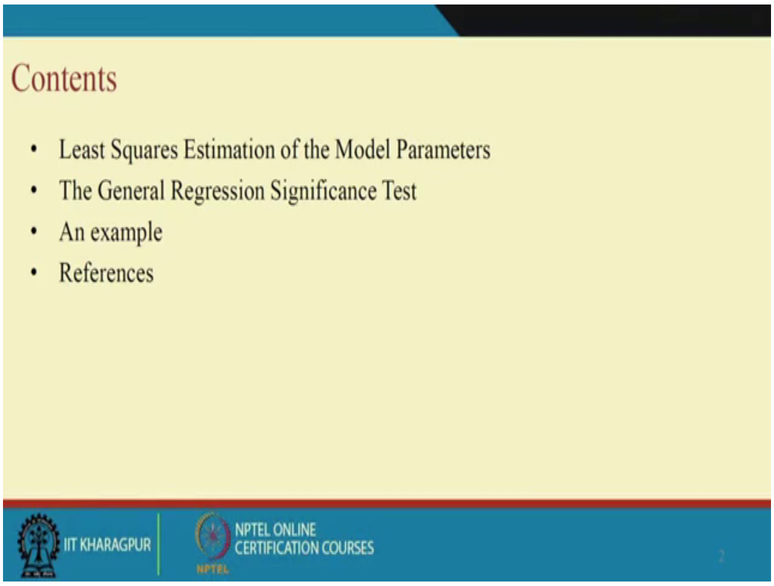
Design and Analysis of Experiments
Prof. Jhareswar Maiti
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture – 24
Regression Approach to ANOVA

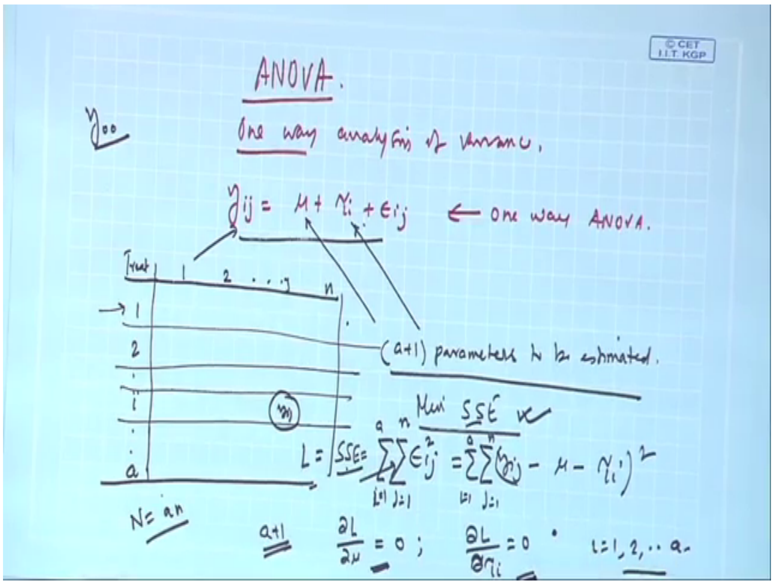
Welcome. Today we will discuss the regression approach to ANOVA.

(Refer Slide Time: 00:24)



You have gone through ANOVA analysis of variance.

(Refer Slide Time: 00:28)

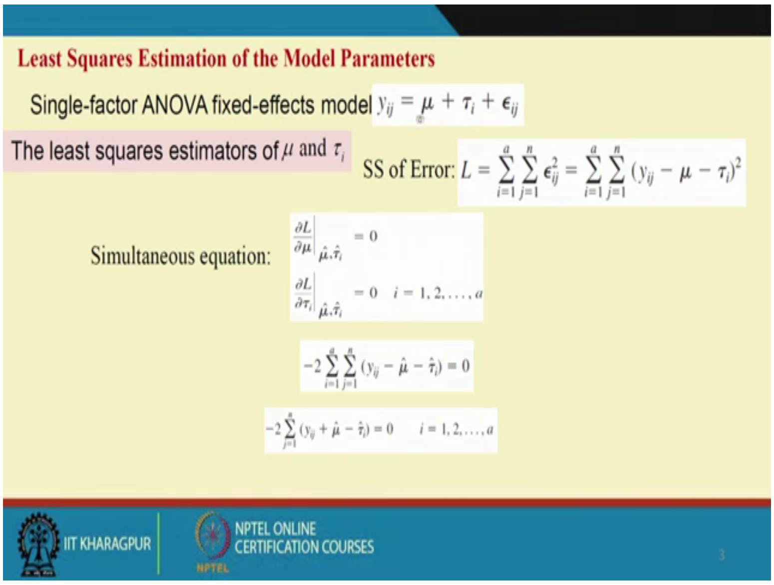


One way analysis of variance, one way we will discuss in term with reference to one way analysis of variance one of the way ANOVA and the model we have given y ij general model is mu plus tau i plus epsilon ij that is for one way ANOVA.

Here it is the general observation and our data was something like this that the factor or the treatment levels 1, 2, i to a levels and these side the observations or replications to n. When we have gone for balanced one and some way are here if it is a somewhere y ij is there and this is the general model. So, how many parameters to be estimated? One is mu another one is a number of tau i. So, total a plus 1 parameters to be estimated. So, in today’s lecture we will see that how regression is used to estimate this a plus 1 parameters and how the SS R, SS E and SS T can be estimated using the regression approach.

Further what I will show you I will show one example and then we will see that the ANOVA results versus our regression result for the model adequacy test, although the parameters level comparison also possible. So, essentially we will see that least square estimation of model parameters, the general regression significant test example and few references.

(Refer Slide Time: 02:36)

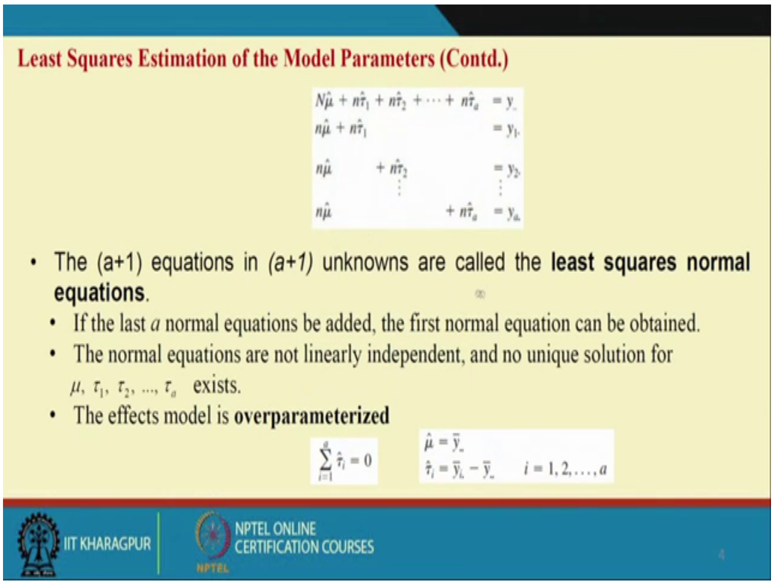


Let us see this model what is the fixed effect one way model is y ij mu plus tau i plus epsilon ij. You require to estimate mu and tau i. So, what is the procedure in least square method? Least square (Refer Time: 02:52) method procedure is you minimize the sum of squares error, sum squares errors SS E, you minimize SS E. So, a minimize SS E. Now in this equation what is epsilon ij is the error term, this is y ij minus mu minus tau i.

Now, if I sum it over i and j j equal to 1 to n, i equal to 1 to a then you sum it j equal to 1 to n, i equal to 1 to a, but what we are saying sum square that mean you first make it square. So, this square and then sum it. So, this is your SS E, this is your SS E. So, we say this is the its denote that this is the function of L, a L is this function not function of L. L is noted by this function, then what you require to do, you require to find out del L by del mu and put it to 0 and del L by del tau i then put it to 0, i equal to 1 to like a. So, what it will give you? It will give you one equation here and here are number of equation. So, you will be having a plus 1 number of equations simultaneous equation which to be solved. So, let us see the slides.

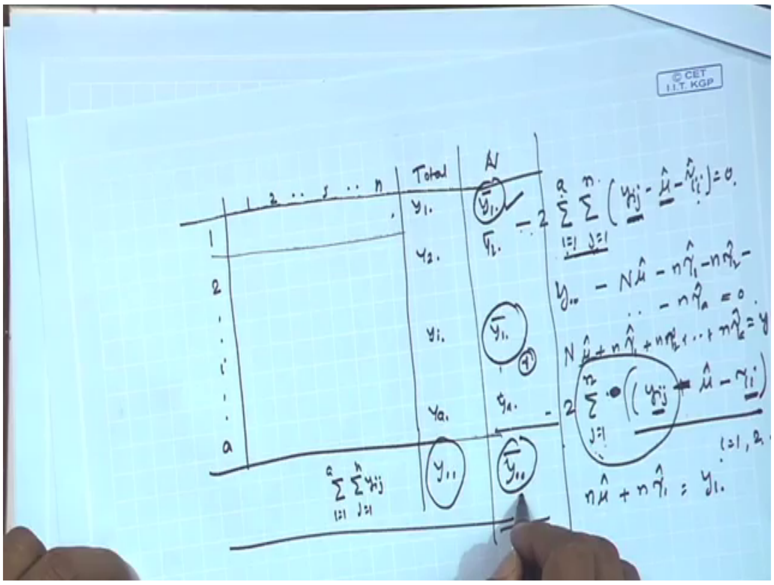
What are the simultaneous equation? So, del L del mu given this is 0 and del L by del tau i given this is 0. So, the first one if you take derivative of these with respect to mu you will get this equation. Now if you do the same with respect to tau i you will get this equation. So, this equation and this equation here i equal to, so how many equations you have? You have a plus 1 equations that is what we are seeing here.

(Refer Slide Time: 05:30)



How this equation is coming n mu all this? Suppose you just think of this equation equal to 0. So, sum total of y ij across j and i will give you n mu will give you this one y double dot there is a grand total and then mu here n times, n mean the total number of observation. So, n equal to a n that n times mu cap and for every tau there will be what is the how many observations are there; for every tau, tau 1 n observation, tau 2 n observation, like tau i n observation, but for mu grand mean total observation is N n equal to a into n. So, as a result when you put the of the number of observations into equation then what happen this will be n, sum of all n y ij which is nothing, but y total. What is y total? y dot dot, that is what we have computed earlier. So, let me repeat this case again.

(Refer Slide Time: 07:01)



I have 1, 2, i, a treatments and you have 1, 2, j, n your applications, then if I write total here this will give you y 1 total, y 2 total, y i total, y a total and if you find out the grand total this is y dot dot, grand total is total of all those things. This is nothing but grand total means sum of j equal to 1 to n, i equal to 1 to i y ij.

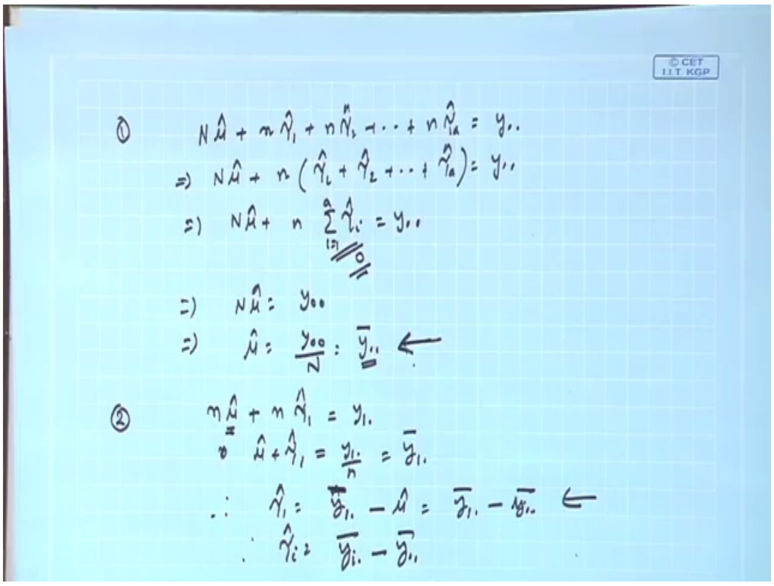
So, what is our first equation? First equation is minus 2 i equal to 1 to a j equal to 1 to n, y ij minus mu minus tau i equal to 2; obviously, given that there tau i cap. So, in this equation what I mean to say if I remove the sum this minus 2 will also be canceled then what will happen if I write this portion this positions, this is what this was n time total of y ij that mean y double dot bar this portion gives you this. Now, minus, minus this 1 i equal to 1 to a j equal to 1 to N in capital, total is N times mu cap. Now, plus this what is happening tau 1 will be for small n types because there are n observation. So, minus n tau 1 cap minus n tau 2 cap like this minus n tau a cap this will become 0, this will become 0. So, then we are getting this equation first equation. N mu cap plus n tau 1 n tau 2 like this small one this is giving you this equation.

Now, come to the second formula. What is the second formula? So, from this formula we have first formula we are getting N mu cap plus n tau 1 cap n tau 2 cap plus n tau a cap this equal to y double dot this is our equation number 1.

From the second formula second formula is equation is y j equal to 1 to n then y ij minus plus mu cap minus tau i this becomes 0, suppose you are making doing that minus minus this. So, what is happening here then this is minus. So, using this equation what you are getting? You are basically supposed for here i equal to 1 to a. If I put i equal to 1 what I will get this will go out. So, it is n times mu n times mu cap plus what you will get here this tau 1 that that will also be n times tau 1 cap equal to y 1 dot because this is the first row this will this quantity will give you first row total.

So, in this manner, in this manner you are getting first row total form the first equation, second equation, second row this. So, you will be having such n a plus one least square normal equations. So, what is the constraint here? Constant is that if you sum up a number of equations starting from the second one starting from this one to the last one that is a number of last 2 equations if you sum up you will get the first equation. So, as a result effectively you have unique a number of equations. But you want to you have to estimate a plus 1 parameters. So, as a result what happened you have a constraint it is your over parameterize model. So, you have a constraint that i equal to 1 to tau i cap equal to 0 this constant. Using this constant, so what happen and all these a plus 1 equation you will be able to estimate the parameters. How?

(Refer Slide Time: 12:52)

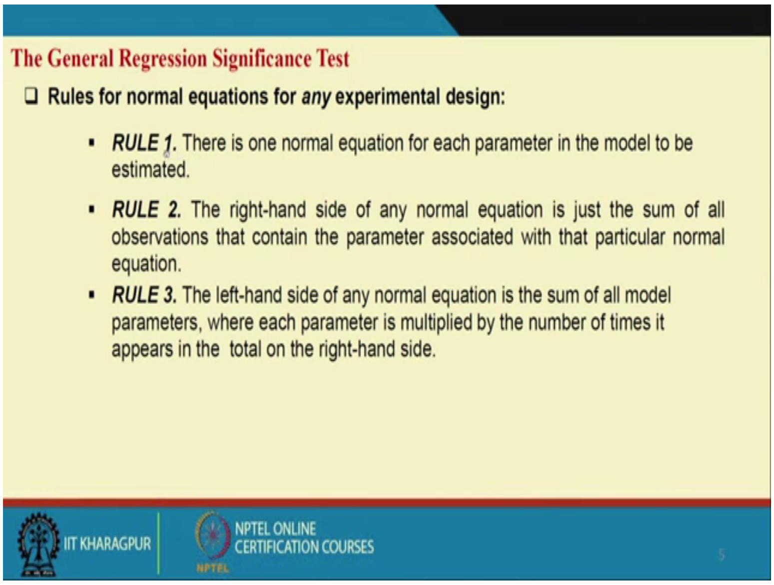


So, what we got? We got N mu cap small n tau 1 cap small n tau 2 cap like this small n tau a cap equal to y dot dot that is the total grand total. Then this is nothing, but N mu cap plus if we take in common this is tau 1 plus tau 2 all caps are there plus tau a equal to y dot dot this is nothing, but N mu cap plus n sum of tau i cap i equal to 1 to a equal to y double dot now this is 0 because of the constraint. So, we can write N mu cap equal to y double dot other way mu cap is y double dot by n; that means, grand total by total number of observations this is nothing, but y bar dot dot and this is what we have done in ANOVA also that is what is the grand total average this is from average formula that sum of all observation divided by that total number of observations. So, this is fantastic.

Now, if I take the second equation this is my first equation then second normal equation this is n mu cap plus n tau 1 cap equal to y 1 dot, then a second one mu cap is known because already mu cap is estimated. So, if I write la now that that mean mu cap plus tau 1 cap equal to y 1 dot by n this is nothing, but y 1 dot average. So, as a result tau 1 cap is nothing, but y 1 dot average, y 1 dot average minus mu cap equal to y 1 dot average minus y dot dot that is average. So, here y 1 dot bar this average. So, this is what you have seen earlier also. That we said that for every row average if you can recall we found out that overall average and then the tau then we found out the average here, this is 1 dot average, y 2 dot average, y i dot average, like y a dot average then we say that what is the effect of first level that these average minus this average.

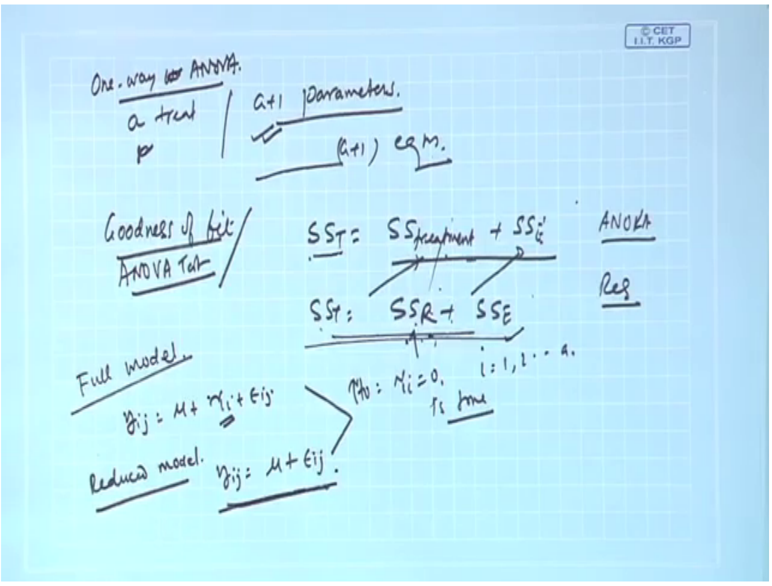
What is the effect of ith level? y i average minus y that this grand average. So, that means, in general from this equation we can say tau i cap equal to y i dot that is the average ith average minus by this average this is the estimate you have seen earlier; now, from here if I know that all the tau and mu estimates then the residuals also will be known from the primary equations y ij equal to mu plus tau i plus this.

(Refer Slide Time: 16:32)



So, this is the rules. Now, for normal equation that there are three rules. Rule one is there is one normal equation for each parameter in the model to be estimated. So, we have in one factor one way ANOVA, we have a levels for a treatments levels and grand mean.

(Refer Slide Time: 16:55)



So, as a result a plus one parameters to be estimated that I told you I am repeating the same thing. So, that for every parameter there will be a normal equation a plus a 1 equations will be there and it is seen. Then the right hand side of any normal equation is just the sum of all observations that contain the parameter associated with that particular normal equation. See here, right hand side is what sum of all observations because this is a this equation with respect to mu this is a equation right hand side sum of all parameters, sorry sum of all observations sum of all observations second one you see with respect to tau 1 sum of the observations. So, like this, the right hand side of n normal equation is just sum of all observations that contain the parameter associated with that particular normal equation.

Rule 3, the left hand side of any normal equation is sum of all the model parameters where each parameter is multiplied by the number of times it appears in the total on the right hand side. You see, suppose you this one you consider this one if you consider you see that this is y 2 bar it is n times the right hand left hand side a n times it appears for a particular row number of n times the number of observations you have got n application is there. So, these parameters in this equation the parameters are multiplied by that number of observations that is n. So, in the regression approach you are (Refer Time: 18:57) there will be as many equation as many parameters we have. And the general rule is that three rules are there, if you apply this three rule you will be able to get all the equations even if you do not go by the least square that traditional mathematical procedure.

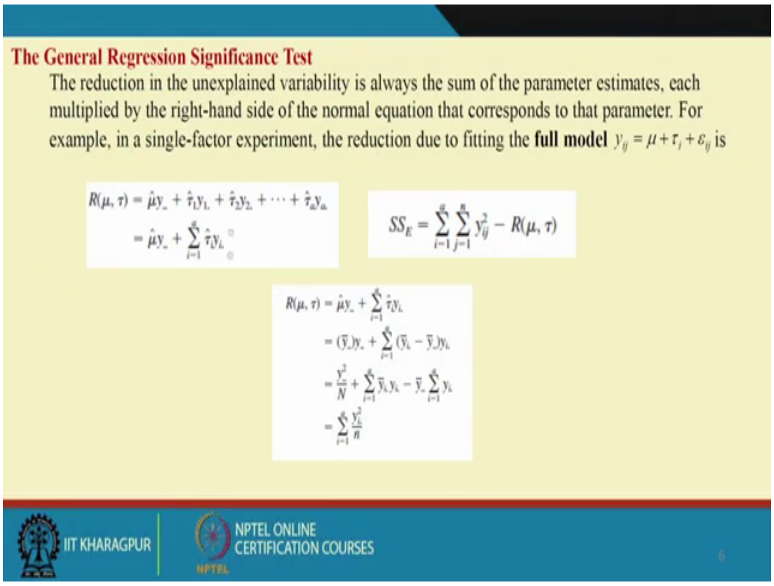
Now, once you have the general equation the normal equations you are in a position to estimate all the parameters which are nothing, but the way we have seen in one way ANOVA lecture. Similar, same thing is coming to obvious. Now, I will show you the important one there that is means that goodness of fit measures, goodness of fit as well as I can say that the test ANOVA test, overall test. So, what happened we have seen that the corner we ANOVA partitioned the total sum square total into sum square treatment plus sum square error.

Now, using regression approach how do you compute these things? SS E and SS treatment if we say SS treatment in if and in regression you have seen this is in ANOVA one way ANOVA and regression what you have seen you have seen that SS T equal to SS R plus SS E. What is the SS R? Sum square regression. So, this sum square regression and sum square treatment, ultimately it is a partitioning the total variability into 2 parts. So, they will be almost same maybe because of the approach. There will be some variability is some difference in the calculated values may be am, but approaches are same; that means, SS R here is nothing, but equivalent to a SS treatment.

So, let us see then using ANOVA regression approach for how to compute that SS R SS E, SS E and we get the earlier results what we have already used. So, here we are using, we will be using to the concept of full model full model mean if hours is a one way ANOVA case then y ij equal to mu plus tau i plus epsilon ij that is full model.

Another one is reduced model. Reduced model means suppose then this factor is not contributing in any way then this model can be written like this mu plus epsilon ij when h 0 that tau i equal to 0, i equal to 1 to a is true it is true then this model is nothing, but this is reduced model. That means, the treatments are not at all contributing to the response or that the independent variable in terms of regression is not contributing to the overall variability of the y the response variable. So, we use these 2 and then we will apply certain rules to find out that the sum squares for total, sum square for error and sum square for treatment or regression. Let us see what is this rule.

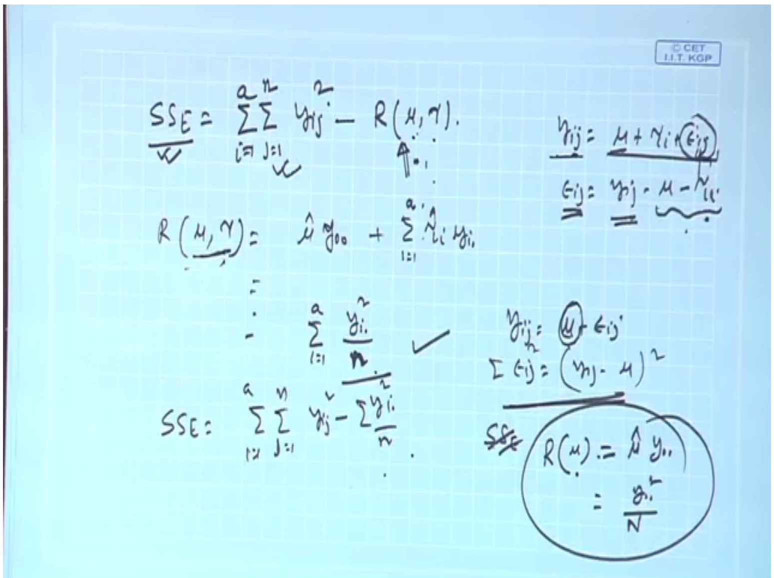
(Refer Slide Time: 23:04)



Rule is that the reduction in the x unexplained variability is always the sum of the parameter estimates in the sense that each multiplied by the right hand side of the normal equation that corresponds to that parameter. For example, in a single factor experiment the reduction we have seen already the single factor experiment. Now, if we say the reduction is R mu tau then this is basically how many parameters are there here mu and a number of tau a plus 1 parameters. You see each of the parameter estimates are multiplied by the right hand side value of the normal equation. So, what is the estimate of mu? That is mu cap. So, what is there in the right hand side of the normal equation? Grand total multiplied by grand total.

What is the second one? Tau 1 right hand side y 1 total; like this tau 2 cap y 2 total tau a cap y a total. So, this is now this correspond equivalent to mu cap y double dot plus i equal to a equal to 1 to a tau i cap, y i dot, y i dot. So, this is what is the reduction of variability from of y reduction over a very variability that is if you use the full model.

(Refer Slide Time: 24:54)

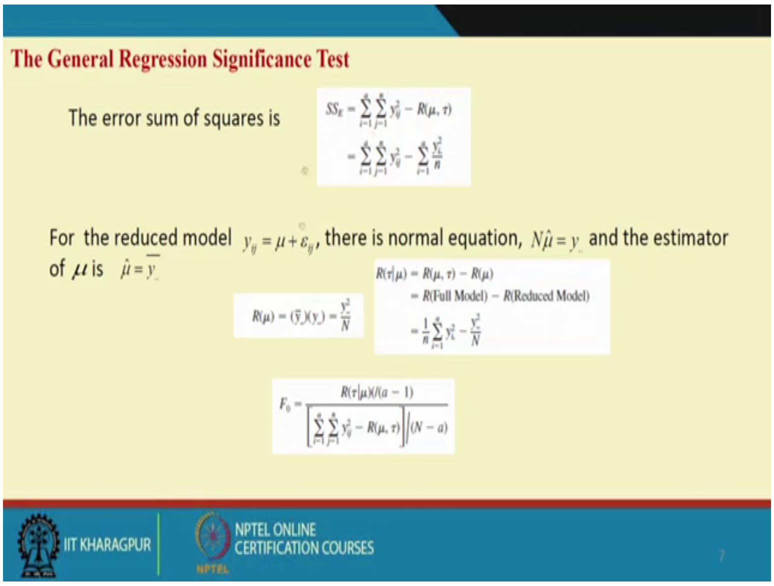


Now, if we assume that is from SS some square point of view; that means, the sum square of all the observations we are not talking about some square deviation some square of all the observation j equal to 1 to n, i equal to 1 to a and if this is subtracted by R mu and tau. What is this? This is what is the reduction of variability if the ANOVA model is used say in them the difference is giving you some square errors. That means, this is the variability in y and this one is here mu is included. So, this one is the variability including mu means what I mean to say y ij equal to y ij equal to mu plus tau i plus epsilon ij, I am taking square of this.

Suppose then and this, this one this one ultimately we want this. So, y ij is a equal to epsilon ij equal to y ij minus mu minus tau i. So, that sense if I square this and then take the sum you will get this and this one is coming here and whatever this, this is what is the model part. So, this model part is coming here now you understand. So, as a result we are saying this is SS E. Now, then what about this R mu and tau? That is we said that mu cap y double dot plus sum up sum of tau i cap y i dot i equal to 1 to a.

So, if you expand these you will get this one you see what you are getting if you expand this you are getting this mu i cap is nothing, but y grand average into this and into this and if you expand this, you will be getting that is nothing but y i dot square by n i equal to 1 to a y i dot square by capital N small n and square by more small n. Now that means, they say if I put in SS E then this is nothing, but j equal to 1 to n i equal to 1 to a y ij square minus sum of i dot square by n. So, this is your SS E.

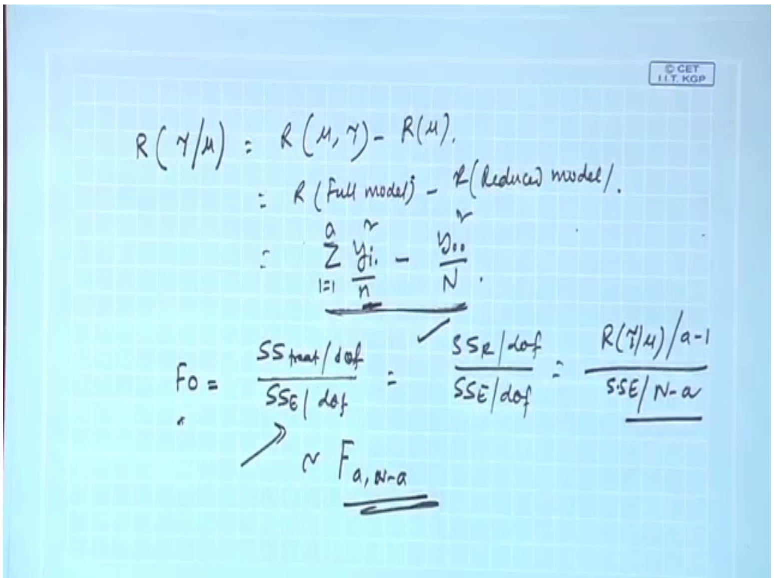
(Refer Slide Time: 28:00)



Now, this you are getting from the full model now if we do the same thing from the for the reduced model mu plus epsilon ij. Then epsilon ij is what? y ij minus mu if you t square you make square it and some of this then SS E what happened you will get from here. But rather we are interested in this R we are interested only R mu here mu and tau because for the full model and here these what is the procedure multiplied by. So, estimate is mu cap multiplied by its total, now this will give you y dot dot square by capital N you mean multiply do this you will get this.

So, as a result what happened now, now if I want to know what is the R value of the x factor. That factor one factor given mu what will happen then, this is nothing, but R mu tau minus R mu.

(Refer Slide Time: 29:10)

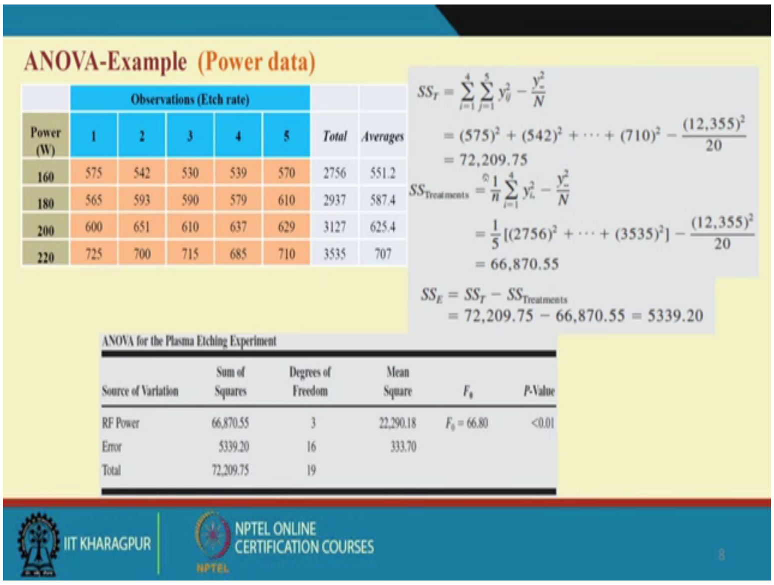


So, that mean R of full model minus R of reduced model. So, what do you got in R of full model we have already seen this is y i dot square by n i equal to 1 to a minus what is the R of reduced model this square by N. What is this value? This is nothing, but the treatment effect contribution sum square treatments sum square R, this is R.

So, anyhow using all those things now what happened, you got SS E using this formula you are getting the that factor contribution or SS R or as a statement by this formula 1 by n of this formula. So, I think what I have done here. So, this one y by n is there fine 1 by n of this formula then you have you can find out that F 0 value, F 0 value is what? F 0 value is our SS treatment y a degree of freedom by SS error by degree of freedom from ANOVA point of view from regression point of view SS bar by degree of freedom by SS E by degree of freedom. This is nothing, but now SS bar a is here R tau given mu divided by a minus 1 and then SS E already you know the formula is given to you. So, I am not writing here further SS E by N minus a. Now, if this follows same thing a N minus a and rest of the things known to you.

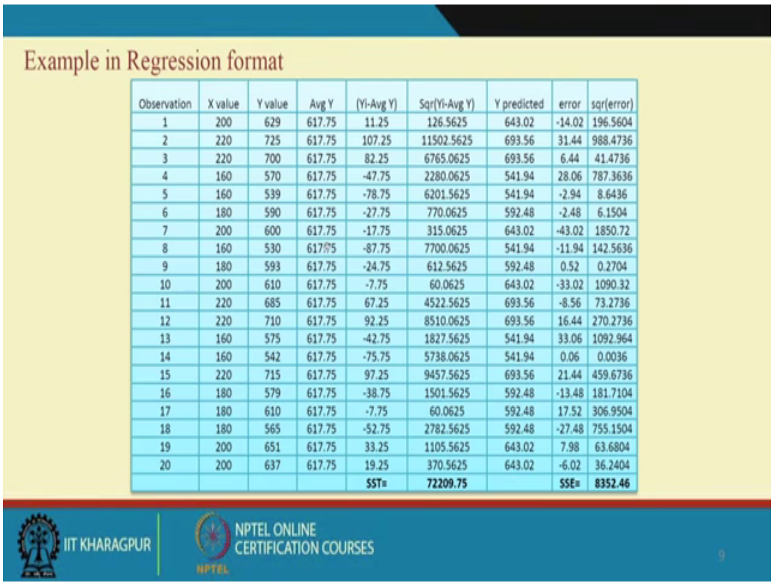
So, very quickly I will show you that in ANOVA for the power each read data we have seen SS T is 72,000 all most, treatment 66, SS error is 53 like this and then this is the ANOVA table and we found out that there is treatment effects.

(Refer Slide Time: 31:46)



Now, we have used the same thing for regression. What we have done all there are 20 observations, then x value for first observation 200, 200 like this, then y values are there, then regression model is fit then y corrected sum square everything.

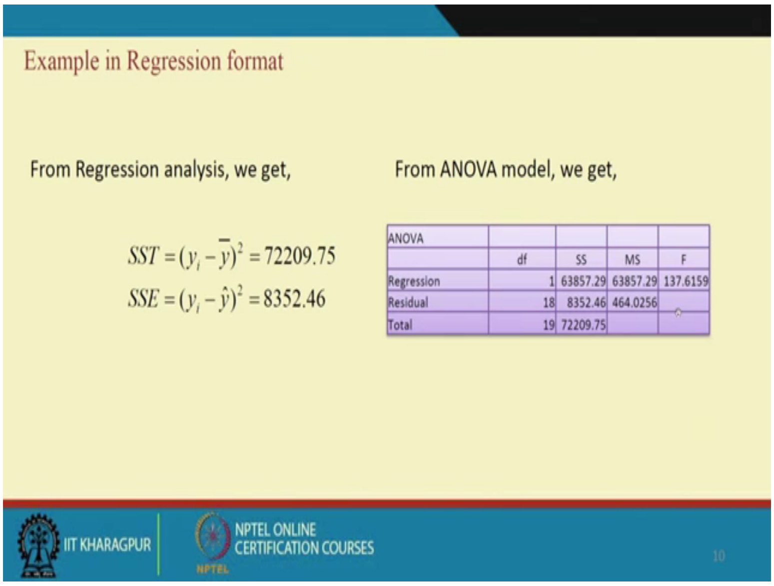
(Refer Slide Time: 32:00)



Now, SS T is 72.209.75 here 209.75, then SS E is 8352.46, but here it is less so maybe that is because of the rounding error and other things may be there. But whatever may be the thing, but ultimately total is same and then is the predicted values we have taken for

2 decimal. So, the error is also put 2 decimal maybe that may be the effect or may be the approach from the approach also there is little different values.

(Refer Slide Time: 32:53)



And then what happened we found out that from regression also the statistics is very high and we are saying that it is significant.

So, from ANOVA as well as from regression both case you are finding that the power has effect on the each eth. So, any one way you can go, but please keep in mind here the degree of freedom each number of parameters to be estimated minus 1. So, intercept plus the model that coefficient regression coefficient for power only for 2 minus 1 and then residual is more and that is basically this how. So, that is MSE is 464, now what about MSE here 333.

So, find out if you can find out that there is, there may be calculation error please tell me in the forum or if you find out that know that there is even further or much closer explanation for the 2 also discuss, all those things will come.

Thank you very much for your patient hearing.