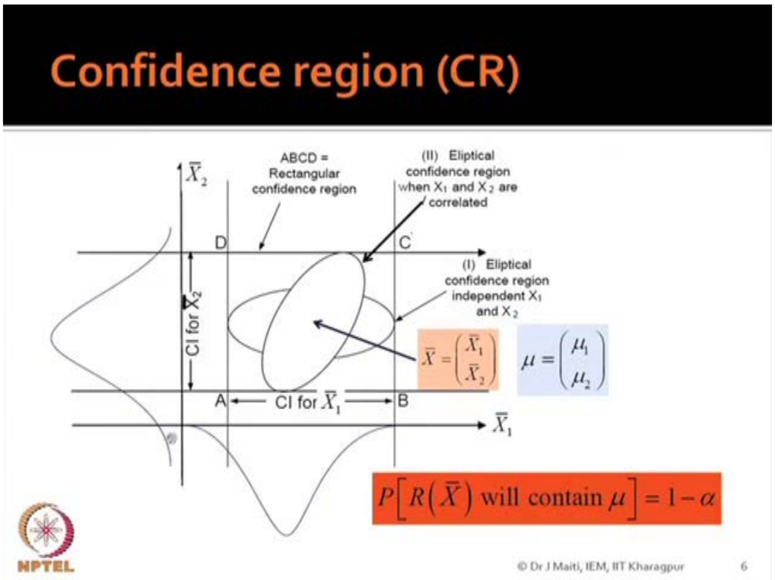


If I know something Z distributed all of you know that minus Z alpha by 2 less than equal to Z less than equal to Z alpha by 2. That is what we have seen in univariate case this is my Z alpha by 2, this is my minus z alpha by 2 and this line will be 0 definitely. Then you put the Z value here so minus Z alpha by 2 less than equal to a transpose X bar minus a transpose mu divided by square root of a transpose sigma a by n. This will be less than equal to z alpha by 2.

So, if you further just mathematical manipulation you do, what you will get here a transpose X bar minus Z alpha by 2 square root of a transpose sigma a by n less than equal to a transpose mu less than equal to Z alpha by 2 plus a transpose X bar plus Z alpha by 2 square root of a transpose this by n. What will happen if I choose a equal to all 0 0 then, jth element will be 1, and then 0 0 like this. What is this, this is 1 cross p this is the transpose 1 cross p. What I am doing the first constant is 0, second constant 0. Only the jth 1 is 1, if you put this into here in this equation what you will get.

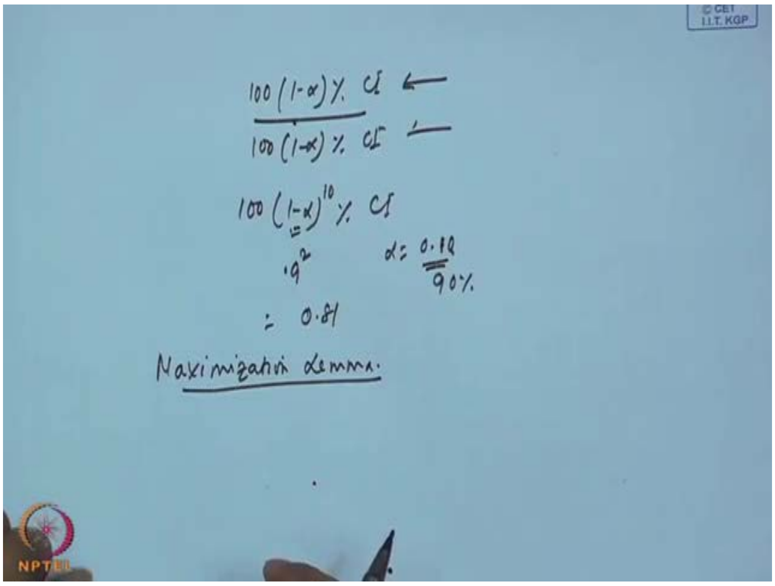
You will get like this X bar j minus z alpha by 2 root over S j j by n sum less than equal to mu j less than equal to x j bar plus z alpha by 2 square root of S j j by n, but there is a problem, what is the problem? Problem is, see if you see this equation, in this equation it is just like a univariate confidence interval because we have noted j only. If I do not write j the way we have defined in univariate case, we have written X bar minus z alpha by 2 then, your S by root n less than equal to your mu less than equal to X bar plus z alpha by 2 S by root n. There is no difference between this two, if there is no difference then ultimately where you are landing now you are going back to this.

(Refer Slide Time: 57:42)



You are going back to individual interval A B for X 1 bar that is for mu 1 and you are A D for mu 2. If I go by this method what will happen ultimately, your problem lies here.

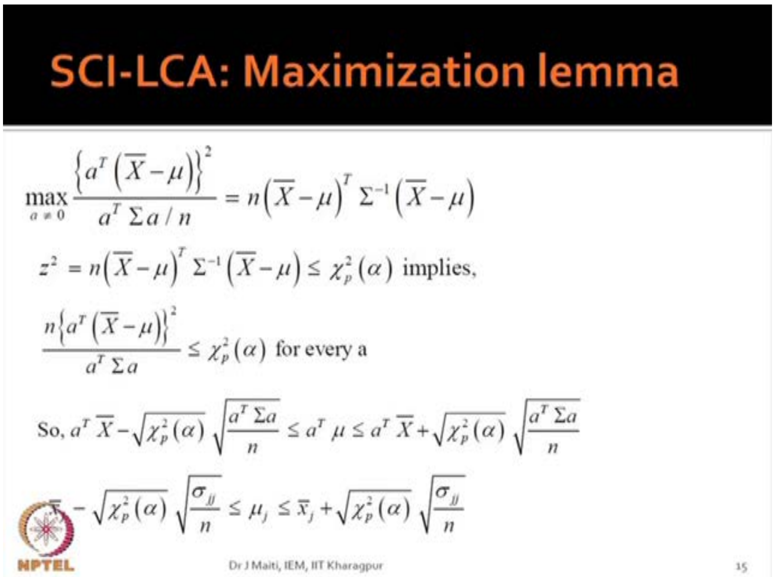
(Refer Slide Time: 58:08)



Problem is suppose you want 100 into 1 minus alpha percent confidence interval. When there is 1 variable you are getting like this, when there is 2 variables again you are getting this. So, second one 100 percent CI. So, like this if there are 10 variables what will happen ultimately you will get 1 minus 100 to the power 10 percent CSI. Considering that all independently we are estimating the CSI. So, if your alpha is 0.10 then it is 0.9 to the power, even for the two variable case, it will be 0.81. So, 81 percent simultaneous interval you are getting confidence interval.

You will not get 95 percent or here 90 percent what you thought of. So, you require to rectify this, if you go by linear combination you require to rectify this that rectification mathematics is also available in Johnson and Wichern. In chapter 2 you go to the matrix algebra portion you will be finding out, there is one beautiful concept called Maximization Lemma it is coming from matrix algebra.

(Refer Slide Time: 59:51)



What is this Maximization Lemma? So, my work is to simplify this and tell you. There are 2 3 pages write up you will be getting whole of that a X bar mu sigma like this. So, do not worry while reading this one, but you please remember that this is what that for non zero a that the maximum value of a transpose X bar minus mu square by this, this is nothing but the Z. This one a transpose X bar minus mu by a transpose sigma a by n that square root is Z this square root square so that is Z square. So, Z square follow this quantity Z square will be this. Now, all of you know what is the distribution of this, what is the distribution of this?

Student: Chi square distribution.

So, in next class we will discuss chi, this one we will continue.