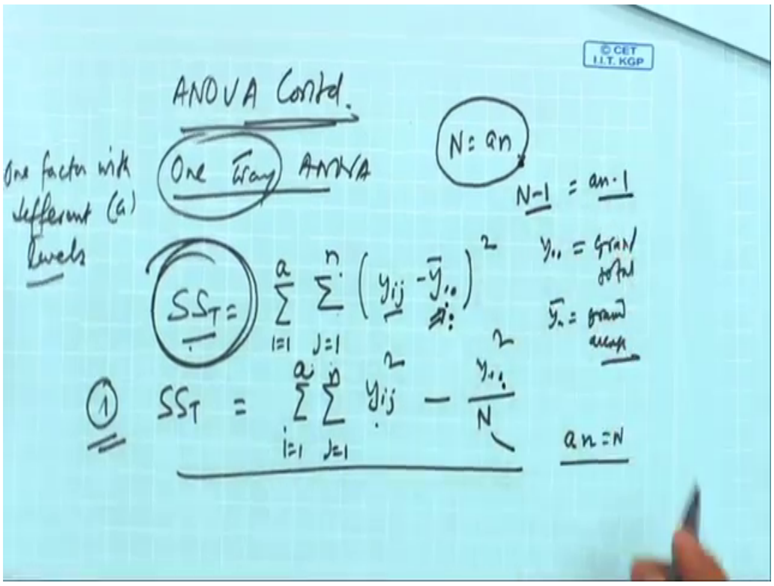
Design and Analysis of Experiments

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

Lecture – 14
Analysis of Variance for (ANOVA) (Contd.)

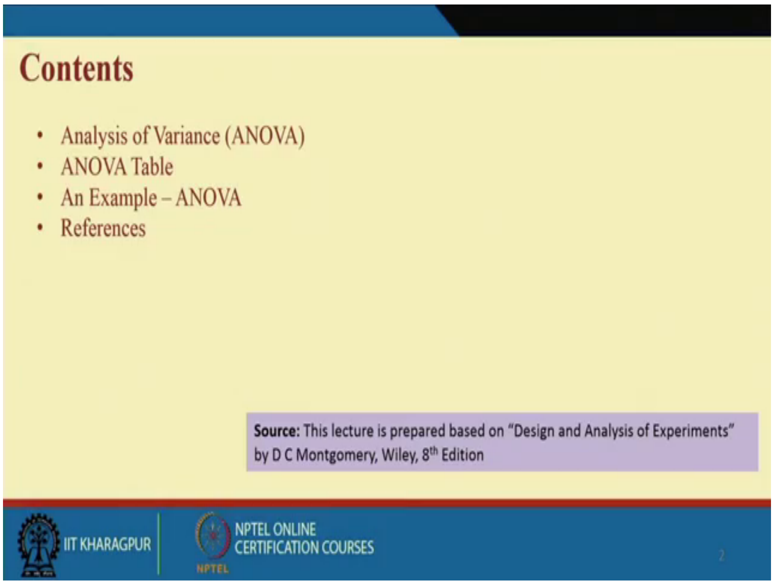
Welcome we will continue analysis of variance. So, I am writing one way analysis of variance.

(Refer Slide Time: 00:21)



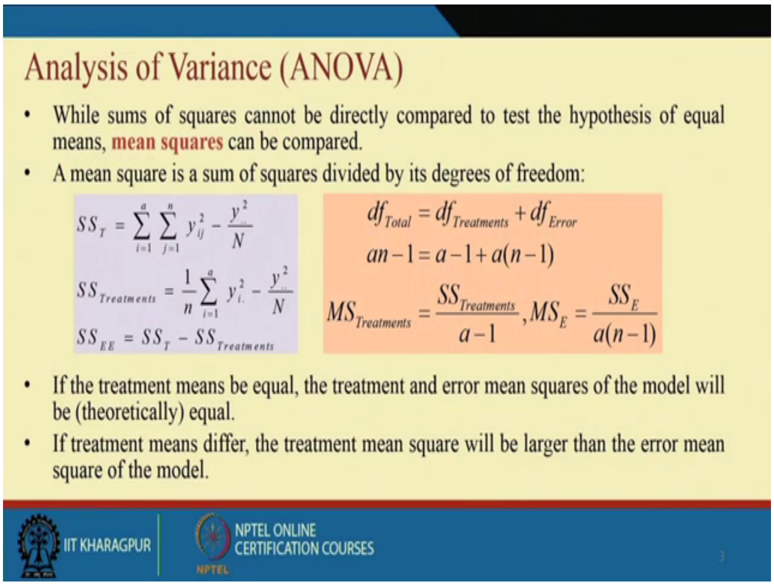
So, the word one way is coming from only one factor with different levels, different means a levels. So, last class what I have discussed that is also this one fact, one way ANOVA.

(Refer Slide Time: 00:45)



So, I will see the same example, the calculation, and also we see that how ANOVA is used for hypothesis testing of the equality of the different level means.

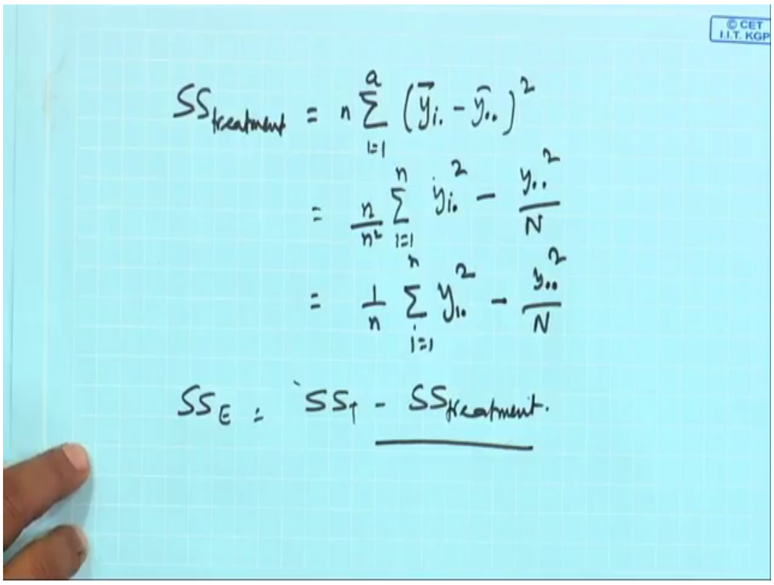
(Refer Slide Time: 00:57)



So, let us see this, last class you have seen what is the formula for computation of SSt. You have a number of levels, you have n number of replication, you have individual observation y i j, everything is subtracted by their grand average, and you make the square, this is what we have given. Now this can be simplified to j equal to 1 to n i equal to 1 to a, then y i j square minus y dot dot square divided by N. So, what are the differences y i j you know this y i j, but what is y i dot dot y i dot dot is.

Sum total grand total what is N here. In this case I have a levels, each level n replication. So, n equal to a n what is y i y dot dot. This is the grand total a total of all y i j, and then what if I write what is this. this is grand average that I have discussed in the last class. So, this is the second one for, if you use SSt equal to the, this makes your computation easier.

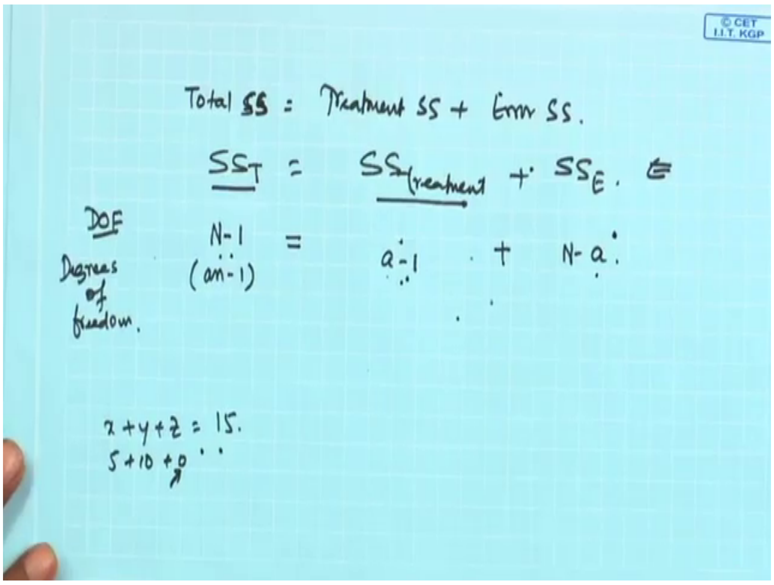
(Refer Slide Time: 02:48)



Now, this formula we will use for SSt. Now similarly you can compute SS treatment, but in SS treatment what I have said, you have n replication against each level, then i equal to 1 to a different levels.

So, every level you have average then minus grand average this square, this is what we have seen in the last lecture, this can be written like this y i equal to 1 to n, then y i dot this square minus y dot dot square divided by n. And here it will be 1 y square will come. So; that means, it is nothing, but 1 by n i equal to 1 to n y i dot square minus y i dot dot square divided by n. So, that mean every row total square take their sum divided by n, and this portion is the correction factor, or I can this y dot dot square by n that will be total, subtracted these two you calculate SSt and SS treatment, then SS e simple subtraction SSt minus SS treatment. So, this is the formula we will be using. So, I am repeating the sum square decomposition.

(Refer Slide Time: 04:31)



The total, total some square SS mean sum square equal to treatment sum square plus error sum square, this is SSt equal to SS treatment. I am repeating this, because this is very important SS e error. Now another important concept is, that what is the degree of freedom.

DOF means degree of freedom, when you are computing SSt what is the degrees of freedom you are enjoying. So, please come to the SS total computation, what I have written here SS in order to calculate SS total, you have n number of observations attend, n number means a N capital N these numb amount of data you have, but for every time you have subtracted this grand mean grand average. So, you have estimated grand mean that mean one parameter is estimated here, which is subtracted from this, as a result what happened as every time you are subtracting this. So, what happened?

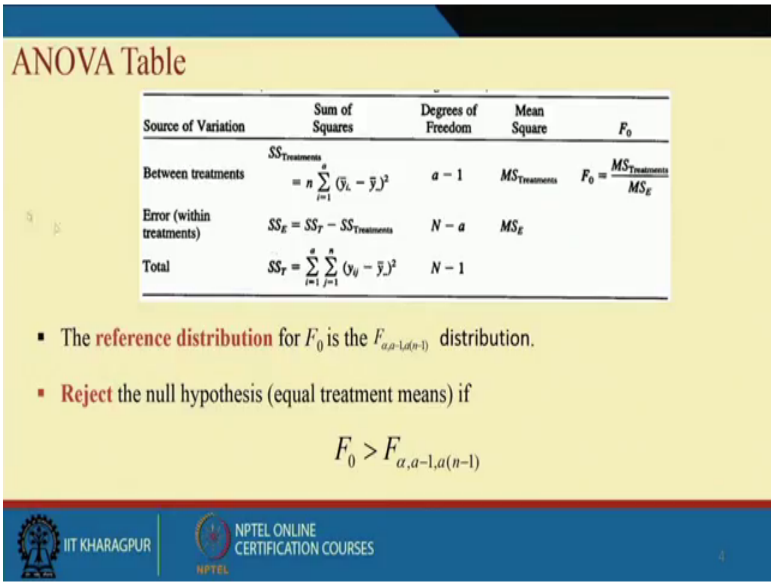
You a eventually in this computing this statistic SS total you have N minus 1 independent observations for computation of these one is lost. So, that mean this equal to a n minus 1. So, what I mean to say that degree of freedom for this is N minus 1 or you can write a n minus 1. So, it is something like this. Suppose I have given you x plus y plus z equal to 15. So, what is the degree of freedom you have; in this calculation you have two degrees of freedom, the reason is if you write this 5 this 10, then this will automatically become 0, same things happens here, same thing happens here. So, long you, as soon as you put N minus 1 number of observations here.

The N th observation capital N N th observation will automatically we calculated. So, you have effectively N minus 1 freedom. So, now, what about SS treatment, how much data you are, you are actually enjoying when you are calculating SS treatment, when you are calculating SS treatment you see that you have a number of observations only this i equal to 1 to a; that is varying.

So, a number of observations, again this grand mean is calculate computed. So, that also coming from this a number of means; so, 1 degrees lost. So, that means you have here a minus 1 degrees of freedom. So, then N minus 1 degrees of freedom here equal to a minus 1 plus this degrees of freedom, also what happened the total degrees of freedom remains for calculation of error already here a minus 1.

Degrees of freedom in here 1; so, what will happen you have N minus a degrees of freedom left to calculate the error and you see N minus 1 equal to a minus 1 plus N minus a. So; that means, not only the total sum square a is divided into two parts, like some sum square treatments sum square error, your total degrees of freedom is also divided into two parts, and total degrees of freedom is equal to treatment degrees of freedom plus error degrees of freedom very important, one second.

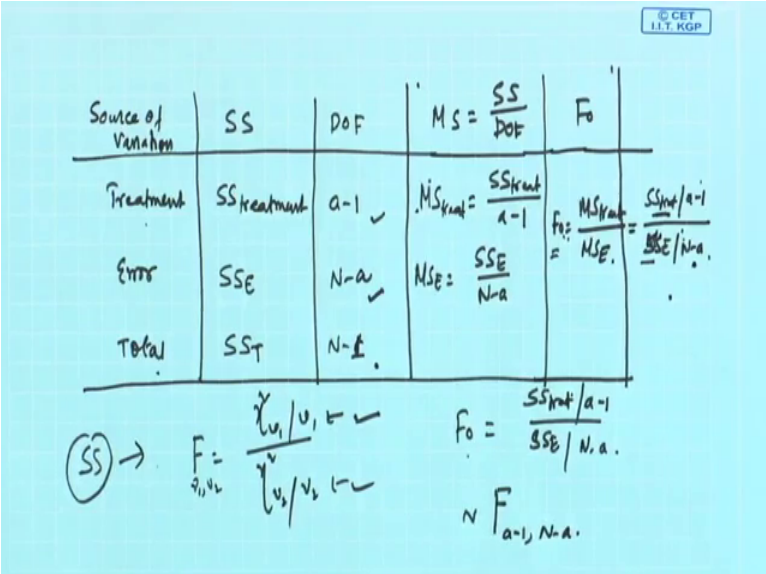
(Refer Slide Time: 08:44)



Now, I will show you that the ANOVA table come to this slide. So, SS treatment SS error by subtraction and SS total by this formula N minus 1 a minus 1 N minus this way degrees of freedom.

So, you calculate another sum square, which is known as mean square. What is this mean square. Mean square is the sum square divided by degree of freedom. So, that mean your ANOVA table will be like this.

(Refer Slide Time: 09:18)



Sources or source of variation number 1 in this example, how many source; one is treatment another one is error, and then the total. Total is nothing, but treatment plus error. Then suppose I want to know what is their sum square. So, sum square treatment, then sum square error, then sum square total, and you know how to compute this sum square treatment, the formula already given to you. Now what is the degree of freedom. Degree of freedom for.

Treatment, is a minus 1 for this equal to N minus 1. So, this will become N minus a. So, total degrees of for while calculating SSt degree of freedom is N minus 1, while calculating treatment degree of freedom is N minus 1, then remaining N minus capital N minus a degrees of freedom for SSt, then what is mean square.

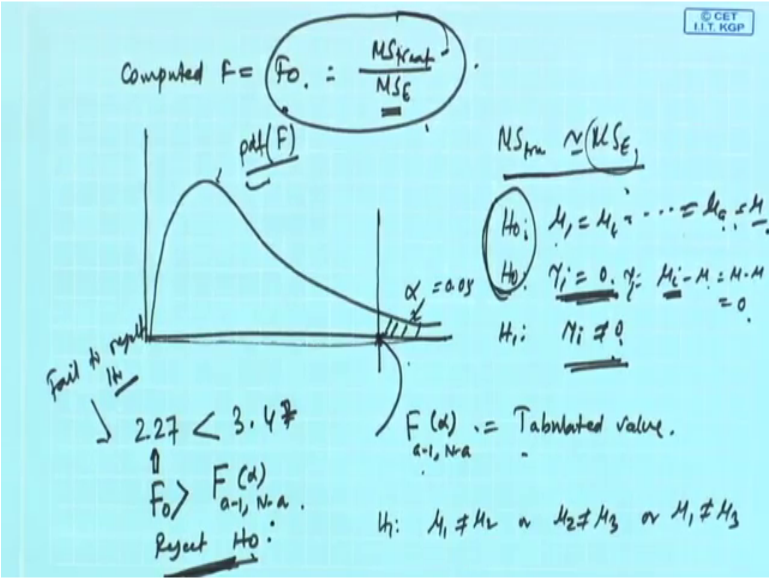
Mean square means that this is nothing, but SS by dof. So, SS means for the respective one. So, I want to know M S treatment. This will be SS treatment divided by a minus 1. You want to know a M S error which is nothing, but SS error by degree of freedom SS error by degree of freedom.

So, in this ANOVA table you see that everything is there, apart from this there is another quantity called F 0. F 0 is nothing, but M S treatment by M S error. So, we are interested to know. also another one which is F 0 equal to M S treatment divided by M S error. so; that means, this is nothing, but SS treatment by degree of freedom by M S treat SS error by degree of freedom. If you recall that in the when I talk about sampling distribution.

Then I have given you that SS, this one follows chi square distribution with N minus 1 degrees of freedom. So, here what happened SS treatment follows chi square distribution with N minus 1 degrees of freedom, and M SS error with N minus 1 degrees of freedom, and this is the ratio of two F 0 either ratio of 2 chi square variable waited by the respective degrees of freedom. If you know that F distribution we say chi square nu 1 by nu 1, then chi square nu 2 by nu 2, where nu 1 is the numerator degrees of freedom, nu 2 is the denominator degrees of the freedom. So, that is why this quantity is also, and then this will be nu 1 n 2 chi square nu is like this. Now you got one quantity F 0 which is that SS treatment by.

A minus 1 degrees of freedom and SS error by N minus a degrees of freedom, then this will follow F distribution with a minus 1 and N minus a; the numerator and denominator degrees of freedom respectively. So, this is what is given here. So, find alpha I will tell you later on, but a minus 1 and N minus a. N minus a is nothing, but a a into a N minus a. So; that means, a into N minus 1. So, now, you have to do the hypothesis testing, what happened.

(Refer Slide Time: 13:43)



You have computed af F computed F computed F is F 0 how do we compute. Once you know M S treatment and M S error, you are in composition to compute, and you know how to compute SS from the data, then you will be able to compute this one.

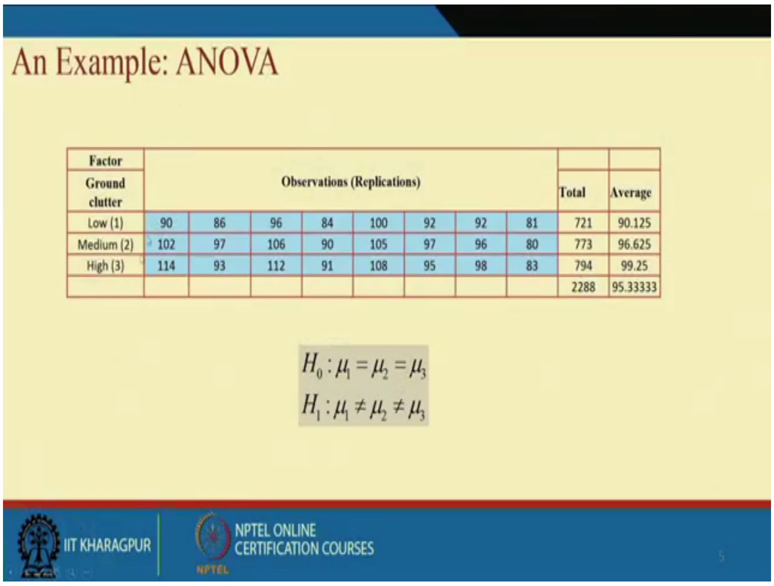
So, this follows theoretically. Suppose this is this following F distribution something like this. So, this is our F distribution pdf of, probability density function of F. So, what do I interested to say that, whether there is difference in the means of the different levels. So, if there is no difference then what will happen MS treatment and MS e will be equal MS treatment, and almost equal to M S e. What is MS e. This is a error, this is a random one, the error, what is occurring here, it is because of random. So, again that M S treatment the different level of a, it is also equal to random error effect means there is no effect; that means, the different means are not different.

The means of different at different levels are not different. Not difference means, they are not different means that different level treatment levels are not different. So, this is my H 0 mu 1 equal to mu 2 equal to mu a. Other way I can write H 0 tau o tau i equal to 0 and H 1 tau i not equal to 0, how tau i equal to. What is tau i. tau i equal to mu i minus mu. now if mu i equal to mu 2 equal to mu a, this will be nothing, but grand mean then, because mean and mu this will be this. So, mu i become mu, then tau i is mu minus mu that is become 0. So, our H 0 is no effect H 1 is there is effect. So, other way I can say if there is no; so, under this null hypothesis.

This quantity, this quantity follows F distribution if H 0 is true, this quantity following distribution. So, what we want then. So, this is the F line. So, any value possible late from the theoretical sets, this may be far away from somewhere we will create a threshold value for a f. So, this one will be alpha. So, then this value is F n minus 1 n minus a into alpha, if that this is known as the tabulated value. So, if your computed value F 0 is greater than or greater than equal to tabulated value F alpha a minus 1 n minus a, then you reject null hypothesis.

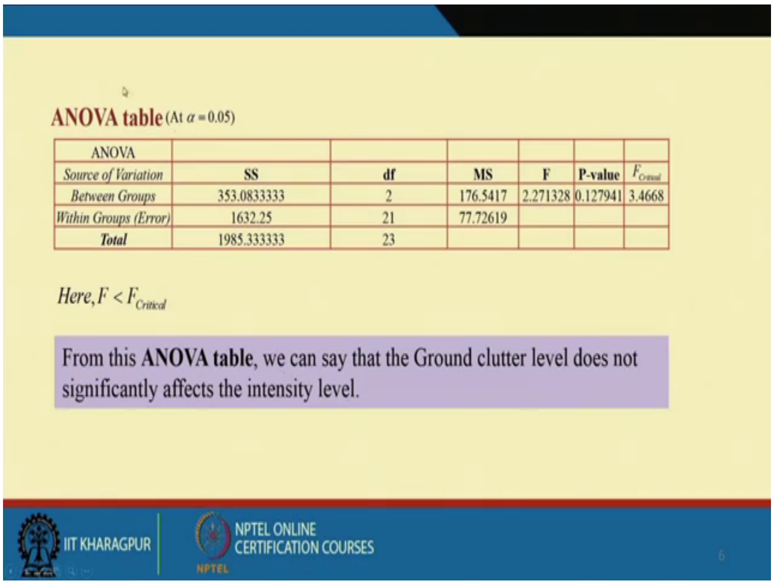
What is the null hypothesis? There is no treatment effect or mean of the response variable at different treatments levels are not different; that is what is F 0? So, from this you are in a position to say that is that there is the null hypothesis, its true or false, if it is false reject null hypothesis.

(Refer Slide Time: 17:26)



So, let us see the example, same example here ground clutter different level is low, medium, high, average our feel three levels low medium high mu 1 equal to mu 2 equal to mu 3 H 1. They are not equal, may be at least 1 is not equal, need not be that all three has not equal, but at least at least H 1 will be at least mu 1 not equal to mu 2 or mu 2 not equal to mu 3 or mu 1 not equal to mu 3. If any one or more of them satisfied that is what is H 1 in there is at least one pair, which is different.

(Refer Slide Time: 18:03)

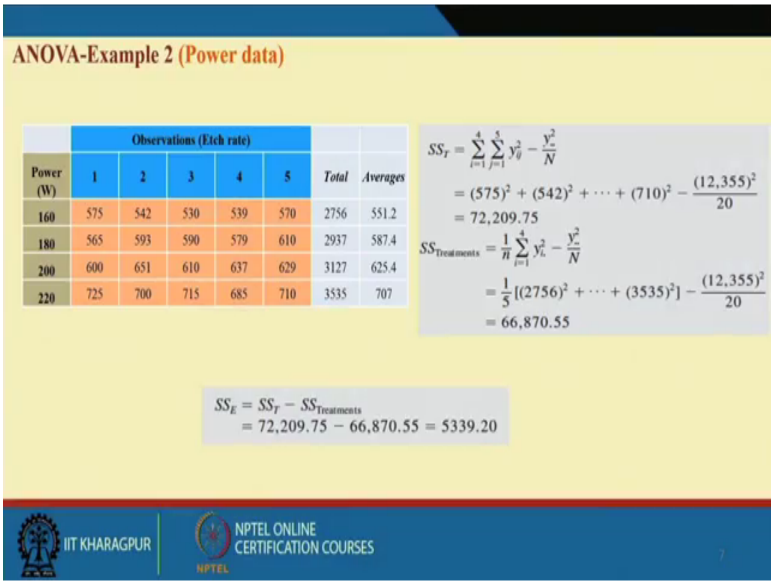


Now, using ANOVA table all those what we found out that F value is 2.27, and if I consider alpha equal to 0.05; that means, 5 percent error, each considered type 1 error is error is 0.05, then the tabulated F value is 3.46. So, computed F value is 2.27 which is less than tabulated F value 3.47. So, as computed value does not exceed the tabulated value, you cannot reject F 0. If you fail to reject our decision here is failed to reject, reject H 0.

So, that mean there is no treatment effect or if you change from data, I mean clutter level from low to high, there is no difference, but you may not except it, because conceptually or from actual field, you are, there is difference. If this is the case then the data, what data, we have used data collection either data collection is from or the data we have used, may be the data from another experiment we used differently.

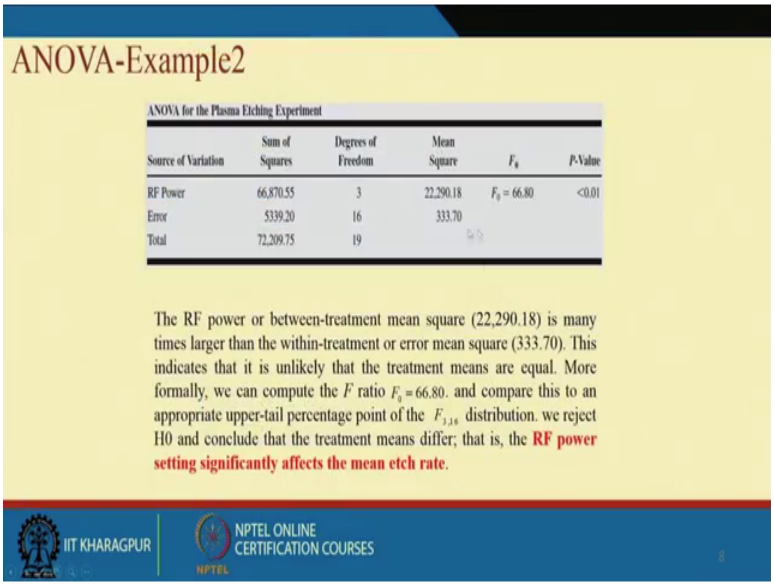
Actually that is what we have done here. If the original data is a basically with two factors, and with that one noise variable that operator, but here intentionally what I am, the same data set we kept, but individual level factor we are here clutter level, we have comparing. So, that may be reason bur, whatever may be the thing the procedure is like this

(Refer Slide Time: 19:55)



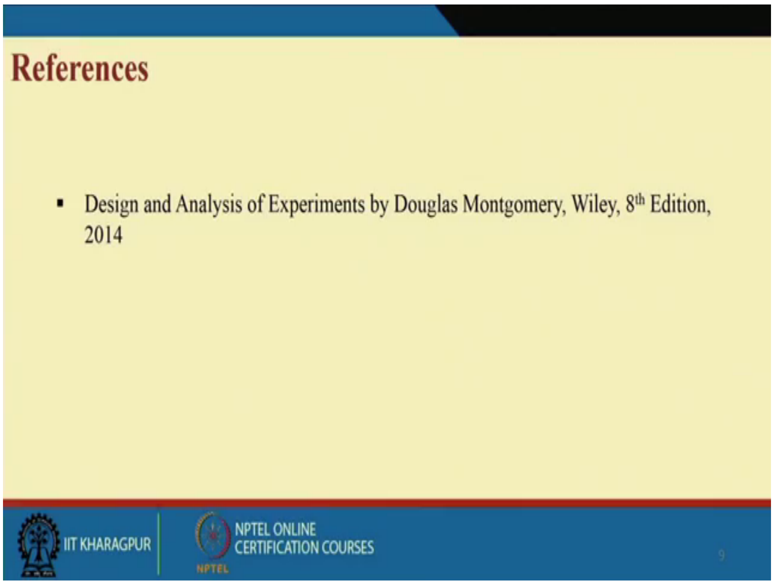
So, now see is another one that power data; that is the insulate case here, what happened when we are doing.

(Refer Slide Time: 20:03)



All kind of energy calculation, and finally, making the ANOVA table. You see we are getting the computed F 0 is 66.80 and tabulated one is again if I consider this one, that a minus 1 is 3 and here error degrees of freedom is 16. So, 3 into 16 degree degrees of freedom, then we will find out that value is that, value is much lower than 66.80. And in fact, the probability type 1 error is less than 0.01, that mean in the second example case, the things are coming somewhere here this side, this much error is there. So, if my threshold value is here, and my actual computed value is falling here. So, that what does it mean, it is satisfying the second one. So, reject H 0 that mean, the different power levels are effecting the mean each rate for the second experiment.

(Refer Slide Time: 21:07)



So, thank you again, and it is again the montgomery book DOE.

Thanks a lot.