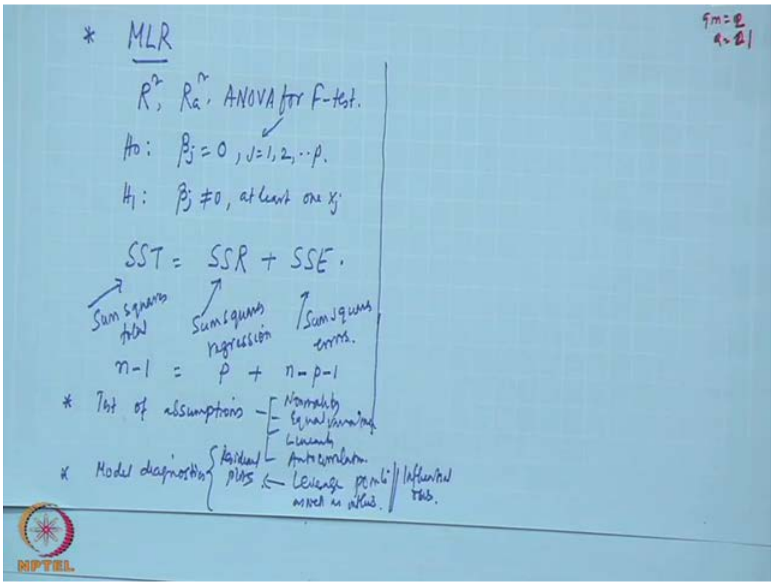
**Applied Multivariate Statistical Modeling**
**Prof. J. Maiti**
**Department of Industrial Engineering and Management**
**Indian Institute of Technology, Kharagpur**

**Lecture - 29**
**Multivariate Linear Regression- Model Adequacy Test**

(Refer Slide Time: 00:29)

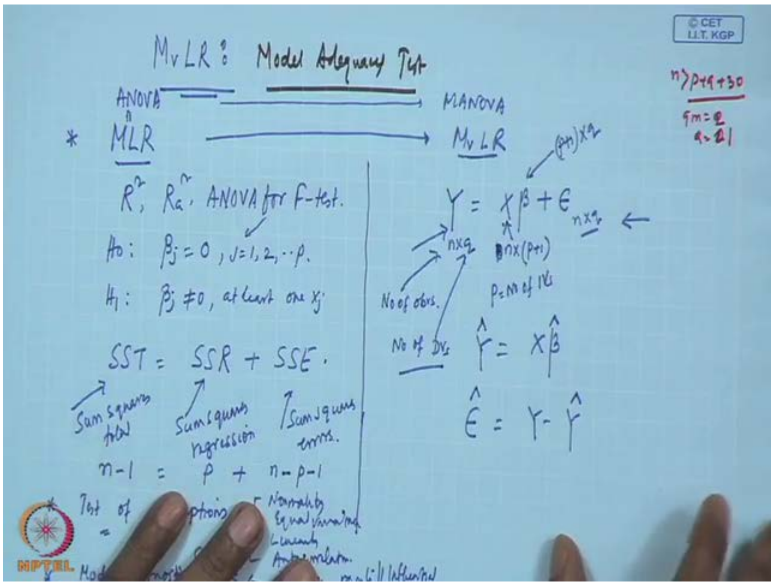



Good evening, we will continue multivariate linear regression. Now, we will discuss about model adequacy test for multivariate linear regression model adequacy test, if you recollect multiple linear regression. So, we have considered R square, R a square we have also developed one table called ANOVA table for F test correct, and what we have done basically we have set certain hypothesis. That hypothesis is beta j equal to 0 for all j equal to 1 to p, and we created alternate hypothesis that beta j not equal to 0 for at least 1 x j that was the hypothesis we have tested. We have used this ANOVA table like this, in addition what we have done also while preparing the ANOVA table, we have found out that the total sum square or sum square total.

We divided or partitioned into two half that is SSR plus SSE, that mean the total or sum square total equal to sum square regressions. Then sum square residual or sum square errors correct, and using this that degrees of freedom was n minus 1 then here p plus n minus p minus 1. So, this one in addition to this is basically all are model fitting, coming to model fitting in addition to this we have test of assumptions. That means we have done normality test, equal variance test, linearity test, we have also tested for auto correlated errors autocorrelation test, you can remember this.

In addition, we also have done something called model diagnostics that is basically coming under residual plots and in multiple regression you find out that several residual. So, different types of residuals like your student residual, press residual or deleted residual then there was standardize residual. Then we have seen that ultimately cooks distance then you defeats, de betas. So, many things are there which are for finding out the leverage points outliers leverage points as well as outliers ultimately we say these are all influential observations. So, can all those things not be done in MvLR you are getting me, which is possible and most of the times the software will give you the outputs. Suppose a space has many tabs has sys tag, statistical whatever software you use, ultimately you will get this outputs.

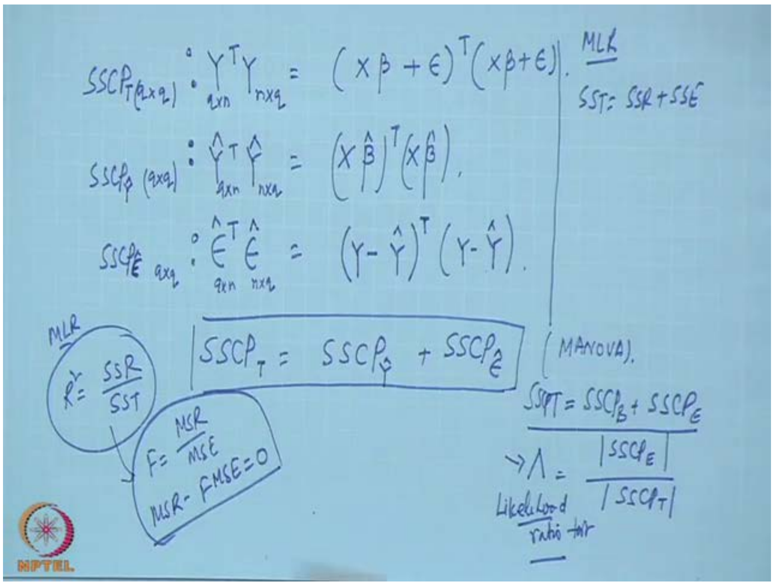
(Refer Slide Time: 04:51)




So, we will see some analog from MLR to MvLR that mean multivariate linear regression and if you do little more imagination then what you will find out that it is just MLR if it is related to ANOVA then MLR is similar to MANOVA, getting me. So, now let us see that in MvLR how do we get that F test, R square, R a square, similar thing and how do we how do we partition the matrix that matrix is sum that total sum square cross product matrix.

All those things and what about the test of assumptions, this is in totality will be discussing within this 1 hour or 50 minutes of time. So, let us write down again in the last class what we have written and so that you can co relate with this we say this is n cross q where n is the number of observations q is the number of dependent variables. Then we say this one is our n cross p plus 1, so our p is number of i vs and your beta is your p plus 1 cross q, now what is q, q is number of d v then this one is n cross q. So, this equation we will be using and once you have estimate that Y cap, this is your x beta cap and also we know that epsilon cap is Y minus Y cap. These are the few things we have discussed earlier and these things will be required.

(Refer Slide Time: 07:27)



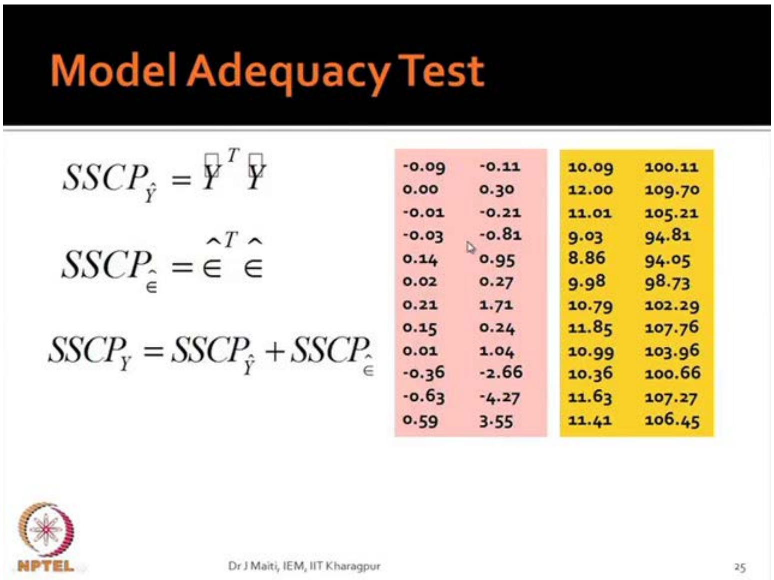
So, what we will do with this things suppose I want to know Y transpose Y, getting me what will be this, this will be x beta plus this transpose into x beta this, we are writing same equation that this transpose this into this. Now, if I write Y cap transpose Y cap then this is nothing but x beta cap transpose x beta cap. Similarly, I can write this transpose epsilon cap which is nothing but Y minus Y cap, transpose Y minus Y cap. Now, I want to know what is this Y transpose Y, what is Y, Y is n cross q then what is your transpose q cross n correct. Then this one if I write like this some matrix you will be getting which is q cross q, similarly here you see is again q cross n, n cross q you will also be getting some matrix here of q cross q.

Here, this one is also q cross n, n cross q you will be getting something here q cross q. So, if I say this is my SSCP total which is a q cross q matrix, this is my SSCP Y cap that is basically what is the that regression point of view, how much we are getting that is also equal band if I say this one is like this epsilon, if I write like this or cap. So, see you will find out in MLR what you found out we found out SST equal to SSR plus SSE. Similarly, here we can write like this, SSCP total equal to SSCP your Y cap plus SSCP epsilon cap can you find any relationship with this.

With MANOVA, we have found out that SSCP total equal to SSCP between plus SSCP error within in MANOVA we have used. Suppose T SSCP total equal to SSCP between plus SSCP error, so because q variable Y variables your matrix becomes like this, suppose I want to create R square can I create R square. Here, in the same manner, here R square is what SSR by SST that is from MLR point of view, but you will not get here because these are all in the matrix domain. But, in MANOVA what you have done we created one Wilk’s lambda that one we say what is SSCP, I think E divided by SSCP total.

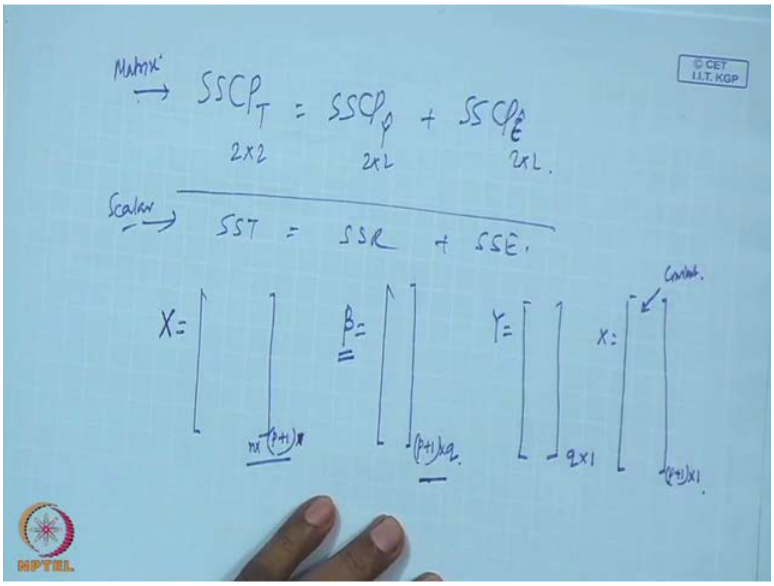
So, this is what is the likelihood ratio that likelihood ratio test can you get similar measure, here in your multivariate linear regression case or not. So, if yes here in multivariate linear regression also we go for likelihood ratio test and here that R square like measures 1 to just exactly 1 to 1 that type of relationship is not possible. But, yes there some other measure which basically where derived, based on that F ratio F statistics point of view that m what is this MSR by MSE in R square. In ANOVA table you find out that F equal to MSR by MSE, so MSR minus FMSE this equal to 0. So, this type of relationship can be used and will find out some measures with respect to this, so in this context.

(Refer Slide Time: 13:33)



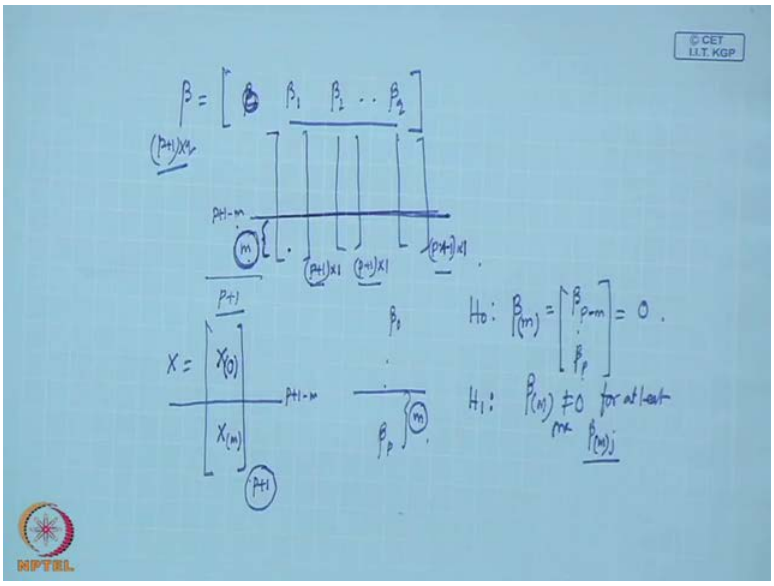

Let me show you this slide, in this slide you see this for Y 1 that is for profit this is for sales these are residuals, getting me, residuals. So, how many how many Y, Y were able 2 Y variable, so what will be the q cross q 2 cross 2, so this data can be used because this is epsilon cap. This data can be used to find out this 2 cross 2 that SSCP matrix, now what this one this is the predicted one which is Y cap so predicted one this predicted one when subtracted from actual Y value, we have got this residual. So, this can be used to find out this SSCP Y, now SSCP Y will be SSCP Y cap plus SSCP epsilon cap. So, repeat this fitted values will be used to find out SSCP Y cap this will be used to find out SSCP cap and already Y values are there you can check also using this formulation, so in this example our q is 2.

(Refer Slide Time: 14:59)




So, all SSCP totals that will be 2 cross 2 equal to SSCP Y cap that will be 2 cross 2 plus SSCP epsilon cap that will be 2 cross 2. So, it is similar two your SST equal to SSR plus SSE, so difference is this is the matrix domain scalar vector and scar domain this will matrix fine we will do some test that test is basically that what I can say. Here, a general test that is what I can say that generalised hypothesis testing, generalised hypothesis in the sense you have x data matrix and which is basically your p plus 1 n cross p plus 1. So, your beta is there, this vector beta vector which is how many beta is p plus 1 cross q correct, and your Y variable. So, if I go by variable Y is having q cross 1 variable, x is having your including the constant I can say p plus 1 cross 1 variable including constant is there, now you may be, so if I write down, here for the Y variable q 1 variable case.

(Refer Slide Time: 17:06)




So, the beta part I can write like this beta is beta 0, beta 1, beta 2, beta q will write, no we will write like this beta 1 to beta q. Now, this beta 0 contains p plus 1 into 1, this also contain beta 2 also contain p plus 1 cross 1 like this that is why it is p plus 1 cross q fine. Now, what I will do here I will say that is sum of the variables, independent variables are not contributing, instead of writing that the j-th variable is not contributing we are trying to find out some of that independent variables are not contributing, getting me.

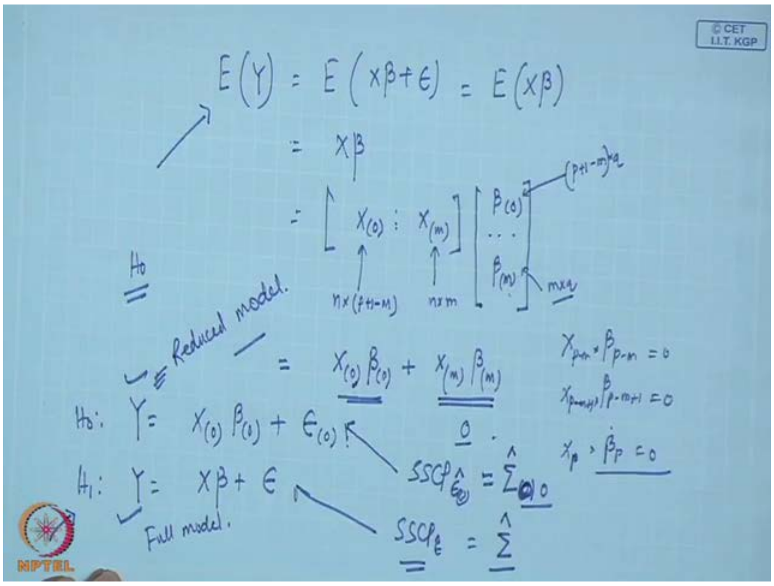
Subset of p plus p variables constant always, here subsets of p variables are not contributing towards explaining the Y variability. So, here you have p plus 1, p plus 1, p plus 1 if I say that this one is starting from this, that is p plus 1 minus n up to this, then this one is n because total is p plus 1. So, we are trying to say that this p plus 1 minus m this many variables are contributing and then reaming m variables are not contributing.

Instead of writing that the particular beta j variable, beta k variable is not contributing or beta j variable is not contributing what you are doing. Here, you are simply saying that the subset of variable is not contributing getting me, if this is the case. So, what I am trying to say that I am creating two sets this one is your x 0 this one we are writing like x within bracket m, getting me. So, that means I have to tell p plus 1 including constant p plus 1, so under h 0 that is null hypothesis h 0, what we are saying that beta m, getting me.

We are basically trying to say beta m which is equal to your that beta, this one is p plus 1 minus m, so that means p plus 1 minus m plus 1, so p minus m, p minus m to beta p I have beta 0 to beta p, I am taking the last m not contributing. So, that means p to p minus m, so this is 0 please understand what I am trying to say. We are trying to say that our p variables let it be 20 p equal to 20, so I am saying that the last that 15 variables are contributing 5 variables as a whole not contributing.

So, I can rearrange those variables, first 15 variables contributing like this, p plus 1 minus m that x 0 and last is not contributing x m which can write like this. So, here also we are writing x 0 within bracket, so that it will not be confused with x 0 that constant correct, then each one is. Here, that beta m not equal to 0 for at least one of the variable, for at least one beta m j that means the remaining variable, what you are discarding they are not to be discarded, they are also contributing.

(Refer Slide Time: 22:32)




So, then under such situation if I want to write down what is the expected value of Y this is nothing but expected value of x beta plus epsilon which is nothing but expected value of x beta the reason is expected value of epsilon that is basically 0 and ultimately this leads to x beta only. So, if you write this one like this that x 0, then x n x 0 and x n x 0 p plus 1, what is this, this one is n into p plus 1 minus m and this one is n into m, clear. So, then same thing I can write for beta 0, beta m and what is this beta 0, this is definitely p plus 1 minus m cross q and this one is your only that how many m cross q.

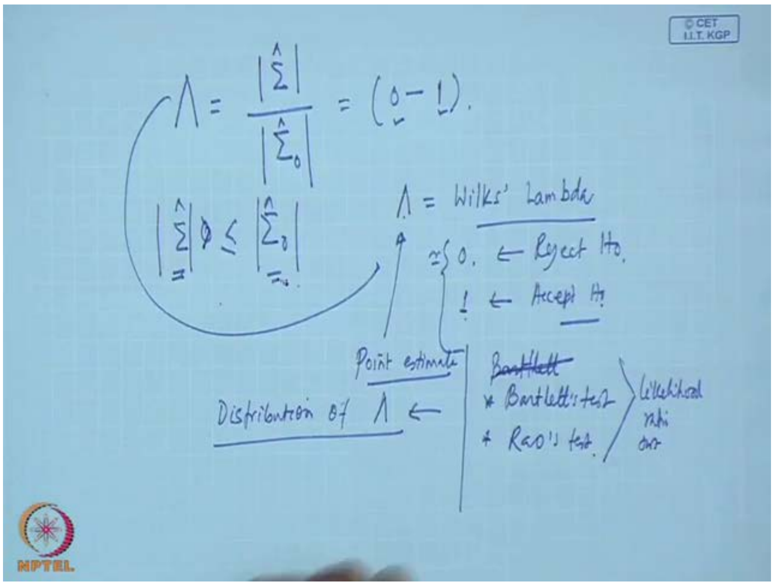
Now, resultant will be basically this will be m cross q and this cross q like this, so you can write down like this one x 0, beta 0 plus x m beta m we are not considering the error term because we have taken expected values of this. If m variables are not contributing towards explaining the Y's variability then ultimately what will happen beta m will 0. So, that means this resultant will be 0 getting me, your x m, beta m, beta m, beta m is basically beta p minus m beta p minus m plus 1 like this beta p everything is 0.

Whatever, when you multiplied this into x p minus m x p minus m plus 1 and x p this resultant quantity will also become 0 what I am saying, here that this quantity 0 means this expected value is dependent. On this we are using 0 to denote that under h 0 this is a situation then under h 0 our value will be Y will be x 0, beta 0 plus epsilon. So, I can write like epsilon 0, so this is under h 0 if we say that beta m is contributing then we will write the alternate one.

Alternate is, no this beta m 0 that you have considered that is false which is basically x beta plus epsilon what you have taken earlier, getting me. So, what will happen, here if I fit this model this model you will be finding out SSCP epsilon 0 estimate and here also you will be finding our SSCP epsilon cap this estimate, getting me. Now, as given by Johnson and Wichern in this book I have seen that, if I say this is nothing but that estimate of this or you can straight away write 0 and this is nothing.

But, this one estimate of this I have taken from Johnson and Wichern you just check, so what will happen then one hand I have taken the total set of independent variables. In the other hand, what we have taken we have said that h 0 is true that means m variables are not contributing, getting me. If h 0 is true then the SSCP from the reduced model and SSCP from the full model will be same, the error term will be same. So, as a result if I say under h 1 this model is my full model and under h 0 what is this model is reduced model, my question is that error SSCP error that matrix that will be moral or less equal.

(Refer Slide Time: 28:24)



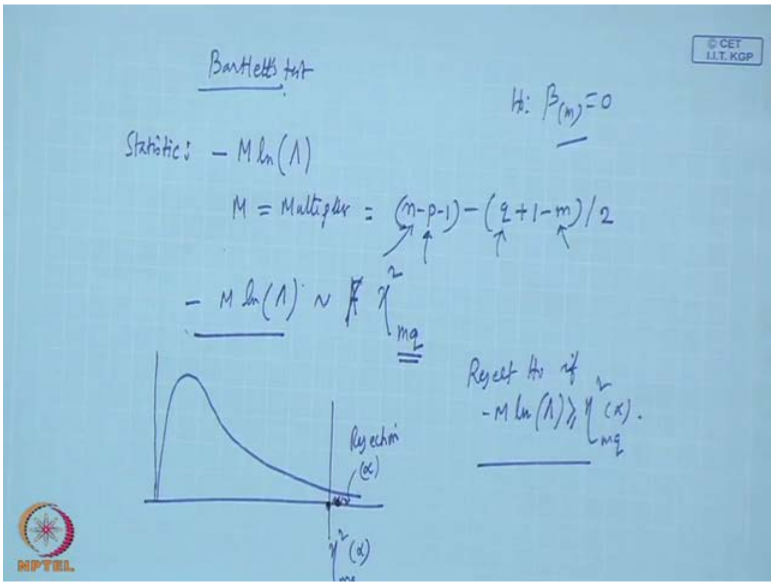

Now, if we create a ratio which is determinant of the that full model SSCP E this of this your reduced model SSCP E getting me, so this ratio will vary will stay in between 0 to 1. So, when it will become 0 suppose obviously when I am going for full model that this one it should be less than equal to because in this reduced model you have eliminated many variables. Their contribution you are not considering that will come to the error term only SSCP E, so obviously now what will happen if this value is sufficiently small sufficiently lowered, much lowered.

Then this value it will be close to 0 if both are equal, it will be close to, so if you find what it indicates if you find out that lambda that this lambda is known is Wilk's lambda that in a MANOVA you have tested this. If this value is sufficiently close to 0 then what it is indicating, it is indicating that sigma capital sigma cap 0 is much more than the full model sigma cap, getting me. So, as a result you reject h 0, but if it is equal to close to it will not get one if it is close to one except h 0 getting me fine.

But, problem lies where problem lies this lambda, but the way you are making these are all generalized variance when you take the determinant of a variance covariance matrix. These are all generalized variance and this lambda it is a point estimate getting me, point estimate what is the guarantee that this point estimate is giving you that what you want exactly there is uncertainty.

So, we require to know the interval estimates, so as result what is required we required to know the distribution of what distribution of lambda this is very, very important correct. But, lambda computation is complex you see that, so much fast from matrix SSCP E matrix that full model reduced model then if it is p is very large, q is also large. Then in that case what will happen, complexity is huge you cannot get straight cart the distribution for lambda. But, researcher particularly statically they have worked on this and they found out that that lambda in different transformation was made for Wilk's lambda. Some distributions where found out the two important test or distributions given by, one is Bartlett I think Bartlett test and another one is Rao's test these are all likelihood ratio test, so what is Bartlett test.

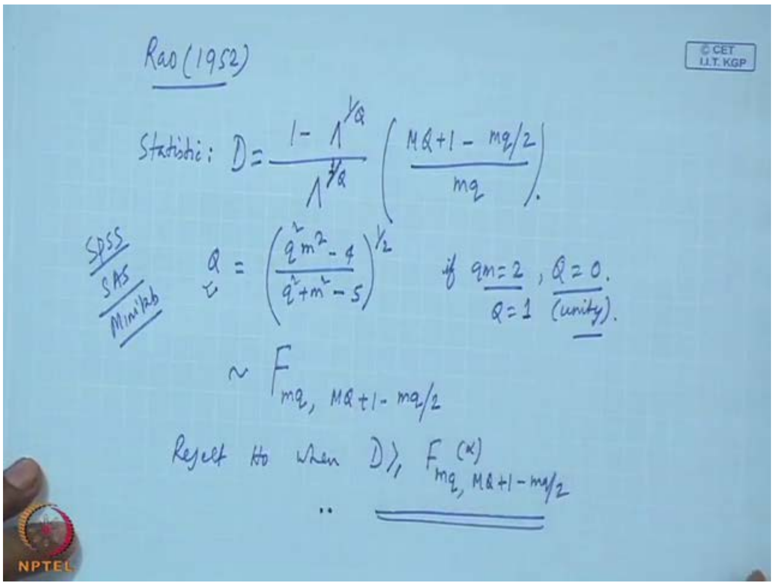
(Refer Slide Time: 33:06)




How to compute Bartlett test Bartlett test says it says that that he create one statistics which is minus M log base e lambda log s b lambda this one statistics created. Where your M is a multiplier this value is n minus p minus 1, minus q plus 1 minus m divided by 2, so you know what is n, total number of observation what is p, number of x variables what is q number of Y variables, what is m that we are saying. Beta m under h 0, beta n equal to 0, so that n variables you have dropped correct, if you say this one then this statistics that mean minus m log lambda this follows F distribution. Also it follows chi square distribution and the certain condition, so chi square distribution that was given by Bartlett yes can be approximately F distribution also, but chi square distribution that n with m q degree of freedom. So, m is number variables dropped from the independent side q is the number of Y variables, so this follows this distribution.

Now, what will you do then you know that this chi square distribution looks like this and our this minus M log lambda this follows chi square distribution. So, for some values that is the rejection, so you find out what is the value is minus M log lambda is falling somewhere in the rejection region or not. So, this point that critical value is you chi square if I say this is alpha then chi square M Q alpha this is your critical value. Now, you will reject h 0 [vocalised noise] if minus M log lambda greater than equal to you can write chi square M cube alpha.

(Refer Slide Time: 36:16)

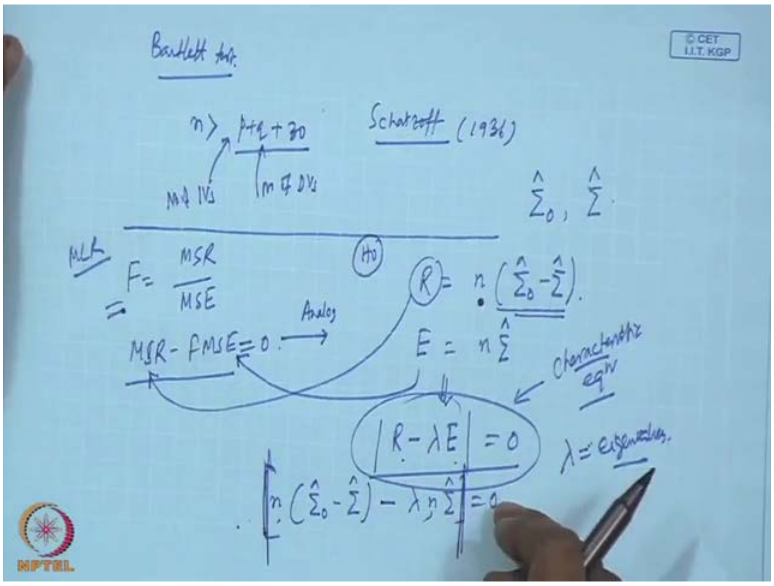



Now, second one is Rao's statistics in 1952, Rao has given similar statistics for Wilk's lambda his statistics is this one. That the statistics what is created it is 1 minus lambda to the power 1 by Q divided by 1 by Q into M Q plus 1 minus m q by 2 divided by m q this is the statistics where M. Already you have computed M is the multiplier earlier you have computed M, so M we computed like this is our M.

So, that M multiplier what is Q, Q is a quantity which is q square small q square plus m square minus 4 divided by small q square plus m square minus 5 whole to the power half this is given by Rao. So, please remember if q m equal to 2 this quantity Q becomes 0, Q becomes 0 if small q into m equal to 2 then q equal to Q equal to 0. Under such condition you have to consider Q equal to 1 unity other cases no problem fine, so you first find out this statistics. So, 1 minus lambda 1 by lambda to the power 1 by q by lambda to the power 1 by q 1 by q, this is not Y this is 1 by q m, q plus 1 minus m q by 2 by m q. This quantity q will be computed like this follows F distribution with m q and M Q plus 1 minus m q by 2 degrees of freedom, so what will be when you reject null hypothesis if I say this statistics.

This total value is D then you reject h 0 when d must be greater than equal to F m q, M Q plus 1 minus m q by 2 into that one alpha value this, not into this is the alpha probability value. So, when you are using this ratio likelihood ratio this statistics this two statistics are useful fine. So, when if you use SPSS, SAS or some other programme like mini tab you can ask Bartlett test Rao test, but Bartlett test is available Rao test I do not think Rao's test I am not seen it may be available.

(Refer Slide Time: 39:58)




Now, question is these test are very, very sensible in the sense what I mean to say for example Bartlett test you say that chi square distribution is followed. But, chi square distribution under what condition it says that if your sample size is small Bartlett test will not work. So, it says that n must be greater that p plus q plus 30 what is p, p is number of i v square is number of d vs and 30 plus 30, this is the case. This one is basically developed by Schatzoff in 1936 prove that for Bartlett test this is the require men then only the chi square distribution be followed I think you understand.

But, there are other test not based on likelihood ratio getting me, there are the test which are not based on likelihood ratio test. These other tests are interestingly suppose in ANOVA we know that F equal to MSR by MSE that means MSR minus FMSE this is 0, now it is equivalent or analog in multivariate domain. So, you all know that your multivariate domain you will not get scalar MSR, MSE all are scalar quantity in MLR multiple liner regression.

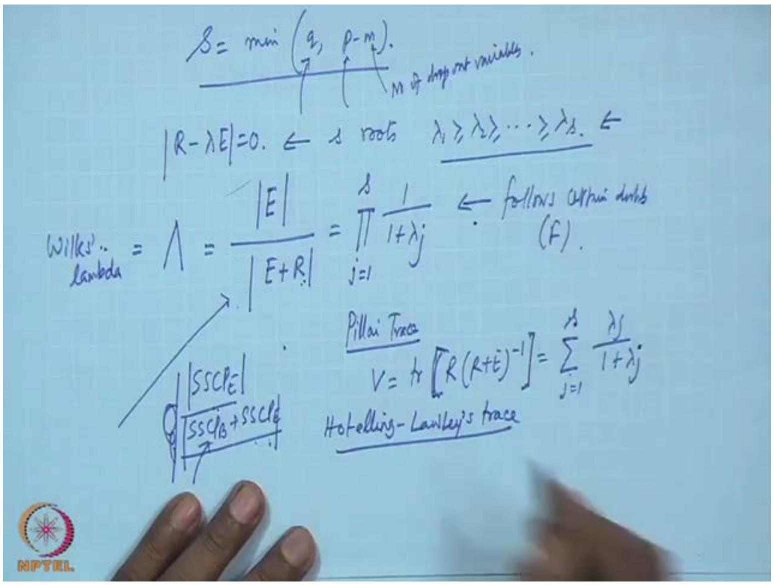
But,, MvLR you will not get this you are getting SSCP matrix and under h 0 beta m equal to 0, we found out two things that we say is that under h 0 you found out this SSCP and under full model you found out this one. Now, you can create one statistics which is basically R, similar to R m T using because of this regression that means the regression. This can be like this you can create like this R equal to this that means what is this difference how much the m variable you are dropping that is contributing you are testing h 0.

Here, also in ANOVA you test h 0, here and then you through f test you say some of the variables are contributing that is in MLR. But, in MvLR also if I say that what is this R is the difference between this and this multiplied by the number of observations and if I say that e equal to n that full model, getting me. Now, if I create like this what is happening, now the R can be similar to MSE and R can be similar to MSR, E can be similar to MSE. But, as they are in the matrix domain their equivalence in terms of your scalar domain you have to determinant.

So, that is why what happen that the statistician have developed that if you take this determinant equal to 0 then you will get very interesting roots characteristic what is this basically if I write it is nothing. But, n epsilon 0 cap, sigma 0 cap, minus sigma 0 that is R minus lambda n sigma 0, sigma cap this one, this is R minus lambda E this is the matrix. Now, if I take determinant here, so ultimately what is happening, this n will go out n will be cancelled because this you are putting is 0.

So, this one is 0 this is the characteristics equation getting me and lambda will give you Eigen values that you know. Now, this Eigen value lambda how many Eigen value will be there it all depend on that what is this order of this matrix. This resultant matrix that is what is the order of this matrix, we always think of the determinant mean full order full range matrix is require to be in.

(Refer Slide Time: 45:46)




Now, suppose that we say that the order is basically s which is minimum of q and p minus m that will be the order, q is number of dependent variable, p is number of independent variable, m is number of drop out variables. So, this is number of drop out variables, so then once you solve this equation R minus lambda E this equal to 0 you will be getting this will be s power that power will be because there are s root will be there. These roots are basically lambda 1, greater than equal to lambda 2, greater than equal to lambda s these are the roots and you will see that these roots when you do this Eigen value decomposition of matrix.

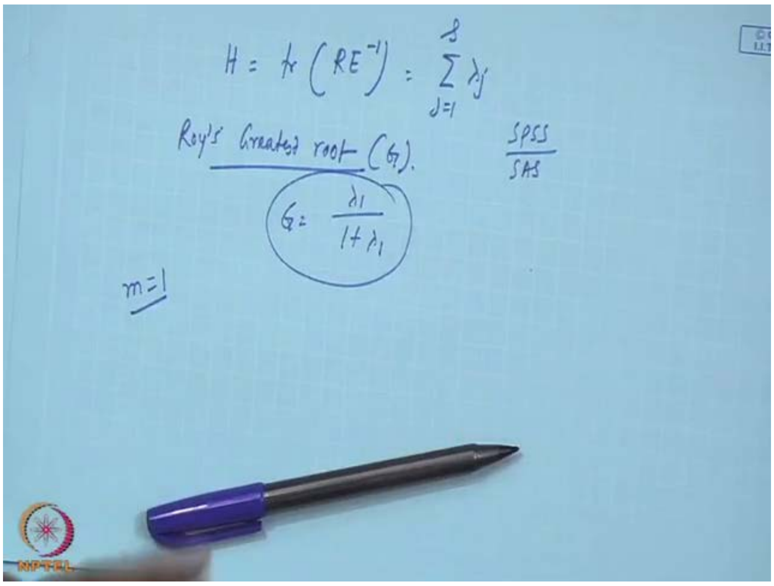
Similar, this one then ultimately what you will find out that p roots are, here it is s root and they will be there. First one is the maximum one followed next max, next max like this, so using this property there are several statistics developed. Several statistics were developed for example, you will find out that Wilk's lambda will become like this, Wilk's lambda will become like this equal to this.

This is nothing but determinant of E I think E you can understand, now the way we have defined divided by E plus R are you getting some clue with MANOVA Wilk's lambda is SCPE divided by SSCP between plus SSCP E. Now, between what happen population contribution, so it is similar to regression contribution error. So, that determinant of this divided by determinant of this we have computed here. So, this one when you define in terms of your Eigen values it will be like this multiplication j equal to 1 to s 1 by 1 plus lambda j.

So, that means what you require to do you require to find out this equation then find out the lambda values you will be able to find out this Wilk's lambda values also small lambda capital lambda. This definitely follows certain distributions and in all cases F distributions are approximated, here some other distributions like chi square. You have seen Bartlett has given chi square distribution Rao has given f distribution for this Wilk's lambda apart from this, there is another statistic which is known as Pillai trace.

Pillai trace statistics what it is this done it takes V equal to trace of R into R plus E inverse, so these are just variant almost similar and under certain condition. But, to get the different benefits it is done like this and this one will become j equal to 1 to s lambda j by 1 plus lambda j you will get also another statistics, here which is Hotelling Lawley trace, Hotelling Lawley's trace another one.

(Refer Slide Time: 50:12)




Then what is Hotelling Lawley's trace, Hotelling Lawley's trace is H which is trace of R E inverse this one is nothing but lambda j, j equal to 1 to s there is another statistics which is known as Roy's greatest root. Roy's greatest root G, so that G is taken only the first Eigen value this one, so all software including SPSS, SAS they report all those statistics. Now, what is this you please understand under H 0 this statistics are developed, so that means most of the time we will find out that F distribution?

Now, if this value is more than or equal to the cut off value from F distribution, so what you will do then in that case you reject the null hypothesis. That means the set of x variables m in number what you have rejected that is not correct, now if I consider m equal to 1 that means individual variables at a time you are testing. So, for every variable you test and then see that weather it is contributing or not like in ANOVA case, individual variable s s you have identified.

(Refer Slide Time: 52:13)

## Other tests

$F = \frac{MS_R}{MS_E}$ $\quad MS_R - FMS_E = 0 \quad |R - \lambda E| = 0$

$H_o : \beta_{(m)} = 0 \quad R = n(\hat{\Sigma}_0 - \hat{\Sigma}) \quad E = n\hat{\Sigma}$

$|(\hat{\Sigma}_0 - \hat{\Sigma}) - \lambda\hat{\Sigma}| = 0 \quad \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_s, s = min(q, p-m)$

$$\Lambda = \frac{|E|}{|E+R|} = \prod_{j=1}^{s} \frac{1}{1+\lambda_j}$$

Pillai's trace

$$V = tr\left[R(R+E)^{-1}\right] = \sum_{j=1}^{s} \frac{\lambda_j}{1+\lambda_j}$$

Hotelling-Lawley's trace

$$H = tr(RE^{-1}) = \sum_{j=1}^{s} \lambda_j$$

Roy's greatest root

$$G = \frac{\lambda_1}{1+\lambda_1}$$

NPTEL J Maiti, IEM, IIT Kharagpur 30

So, with this I want to go to the example.

(Refer Slide Time: 52:15)

## Model adequacy test - example

Multivariate Tests[b]

| Effect | | Value | F | Hypothesis df | Error df | Sig. |
|---|---|---|---|---|---|---|
| Intercept | Pillai's Trace | .835 | 17.693[a] | 2.000 | 7.000 | .002 |
| | Wilks' Lambda | .165 | 17.693[a] | 2.000 | 7.000 | .002 |
| | Hotelling's Trace | 5.055 | 17.693[a] | 2.000 | 7.000 | .002 |
| | Roy's Largest Root | 5.055 | 17.693[a] | 2.000 | 7.000 | .002 |
| Absenteeism | Pillai's Trace | .478 | 3.209[a] | 2.000 | 7.000 | .103 |
| | Wilks' Lambda | .522 | 3.209[a] | 2.000 | 7.000 | .103 |
| | Hotelling's Trace | .917 | 3.209[a] | 2.000 | 7.000 | .103 |
| | Roy's Largest Root | .917 | 3.209[a] | 2.000 | 7.000 | .103 |
| Breakdown | Pillai's Trace | .754 | 10.734[a] | 2.000 | 7.000 | .007 |
| | Wilks' Lambda | .246 | 10.734[a] | 2.000 | 7.000 | .007 |
| | Hotelling's Trace | 3.067 | 10.734[a] | 2.000 | 7.000 | .007 |
| | Roy's Largest Root | 3.067 | 10.734[a] | 2.000 | 7.000 | .007 |
| Mratio | Pillai's Trace | .836 | 17.889[a] | 2.000 | 7.000 | .002 |
| | Wilks' Lambda | .164 | 17.889[a] | 2.000 | 7.000 | .002 |
| | Hotelling's Trace | 5.111 | 17.889[a] | 2.000 | 7.000 | .002 |
| | Roy's Largest Root | 5.111 | 17.889[a] | 2.000 | 7.000 | .002 |

a. Exact statistic

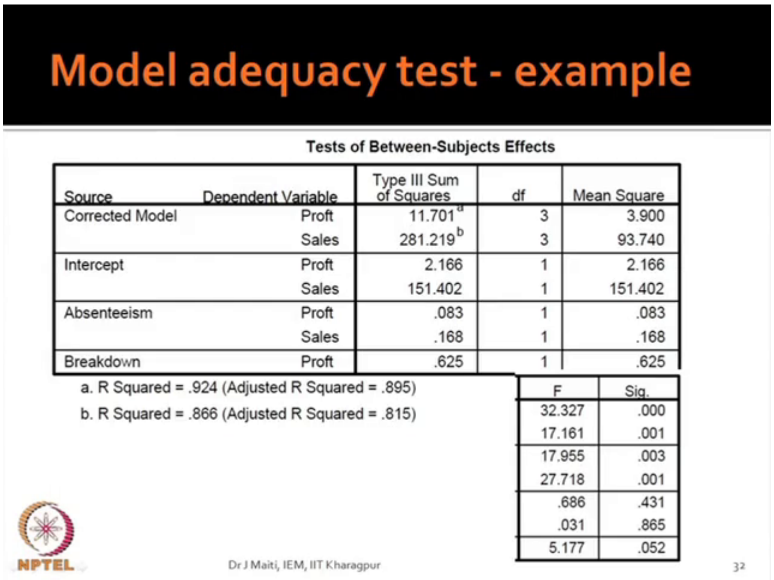
b. Design: Intercept + Absenteeism + Breakdown + Mratio

NPTEL 31

This is this indicate data, so intercept see Pillai's trace, Wilk's lambda, Hotelling trace Roy's largest root all those their values are computed there F trace is taken and then significance level is very high. That mean intercept is there, it is contributing and absenteeism, similarly is it significant not that 5 percent level it is not significant more than little, more than 10 percent level it is significant. Now, you can drop this variable if you think, so what you think it is required, fine then you say break down in every stage.

Interestingly everyone Pillai, Willk, Hotellite, everyone is giving the probability level almost equal, but the value of differing and the degree of freedom will remain same error degree of freedom also remain same. So, this is the case what is not we are saying we are saying, here that you are going for model adequacy test by similar F test like in case of multiple regressions ANOVA. But, here also you are doing this and you are finding out that yes some of the variables are contributing, so multivariate regression is required.

(Refer Slide Time: 53:42)




Tests of Between-Subjects Effects

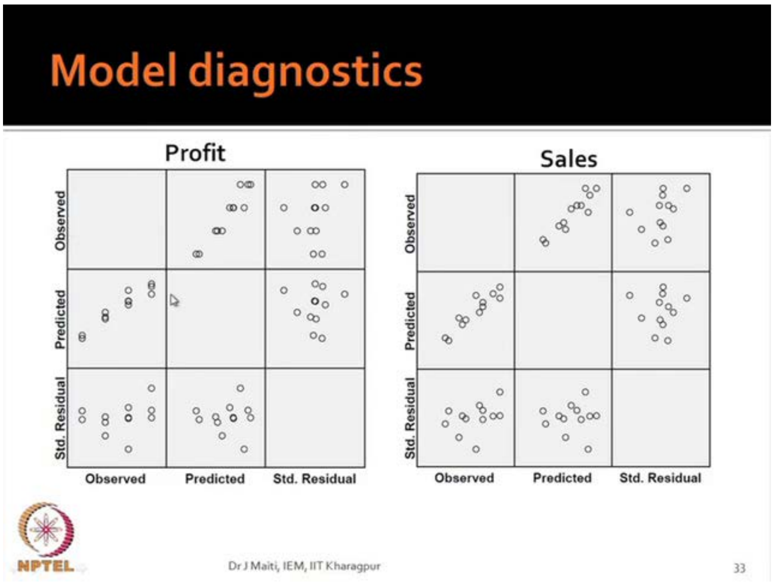
| Source | Dependent Variable | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| Corrected Model | Proft | 11.701[a] | 3 | 3.900 | 32.327 | .000 |
| | Sales | 281.219[b] | 3 | 93.740 | 17.161 | .001 |
| Intercept | Proft | 2.166 | 1 | 2.166 | 17.955 | .003 |
| | Sales | 151.402 | 1 | 151.402 | 27.718 | .001 |
| Absenteeism | Proft | .083 | 1 | .083 | .686 | .431 |
| | Sales | .168 | 1 | .168 | .031 | .865 |
| Breakdown | Proft | .625 | 1 | .625 | 5.177 | .052 |

a. R Squared = .924 (Adjusted R Squared = .895)

b. R Squared = .866 (Adjusted R Squared = .815)

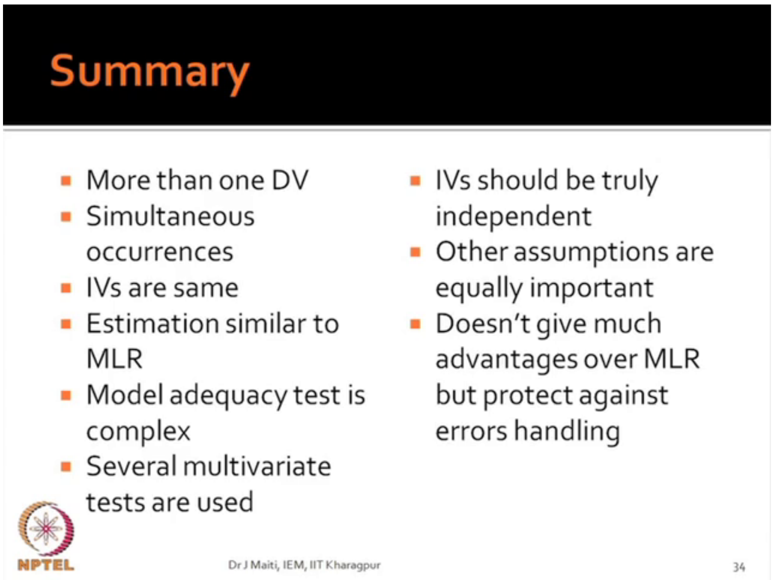
Now, go to the next one, so this is the your total sum of squares under that profit, under sells and you see R square like univariate R square that multiple regression R square. For first one it is 0.924, for the second one that means for profit 92 percent for your breakdown, not breakdown it sells your 86 percent. Also you have tested the significance level, here for different intercept for absenteeism for breakdown, getting me and their significance level, everything is tested.

(Refer Slide Time: 54:33)




Once you accept this you have to go for diagnostics similar to you MLR this diagnostics is require and almost all the software they will give you these types of plots, these plots are basically I say residual plots, now see there is dual plots. So, observed versus predicted it should be straight linear, now observed versus if you see observed versus this one this is observed predicted fine. Now, if you go by residual versus predicted or residual versus observed that should not give any trend, it should be random it is coming like this. So, predicted versus residual also that should also become random, so this for profit this is the case, for sales also similar thing we are getting and in fact from this figure you are not getting any unusual observations.

(Refer Slide Time: 55:43)

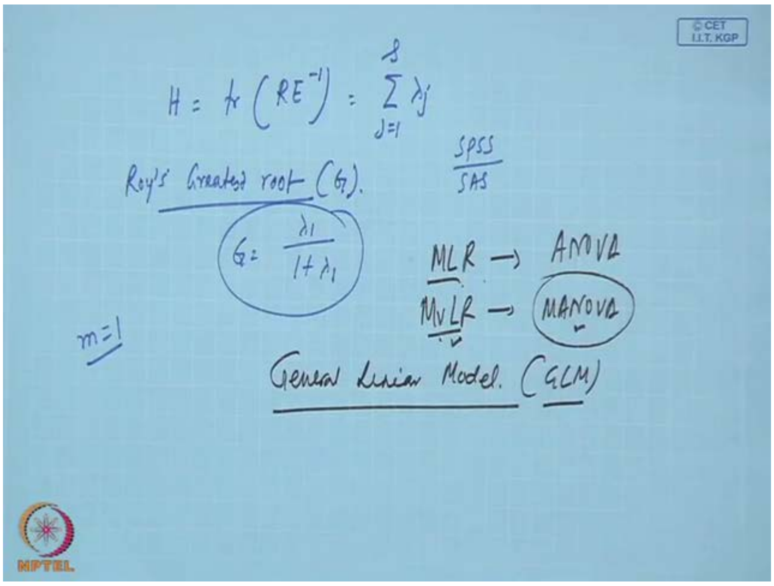



In summary when you go for MvLR you see that when you have more than one d v both more than one d v and all the d vs they occur simultaneously. Then I Vs are same means independent variable are same affecting different d vs estimation of this is the condition when you will go for this the multivariate normality that covariance structure equal. Other things that assumptions you have to test then estimation, from estimation point of view estimation similar model adequacy test is complex because you are going for likelihood ration test.

There are different test identified that Pillai trace, Wilks lambda, Roy's greatest root and then Hotelling Lawley's all these things is there. Several multivariate test that why is used i vs must be truly independent other assumption also equally important does not give much advantage over MLR, but protect against error handling. So, if you go separately for regression estimate point of view no problem, but adequacy test point of view there will be problem if you go for separate MLR, getting me.

So, this is the advantage and I think when you have multiple Y variables and you are sure that they are following, you are able to prove that is assumption are not violated or reasonably following. Then what you have to do it is better you go for MLR, MvLR multivariate regression first and then you test through different test you are finding out yes these are contributing then you will be if you go for single MLR.

(Refer Slide Time: 58:00)




That means multiple linear regression instead of MvLR absolutely no problem it is similar to that ANOVA and that analog is MANOVA you must go for MANOVA when you have more than one dependent response variables. Then you see that collectively they are reflecting or not, if not then go for ANOVA no problem, similarly MvLR and this. So, suppose this MANOVA and MvLR all those things comes under general linear model if you go through a SPSS or SAS there will be a one link GLM. Any question, thank you very much, next class we will discuss PCM principle component analysis.