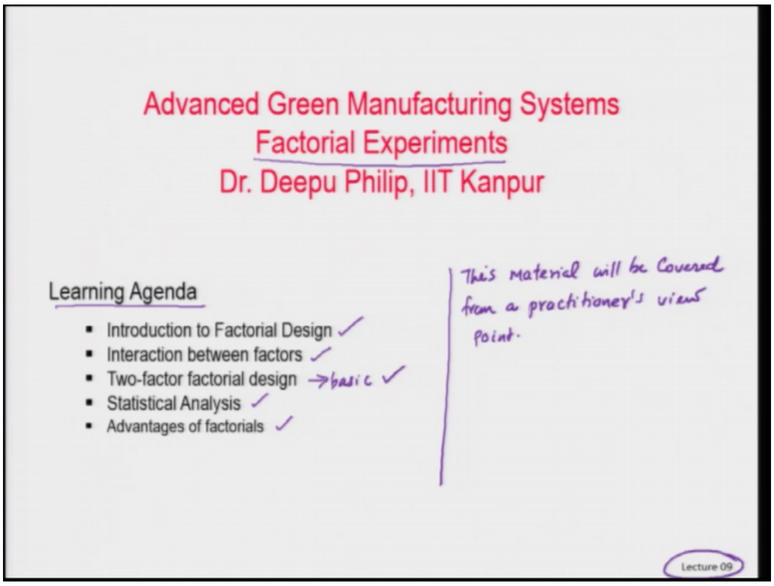
Advanced Green Manufacturing Systems
Prof. Deepu Philip
Dr. Amandeep Singh Oberoi
Department of Industrial & Management Engineering
Department of Mechanical Engineering
Indian Institute of Technology, Kanpur

Lecture – 35
Introduction to Factorial Experiments

Good evening students, welcome to yet another lecture of Advanced Green Manufacturing systems. And as you are we are in the completion of this course, we are getting into little bit more advanced topics. In the previous class I have covered analysis of variance and how we can be used to compare multiple level multiple means and those kinds of things. So, today what we are going to do is we are going to get into one interesting topic called factorial designs or factorial experiments.

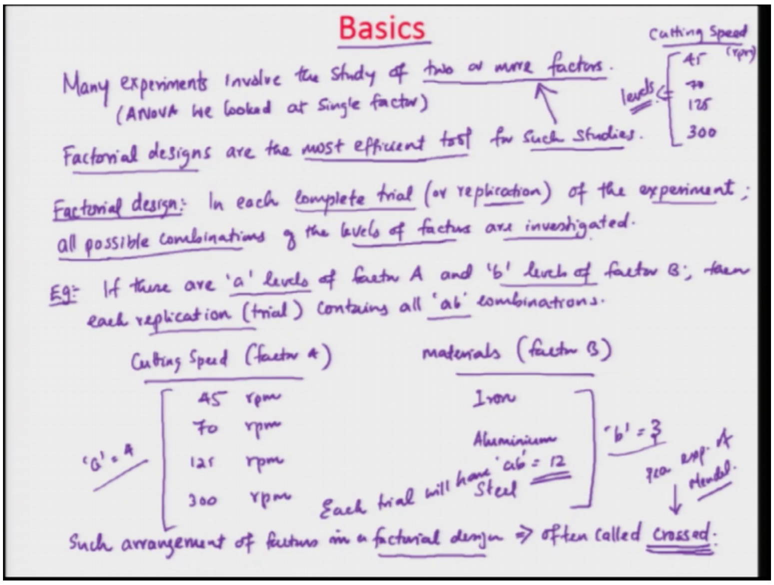
(Refer Slide Time: 00:45)



And these experiments this factorial experiments are very important for our course, specifically because of the fact that this experiment will be used in industry for us to do many of the many of the experiments by which we design where we convert the normal factory into a green factory. So, today's learning agenda is an introduction to the factorial design, the interaction between the factors, and we talk about the basic this is the basic two factorial two-factor factorial design, then how do we analyse the two-factor factorial design using various statistical tools, and what are the advantages of using a factorial, we will go through that.

And this comes under the set 9 lecture set 9 of this course. And please remember that this course as I said earlier this the this material will be covered from a practitioner's view point. So, what we are going to do is we are not going to go into the derivations equations and those kinds of things, but more than that we will just look into what is the basic how do we apply this concept and that aspect of it.

(Refer Slide Time: 02:13)



So, let us first get through the basics, the fundamentals that you need to understand to do this properly in your in a way in a process to convert the green factory. So, the main thing is we know that there are many experiments ok, many experiments involve the study of the study of two or more factors. You might have remembered that in ANOVA, we were looking at the single factor of ANOVA. So, I was like ANOVA, we looked at single factor ok.

So, here we are two the factorial we are looking at two or more factors as the important aspect of this study right. So, there many places where you want to study two or more factor ok. So, why factorials? The factorial design or factorial designs ok, they are the most efficient tool are the most efficient tool, most efficient tool for such studies for such studies. What are the type of studies? Two or more factors such studies is the two or more factors.

The reason why we use factorial designs is because they are the most efficient tool. Compared to other any other tools that is available to conduct a two or study two or more factors the factorial designs are the most efficient tool to do that. So, what do you mean by factorial design ok, how do we define factorial design? The factorial design is that in a simple way to say it is that in each complete trial, remember or trial is also called as a replication, remember we had we were studying this term trial and replication there being one and the same.

So, in each complete trial or replication of the experiment of the experiment ok, every time you conduct an experiment, every time we conduct a complete trial of the experiment or a replication of the experiment, all possible combinations all possible combinations of the of the levels of factors levels of factors are investigated ok. So, what we are doing here is, whenever we do one complete trial or replication of the experiment, we do study all possible combinations, all possible combinations of levels of factors are investigated.

So, for examples, as said you take the cutting speed of a machine let us say 45, 70, 125, 300 or something like this in rpm ok, so then this is one factor and then you can think about it different materials other factor. So, these are the different levels of the factor ok. So, this is the different levels or values that which the cutting speed of the factories used. So, when we say levels this is what we talk about.

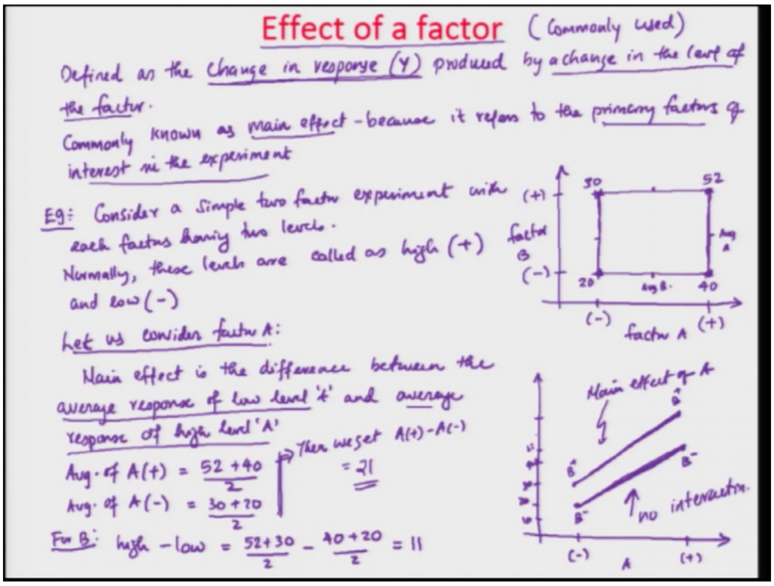
So, then so the one way to think about is this example is that if there are a levels of factor a, there are a levels of factor a, and b levels of factor b factor b, then each replication each replication or what we call as a each trail ok, each replication or trial contains all a b combinations ok. So, this is what actually the factorial is about. If we have a levels of a factor and b levels of another factor, then each replication will have all the, a b combinations of the system or the factors to study right.

So, like for example, if the cutting speed is a factor cutting speed let us say we call just factor a ok, and let us call it as materials as factor b ok. If you see cutting speed has levels of 45 rpm, 70 rpm, 125 rpm and 300 rpm ok. So, these are the four levels. So, a is equal to 4 ok. And let us say we call materials as iron v, and let us say aluminium and steel let us say three things.

So, then the factor of the level of factor is b is equal to 3. So, then what we are saying is the experiment each trial will have a b which is 12. So, this, a b is 12 in our case. So, we will have all that possible 12 combinations we will be able to test in these factors. So, such arrangements, such arrangement of factors in a factorial design ok, this is often called as crossed. So, this is something similar to remember experimental design said it is like the you know pea experiment of Mendel Gregor Mendel who study the genetics.

And so crossing is an example, and he did lot of trials about their red white pea and pea you might have studied all those things. So, the crossing so because this sign of experiments as it origin from agriculture. So, this is one of the reasons you if you say if you say that I am using a factorial design that is crossed, (Refer Time: 09:16) where crossing a body is that all the options of a and b, all possible combination of a and b are being tried in this regard so that is what somebody mentions the term crossing implies.

(Refer Slide Time: 09:29)



Now, let us get into the second part or second concept called as the effect of the factor ok. And this term effect of the factor is it is commonly used. And you will say that I am doing this experiment this factorial experiment is to study the effect of a factor. So, what do you mean by effect factor of a factor? The question here is what do we mean by it, ok, what do we mean by it ok?

So, the answer to that problem is this is defined as the effect of a factor is defined as the change in y or the change in response ok. When somebody says response what we are talking about this y, change in y produced by a change by a change in the level of the factor in the level of the factor.

So, what we are saying is that when we change when we bring a change in the level of the factor, when something in the factor level is changed, how does it how does that reflects in the output or the y as the response variable, how does the response variable gets changed because of the change in the factor level is the what we call as the effect of a factor ok; commonly known as commonly known as main effect, normally call this as the main effect. Why do we call it as a main effect because it refers to the primary factors, it refers to the primary factors primary factors of interest in the experiment.

So, the main effect, we normally call it as main effect, because the primary factors of interest in the experiment are being studied. So, this cutting speed, this is a primary factor. Materials this is another primary factor, ok. So, same way we call this is a main effect or the change in the level of the factor, how does it create a change in the response of the y or the response variable is what we study and that kind of a study is commonly called as a main or is called as a main effect.

Now, let us consider an example, let us take an illustrative example for this case. So, they consider simple two-factor experiment, simple two-factor experiment in factor experiment with each factors having two levels having two levels ok. So, normally these levels are called as called as high ok, it is denoted by a plus sign positive sense and low denoted by a minus sign or negative sign. So, in a one way to respond reply the you know or depict this is you can put it in graph, you have something called as a factor a, and the other axis you have something called as factor b ok.

And we have something called as a this is the minus sign ok, not minus 1, minus the low of factor a, and this is the plus the high of factor a. Here is the low of factor b; and here is the high of factor b. So, then this particular point this intersection point is the low of factor a at the low of factor b. This point is the low of factor a and the high of factor b. So, you can think about it as a line here. Similarly, this point is the high of factor a and low of factor b, you can think of it this way. And this point here is the high of factor a and high of factor b ok.

So, then we can think about it as a scenario where this basically says the response of ok, this image shows the response of how the change in y, the value of y will change with the value of x. So, I am just using some numerical numbers here let us say the value of y here is 20, and let us say the value of y here is 30, then the value of this here is 40, and the value here is 52, let us say we do this for example, right.

So, then let us do let us consider factor a, let us consider factor a ok. So, in the for this particular case, we are considering the specific factor called factor a right. So, the main effect main effect is the difference is the difference between average response between the average response average response of low level a low level a and the average response of average response of high level a ok.

So, what is the low level a ok. So, the low level a, the average this the difference between this difference between the average response of low level a and the average response of high level a. So, what is the average response of low level a. So, the low level a has the high level a has two values 40 and 52 the low level a has two values twenty and thirty. So, the average of high level ok, average of a high level is equal to 52 plus 40 divided by 2 because are two values the average will be this right. And then the average of low level a minus will be it will be 30 plus 20 divided by 2 right.

So, if you think about it, what we are actually doing is, we are saying this is the average, the average is between here somewhere here right. And here is the other average as part of this. So, you are finding the difference between both of these averages ok. So, if you do this, then what we do is, then we get a plus minus a minus will be, we will give you get this will be 20 plus 30 by 2 is 50 by 2 is 25 then a 52 by 40 divided by 2 that will be 92 by 2. And if you do that, then it will give you this is 25, so that will give you 21 right. So, 21 will be answer for this particular case.

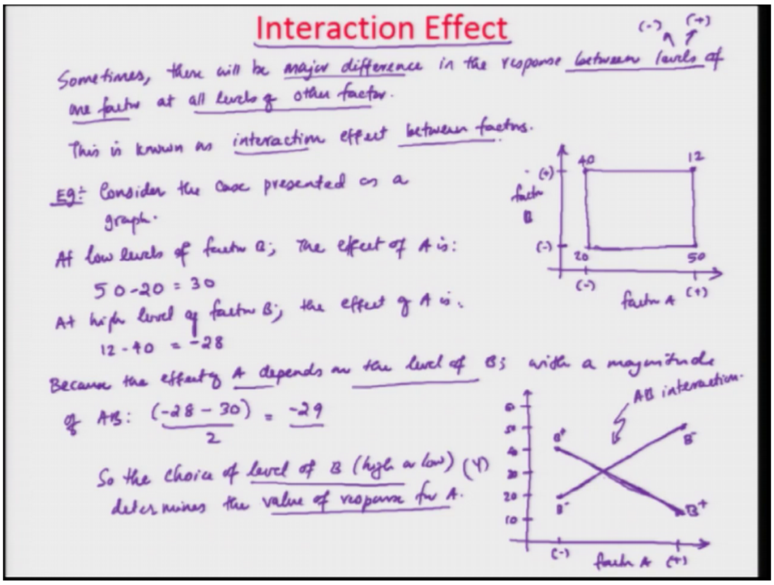
So, what we can actually think about doing it similarly if we do for b ok, for b the low levels of b. Now, let us take factor of b. For b the low levels will be b the low level is 20 and 40; and the high level is 30 and 52. So, for b, it will be the high minus the low, which will be equal to it is high is 30 and 52, that will be 52 plus 30 divided by 2 minus then the difference. So, will be 40 plus 20 by 2 which will give you something called 11. So, the average of the b is, so this is the average a in the average of the b is between these values ok. This is the average of b right.

So, if you want to think about it, in another way to represent this diagrams or these response factor is, you can think about it as a scenario where you have values or the low and the high ok. This is a; this is the low and the high, this is b right. So, the at the lower value of a, this is how in the case of a it will work ok. This is for the low value of b and low value of b. And similarly for the high value of b it will be behave something like this b plus and b plus.

So, this is what we can call it as the and instead of this we let us do this way in a different fashion let us put numerical values here 10, 20, 30, 40, 50 stuff like this right. So, then in this case what actually happens is for a, particular a, the lower values of a for two different levels of b is like this.

And same way for the high levels of b, for the high level b, this is how the values of a actually changes right. This is another way of thinking about the response or how does the change in response actually happens for a or two different levels of b ok. Hopefully, this actually make sense to you guys. So, this is what we can call it as in a way main effect of a for two different levels of the other factor, assuming that the other levels are the factories staying constant all right.

(Refer Slide Time: 20:35)



So, then there is one more effect that we need to talk about which is known as the interaction effect ok. And the interaction effect is another important term, it is usually happens along with the main effect. So, the interaction effect is also called as so sometimes there will be difference, there will be a major difference will be major difference in the response between in the response between levels of one factor, one factor at all levels of levels of other factor.

So, what will happen is when the response between one levels between the levels between the levels is levels mean implies low and high ok, between the low and high levels with or can be multiple levels also more than that, between the levels of one factor at all levels of other factor there will be a major difference ok. This major difference this is known as this is known as interaction effect interaction effect between factors ok. So, what we call this as the interaction effect. This major difference in the responses between levels of one factor is what we call as the interaction effect.

So, let us consider the same case or let us consider another case. Example, consider the case presented in the graph presented as a graph ok. So, we have two factors again ok, ok. So, here is factor a, and here is factor b. Here is the low of factor a; here is the high of factor a; here is the low of factor b; here is the high of factor b. So, now, what we have is same way we have a rectangle in this regard you know that part right. So, these points are connected that way right, and then these points are connected this fashion and that is pretty much it. And so let us say the at the low and low it is 20, and here it is let us call it as 40, and here is 50, and this is 12 ok.

So, in this particular case, what happens is at low levels of factor b, low levels of factor b ok, the effect of a the effect of a is it is given by what the low level of b, this is the low level of b. It is given by 50 minus 20 which is equal to 30 right. Then at high levels level of factor b, which is this one high level of factor b, the effect of a is, what is effect of a the high level is 12 and 40. So, it is 12 minus 40 which is equal to minus 28, negative 28 ok.

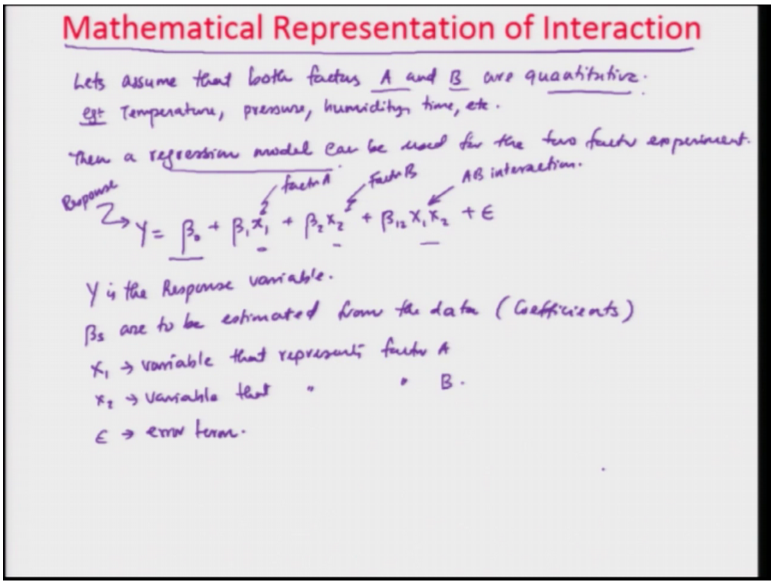
So, then what happens here is because the effect of a, effect of a depends on the level of b. Because in the previous case it was a same, it was still place of the same sign, but whereas, in this case what actually happen is it is a considerable, it actually changes the direction this regard ok, because the been because the effect on a, effect of a depends on the level of b then with the magnitude of what will be the magnitude, it will be the difference magnitude of a b. This is the a b. It will be minus 28, the difference of this minus 30 divided by 2 which will be minus 29 right; so 30 plus 28, 58 by 2 it will be minus 29.

So, the a b effect in this regard what we call it as is because the way this actually a happens. So, if we represent this, this is factor a ok, it has a low value here and a high value here ok. We are represent these levels here ok. And then we have let us say 10, 20, 30, 40, 50, 60 like this and this side right this is the y or what we call as a response. So, then for the low levels of a ok, for sorry the low levels of b this is a 20 ideally right. And for the at b at low, the high levels it will be at 50 ok. So, this is 50 ok. So, this is low b, low b, this is how it is behaving.

Now, for the high value of b, the first one is at where is it high value of b the first one is at 40. So, this is b plus this is this particular value and the second one is 12 right. So, it is somewhere here. So, then this is also for the high values of b. So, this effect is what we call as the a b interaction. The previous case you can see that no interaction ok.

Here we can see there is an interaction happening because it is behaving in the opposite direction right. So, the choice of level of b high or low determines the value of response for a ok. So, the value of response of a to some extent is dependent on the value of the level of b. So, if b is at the low level, a is showing a growing increase; if the b is at high level, then a shows a redact reducing behaviour ok. I hope this kind of makes sense to you guys ok.

(Refer Slide Time: 28:17)



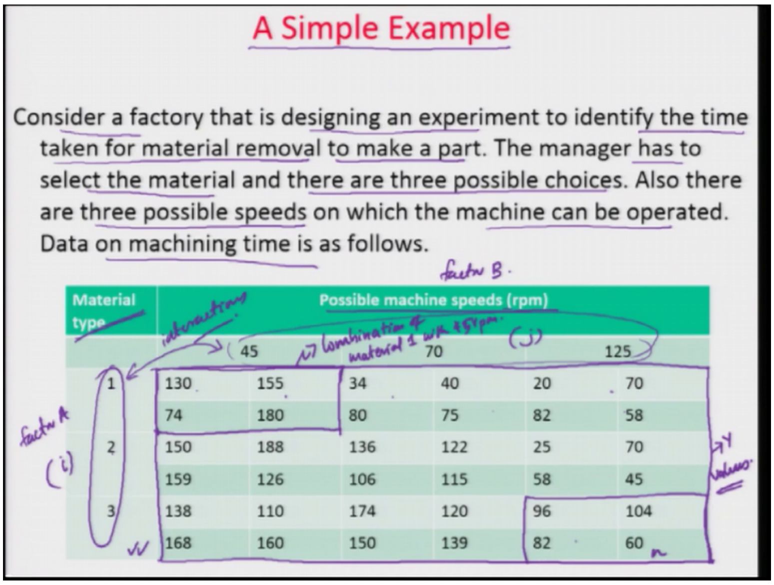
Now, with this what we will do is, we will move towards the how do we represent ok, the mathematically how do we mathematically represent an interaction or how does an interaction actually gets represented right. So, the most important thing is let us assume let us assume that both factors a and b are quantitative ok. Some examples, examples, ok, it can be temperature, pressure, humidity, time, etcetera. So, both the factors a and b both are quantitative factors. Then we can say that then that a regression model, a regression model can be used to can be used for the two-factor experiment, because you know that for quantitative values, we can use a regression. So, we can think about using it as a regression model.

So, the regression model can be written as y is equal to y is your response the output variable, y is equal to can say beta 0 that is your constant plus beta 1 x 1 plus beta 2 x 2 [noise] plus beta 1 2 x 1 x 2 plus error ok. So, you can think about x 1 as factor a, x 2 as factor b, and x 1 x 2 as the a b interaction ok.

So, y is the response variable ok; betas are to be estimated from the data ok, these are the coefficients all right. And variables x 1 variable that represents factor a, x 2 variable that represents factor b. Then epsilon is the error term. So, what we do is we calculate the beta 0; this is the mean of means ok. Then from there what is the change in value due to the value what is the change in the mean value due to the presence of x 1, then x 2 then the interaction effect. So, this is one way of mathematically modelling the interaction in this level fine.

So, let us continue this and let us stop the theory now here and try to do work an example and understand the problem using an example.

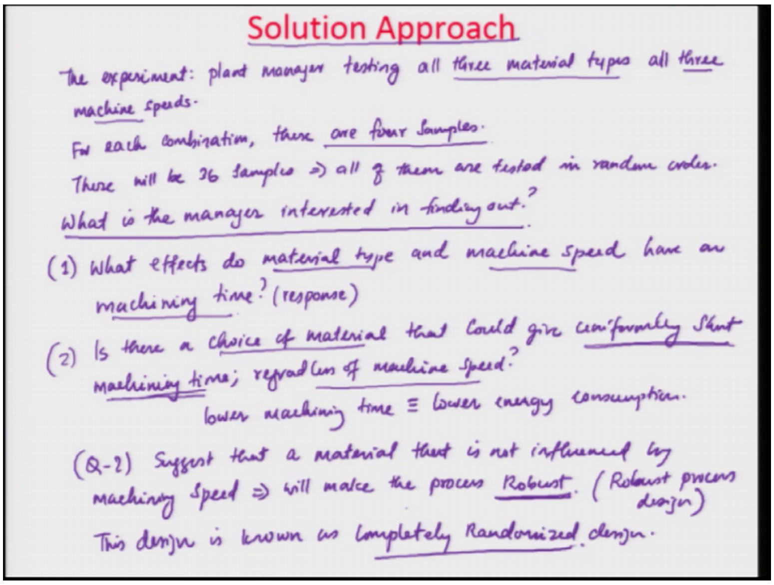
(Refer Slide Time: 31:53)



So, here is a simple example. And this example is somewhat related to the green manufacturing system. So, consider a factory ok, there is a factory available that is designing an experiment to identify the time taken for material removal, how long it takes to remove the material to make a part. So, the factory is making a part and it is trying to identify the time taken to remove the material. The manager has to select the material, so to select the appropriate material. And there are three possible choices. So, materials there are three of them are available three possible choices. Also, there are three possible speeds ok, there are three possible speeds on which the machine can be operated right. So, data on the machine time is as follows. So, the machining speeds are available right here, and the material type are available right here.

So, let us call this as factor a; let us call this as factor b right. So, the factor a has three things which is the material type 1, 2 and 3, they are three type of materials. And material speed is you can think about as 45, 70 and 125 right, so that is other three levels of the rpm that it is there. And these values in between here these things these values, they are the observations this are your y values observed that is your responds the machining time, it can be the machine time. Let us assume that this machine time is in seconds or minutes or something like that does not matter. So, these are the numerical values that we got for the particular problem. Hope this example actually make sense to you.

(Refer Slide Time: 33:41)



How do we solve this problem or how do we approach at solving this problem? And the solution approach of these problems of is important because this results this influences your analysis or the result. So, the experiment consist of what is the experiment here, the experiment is the is a plant manager testing is the plant manager he is testing the experiment consist of the plant manager testing all three material types, all three material types, all three material types against all three all three machine speeds.

So, in this particular case, what we are trying to do is, all three materials types are being tested by against all three machine speeds ok. So, for each combination, there are four samples. So, how does how do we get that? This is you can see that this is the combination of these four samples ok, a combination of material 1 with 45 rpm ok, rpm of the machine speed ok. Similarly, for each case you can see that if you take these case, you can see these four things is the time for material 3 with 125 rpm ok, so that is what we are talking about. For each case there are four samples ok.

Then, so how many samples are there. So, this is 4, this 4, this 4. If you think about it, we will have how many samples will be there, so there will be 36 samples ok. And all of them are tested in random order all of them are tested in random order ok, that is what we are doing in this case.

So, the other question is what is the manager interested in manager interested in finding, what is the manager interested in finding out this is the big question. Why would the manager do an experiment like this? Ok, so one thing that the manager might be interested in finding if finding out is what effects do material type material type and machine speed ok. What effects do material type which is your factor a. So, material type is your factor a, and machine speed factor b.

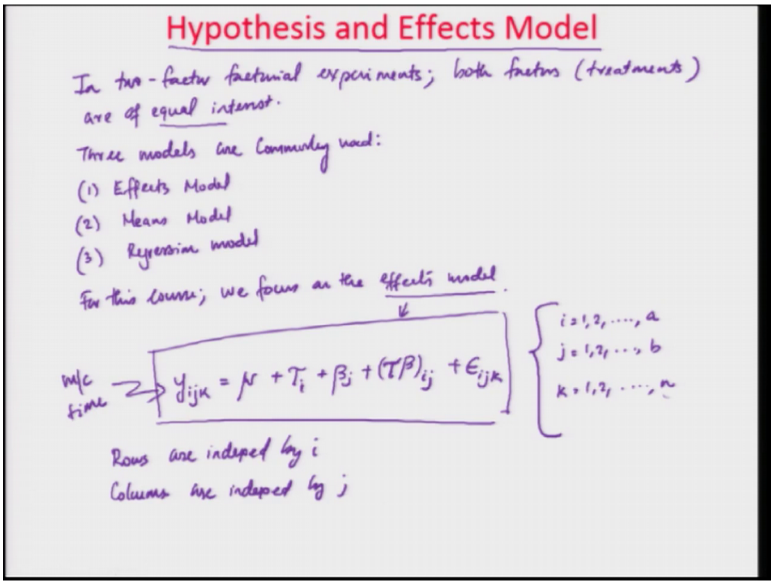
So, material type and machine speed together half on machining time that is one question the manager is interested in finding out. What effect do the material type and machine speed will have on the machining time? This machining time is your response right.

Second part is he is might be also interested in finding out that is there a choice of material, is there a choice of material that could give uniformly short uniformly short machining time; regardless of machine speed? So, what is interested in doing here he is trying to find out is there a material that will uniformly give a short machining time; regardless of machining speed. Why because lower machining time equal to lower energy consumption or equivalent to lower energy consumption. So, he might be interested in reducing the energy consumption. So, he want a lower machining time.

So, he is trying to find out which material which material might give him uniformly short machining time; regarding regardless of whatever be the machining speed ok/ So, the question two, the question number two that the manager wants to ask suggest that a material that is not influenced by that is not influenced by machining speed ok, will make the process robust. Robust means whatever be the speed of the machine, it does not matters because it will not inflect the time taken to produce it So, this is also what we studied earlier is that robust process design.

So, the mat the manager wants to choose a robust material, so that the machine speed is not affecting the machining time it is a hypothetical case ok. Also this design this particular design this design ok, this design is known as completely randomised design, completely randomised design ok, because all the 36 observations are totally randomized. And the data is collected according to the random of random setting of each material and the machining speed.

(Refer Slide Time: 40:31)

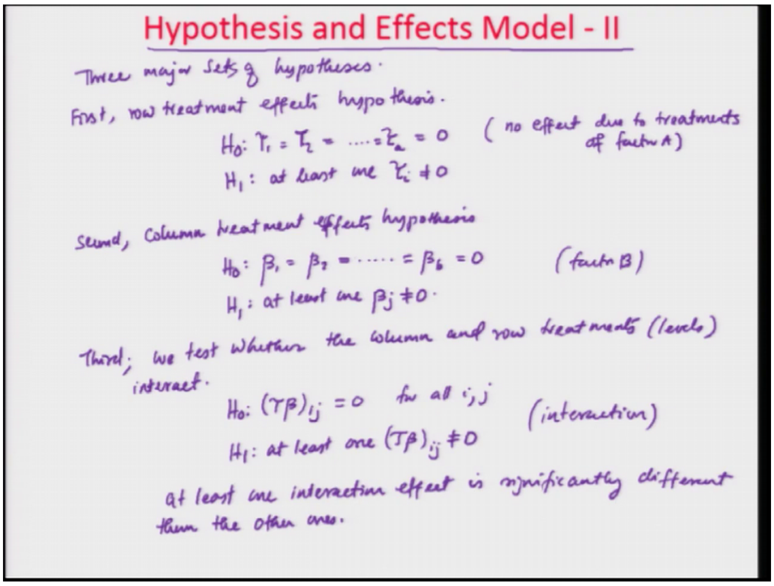


So, what the hypothesis we are trying to tests here and what is the effect model that we are looking into ok. So, in two-factor, there are two factors here in two-factor factorial experiments, both factors also known as treatments, in another term that is used both factors are of equal interest. So, we have equal interest on both the factors typically. Three models are commonly used; people use three models. Number 1 the effects model; number 2 is the means model, and number 3 the regression model ok. All three models are used ok.

For this course we focus on the effects model, on the effects model ok. So, what is the effects model right, how do we write the effects model? The effects model can be written as y i j k ok, y is what is called as our response variable is equal to mu plus tau i plus beta j plus tau beta i j plus epsilon i j k where i j k, where i varies from 1 to all the way to a, there are a levels of factor a, j varies from 1 to all the levels of b which is for factor b, and k varies from 1 to all the way to little n little, n will be the number of replications or number of observations ok. So how many observations are there that is why ok. So, rows are indexed by i; columns are indexed by j.

So, if we go back in the previous example, these are your i; here is your j; and these ones each individual observations here is your n little n. So, this model this particular model ok, this is the effects model, we call it as the effects model. The reason is there the machining time, this is the machining time machine time is a function of a linear combination or a combination of mu the average plus the tau i tau i is a average another value that is for because of the factor a because i is for factor a and beta j is for the factor b be which is your machining speed the speed and tau beta this is your interaction the interaction of both the factors factor a and factor be i j together, and then some error all right.

(Refer Slide Time: 44:31)



So, now let us talk about the what the hypothesis ok, the second part of this, because we just only talked about what is the effects model ok. So, what are the hypothesis we test ok. Three major set of hypothesis major sets of hypothesis ok, hypothesises. So, the first hypothesis, first row treatment effects hypothesis ok.

So, first what we do is we treat we do the row treatment effect hypothesis, where we are saying is that H 0, the null hypothesis is that tau 1 equal to tau 2 is equal to not t tau 2 is equal to all the way to tau a is equal to 0, that means, no effect due to treatments or levels of factor a right. And the alternative hypothesis H 1 says that at least one tau i is not equal to 0 ok. So, basically saying that there is the one of them is different significantly different than the other ok.

Second column treatment effects hypothesis. What we do next is, we do the column treatment effects hypothesis, where the null hypothesis will be H 0 is equal to beta 1, which is equal to beta 2, which is equal to all the way to which is equal to beta b little b is equal to 0 ok, where we are saying that no effect due the treatments of factor b. And alternate hypothesis is equal to at least one beta j is not equal to 0 that we are saying these are one of the beta j is significantly different from all other betas.

Third we test whether the column and row treatment treatments or what we call as levels and treatments levels interact ok. On the top of the main effects, we also test the interaction effects, so where we are saying is there H 0 the null hypothesis is that tau beta i j is equal to 0 for all i comma j. So, for all i and j, the tau beta i j is equal to 0 that is what we do one. Then the alternate hypothesis here is that at least at least one tau beta i j is not equal to 0.

So, we are saying that at least one of the interaction effects is significantly different. So, which means at least one interaction effect is significantly different than the other ones ok. So, this is the main aspects of the hypothesis testing in this regard.

And the next part of this question is all this problem we solve about how do we analyse this whole thing in (Refer Time: 48:56). So, if we go back to this capture this example, what we were actually trying to do is, we were trying to first find what is the effect of these three on whether any one of these material as a significant significantly different performance compared to the other two. And then the another one that you are trying to do is any one of these ok, these situations this speeds will have a significantly different performance or significantly different behaviour on the reduction of the machining time that is what we are actually trying to find out in this regard.

And then third part is that is any of these interactions ok, do they actually do influence, ok, is any of these interactions is this is the. So, this is for the factor a, this is for factor b, this is for the interaction ok. We have these three hypothesis that we would be interested in actually studying and modelling on with this ok.

And in the next class what we will actually do is we will actually do the statistical analysis, how do we analyse this model ok. What is see how do we set this up, how do we do the numerical calculation. And once we get the numbers, how do we study or how do we make sense out of those numbers, and what can we conclude out of those numbers that will be the next lecture out of this course and thank you for your patient listening. And in the next class, we will work on how this statistical analysis of this model is being done.

Thank you very much.