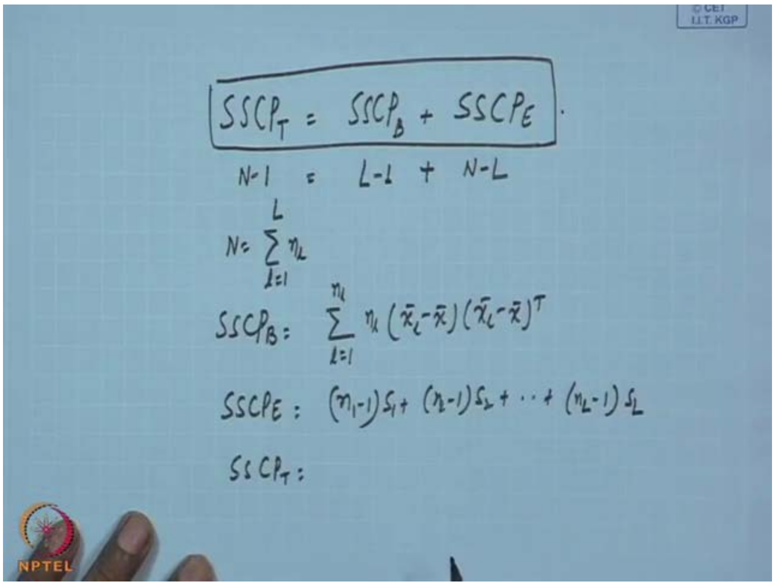
Applied Multivariate Statistical Modeling
Prof. J. Maiti
Department of Industrial Engineering and Management
Indian Institute of Technology, Kharagpur

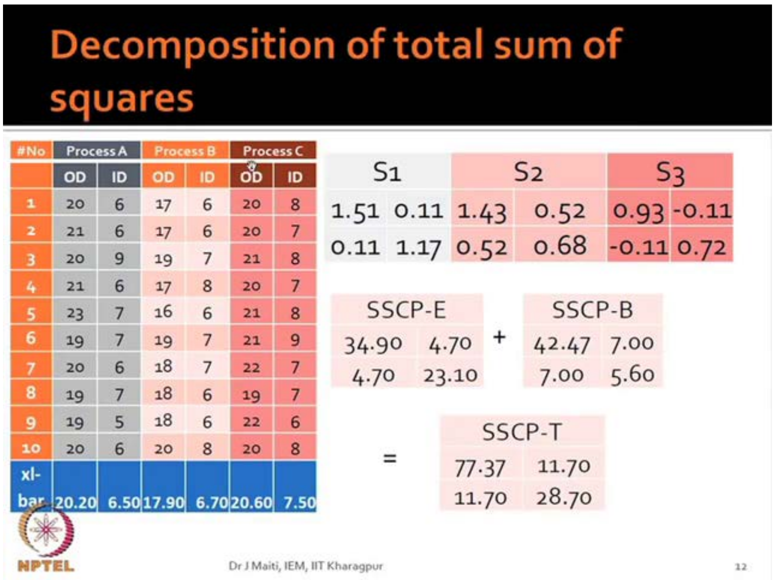
Lecture - 17
Multivariate Analysis of Variance (MANOVA) (Contd.)

(Refer Slide Time 00:30)



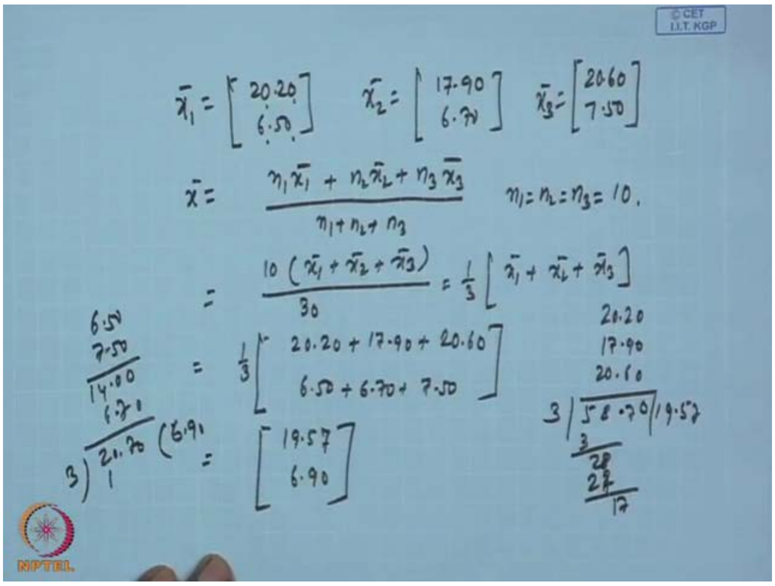
Good morning. We will continue MANOVA multivariate analysis of variance. What we have seen in last class that we have decomposed SSCP total into two parts, SSCP between population plus SSCP error. We have also discussed that the degree of freedom for SSCP total is N minus 1, where N is equal to l equal to 1 to capital L n l, and there are l populations. So, the degrees of freedom for SSCP B is L minus 1 and rest the total of SSCP B and SSCP E s degrees of freedom will be n minus 1, so the rest will be N minus L. And as I told you in last class that SSCP B will be computed using this formulation l equal to 1 to n l n l x l bar minus x bar and x l bar minus x bar transpose. SSCP E will be computed using this formula n 1 minus 1 s 1 n 2 minus 1 s 2 like up to n l minus 1 s l. Then SSCP T will be the sum of this two this equation.

(Refer Slide Time 02:26)



Now, let us see one problem what we have discussed in last class that there are three processes A, B and C, and we have sample ten observations from all the processes and two variables outer diameter and inner diameter. You see the mean values for process A is 20.20 and 6.50.

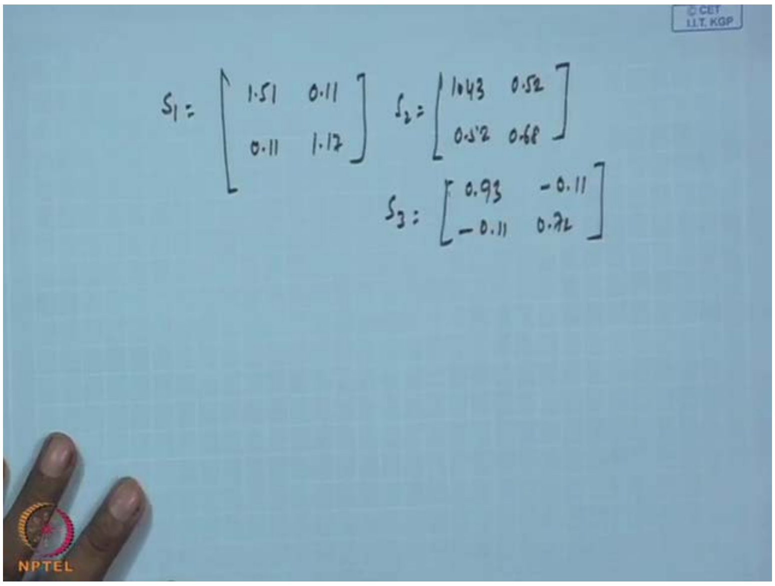
(Refer Slide Time 03:05)



So, if I write the mean values here, then for process A, that mean value will be 20.20 and 6.50. Similarly, for process B, it is 17.90 and 6.70 and what about for process C? That is x 3 bar. This is your 20.60 and 7.50. Now, what will be your x bar? x bar will be n 1 x 1 bar plus n 2 x 2 bar plus n 3 x 3 bar divided by n 1 plus n 2 plus n 3. Now, we have n 1 equal to n 2 equal to n 3 equal to 10.

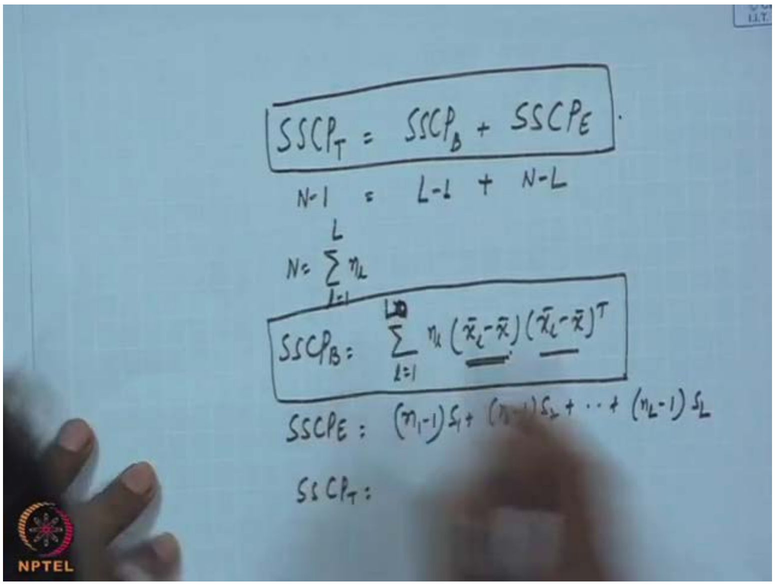
So, we can write this one as 10 x 1 bar x 2 bar x 3 bar by 10, which is 1 by 3 into x 1 bar plus x 2 bar plus x 3 bar, so what will be this quantity? Then 1 by 3 x 1 bar is 20.20, 6.5, like this, your case will be 20.20 plus 17.90 plus 20.60. Second one will be 6.50, 6.70 plus 7.50. If you add all those things 20.20, 17.90 and 20.60, quantity will be 17, 8, and 58.70 divided by 3; it will be 1, 9, 27, then 17, then 5 almost seven. So, this quantity will be 19.57 and second one will be 6.5 plus 7.5. This is 14 plus 6.70. This is 20.70 divided by 3. This will be 6.90. So, this is 6.90.

(Refer Slide Time 05:53)



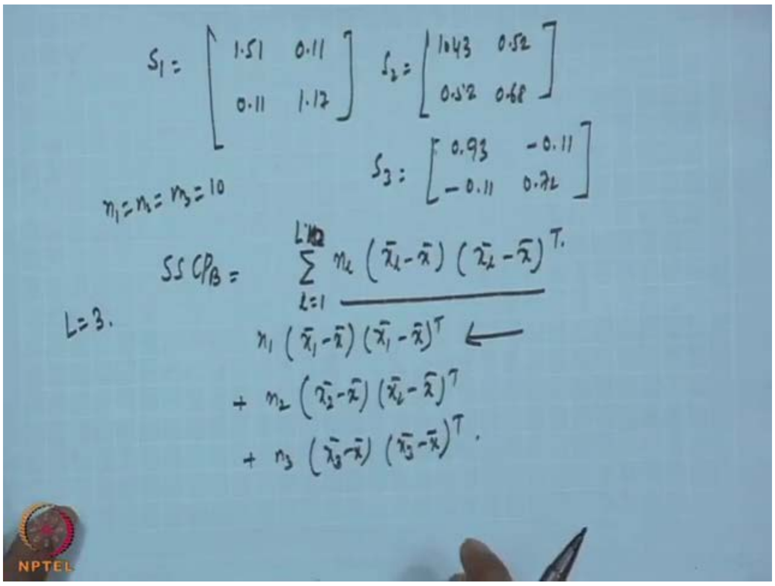
What more we want? We want to know what is s 1? And in last class, we have computed s 1 is 1.51, 0.11, 0.11, 1.17. Your s 2 is 1.43, 0.52, 0.52, and 0.68 and s 3 is 0.93, minus 0.11, minus 0.11 and 0.72.

(Refer Slide Time 06:38)



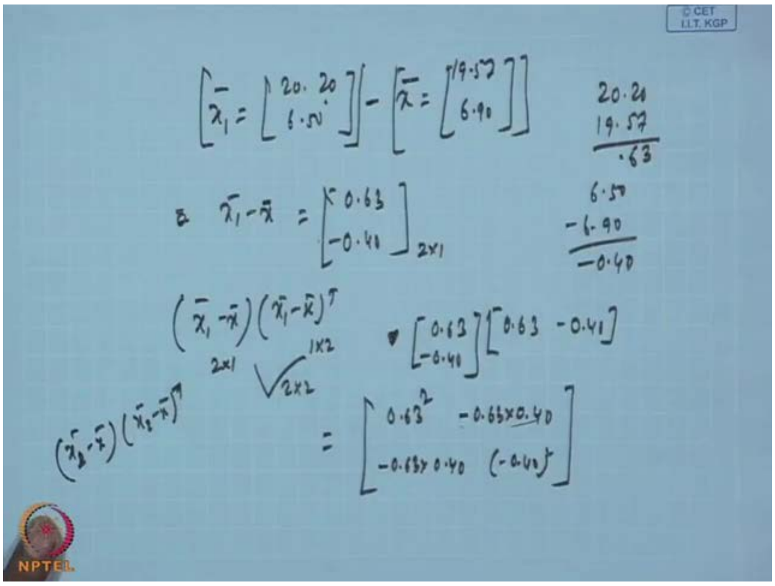
Now, we want to calculate SSCP B, first SSCP B, if you require to calculate SSCP B, you require x l bar minus x bar and x transpose to be multiplied. So, I am writing the formula first, SSCP B as n 1 equal to n 2 equal to n 3 equal to 10. So, I can write like this your l equal to 1 to n l. n l is basically n 1 to all those things n l that is the formula x l bar minus x bar x l bar minus x bar transpose. So, we will first consider a thing, we have done a mistake l equal to 1 to L, capital L.

(Refer Slide Time 07:43)



This is the capital L. So, our L equal capital equal to 3. So, you will be getting 3 sets of values here. First one n 1 x 1 bar minus x bar x 1 bar minus x bar transpose plus n 2 into x 2 bar minus x bar plus x 2 bar minus x bar transpose plus n 3 into x 3 bar x bar and x 3 bar minus x bar x 3 bar minus x bar transpose. I will show you one calculation first one and then similarly, other also will be computed.

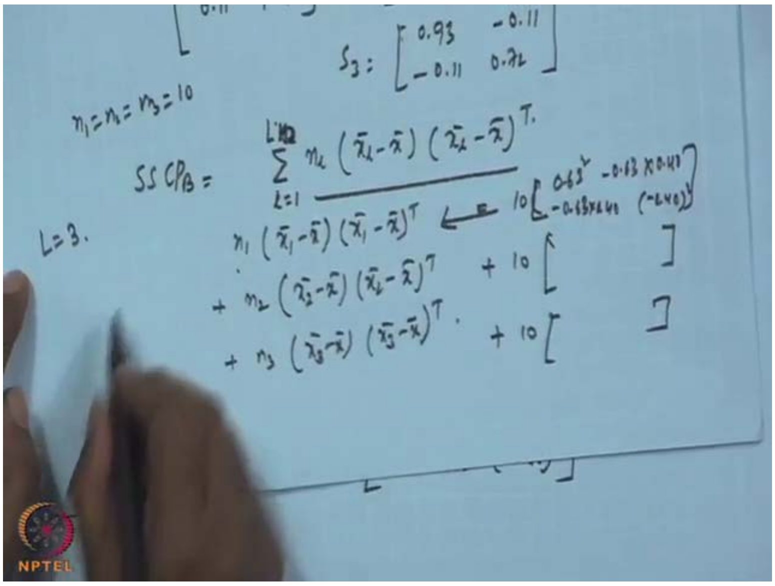
(Refer Slide Time 08:26)



Now, our x 1 bar, if you will see what our x 1 bar here, you see x 1 bar is 20.20 and 6.50, so 20.20, 6.50 minus, minus, if we want to do x bar, so I can see this minus x bar equal to our x bar value is this 19.57. This is also a matrix of order 2 cross 1. So, then this quantity, we can write x 1 bar minus x bar is 20.20 minus 19.57, so this is 3, 6, I think 0.63. Second one will be 6.50 minus 6.90, this will be minus 0.40, 0.40.

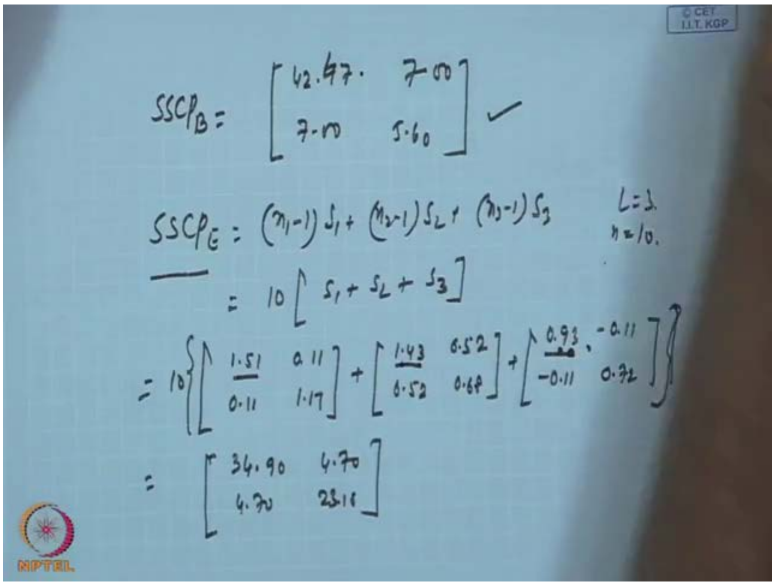
So, what do you require? You require to calculate x 1 bar minus x bar into x 1 bar minus x bar transpose. This is 2 cross 1. This is 1 cross 2. Your resultant matrix will be 2 cross 2 and that is what you want. So, I can write like this now 0.63 minus 0.40 into 0.63 minus 0.40. The resultant matrix will be 0.63 square minus 0.63 into 0.40 minus 0.63 into 0.49, then minus 0.40 square free see that as a square matrix 2 cross 2 and symmetric matrix because this portion here this equal to this. In similar manner, you will be getting x 2 bar minus x bar x 2 bar minus x bar transpose and x 3 bar x bar this transpose.

(Refer Slide Time 10:50)



So, now what will be the resultant value? Resultant value is this n 1 equal to 10. So, this quantity will be 10 into we have already found out the value 0.63 square minus 0.63 minus 0.63 into 0.40 minus 0.40 square plus 10 into whatever matrix value you have get plus 10 into this. So, resultant the ultimate thing that SSCP, SSCP what you will get here SSCP will be?

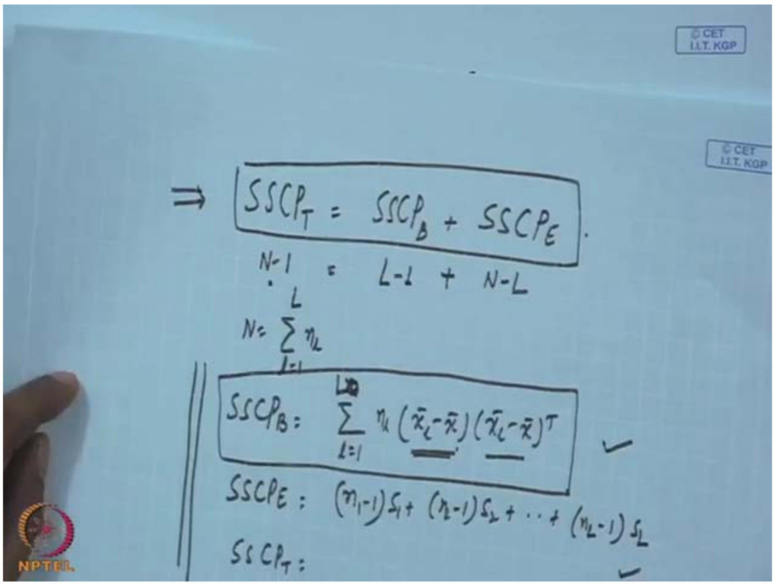
(Refer Slide Time 11:38)



SSCP B will be, we have computed that will be 42.47 and 7.00, then 7.00 and 5.60. So, SSCP B is calculated like this. Now, we want to calculate SSCP E. I told you this will be n 1 minus 1 s 1 plus n 2 minus 1 s 2 plus n 3 minus 1 s 3. In this case, L equal to 3 capital L. So, we will go by like this n 1 into n 3, all equal to n equal to 10. This is nothing but 10 into s 1 plus s 2 plus s 3.

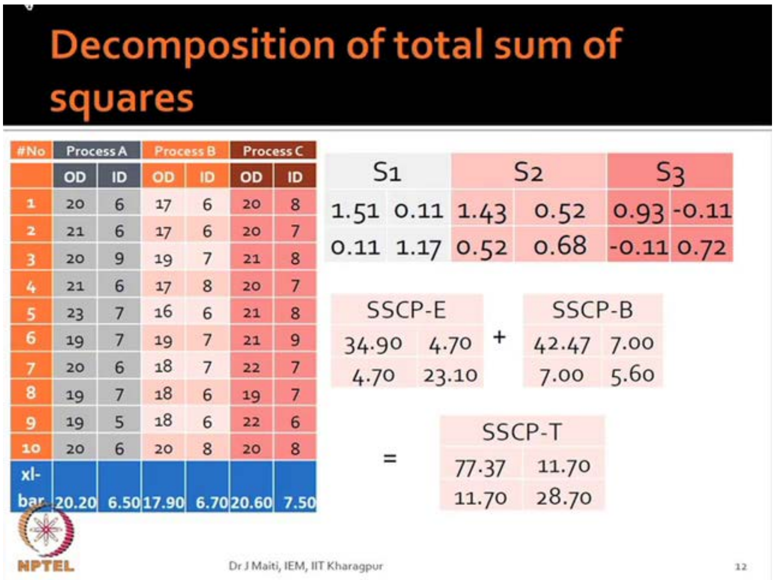
So, I will write down, now equal to 10. What is our s 1? You see our s 1 is this s 2 is this s 3 is this rather this 2 cross 2 matrices so you write down. So, s 1 is 1.51, 0.11, 0.11, 1.17, so plus 1.43, 0.52, 0.52, 0.68 plus 0.93 minus 0.11 minus 0.11, 0.72. This is your SSCP E. If you see that 1.51, 1.43 and 0.93 this you, if you add what are the values you will be getting that 15, 14, the 29 plus 9, almost 33 point something you will be getting. So, this resultant quantity will be your 34.90, 4.70, 4.70, and 23.10. So far what we have discussed if I recapitulate again, I think this is the computation part.

(Refer Slide Time 14:11)



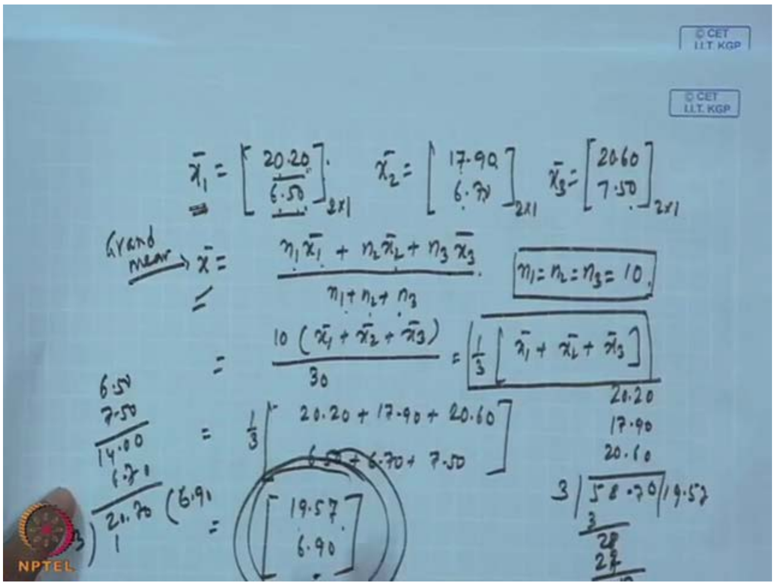
We are considering that we said that in a MANOVA, in one way MANOVA, the SSCP total decomposed into SSCP between population and SSCP error. We also seen that the degrees of freedom is N minus 1 for SSCP total, L minus 1 for SSCP between and SSCP E is N minus L. These are the formulation, computational formula for SSCP B and SSCP E.

(Refer Slide Time 14:45)



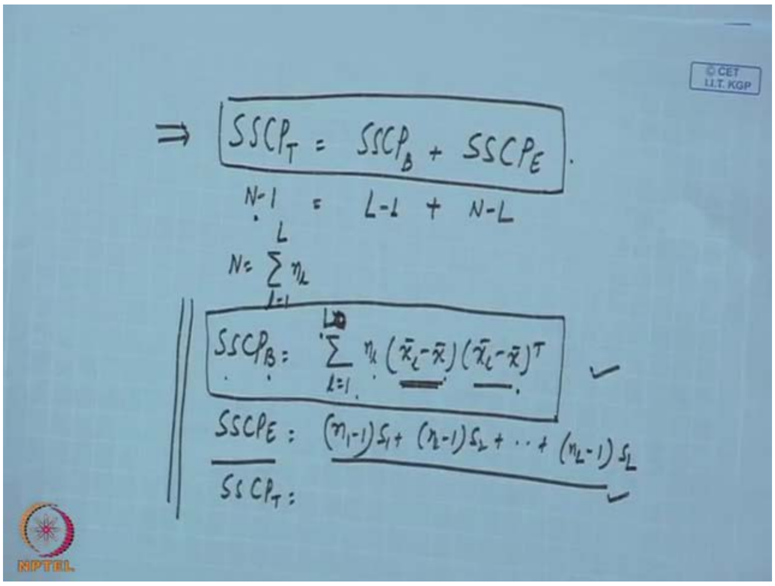
Then, I have given you this, this one, this is the problem. What I said that let us compute SSCP B and SSCP E for the data sheet given for this. What we have done? You have to first find out what are the mean vectors and these mean vectors are like this.

(Refer Slide Time 15:10)



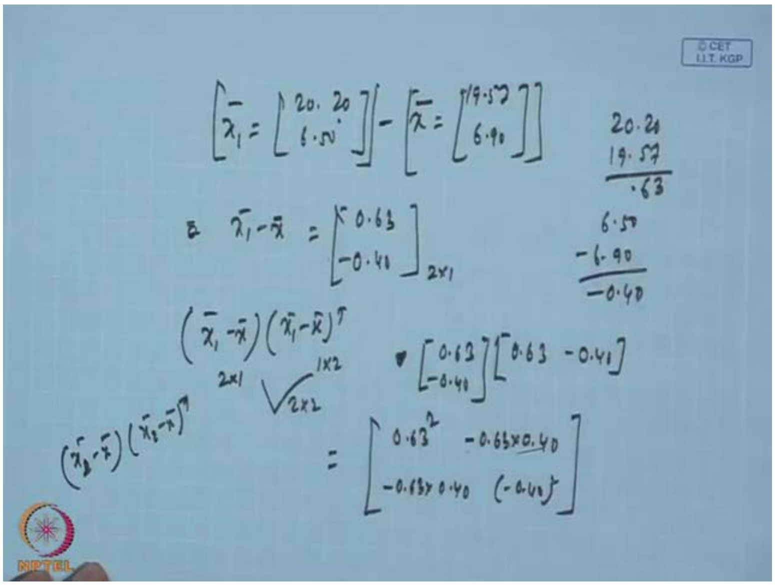
This means vectors are x 1 bar equal to 20.20, 6.50. This is a 2 cross 1 vector. 17.70, 6.70, this is also 2 cross 1 and x 3 is 20.60, 7.50. This is also 2 cross 1. Using these 3 vectors, mean vectors, you have computed the grand mean. We have already computed the grand mean. The grand mean is n 1 x 1 bar plus n 2 x 2 bar plus n 3 x 3 bar by n 1 plus n 2 plus n 3 and in our case it is equal sample in each case. So, finally, it goes to this one that x 1 bar is 1, 1 by 3 times the sum of the 3 mean vectors and as resultant value is this. So, our grand mean is 19.57 and 6.7, 6.90.

(Refer Slide Time 16:09)



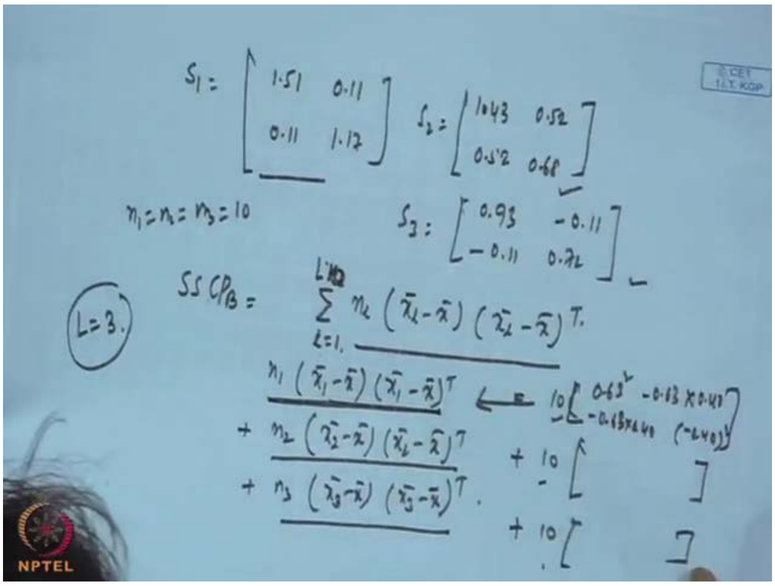
Then, we computed SSCP B. SSCP B using these formulations, SSCP B is l equal to 1 to l capital L n l into this and SSCP E is this value.

(Refer Slide Time 16:30)



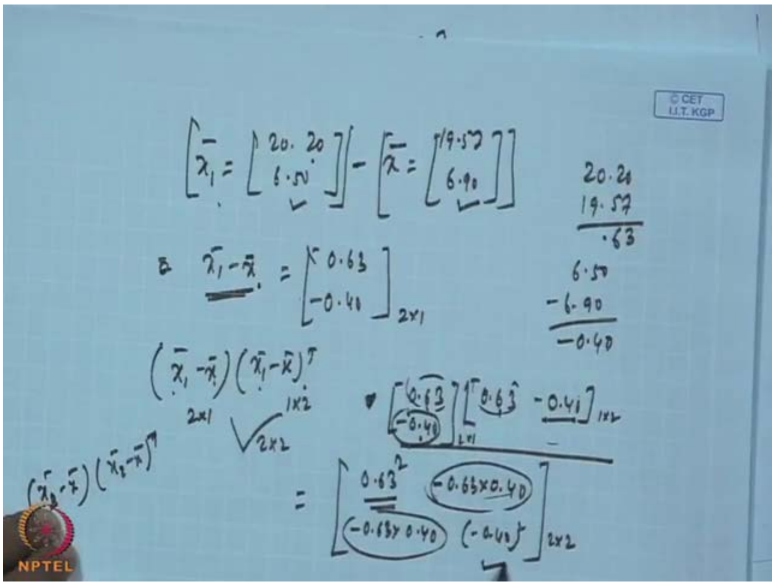
Now, from computation point of view what we have seen that I have shown you one computation because the SSCP B part if you see, it is basically sum of 3, 3 different elements. Those 3 different elements are like this.

(Refer Slide Time 16:49)



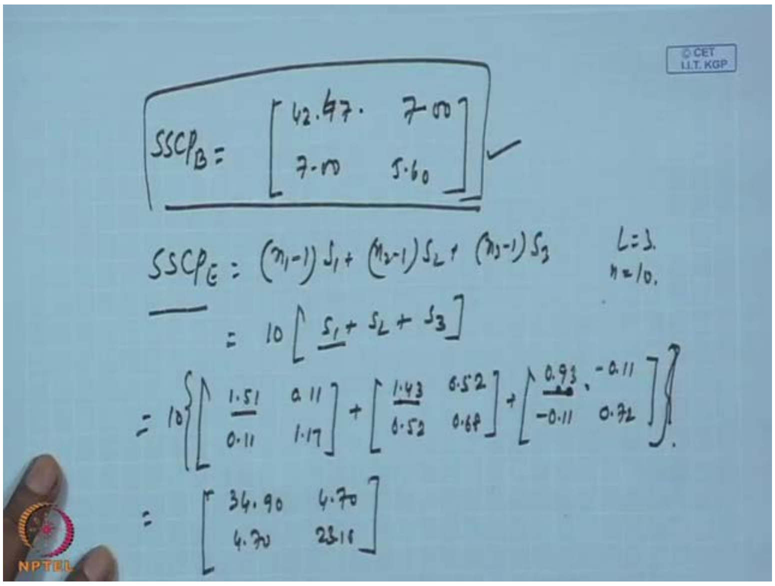
Three different element s that is n 1 into this because see this sum total of this l equal to 1 to capital L. So, our capital L is 3. So, if I put l equal to 1, then this is the quantity if I put l equal to 2, then this will be the quantity if I put l equal to 3, then this will be the quantity. Our case is equal size case, so n 1, n 2, n 3; I am putting 10, 10, and 10. Then you are computing this one. How we are computing this x 1 bar minus x bar x 1 bar minus x bar transpose?

(Refer Slide Time 17:27)



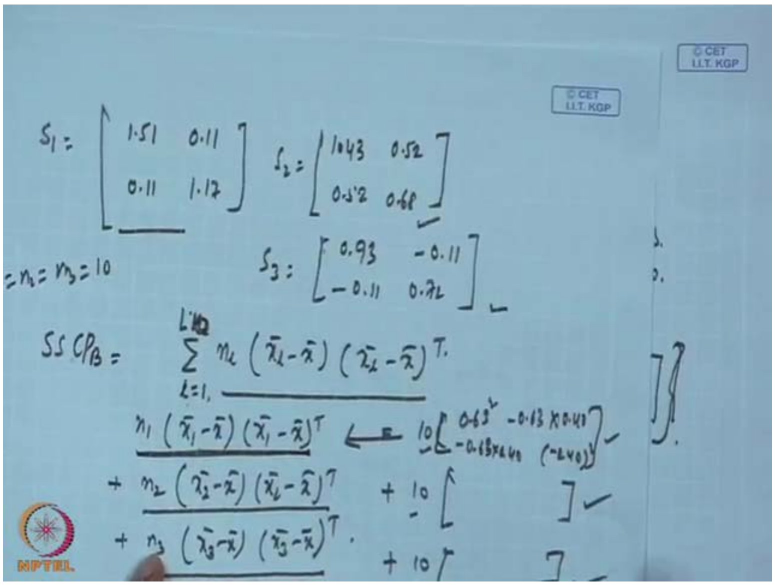
That one computation formulation I have shown you. This formulation is like this, first your x 1 bar is this, x 2, x bar is this, the difference will give you x 1 bar minus x bar, then x 1 bar minus x bar and its transpose. This is what you have written. So, it is a 2 cross 1 matrix. This is 1 cross 2 matrix. You will be getting resultant 2 cross 2 matrix, 0.63 into 0.63, 0.63 square, 0.63 into minus 0.40, this will be the value. Then minus 0.40 into 0.63, this is the value and minus 0.40 into minus 0.40, this is the square value.

(Refer Slide Time 18:11)



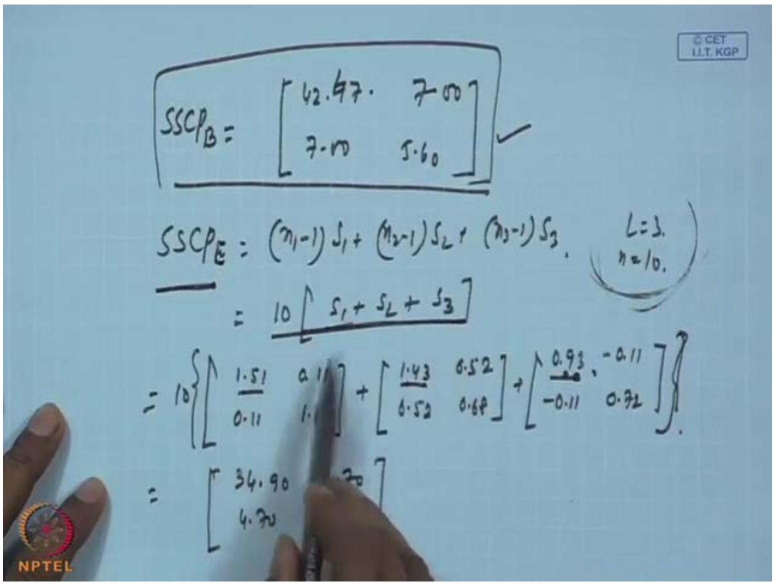
The resultant values are like this. So, for the first one, s 1, this SSCP resultant values is this where all three parts are added.

(Refer Slide Time 18:29)



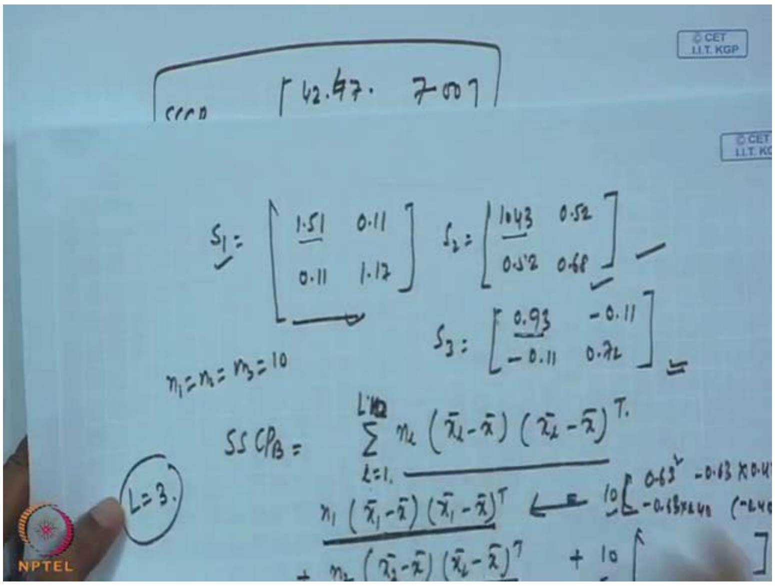
All three parts are added, parts related to this component, this component.

(Refer Slide Time 18:32)



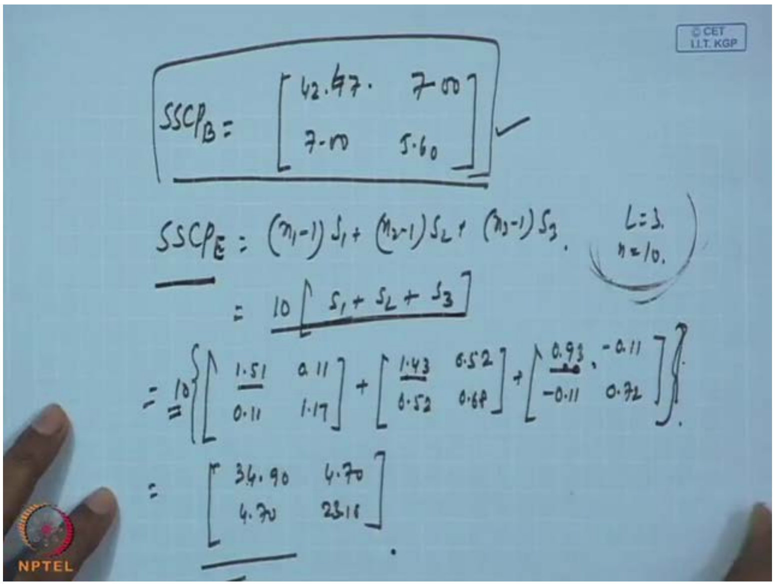
This component when added, you will be getting like this, then what you require to know? We require knowing what will be the sum square cross product error matrix the formula is. It will be the sum total of n 1 minus 1 s 1 n 2 minus 1 s 2 n 3 minus 1 s 3 because we have three population n equal to 10. So, our value is like this.

(Refer Slide Time 18:59)



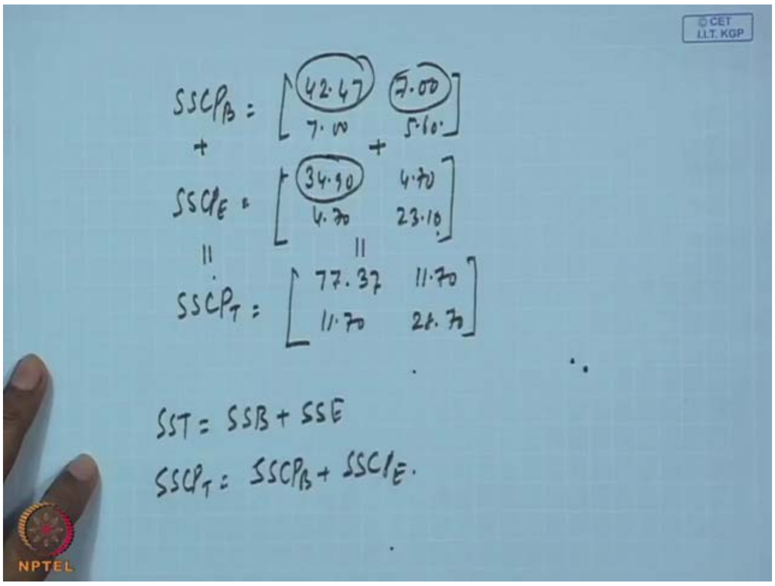
We have started with this that our s 1 is this one, s 2 is this matrix, s 3 is this matrix. So, you are adding 1.51, 1.43 plus 0.93.

(Refer Slide Time 19:10)



This will be nothing but 3.49 and if you multiply by 10, it will be 34.90. In similar manner, you are getting SSCP E.

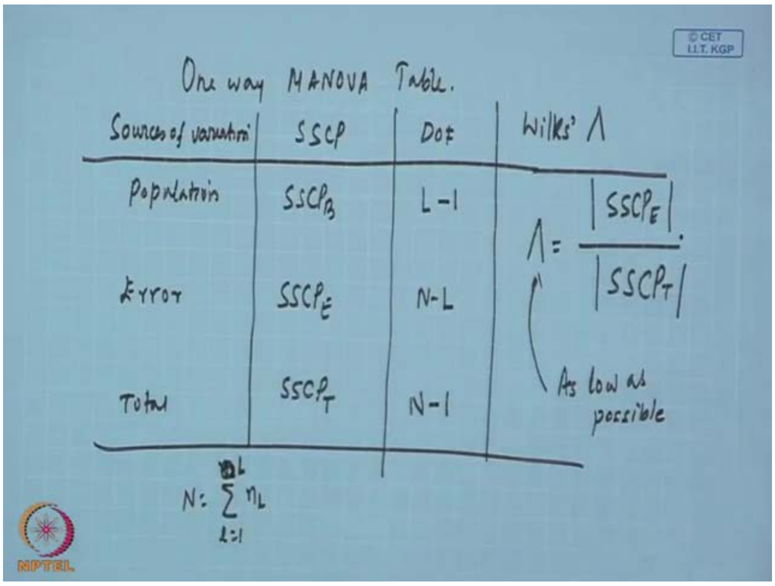
(Refer Slide Time 19:23)



So, what we have now with us? We have SSCP B matrix, which is 42.47, 7.00, 7.00, and 5.60. You have SCP error matrix, which is 34.90, 4.70, 4.70, and 23.10. So, you know your SSCP total now, this plus this resultant will be 77.37; now 7 plus 4.7, 11.70, and 11.70. Then 23.10 plus 5.60, it will be 28.70. This is the decomposition of total matrix, this equal to this plus this or this equal to this plus this.

So, in MANOVA, you have to decompose SST sum square total sum square between population and sum square error and in MANOVA, what you are doing, SSCP total is decomposed into SSCP between plus SSCP error. So, because of more than one variable, your computation will be in the matrix domain. Once you decompose the SSCP T into the parts, you are in a position to calculate the MANOVA table.

(Refer Slide Time 21:15)

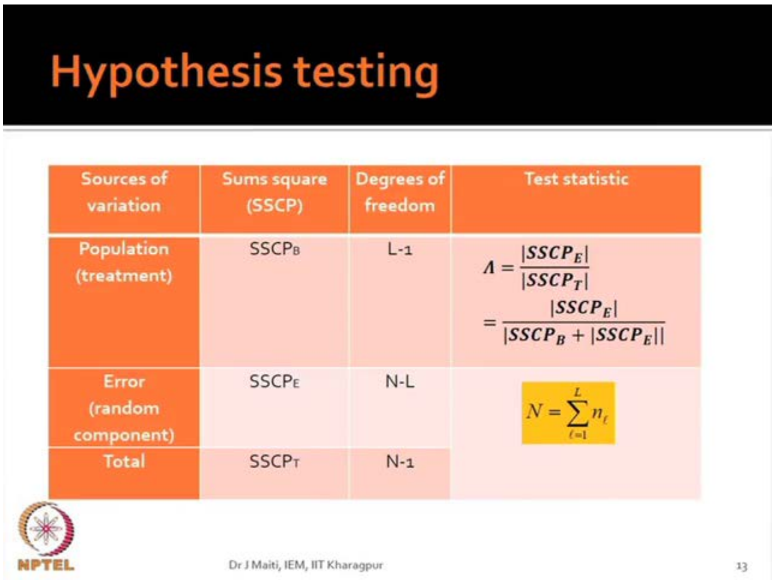


Please remember this is one way MANOVA. We were talking about one way MANOVA because we have taken only population of some kind one kind that population one to population and not some other kind of factors are considered in this table. You have to write down what is the source, similar manner sources of variation, so definitely your population is one source. Then error is another one, then total is coming, then you are writing SSCP matrix. So, our case is SSCP between SSCP error and SSCP total. Then you will be writing degree of freedom and I told you that N is sum total of l equal to 1 to capital L n l; this is the case.

So, in that case, our degree of freedom for total will be N minus 1, for SSCP B, L minus 1 and this will be N minus L. Then in MANOVA, you find what the mean square between mean square, error is, all those things here we will not use like this. We will create a matrix called Wilks lambda. Wilks lambda, this is basically a ratio; this wiki wills lambda ratio will be computing where Wilks lambda is determinant of SSCP error by determinant of SSCP total and what you want? We want that the SSCP error, this determinant must be as low as possible compared to the total.

So, this quantity will be as low as possible or otherwise we can say as small as possible, but there is certain distributions available using Wilks lambda. If you go by multivariate books, you will be finding out that different combinations, different sampling distributions are available.

(Refer Slide Time 24:17)

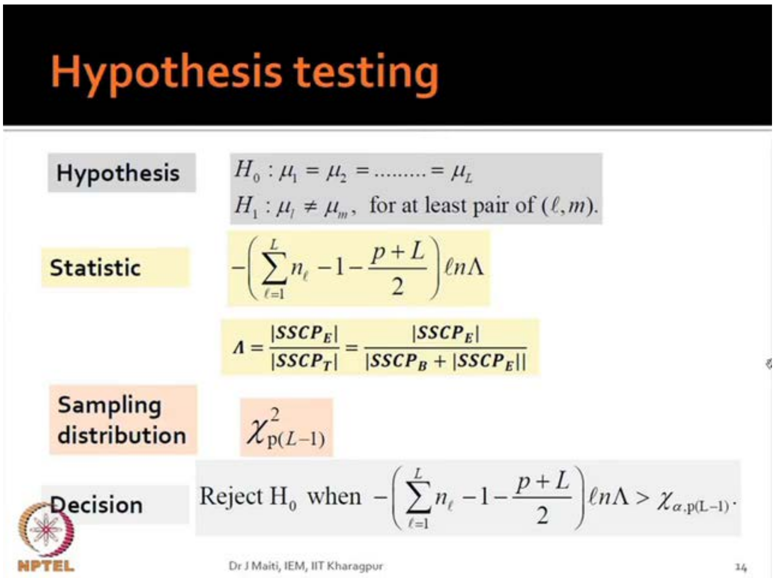


Hypothesis testing

Sources of variation	Sums square (SSCP)	Degrees of freedom	Test statistic
Population (treatment)	$SSCP_B$	L-1	$\Lambda = \frac{\lvert SSCP_E \rvert}{\lvert SSCP_T \rvert} = \frac{\lvert SSCP_E \rvert}{\lvert SSCP_B + \lvert SSCP_E \rvert \rvert}$
Error (random component)	$SSCP_E$	N-L	$N = \sum_{\ell=1}^{L} n_\ell$
Total	$SSCP_T$	N-1	

We will show you a general one in this class.

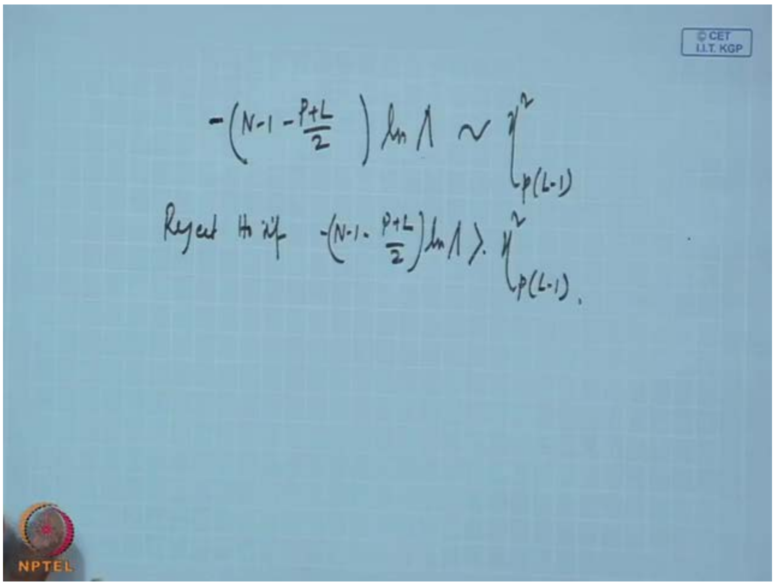
(Refer Slide Time 24:30)



Now, with using whatever we have developed so far, based on this, we want to test the hypothesis. What are the hypotheses? First hypothesis is null hypothesis and alternative hypothesis in our case. Null hypothesis is there is no difference in the mean vectors for all above lessons and alternative hypothesis is at least one pair of populations is different in terms of their mean vectors. We will use a statistics. This statistics is minus within bracket l equal to 1 to L n l; this is capital N minus 1 minus p plus L by 2 log lambda.

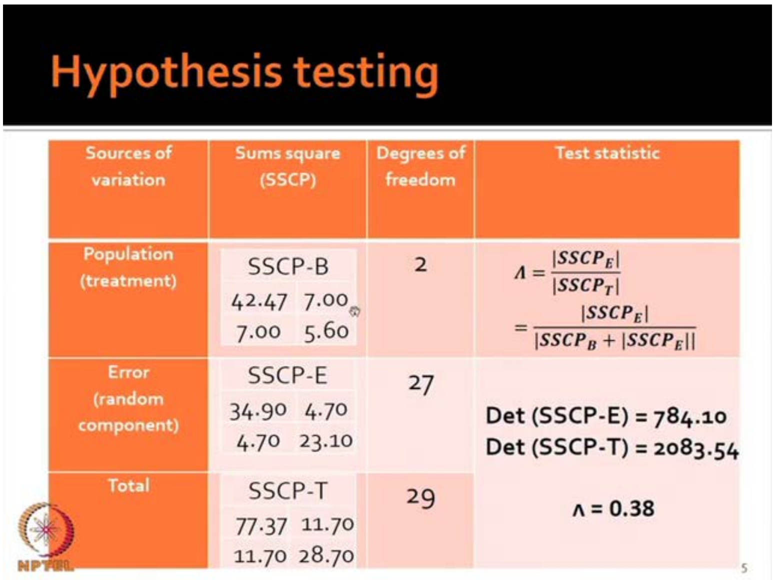
Lambda, you know that how to compute lambda. Then this statistics follows chi square distribution with p into L minus 1degrees of freedom. Then what is our statistics?

(Refer Slide Time 25:32)



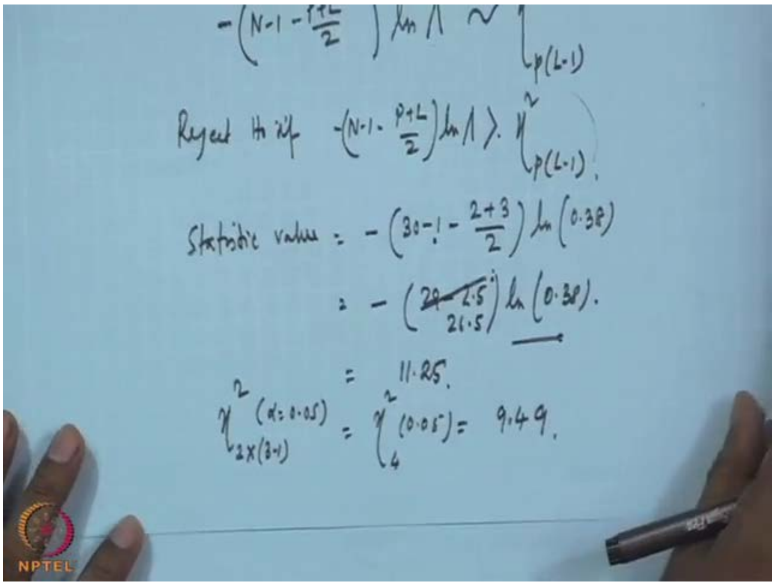
Our statistics is that N minus 1 minus p plus L by 2 log lambda lagging lambda. What we are saying this follows chi square p into L minus 1. What will be your decision? The decision will be if I think this is minus, 1 minus is there, because Wilks lambda this is a low value keeping minus here making it positive. So, our hypothesis is reject H 0 that may no difference among the population mean vectors if minus N minus 1 minus p plus L by 2 log lambda, this value greater than chi square p L minus 1. We will reject null hypothesis under this condition.

(Refer Slide Time 26:51)



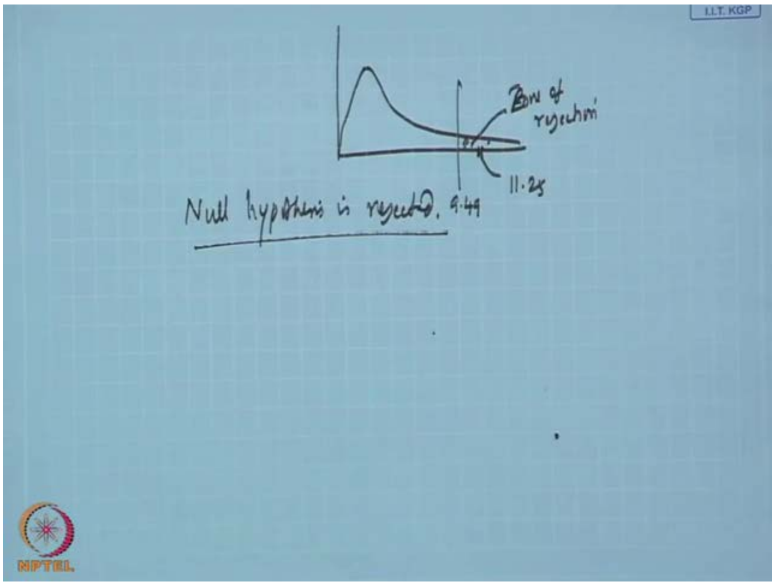
Now, we will see that for the problem given, what is the status? You see the slide. So, SSCP B, you have already computed 42.47, 7 like this. SSCP E is also computed. SSCP total is also computed. Our capital N is 30. So, these degrees of freedom, I hope there will be no problem, 29, 27 and 2, and then test statistics is lambda. If you would, if you compute determinant SSCP E, this one 34 into 23.10 minus this square, you will be getting 784.10. Total case, total case will be getting this one, 77.37 into 28, this will be getting this one. So, your determinant of SSCP E is 784.10, determinant of SSCP total is 2.83.54. Now, if you take the ratio, you will be getting lambda equal to 0.38. Now, what will be the log lambda?

(Refer Slide Time 28:11)



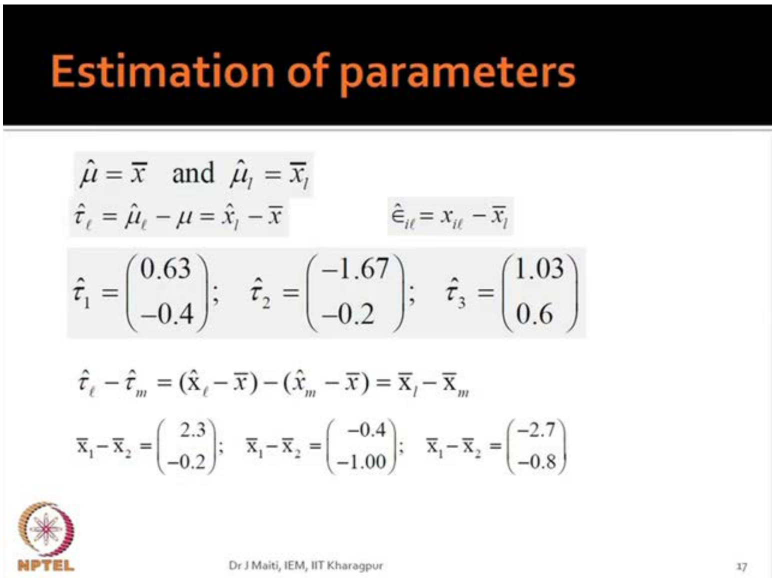
Our statistics value will be minus N is 30 minus 1 minus p is 2 plus L is 3 by 2 log of 0.38, this one will be minus. This is 29, 29 minus 2.5. So, that means the resultant will be, this resultant will be 26.5 into log 0.38. Find out this value. I think the resultant quantity, this will be 11.25. Now, we require finding out, what we require to find out chi square p L minus 1. p is 2, L minus 1 is also 3 minus 1 that is 2 and you have to take a alpha value. Let alpha equal to 0.05 and this one is chi square 4, 0.05, this value is 9.49.

(Refer Slide Time 29:42)



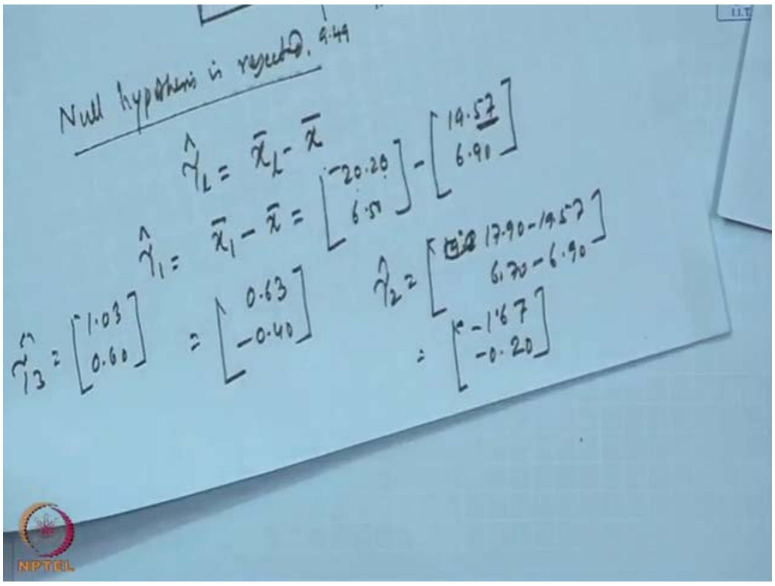
The statistics value is more than the tabulated value. The case is coming like this. This is my tabulated value, 9.49 and my computed value is here, 11.25. This is the zone of rejection. So, null hypothesis is rejected or some space. Now, this null hypothesis is rejected. There is alternative hypothesis that there is some difference among the population.

(Refer Slide Time 30:32)



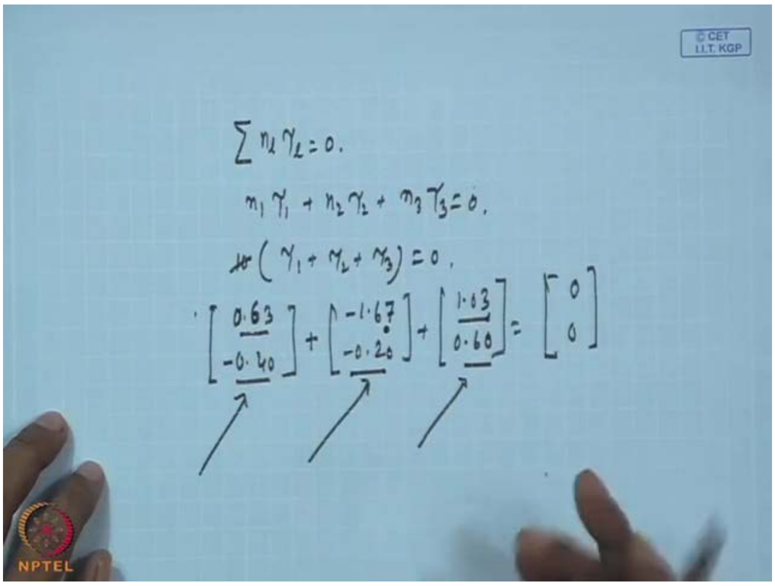
This means as null hypothesis is rejected, so we can say that population effects are there. Now, we want to know the population effect. First, you see the slide. I think if you can recall my last class, then you will find out that we have computed population effect in terms of tau. Tau 1 is the population first population effect, tau 2 second population effect, tau 3 third population, so we can calculate tau. How we calculate tau l?

(Refer Slide Time 31:08)



Tau l is x bar l minus x bar. So, if you want to compute tau 1, if I say this is the estimated value, then 1, one then you have to write down what is x 1 bar minus x bar. I think we have already seen the x 1 bar and values are there with you, x 1 bar and x bar values. If x 1 bar is 20.20, so your value is 20.20, 6.50 minus x bar is, what we got x bar, 19.something, I think 19.57 and 6.90. What will be this value? Then, this value 20.20 will be 0.63 minus 0.40. So, similarly, tau 2 value, you will be getting tau 2 value also find out in the same manner that 19. Tau2 value will be 17.90 minus 19.57 and 6.70 minus 6.90, this value will be, and tau 2 value will be minus 1.67 minus 0.20. Similarly, your tau 3 value will be 1.03, 0.60.

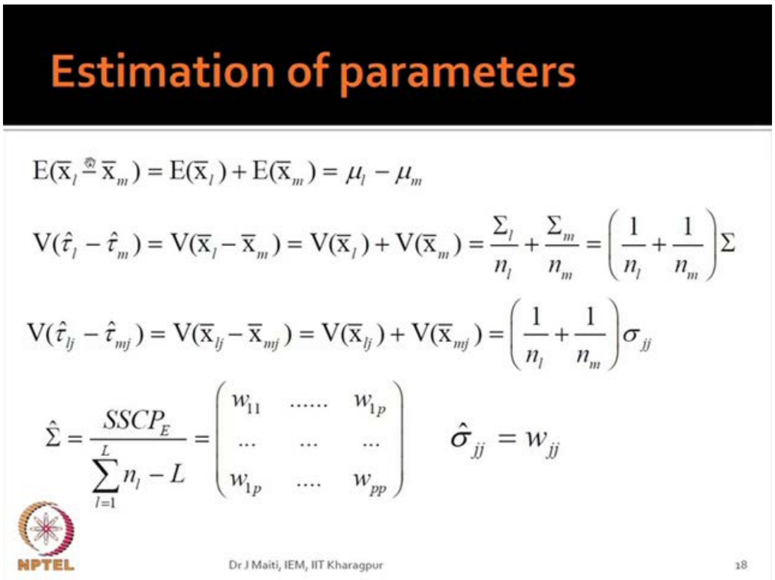
(Refer Slide Time 33:17)



I told you in beginning of ANOVA as well as in MANOVA, we say that the parameters, populated parameters are basically that tau l basically are having a certain property that property is sum total of n l tau l, this will be equal to 0. So, that means in this case, n 1 tau 1 plus n 2 tau 2 plus n 3 tau 3, this will be 0. Our n 1 n 2 n 3 equal that is 10 plus, so tau 1 plus tau 2 plus tau 3 that equal to 0.

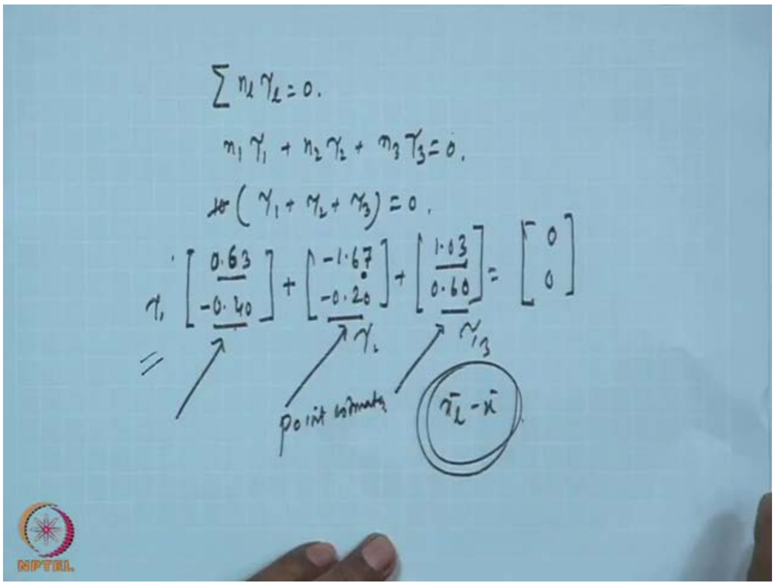
Now, see what is given here. What is tau, 1.03 this, not tau 1, this is tau, 0.63 is tau 1 minus 0.40, so I can say cancel 10 plus 2 is minus 1.67 minus 0.20 plus tau 3 is 1.03, 0.60, what the value is coming? You see 0.63 plus 1.03, this is 1.66 minus 1.67, it is 0 basically that because of rounding error, this space comes, so this is 0; similarly, 0.40 minus 0.20 that minus 0.60 plus 0.60. So, we found out now the effects of each of the population and well also found out collectively there is a difference. Now, we want to find out which variables are making the difference.

(Refer Slide Time 35:23)



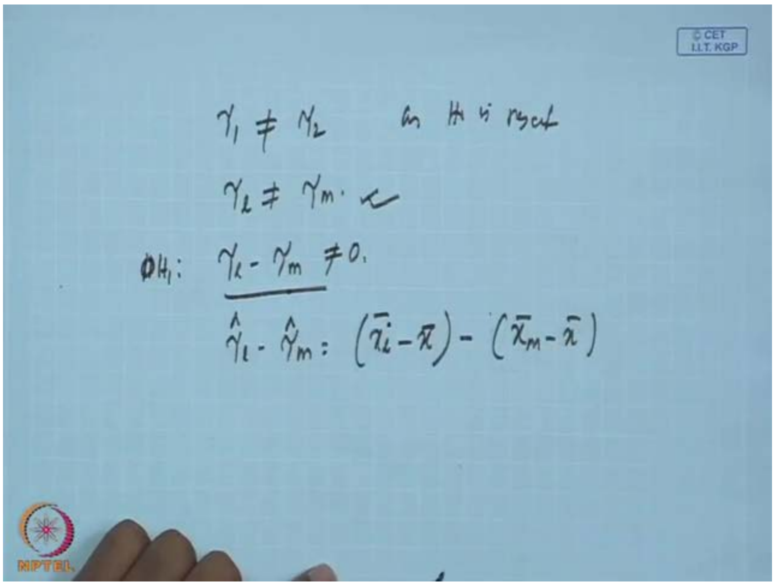
That is our next objective. So, in order to do so I will first show you the slide; here I am showing you something more. What is that?

(Refer Slide Time 35:45)



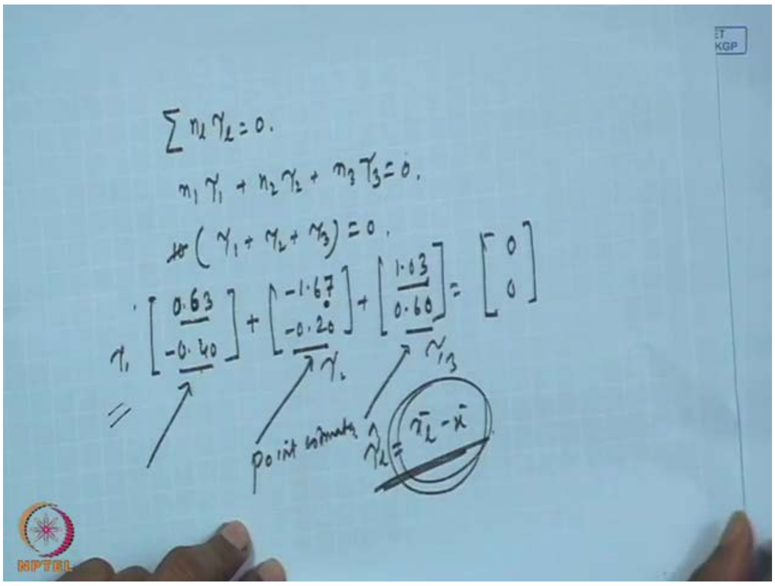
You have already found out that this is your tau 1, this one is tau 2. This one is tau 3. Other one is tau 3. So, there are all known as point estimates. Now, we want to know the interval estimate of all those things. So, if you require, you want to find out the interval estimate, then you will definitely require finding out that x l bar minus x bar, what will be the estimate, all those things, so we have seen earlier.

(Refer Slide Time 36:24)



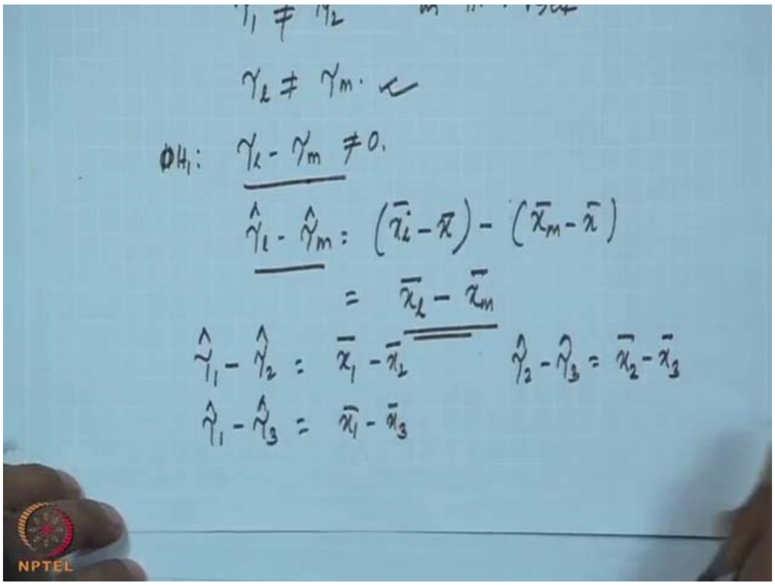
In this class, I will show you that as we are saying that tau 1 may be different from tau 2 as collectively H 0 is rejected. So, I want to create a situation that tau l not equal to tau m, then tau l minus tau m that not equal to 0 that is your alternate hypothesis. So, what I said earlier that for individual tau l, you can go for interval estimation. Also, we are here going for the interval estimation of difference because that is what we want; H 1 is accepted. Now, tau l minus tau m cannot be written like this that x l bar, I am taking the estimate, x i bar minus x bar minus x m bar minus x bar. Let us see.

(Refer Slide Time 37:44)



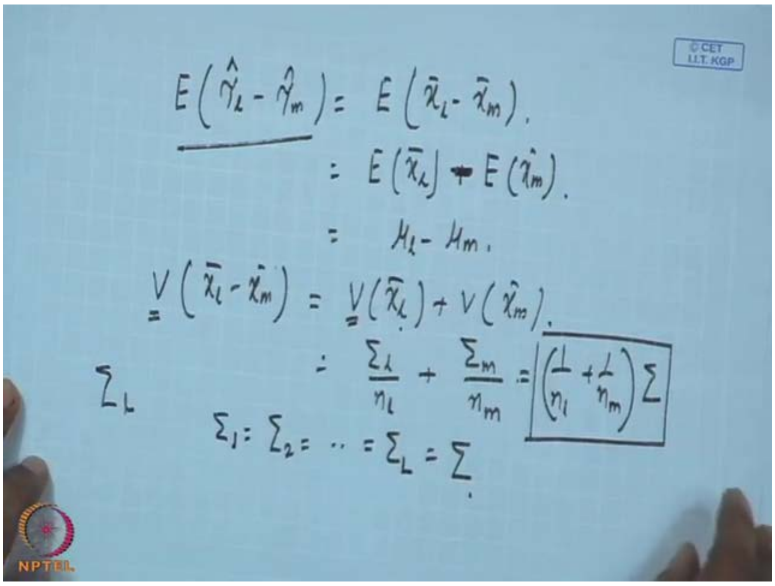
What I said that tau 1 l base every population mean minus the grand mean.

(Refer Slide Time 37:51)



So, then this one is tau x l minus tau x m x bar, x bar is cancelling out. So, if I want know the point estimate of these, this is nothing but this one. So, you have already seen what this x 1 bar is. Suppose I want see the value tau 1 cap minus tau 2 cap, then it will be x 1 bar minus x 2 bar, yes or no? So, similarly, tau 1 cap minus tau 3 cap equal to x 1 minus x 3 bar and tau 2 cap minus tau 3 cap is equal to x 2 bar minus x 3 bar and using the matrix values, you can find out what are the difference value, which is not a big problem. You can easily find out. So, once you get these values, it will be point estimate.

(Refer Slide Time 39:10)

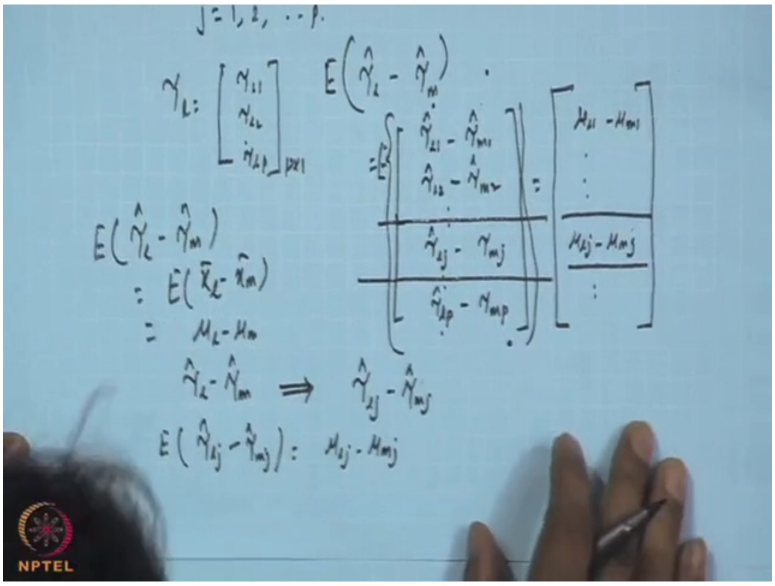


Now, I want to find out the interval estimate of the same. So, what you require to do for interval estimation? You require to know what is the expected value of suppose tau l estimate minus tau m estimate; this is nothing but expected value of x l bar x m bar what you have seen earlier. We also know that this nothing but expected value of x l bar minus x m bar what you seen earlier. We also know that this is nothing but both expected value of x l bar minus expected value of x m bar because the two populations are independent. So, this will be mu l minus mu m.

Similarly, you require computing variance component of x l bar minus x m bar, which will be variance of tau l minus tau m bar tau l minus tau m cap. Now, variance of this is equal to variance of x l bar plus variance x m bar. We have seen earlier this one. So, this is what the variance of x l bar is. x l bar is vector quantity. So, it is coming from the population l. Now, population l covariance matrix is capital sigma l and it is the mean covariance of the mean value. So, it will be sigma l by n l plus capital sigma m by n m.

Now, you see that I have, although I have written here variance, but you can write also co variance. Whenever there are more than one variable, we say variance structure means variance plus co variance structure. Now, in MANOVA, the assumption is all the population variable, it is equal this equal to capital sigma. So, then we can write this one as n l plus 1 by n m into sigma, because sigma l equal to sigma m equal to sigma because all the population have equal covariance. So, your variability part is taken care of by this covariance part. I think in vert link square time you have also seen this one.

(Refer Slide Time 42:31)

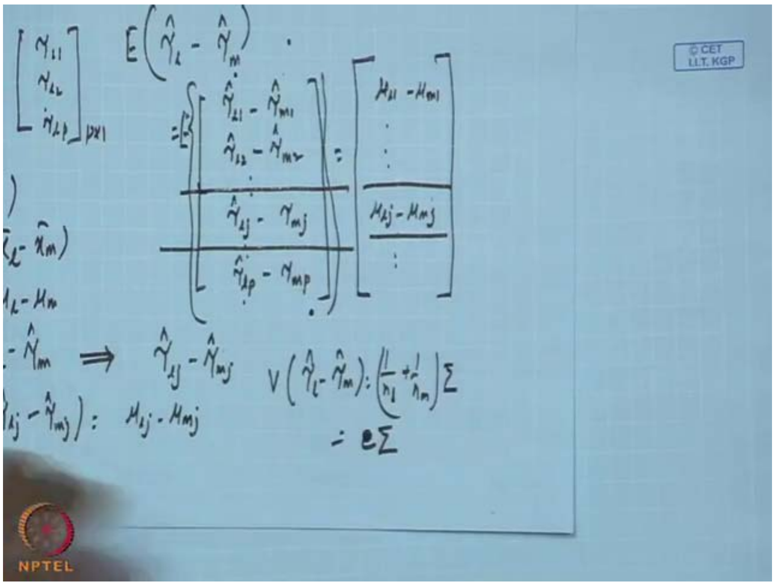


Suppose we are interested to this is from the overall that population f x point of view, you are finding out the difference. Now, there is j equal to 1 to p variables. As I told you that I want to know what are the variables making the effect, so that means tau l is nothing but your tau l 1, tau l 2 like this tau l p, p cross 1. When I say that tau l cap minus tau m cap that means I say here this is tau this l 1 minus this tau m 1 tau l 2 minus tau m 2. So, like this, I can get one point where tau l j minus tau m j, then slowly up to the last variable tau l p minus tau m p.

If I say that tau l minus tau m not equal to 0 that means I am saying that this vector tau l 1 to tau l p is not equal to tau m 1 to tau m p. So, we have seen the variance covariance structure of this one. What I can say that the variance structure of tau l cap minus tau m cap that you have seen this is the covariance structure expected value also you have seen, so I want to know. Also, suppose if I take a particular variable here tau l j minus tau m j, what will be the covariance variability and mean value for this? So, when I expected value of tau l cap minus tau m cap, it is nothing but expected value of x l bar minus x m bar, which we say it is nothing but mu l minus mu m.

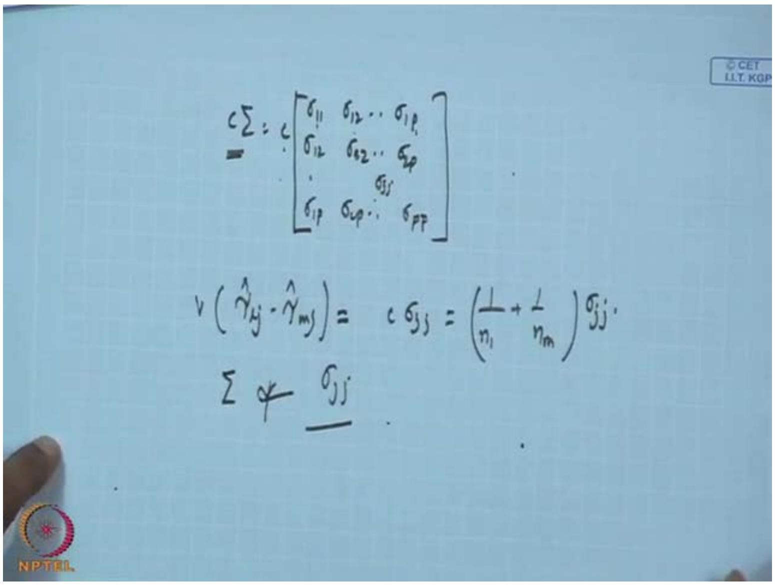
Then, this one is nothing but if we write down the expected value of this, expected value of this or expected value of this totality, then what you will get here you will get here, mu l 1 minus mu m 1, like this here, you will be getting mu l j minus mu m j. Yes or no? So, for a particular variable case, the difference if I see and then the expected value if I want to find out that mean what I am interested. Now, instead of tau l cap minus tau m cap, I am interested now with a variable that is tau l j minus tau m j. So, expected value of tau l j minus tau m j, this is nothing but mu l j minus mu m j; this is third quantity.

(Refer Slide Time 45:53)



Then, what will be the variability part here? Variability of tau l minus tau m; that one you have seen as 1 by n l plus 1 by n m into sigma. So, we will find out the value. So, if I write down that this one, this one, this quantity as c, so it is basically c sigma, c sigma.

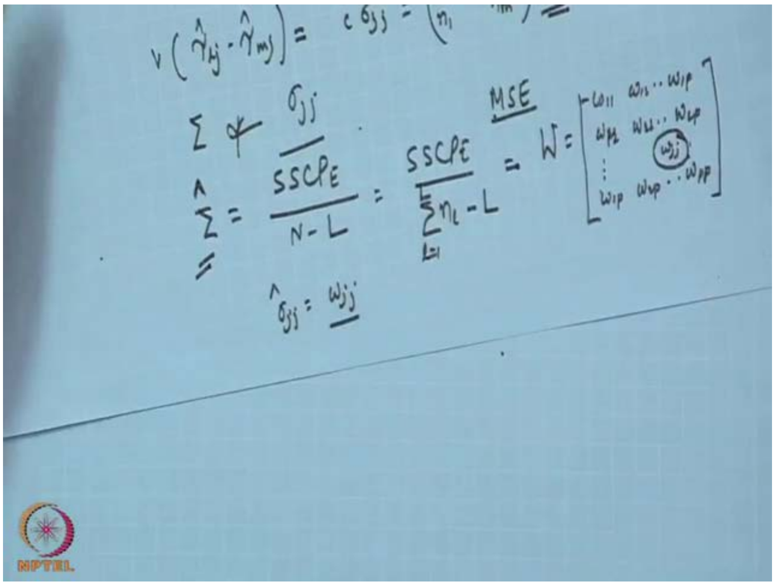
(Refer Slide Time 46:26)



Then, what is happening? Then you are getting this co variance; c sigma is something like this c into sigma 11, sigma 12 that sigma 1p, 12, sigma 22, sigma 11, sigma 12, sigma 2p, where we have started that c sigma is c into sigma 11, sigma 2p. Then sigma 2p is coming like this sigma 1p, sigma 2p, then sigma pp, somewhere in between there will be sigma jj, we have taken this, sigma jj, somewhere in between there will sigma jj.

So, for this variable, what we have taken our case is like this; for this tau l j cap minus tau m j cap that is mu l j minus mu m j, this one the variability part tau l j minus tau m j, this will be definitely c sigma jj. There is absolutely no problem for you. That means this is 1 by n l n m into sigma jj, but please remember, we do not know sigma. What is the capital sigma? We do not know. That means we do not know, we do not know sigma jj also.

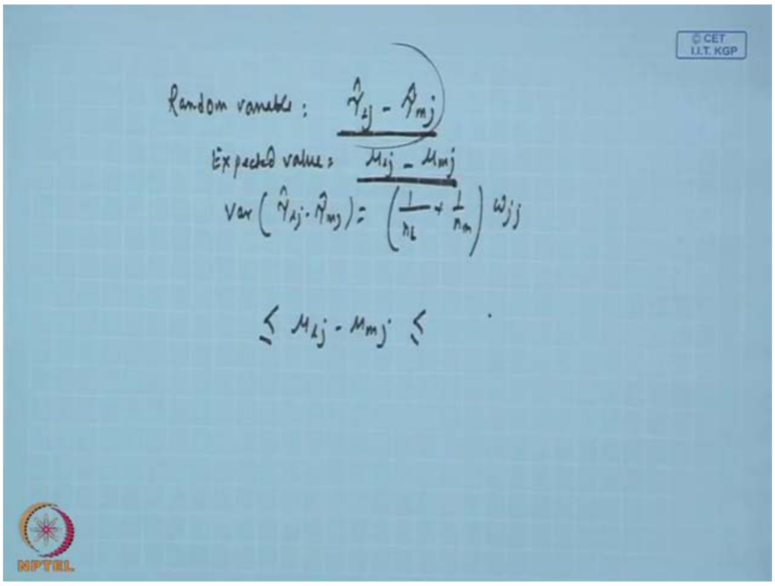
(Refer Slide Time 48:11)



So, what will be the estimate of sigma cap? In ANOVA, we have seen we said that I think you can remember that MSE we talked about that MSE is the estimate of sigma square. So, here also you have found out SSCP E, if you divide it by degrees of freedom, this is the estimate of sigma.

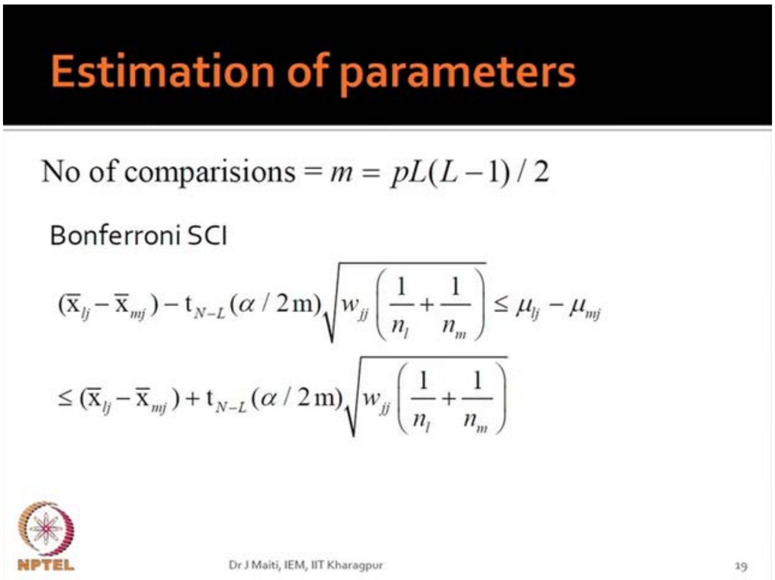
So, SSCP E divided by sum total l equal to 1 to capital L n l minus L, this is your estimate. Suppose if we write down like this as W, which is which is nothing but w 11, w 12, then w 1p, w 12, w 22, w 2p, so like this, you will be getting w 1p, w 2p, w pp in between somewhere w jj. So, this w jj as I told you that sigma cap is estimated like this, now you are getting w jj also; this w jj will be estimate of sigma jj. So, you can say sigma jj cap is equal to w jj.

(Refer Slide Time 49:58)



So, essentially what you got? Now, you got very interesting things; one is your random variable that is tau cap l j minus tau cap m j and its expected value also you got, mu l j minus mu m j that is the mean value. You also got the variance of tau l j minus tau m j cap that one is 1 by n 1 plus l n l plus 1 by n m into w jj. So, you have everything now. Now, what you want to know? We want to know the interval estimate of this that is what we have started. So, interval estimates of, again please remember, this is a random variable, the mean interval estimate for the mean we want to compute not that interval estimate of this. With the help of this, we want to compute the interval estimate of its mean value. So, what you want l j minus mu m j, you want to create an interval.

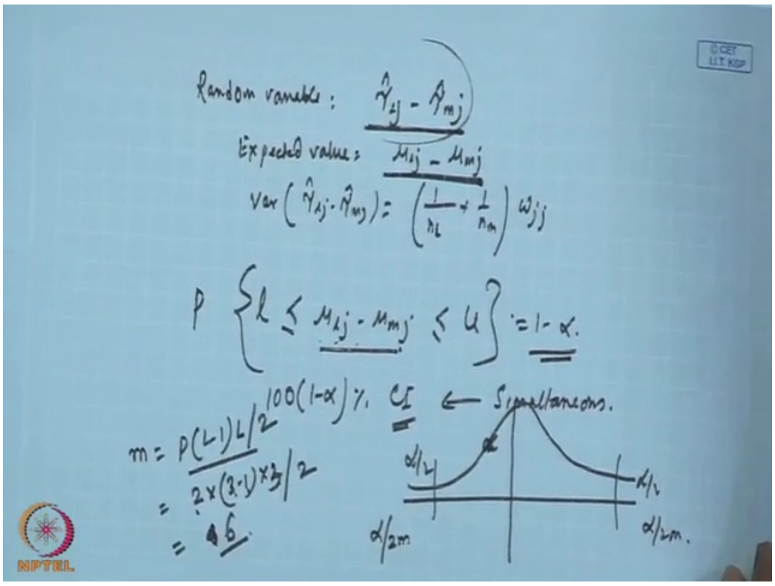
(Refer Slide Time 51:34)



So, what will be that interval? You see that we will approach will go by Bonferroni approach. Why we are interested in Bonferroni approach? Here, you see this slide, see the slide here. Now, how many comparisons are possible? How many comparisons are possible from the variable point of view? You see there are L populations. You are comparing two at a time.

Then, L into L minus 1 by 2 will be the number of comparisons because if there are 3 populations, 3 into 3 minus 1 by 2 that 3 comparisons. If there are 4 populations, then 4 into 4 minus 1 by 2 that is 6 comparisons possible, 6 comparisons possible will 6 mean vectors comparisons. Now, we are comparing again the variables. So, the p values here coming, how many variables are there? p variables are there. So, in total, you have m comparisons, but what you want?

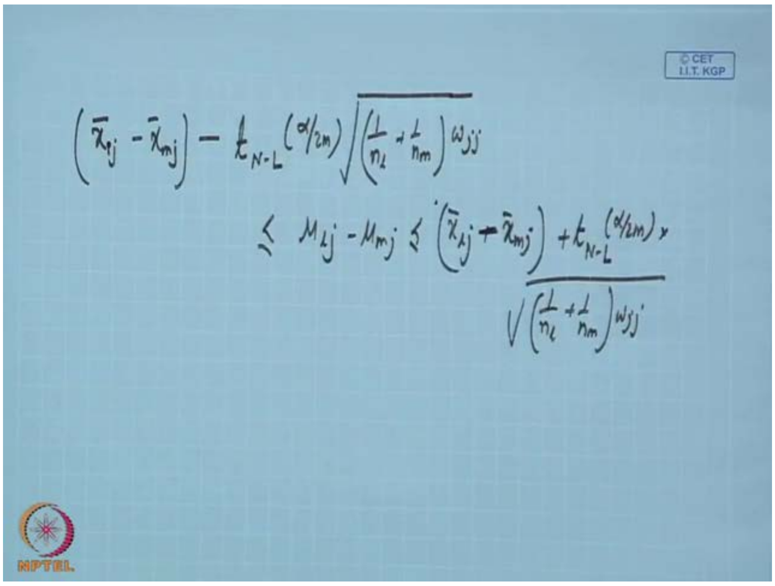
(Refer Slide Time 52:49)



You want an interval here, l and u in such a manner that this will probability of this less than l less than is equal to, this equal to 1 minus alpha in such a manner that you will achieve 100 minus 1 per cent confidence interval that should be simultaneous. That should be simultaneous. Now, we have seen earlier. Now, there are two variables, two variables, two population cases that there are two approach maximization lama, your Bonferroni approach and Bonferroni approach is easier. So, we are considering Bonferroni approach. In Bonferroni approach, what happened odd? It is basically first saying that how many comparisons are there which means how many mu l j minus mu m j that is m?

m is equal to p L minus 1. L in our case, it is p is equal to 2, L minus 1 will be or 3 minus 1 into 2 divided by 2 divided by 2. So, it is 4 p is 2, 3 minus 1, this 2, so 4 comparisons are possible. 3 comparisons, p is, p is how much? p is 2, L is three. So, this is 6 comparisons. L is 3, so 6 comparisons. Simultaneously, we are to make sure our alpha it is a two tail case t distribution. We will be using two tails, this side alpha by 2, this side alpha by 2, but what happened? We have 6 comparisons that Bonferroni says you divide it either equally or with certain vertex, you are dividing actually. So, that means alpha by 2 m and this side also alpha by 2 m, so then this is the case.

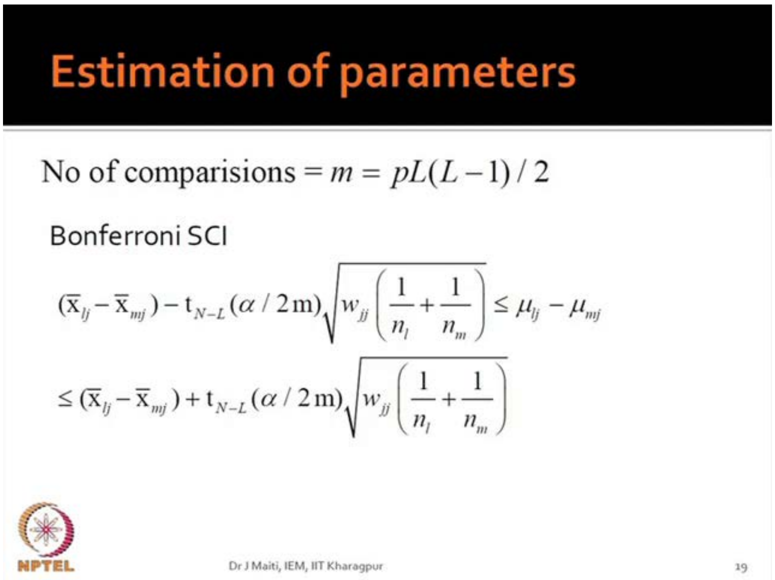
(Refer Slide Time 55:12)



So, if this is the case, then what you are required to know? Now, that means we are creating a t statistic here and t statistic; the degree of freedom will be N minus L. You are considering alpha by 2 m k because m, m comparison is there and you are also multiplying this with what the variable be component, what is this variable be component 1 by n l plus 1 by n m into w jj. This one will be subtracted.

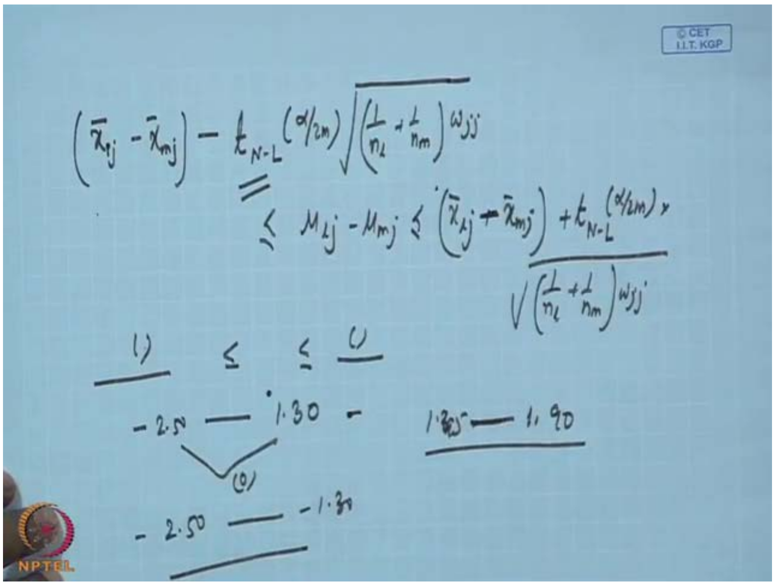
That portion would be subtracted from x l j bar minus x m j bar, this minus this. This is same manner the way we have done earlier. Also, we know the distribution, and then the critical value and we multiplied with the variability. This will be less than is equal to mu l j minus mu m j less than is equal to x bar l j minus x bar m j plus t N minus L alpha by 2 m into the variance part. So, in this manner, you have to find out the difference.

(Refer Slide Time 56:56)



That is what I said the confidence interval for the differences. Then you find out who is difference content of mu that means under null hypothesis, there is no difference; so mu l j minus mu m j will be 0.

(Refer Slide Time 57:16)



If you find out any interval, any interval this side or that side, this value plus this value, it contains 0, then that variable is not differentiative. Suppose the difference will come for a particular variable, let it be minus 2.50 to 1.3, then see within in between, there is 0. So, this variable is not creating the difference. Suppose some value is like this 2.50 to minus 1.30, this is a differentiating variable or 1.45 to 1.90, that is correct, that is also different; 0 is not there. This means null hypothesis is not satisfied with this interval because we have created t distribution t values mean that 0 will be there in between.

I think this is what is one way MANOVA and you have seen in totality like this that you are comparing several population means. Then fine of coming to the conclusion that population means differ, then you are trying to find out that which of the variables are causing this difference. You are going for the interval estimation of each of the variable for different pair of comparisons, and then find out which of the variables are creating the difference. Next class, I will show you one case study.

Thank you.