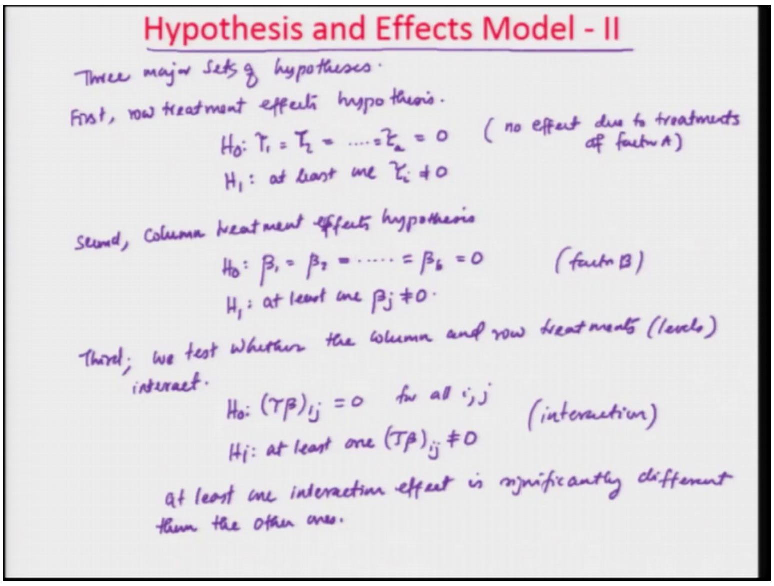
Advanced Green Manufacturing Systems

Prof. Deepu Philip
Dr. Amandeep Singh Oberoi
Department of Industrial & Management Engineering
Department of Mechanical Engineering
Indian Institute of Technology, Kanpur

Lecture – 36
Statistical Analysis in Factorial Experiments

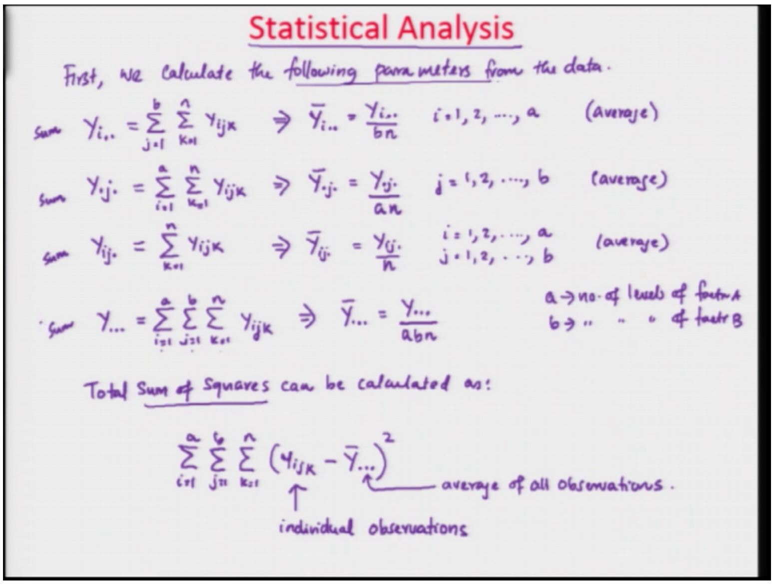
Good afternoon students. Welcome to the continuation lecture of green manufacturing, on which we were studying factorial designs and today, what we will do is; we will continue from where we left off in the previous class.

(Refer Slide Time: 00:32)



So, if you look into this screen, then you can see that, we talked about the three major set of hypothesis, where we just mentioned about the treatment effects hypothesis, which is the row treatment effects, which is the work, which we denoted by the variable tau and then the column treatment effects, which is the beta and then we talked about the column and row or the interaction levels, where is tau beta ok. We talked about all these three hypothesises.

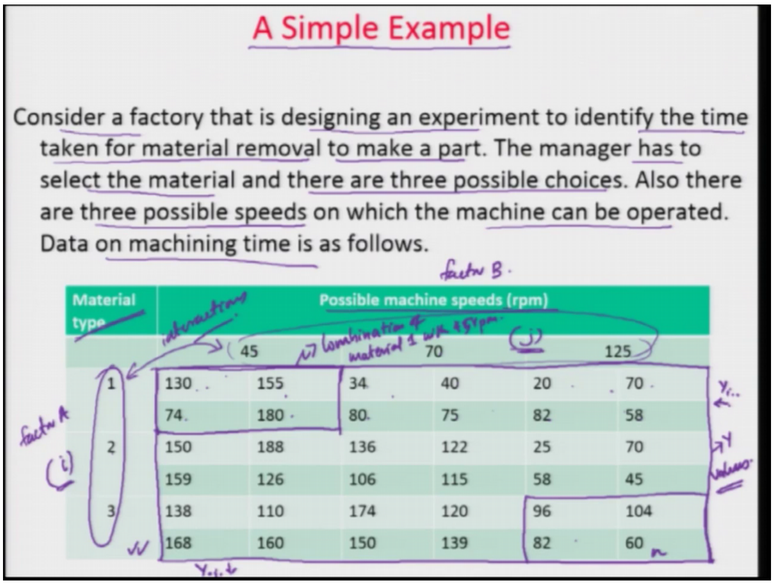
(Refer Slide Time: 00:58)



And now, what we are going to see is how do, we do the statistical analysis of the same to after the hypothesis, how do we do the statistical analysis. So, the first thing we need to do is; first we have to calculate the parameters from the data. First, we calculate the following parameters, the following parameters to the data. So, these parameters are to be calculated from the data ok, these parameters are to be calculated from the data. So, the first parameter that we need to calculate is what we called as Y i dot dot. You have already seen this dot dot business earlier in the anova.

So, this Y i dot dot is given by the equation sigma j is equal to 1 to b. So, this is 1 to b is the column. So, remember, if you go back in the presentation we told about this is the j.

(Refer Slide Time: 01:59)



So, we are talking about j is the coloumns and i is the row, remember that right. So, j is equal to 1 to b for all th values of this and we say sigma k is equal to 1 to n, k is the count for all the individual observations. We are calling it as Y i j k. So, what we are saying is that, for all the values of j equal to 1 to 1 and for all the values of k equal to 1 to n, we are summing the values of the j s and k. So, the j and k will vary.

So, the Y i dot means; we are actually doing the these, the i will remain the same. So, this will be i. So, you will sum all of these k s for each values of this j ok. So, that will give us the Y i j dot. So, then from here, we can calculate Y i dot dot bar that can be calculated, which will be is equal to Y i dot dot which is sum a summation of this divided by b times n b n for so, we can say i is equal to 1 2, all the way to a. So, for each row we will calculate Y i dot dot bar. So, to cacluate Y i dot dot bar, we require Y i dot dot and the Y i dot dot is calculated by this particular equation.

Similarly, the next thing that we calculate is Y dot j dot, this is the next thing that we will calculate as part of this next parameter and the Y dot j dot is calculated as here, the index will be on i first, summation of i is equal to 1 to a for all rows. Now, we are doing and we are going to do the next one is k is equal to 1 to n. So, for all individual observations, we are doing it as Y i j k.

So, the index is on i and k with this line. So, i and k will vary, j will remain constant. So, which implies Y dot j dot bar, the average of that, which will be equal to Y dot j dot divided by a times n where we say j is equal to 1 2 all the way up to b. So, in this case, what we are doing is the previous scenario as we were doing for each individual columns ok. So, we were doing the Y i dot dots for each individual values. We will doing be the Y i j dot will be do, be done for each each one of these columns where as this will be done for each individual rows right. So, with that so, we calculate this second parameter. So, first we calculate the summation value.

So, these are the so, this will give you the sums ok. This is the sum, this is also sum. This gives you the average, this also is a another average or mean value right. So, then that we do is, we calculate the third one, we are going to do is Y i j dot. So, what we are going to do is i and j are remains a same and the k will vary. So, this is is equal to, summation of k is equal to 1 to n Y i j k. So, what we are doing is for i and j, for all i and j they will remain the same, but we are just going to do the summation on the k.

So, then from here, we can calculate Y i j dot bar ok, the average of this. So, this is a sum ok, using the sum we calculate this average and this is calculated as Y i j dot divided by n and here for this case, it is meant for all i is equal to 1 2 3, all the way up to a and j is equal to 1 2 3 all the way up to b ok, this is meant for this particular case. So, this is also another average ok, anything about, let us say third average. Then the fourth parameter that we are going to calculate in this case will be, in this particular scenario will be Y dot dot dot. All the three things.

So; that means, it is a summation, it is a summation of all the three value. So, you have sum of i is equal to 1 to a, then you sum on the column js equal to 1 to b and then you sum the individual observations the applications within this k is equal to 1 to n and we do it is Y i j k v. So, what we are doing here is the Y dot dot dot is the sum, the sum of all observations, then where you are indexing on i j and k and summing it up. So, this will give you definitely another average, which is the Y dot dot dot bar ok. So, this is the grand average or the big average of averages, which will be Y dot dot dot divided by a times b times n.

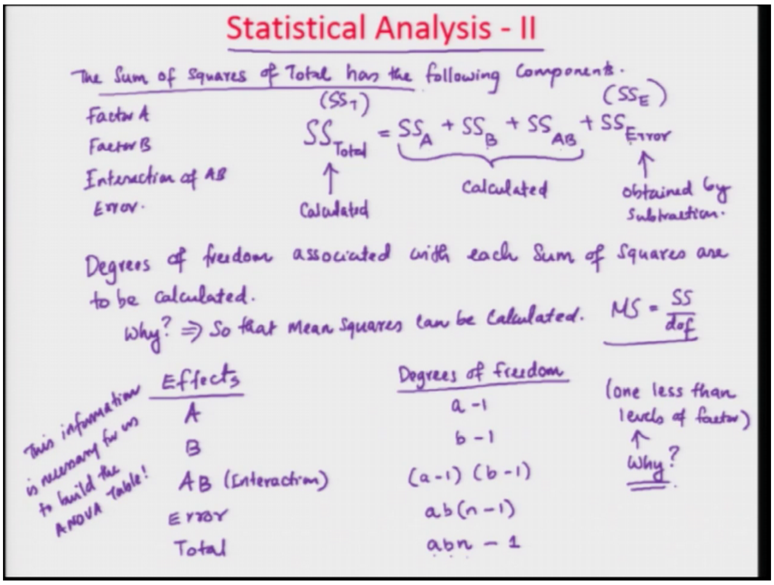
So, what happens if a is the so, a is the number of levels of factor A ok, b is the number of levels of factor B fine. So, a and b are the different type of levels that we talk about. So, all these parameters are mandatory for us to be calculated. So, that we can calculate the the rest of the major things. So, the first thing that we need to calculate is after this is done is, the total sum of squares. Remember, we were calculating the sum of squares in the case of ANOVA and total sum of squares can be calculated as, can be calculated as ok.

So, the total sum of squares in the ANOVA we were using it, using a particular mechanism, in this case the total sum of squares will be equate given by the equation sigma i is equal to 1 to a sigma j is equal to 1 to b and then their replication sigma k is equal to 1 to n n replications Y i j k. These are the individual observations minus Y dot dot dot bar whole square.

So, these are the individual observations observations and whereas, in this case this is the mean ok, which is the average of all observations or the grand mean ok. So, what we do is individual observations minus the average of all observations, find the deviation is square them and sum across all these values, then you will get the total sum of squares in this particular case.

So, I think you guys understand what I was talking about in this case, where we first calculate different sums, the row sum, a column sum and then individual what we called as the replication sum, then we also calculate the grand sum and using which we calculate different averages fine.

(Refer Slide Time: 09:48)



Now, the second part of this once, we calculate all these values remember we had something called as an ANOVA table. So, same way we will be creating another ANOVA table here. So, to do that ANOVA table there is some more things that we need to do ok. So, the sum of squares of total ok, also another way is that the sum of squares of total, which means the total sum of squares has the following components so remember. So, in this example we are doing a components. So, for this particular case we have the, we have a factor a we have A factor B ok, then we have a interaction of AB right and then there is error, we had all the, we mentioned all this kind of things.

So, the sum of squares of total, which is also known as SST is the sum of sqaures has the component due to the sum of square of a factor, a sum of squares due to A factor B and sum of factors due to factor A B and aware sum of squares of error ok. So, this is how we actually calculate or the total sum of squares has this particular components ok. Typically, what we do is we calculate. So, this is calculated and this is also calculated and this is obtained by subtraction.

So, when we calculate the sum of squares of total and these three are calculated, then the difference of these three with this from most total will give you the sum of squares of error right. So, that is how we calculate this. When we do the ANOVA table, it will become little bit more clear to you and what we also do is we also need to know the degrees of freedom, of freedom associated with each sum of the squares, associated with each sum of squares are to be calculated ok. We need to calculate this, why? Why do we need to calculate this degrees of freedom? So, that mean squares can be calculated ok.

How do you calculate mean squares ok. So, the by definition the mean square M S is equal to sum of squares divided by degrees of freedom. So, this is our fundamental equation on which we actually make the the sum of squares, we conect the sum of squares with the mean sqaures. So, the main effects in this case is the effect for us, in this regard is is due to factor A, then factor A B factor B and then factor A B ok, which is the interaction ok, then we have is the error and then we have is the total ok. So, these are the effects that we have all the effects, that are associated with this then we also have is the degrees of freedom ok.

So, what are the degrees of freedom in this regard ? The major degrees of freedom in this regard is the effect of A is given by a minus 1 ok. So, this is 1 less than levels of factor.

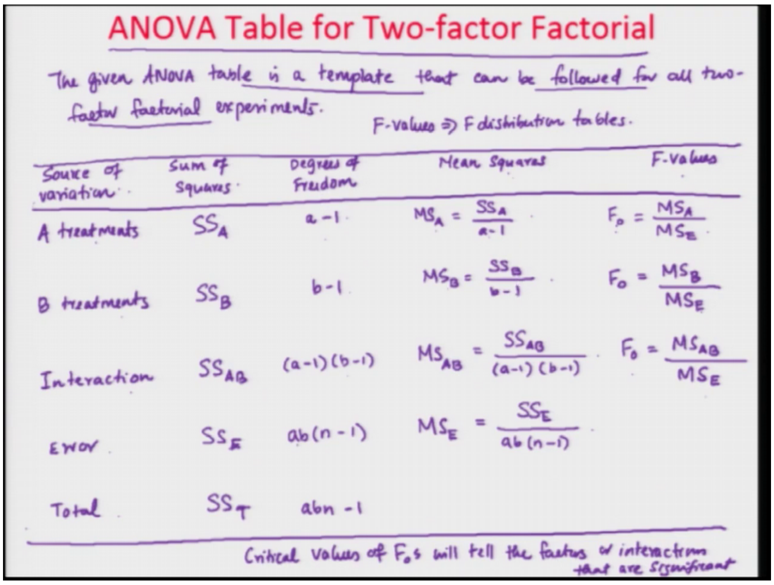
So, the factor is that number of factors is A. So, you reduce 1 out of it is 1 less than the number of levels in this regard. So, you get a minus 1 right then the B is given by b minus 1. There are small b, this many number of levels. So, 1 out of minus out of this will give you the degrees of freedom. The A B interaction is given by a minus 1 multiplied by b minus 1 right.

And the question obviously is that, why it is always 1 less than that? That question please, you know I we have ensure. We have explained in the previous one when we did the analysis of variance, because the reason is that if you have n observations you only need to specify n minus 1 of them. Because, the n th 1 can be always obtained by by specifying the n minus 1 th and the average of it right then the; so, we get the three interaction of sum of squares of degrees of freedom are also this way.

Then the other degrees of freedom is the error which is calculated by a b times n minus 1 n is the number of replications in this regard. So, n minus 1 times a times b will give you the error degrees of freedom and the total degrees of freedom is given by abn all of this multiplied by minus 1. So, this is the a b n is the number of levels of a number of levels of b and number of replications all of them put together and multiply them and then minus 1 will give you the degrees of freedom in this regard.

So, this degrees of freedom are important for us to. So, this information is necessary, information is necessary for us to build the ANOVA table. ANOVA table is required to analyze the factorial effects of the factors and their interactions.

(Refer Slide Time: 16:08)



So; obviously, then we will move to the how do we do create the or how do we build the ANOVA table for the two factor factorial experiments. So, the given ANOVA table is a template is a template that can be followed that can be followed for all two factor factorial experiments.

So, please remember this is a template ok, this ANOVA table is template template means; you can take this template and it can be followed for any of the two factor factorial experiments, but understand that instead of these values you will have to put appropriate, names of the factors and other aspect in this regard. So, the table is typically drawn in this fashion. I cannot draw straight line. So, please bear with me. So, this is the the titles will go here.

So, the first column, what we call it as the source of variation. The first column will give you the source of variation, then the next one we will call it as the sum of squares that is the second one, put the third one is the degrees of freedom right, then the next one, what we do is what we calculate is mean square mean squares right and then what we do is finally, we calculate the F values or the F distribution values. Remember in this regard one of the things that you need to remember is F values F distribution tables.

Please study, what is an F distribution its a which belongs to the chi square family of distributions. Anyway, we will talk about this later. So, the source of variation the major sources of variations are written here, A treatments. So, whatever is the factor A and the type of treatments we write there. Then we write the B treatments, let me give some space in this regard, because this is a template. So, we can do with that way the B treatments then we have is the interaction interaction, these are the another source of variation.

Then what we have the third source of variation is the error and the last source of variation is the total. So, we will write all these things right here and this is the table template in this regard. So, then the sum of squares are to be done. So, here is SS of A and this is is the SS of B.

We do not write this here, but SS of A is calculated where we will be actually using SS of A to be calculate. I will give you the equations for how to calculate this later, but using these factors that we have seen you will be calculating these sum of squares in this regard and remember sum of squares total A B AB and error are the factors of it.

So, individual sum of squares will actually go in this regard, the third case it will be SS of AB will go here. The numerical value will go here actually, the error will be SS of error will go here and the SS of total will go in this regard, SS T is the total it is not SS T is not treatments SST is the total sum of squares of total ok. Then the degrees of freedom remember the individual degrees of freedom I have mentioned to you what it is.

So, that degrees of freedom are also put in this table as factor of it so that anybody can understand what is going on. So, for A that will be a minus 1 as I mentioned earlier, because there are A of a such treatment levels for factor A, for B it will be b minus 1 and for A B it will be a minus 1 multiplied by b minus 1 ok.

So, then the sum of squares of error will also be ab multiplied by n minus 1 and the last one the total will be a b n minus 1 right. So, we can see that if you sum all of these values, then you will get the total sum of squares and the degrees of freedom. They will all sum nicely with each other right.

So, then once it is done we have the sum of squares and degrees of freedom. The next thing for us to do is the calculate the mean squares. So, the M S square mean squares of A treatment is is equal to sum of squares of A divided by a minus 1. So, if you take the sum of squares of A divide by it's own degrees of freedom you get the mean squares of A. Similarly, the mean squares of B is calculated by sum of squares of B divided by b minus 1. So, sum of squares of B is already calculated here, b minus 1 is the degrees of freedom.

So, using that degrees of we can calculate the mean squares of B right. Then the next one is mean squares of A B or what we can call it as I am calling A B in this case, which is the interaction, which is equal to sum of squares of A B, we calculate the sum of squares of the interaction then divided by the degrees of freedom which is a minus 1 multiplied by b minus 1 right.

So, once you divide the interaction sum of squares with it is on degrees of freedom, you get the mean squares of the interaction. Then the last mean square, we can calculate is the mean squares of errors MSE right and MSE is calculated as sum of squares of error SS E divided by what we call as the error degrees of freedom that is a b times a b multiplied by n minus 1.

So, with this system. So, we calculate the mean squares of errors we do not calculate the mean square of total, because absolutely that is not required in this regard. So, this for mean squares the mean square due to the main effect of factor A, mean square due to the mean effect of factor B, mean square due to the interaction effect and then the mean square due to the error. All those four mean squares are calculated right, then after this is done, the next thing that we end up doing is we calculate the F values.

So, there are three F value values, there are calculated the first F value is calculated by MSA the mean square of A divided by mean squares of error ok. So, we use mean squares of error to calculate the F value for the treatments. So, how the F value will did now, determine? How is significant or how important or how much of an effect is created by the air treatment levels in this regard. Similarly, we calculate the next F value, which is equal to M S of B this is due to the effect of the B factor and divided by MS error right. So, the mean square of error the ratio of MSB to MS error will give you the F value the critical F value for the B treatment. So, it will actually.

If this value is significant then we will know that the factor B has a direct impact on the outcome of the Y right and then the last critical value we calculate is the F 0 for interaction which is given by MS of A B this interaction divided by MS of error right. So, the A B mean squares of A B divided by mean squares of error will tell you how much significant the interaction is. So, these three critical values the critical values of F critical values of F, F 0s will tell the factor, factors or interactions that are significant.

So, if the F value of this A and interactions comes to be significant, then we know that we need to focus on A and as well as the A B interaction and then we can ignore B, when we are actually doing this kind of things. So, remember once again this table has 1, 2, 3, 4 and 5. So, 5 columns and the first column is source of variation, which tells you the A treatments, the B treatment, interaction, the error and the total as the individual factors that come.

They are constitute the sum of squares and remember the sum of squares of total is given by and this is the SST sum of squares of total and this is the SS E error right. So, sum of square are total is the sum of squares of A B, A B and error right. So, all those factors A B A B error ok, which will give you sum of squares of total. So, you can think about this as this is equal to sum of all these ones ok, then the next one we have also seen previously is the degrees of freedom, how the degrees of freedom it is basically, 1 less than the number of observations that we have available there.

So, a minus 1 b minus 1 a minus 1 times b minus 1 a b times n minus 1 then a b n minus 1. So, these were the values that we calculated, then the mean squares were calculated as the mean squares of A, which is the sum of squares of the A divided by degrees of freedom. Mean squares of factor B, the means square due to derived factor is some of squares of B divided by the b minus 1, which is degrees of freedom, then the mean squares of A B is sum of squares of A B divided by the interaction degrees of freedom, which is a minus 1 times b minus 1 and mean square of error is calculated as a means squares of error divided by the error degrees of freedom.

So, these mean squares are calculated, then using the mean squares of error, which is the common to the all the denominator. The critical F values for the factor A, then for factor B and for the interaction A B is calculated. These critical F values helps us to determine how to make the decision ok, decision and which factors are to be focused on. So, this brings us to the end of this presentation, but the most important thing that we have to understand out of this is now, we will take this problem and we will do the numerical analysis, but these analysis of variance table that we just showed in the previous slide need to be clear in your minds.

You should be able to understand and recognise it and should be able to read it and you should be able to produce it and I will show you in the next session, how to calculate these numbers and once these numbers are calculated, you can actually use this numbers to study the which factor is critical, which interactions are critical and can go from there. So, thank you for your patient listening and we will see you soon.

Thank you very much.