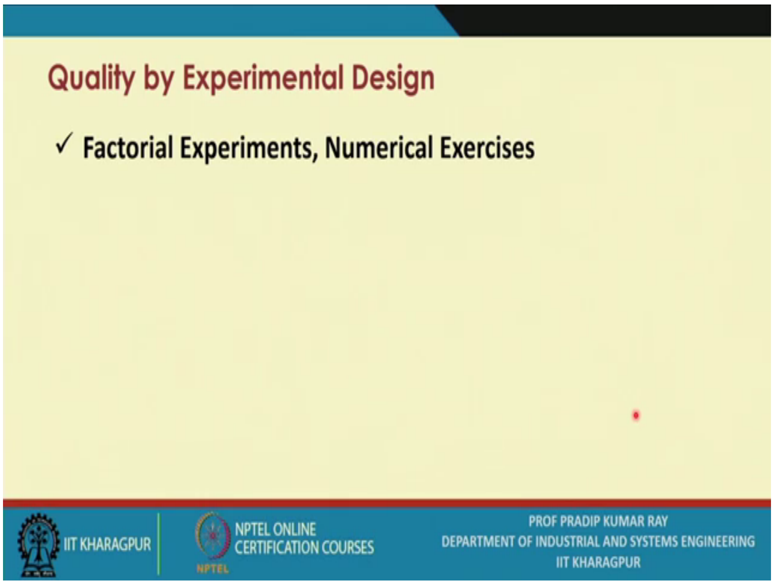
Quality Design and Control
Prof. Pradip Kumar Ray
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture – 54
Quality by Experimental Design (Contd.)

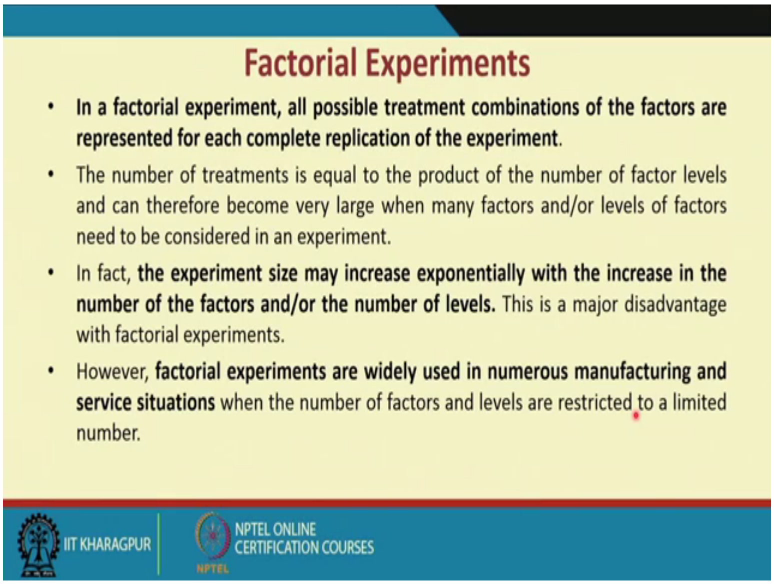
So, under quality by experimental design, during this session, I am going to discuss an important topic called factorial experiments.

(Refer Slide Time: 00:26)



And I will discuss the details of the factorial experiments which are widely used and one or two numerical exercises also, I will try to discuss.

(Refer Slide Time: 00:42)



Now, the factorial experiments are used extensively in the quality for experimental design..

So, in a typical factorial experiment, what we will try to do? We try to consider all the possible treatment combinations, we have already defined, what is a treatment combination and in a in a given experiment, you may have a large number of treatment combinations. So, what you need to do in a factorial experiment, all possible treatment combination combinations of the factors are represented for each complete replication of the experiment, we have already mentioned that what is this replication. So, for each complete replication, you need to consider all the possible treatment combinations of the factors.

The number of treatments is equal to the product of the number of factor levels and can therefore, become very large when many factors and or levels of factors need to be considered in an experiment; that means, when you suppose you have say 5 factors and each factor at two levels.

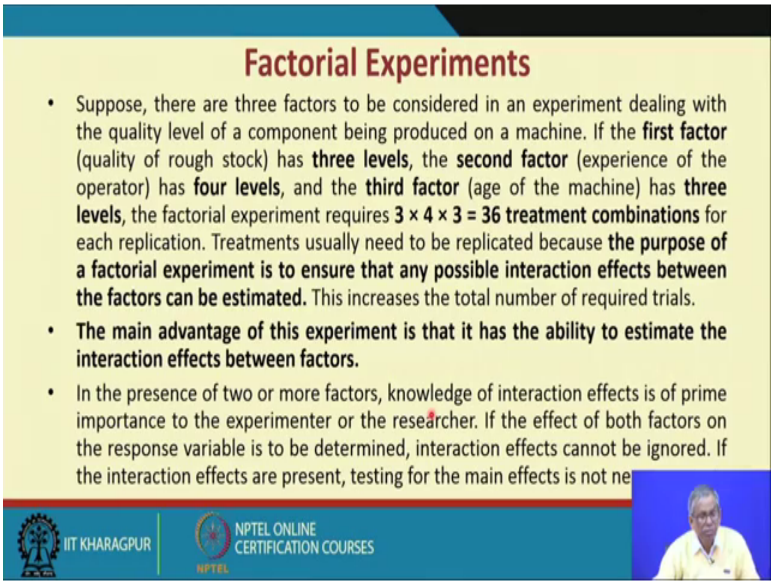
Now; obviously, you know the number of treatment combinations that will have will be far more then when you consider say the two factors; each at two levels ok. So, so, there are cases when you deal with a complex situation, you may come across not only one factor, not only two factors, you may consider the several factors. So, for those cases, what happens if you got for factorial experiment? The number of treatment combinations for each replications that you have to consider could be very high..

In fact, the experiment size may increase exponentially with the increase in the number of factors and or the number of levels, is it ok, there are cases where you know you consider say many factors say 5, 6, even 10 factors and each factor may have the different levels; that means, a factor one may have say two levels whereas, the factor 2 may have 3 levels or 4 levels. So, those; so, this case, we say that is essentially, it is a mixed design and the mixed design is a complex design and usually as a first approximation, what you try to do; that means, you opt for a simple design, even if may not be that accurate. So, you opt for say less number of factors and the first you know the consideration is for each factor I will have the same number of levels..

So, , but there are certain cases where you know the you cannot avoid consideration of say large number of factors. So, this is a major disadvantage with factorial experiments; that means, you must be very very careful, normally, what is done that you focus on only the important factors and for that there could be some screening experiments at the initial stages; that means, how to have say many factors affecting the output response variable you know to consider those factors which are very very critical.

Factorial experiments are widely used in numerous manufacturing and service situations when the number of factors and levels are restricted to a limited number. So, this is the point to be noted; that means the number of factors and levels are to be restricted to a reasonable number.

(Refer Slide Time: 05:05)



So, there is there could be different examples and so, one particular example, let me elaborate and you will come to know that what are the factors affecting the number of treatment combinations? Suppose, there are 3 factors, we considered in an experiment, this is very very common and these 3 factors dealing with the quality level of a component being produced on a machine; that means, ultimately the response variable which you consider in the experiment is closely linked say the quality characteristics of the particular product or the process.

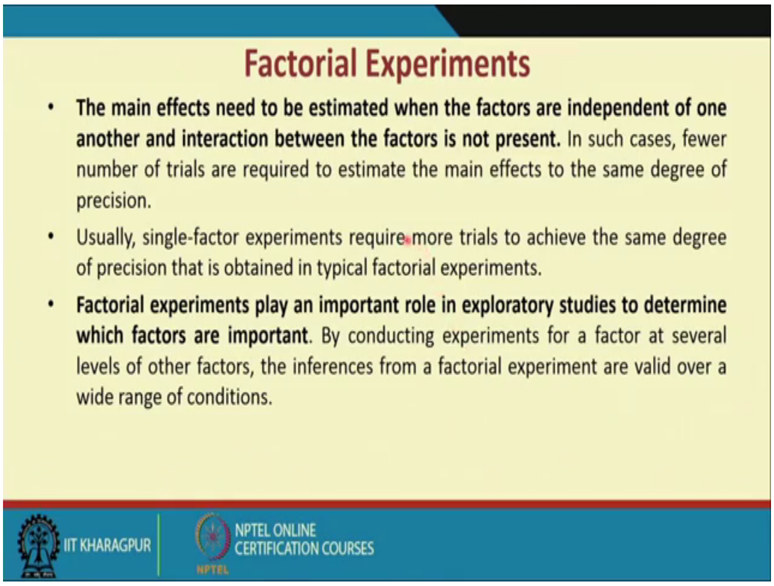
If the first factor say the quality of rough stock has 3 levels, the second factor experience of the operator has 4 levels and the third factor age of the machine has 3 levels so; that means, we are considering 3 factors the first factor is having 3 levels, the second factor is having 4 levels and the third factor is having 3 levels. So, the factorial experiment requires 3 into 4 into 3; that is 36 treatment combinations for each replication.

So, the treatments are usually need to be replicated because the purpose of a factorial experiment is to ensure that any possible interaction effects between the factors can be estimated, is it ok. So, you know the principle of replication already we have mentioned like a principle of randomization; so, because of you know the replications what happens that this will have a direct bearing on the total number of required trials.

The main advantage of these experiment; that means, the factorial experiment is that it has the ability to estimate the interaction effects between the factors, I have already explained what is an interaction effect and what is the main effect, is it. So, in the presence of 2 or more factors, knowledge of interaction effects is a prime importance to the experimenter of the researchers. So, now, you can prove it with several examples. So, many a many a time, you will find that your main consideration is the interaction effect and it has been found in many cases that interaction effect is very very critical or interaction effect is very very significant. .

If the effect of both factors on the response variable is to be determined, the interaction effects cannot be ignored; that means, if there are two factors; factor A and factor b. So, interaction effect is a b. So, a b you are the effect of a b on the response variable may be very very significant and that is why it cannot be ignored if the interaction effects are present the testing for the main effects is not necessary.

(Refer Slide Time: 08:22)



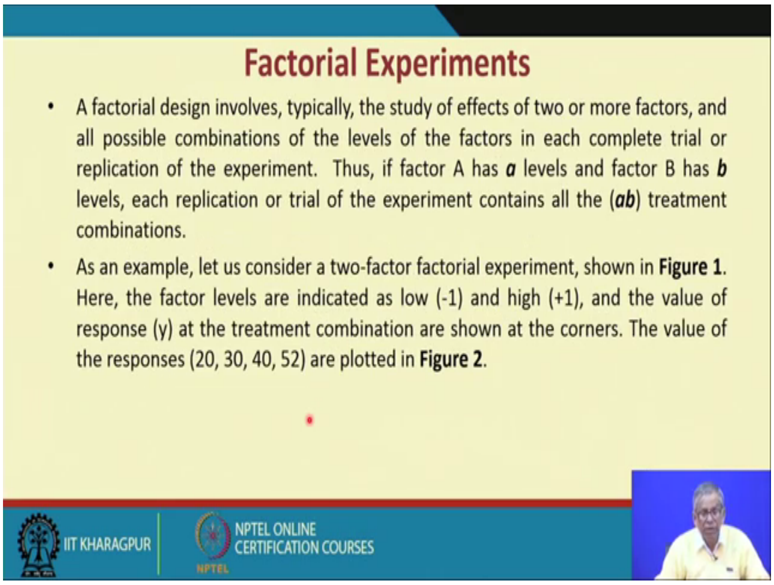
Now, the main effects need to be estimated when the factors are independent of one another and the interaction between the factors is not present is obviously; that means, what you say that the effect of a suppose there are two factors factor A and factor b. So, on y y is the response variable.

Now what you are assuming that you have to prove with the data and its analysis like say the effect of a on y is significant. Similarly the effect of b on y is significant, but effect of a into b; that means, the interaction effect may not be significant with respect to the y in such cases, fewer number of trials are required to estimate the main effects to the same degree of precision, is it ok. So, you have an advantage; that means, fewer number of trials you may have in your experimentation.

Usually single factor experiments require more trials to achieve the same degree of precision than is obtained in typical factorial experiments, what we are saying that the you know in many cases we go for an experiment which is referred to as one factor at a time approach, is it ok, we have already mentioned now in comparison with one factor at a time approach; obviously, the factorial experiment is much better, is it ok, you can prove with the data and its analysis.

Factorial experiments play an important role in exploratory studies and these exploratory studies; normally we carry out during the prototyping stage of a product. So, it is it plays an important role in exploratory studies to determine which factors are important by conducting experiments for a factor at several levels of other factors the inferences from the factorial experimental valid over a wide range of conditions; that means, while you carry out a factorial experiment you must keep in mind.

(Refer Slide Time: 10:37)



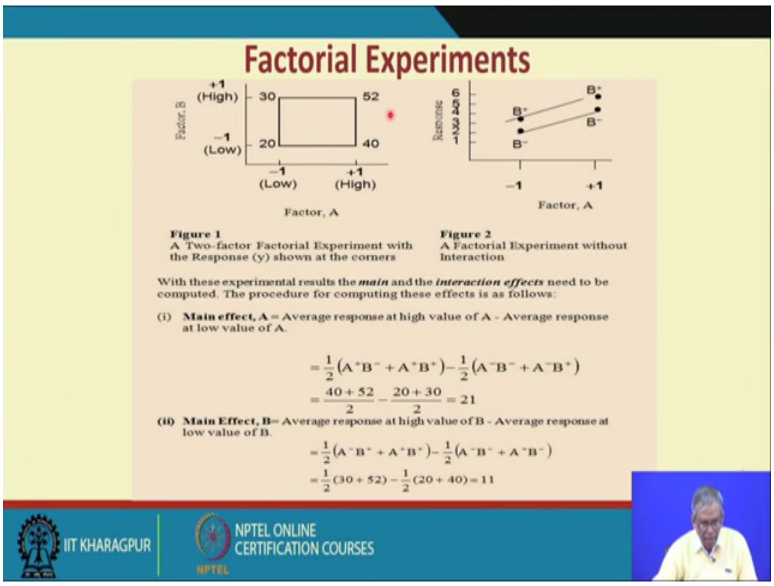
That whether it is the conditions, which you create whether this condition is representative of the actual or the scene or actual situation or not ok; so, ah; so, your knowledge of the actual systems is very very crucial in creating an experimental design.

A factorial design involves typically the study of effects of 2 or more factors. Now, let me elaborate with certain examples. So, the study of effects of 2 or more factors that is that is there in any factorial design and all possible combinations of the levels of the factors in each complete trial or replication of the experiment, is it. So, if factor A has alpha levels and factor B has b levels; that means, a levels not alpha a has a number of levels and the factor B has small b number of levels each replication or trial of the experiment contains all the a into b treatment combinations ok.

As an example just let us let us consider a two factorial experiment that is the simplest one shown in figures, I will show you the figure in the next slide, here the factor levels are indicated as low minus 1 and high that is plus 1 this is just a notation; that means, there are two levels of each factor the factor could be at low level our factor could be at the high level. So, the notation for low level is minus 1 and notation for high level is plus 1 and the value of the response y as a single response we are considering at the treatment combination and shown at the corners.

That means, it is the corners, we will show you the figure this concept is made very very clear, the value of the responses, these are the values you obtain few experimentation; that means, 20, 30, 40 and 50; two see these 4 possible values of y, you get and these responses are plotted in figure 2.

(Refer Slide Time: 12:53)



So, what is this? The figure one; that means, this is the figure 1; that means, it factor factorial experiment with the with the response y shown at the corners; that means, you have factor A and you have factor B. So, what you have considered that this all these two factors are very very very important in explaining the value of y as well as its variability. So, the factor A may have two levels that is minus 1 and plus 1 this is low this is high. Similarly factor B may may have two levels like say the low and high low is minus 1 and high is plus 1. So, what you have this is the first combination that is low versus low. This is high versus low this is high versus high and this is low versus high. So, all 4 combinations you have considered.

Now when you do the experiment with this combination you get a value of y as 20 with this combination, 30 with this combination, 40 and with this combination 52, is it ok. So, this is the explanation of this figure.

In the next figure, figure 2, what we have considered; that means, this is at the minus low level this is high level. Now, when you consider the bth low level, you have these values and we at v at low level, you have these 2 values. Similarly, when you consider against the low value of factor A, you consider high value of factor B and against high value of factor A you consider the high value of factor B.

What do you find that there is no intersection between these two lines with? In fact, they are parallel. So, what do you conclude that that there is the new interaction between the factors is it ok. So, this way you establish the fact that here is the case where there is no interaction between these two factors factor A and factor B.

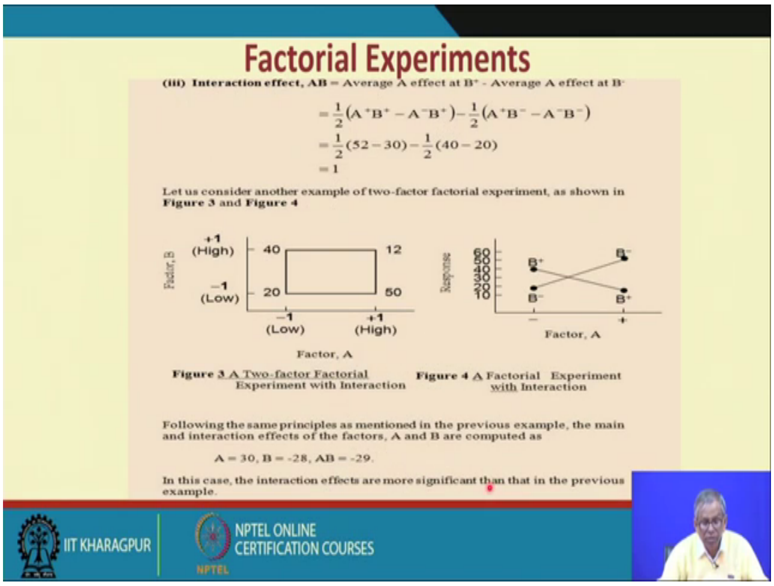
Now, with these experimental results the main effect and the interaction effects need to be computed this point already we have mentioned the procedure for computing these effects, you can just follow these steps, ok, I mean understanding the procedure. So, now, you have to compute the main effect A. So, how do you compute this average response at high value of A minus the average response at low value of A, is it ok? So, this is ah. So, how many average response at high value of A; that means, this is one value high value of A and low value of B high value of A, high value of B.

So, these are the two combinations right. So, you take the average that is why, it is multiplied with 0.5 or half minus average response at low value of A. So, you have low value of A, with low value of B and similarly low value of A, with high value of B. So, these two combinations you consider and take its average and corresponding values, you already got on this the experiment and you substitute you use these values and get the main effects of A as 21 ok.

So, same the procedure you follow to determine the main effect B; so; that means, the average response at high value of B minus average response at low value of B. So, when how do you consider how do you compute average response at high value of B again the same procedure you employ; that means, B at high value B at high value and you consider A at low value A at say you know the high value.

Similarly, here one combination B at low level A at low level B at low level A at high level ok. So, these values you get from the experiment after you conduct the experiment and then you substitute these values here, you use these values to get a value of the main effect B as 11, is it ok.

(Refer Slide Time: 17:38)



I think the procedure is a clearly understood.

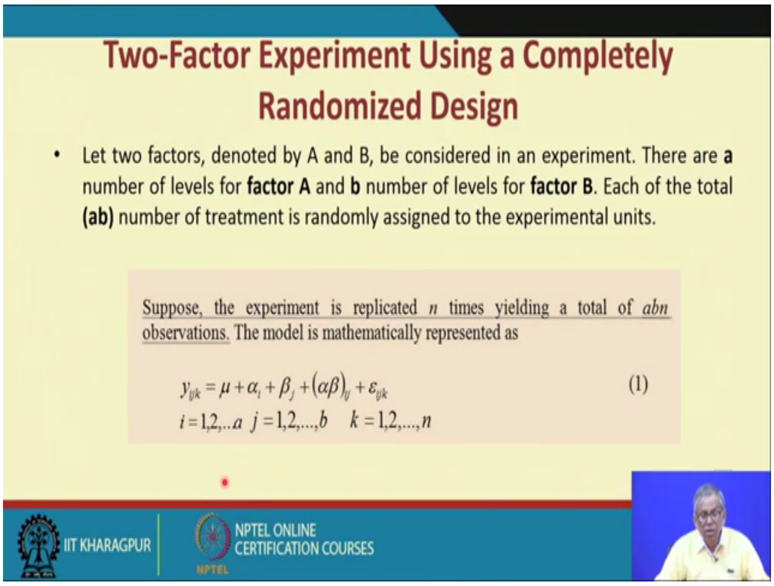
Now, how do you compute the interaction effect the notation is a b; that means, average a effect at b plus minus average effect at b minus. So, here a plus b plus minus a minus b plus is it and this is a plus b minus and a minus b minus. So, any all these values you get through experimentation and ultimately you get a value of a b that is the interaction F in that is one.

Now, let us consider another example of two factor factorial experiment as shown in figure 3 and figure 4, is it ok. So, as you find here in the previous case that is the interaction effect is minimum just one and there could be cases where the interaction effect could be very very significant. So, here is the case like say you consider here is a case where interaction effect would be very very significant. So, here consider factor A and factor B. So, factor A at low level as well as the high level similarly factor B is at low level as well as the high level. So, these are the possibilities and the value of y for all these 4 combinations you get that is 20, 40, 12 and 50 getting my point.

Now, how do you conclude that there is an interaction effect; that means, the factor A low level high level and what do you consider; that means, you consider B at low level B at low level, is it ok. Similarly, B at high level B at high level and what do you find when you draw these two lines? There is an interaction intersection, is it so; obviously, if you find this sort of plot immediately, you will conclude that there is an interaction effect and this interaction effect is very very significant, is it ok.

So, the following the same principles as mentioned in the previous examples. So, please refer to the earlier example the main and interaction effects of the factors A and B are computed as follows. So, what you find; that means, A is 30, B is minus 28 and A B is minus 29, whereas, in the previous case, it was simply just plus 1.

(Refer Slide Time: 20:15)

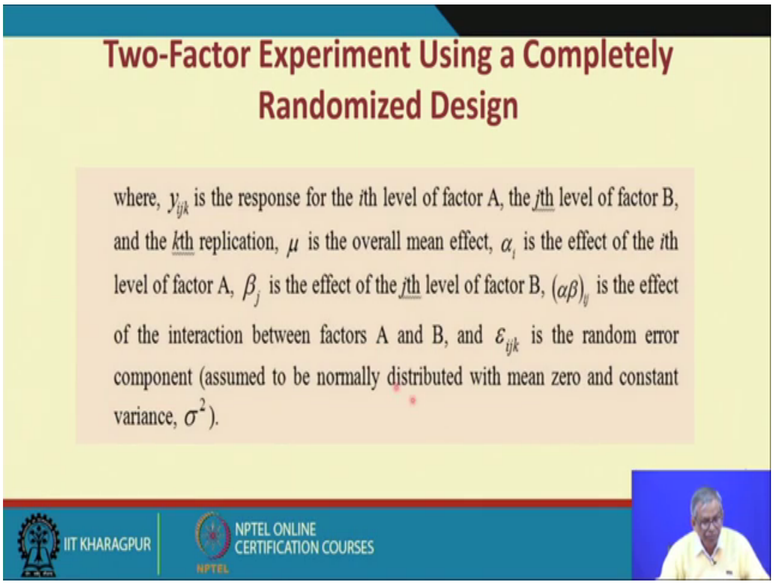


So, in this case, the international effects are more significant than that in the previous examples. So,; so, what we are insisting on that when we whenever you carry out say the say a factorial experiment make sure that you could analyze or you could measure you could estimate the interaction effect and then you conclude that whether the interaction effect is significant or not.

Now, let us move to a particular say the experiment which is referred to as a two factor experiment using a completely randomized design if you recall that we have already in the previous sessions we have referred to the a completely randomized design. So, the basic steps are known to you ok, now here what we are trying to do we are we are using a two factor experiment and with a completely randomized design let two factors denoted by A and B like in the previous case we considered in an experiment.

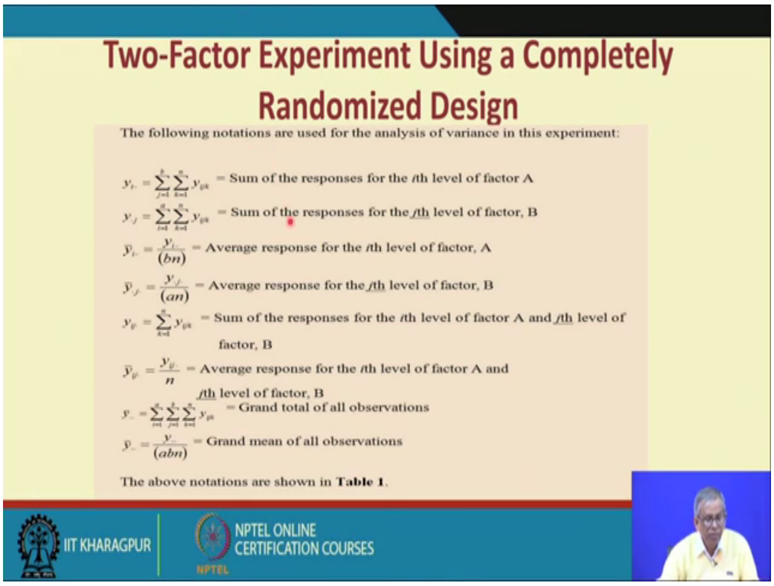
There are small a number of levels for factor A and the small B number of levels for factor B, already we have mentioned each of the total A B number of treatment is randomly assigned to the experimental units that is following the principle of randomization suppose the experiment is replicated in times yielding a total of a B into n observations. So, obvious the model is mathematically represented as; that means, yijk y is the response variable that is mu plus alpha i plus beta j plus alpha beta ij, these are the coefficients. In fact, plus epsilon ijk that is the error term. So, i equals to 1 to a j is equals to 1 to b small b and k is one two up to n; that means, n number of replications.

(Refer Slide Time: 22:13)



Now, so, all these explanations are given over here like yijk is the response for ith level of factor A the jth level of factor B and the kth replication mu is the overall mean effect alpha i is the effect of the ith level factor A beta j is the effect of the jth level factor B and alpha beta subscript ij is the effect of the interaction between factors a and b. So, these are the coefficients. In fact and epsilon ijk is the random error component assumed to be normally distributed with mean zero and constant varia variance sigma square, is it ok.

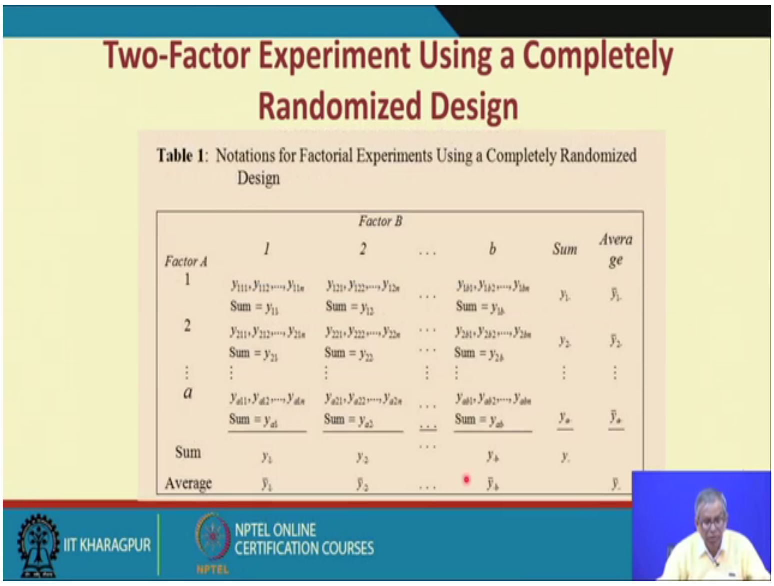
(Refer Slide Time: 22:58)



So, the two factor experiment using a completely randomized design all these notations please go through all these notations; that means, this is these notations we have used and for the ANOVA; that means, for analysis of variance as we have already mentioned that any such experiment you carry out ultimately, you construct the ANOVA table and by looking at the ANOVA table, you conclude whether the model is appropriate or not is it. So, that is the basic purpose.

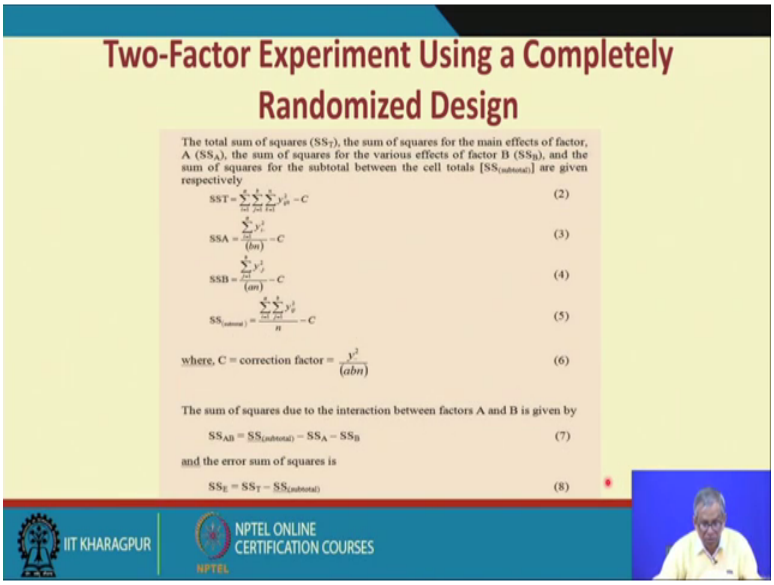
Now, these are the notations we have used like say y i dot dot some of the responses for the ith level of factor B. Similarly, y dot j dot some of the responses for the jth level of factor B then this is the average response for the ith level of factor A similarly this is the average response for the jth level of factor B yij dot this is the sum of the responses for the ith level of factor A and jth level of factor B ok. So, these are the standard notations, we use in any design of experiment this is the average this is y triple dot that is the grand total of all observations and this is the grand means of all observations; that means, y triple dot bar, is it ok. So, this is this is the ultimate expressions you have that that is referred to as the grand mean a lot of observations. So, above notations are shown also in table ok.

(Refer Slide Time: 24:44)



So, the how many factors you have; that is small a number of factors for factor capital a similarly for factor capital b you have small b number of say the levels and then you have all these the values you get and individually for in this cell you compute the sum ok, the notations are this ok, you just follow the kinds of notations we have been using that is this is the individual row sum right and then the average you calculate and similarly this is the column the summation and against each column you have this average and the corresponding notations you have. So, ultimately the grand average you get that is y double dot n bar right.

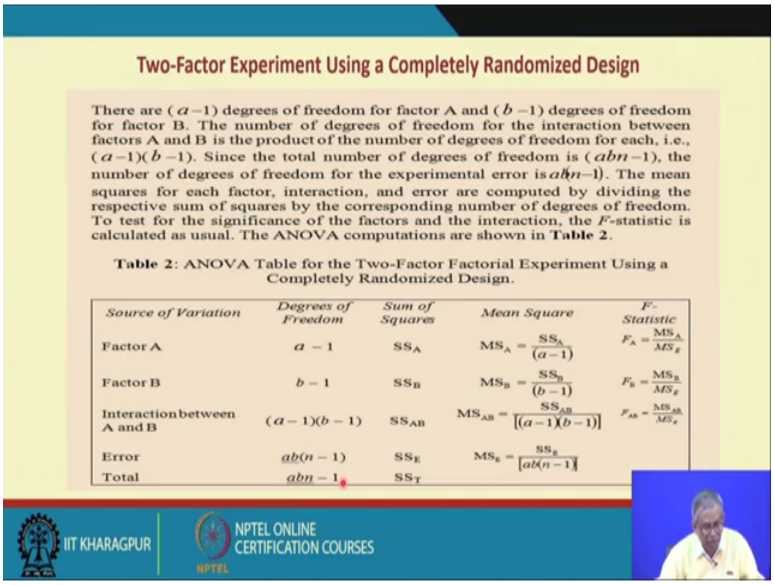
(Refer Slide Time: 25:40)



So, so, you have to construct the table and then what you do you follow the standard procedure that is the total sum of square you have to compute the sum of squares for the main effects of factor A SSA and the sum of squares for the various effects of factor B that is SSB and the sum of squares for the subtotal between the cell totals that is ss subtotal are given respectively. So, this is the total sum of squares.

So, this is the expression this is the sum of squares for factor A, this is sum of squares for factor B and this is sum of squares for the subtotal and you have this correction term that is referred to as capital C and this is the correction factor is y triple dot square divided by abn the sum of squares due to interaction between factors a and b is given by this; that means, SS AB is equals to A B is the interaction effect. So, SS subtotal minus SSA minus SSB and the error sum of square is SSB equal to SST minus SS subtotal.

(Refer Slide Time: 26:53)



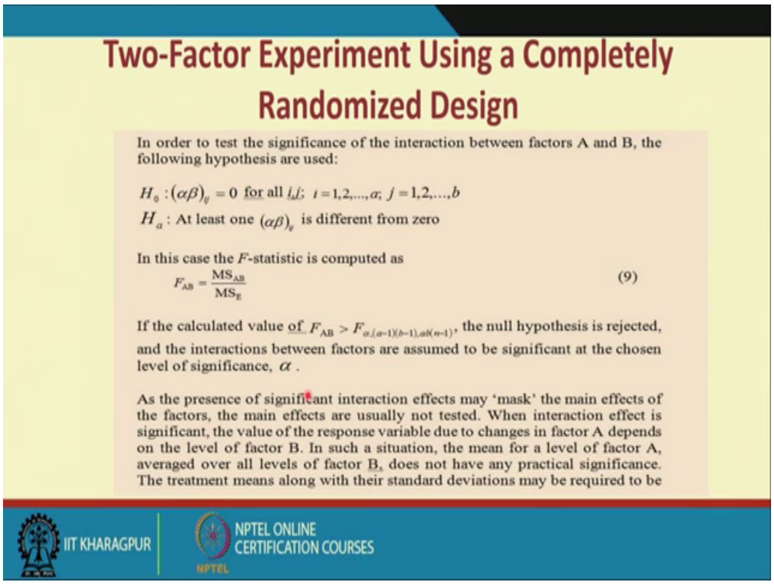
So, this is these are the standard formulations you have.

And now you have constructed the ANOVA table. So, the I say that you just go through all these points right already you know while we refer to you know the completely randomized the design we have referred to all these all these approach approaches and. So, the factor A could be the source of variation factor B also could be the source of variation in y and the interaction between A and B that could be also the source of variation there must be either and there could be the total.

So, you have you need to identify the degrees of freedom. So, a number of levels is there. So, that is why it is a minus 1 degrees of freedom there are small b number of say the levels for factor B that is why the degrees of freedom is b minus 1. So, interaction between A and B degrees of freedom is a minus 1 into b minus 1 and for the error it is a b into n minus 1 and the total is a b n minus 1, is it ok.

So, you compute the sum of squares applying the formula and then you compute the means square say mean square for A mean square for A mean square for A B mean square for say you know the error term. So, then you compute the F statistic for a F statistics for b and F statistics for A B.

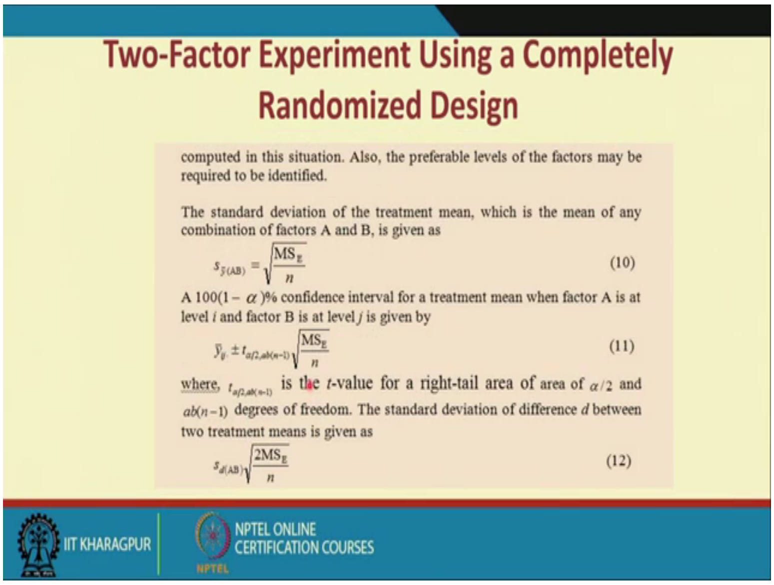
(Refer Slide Time: 28:27)



So, now you go for the hypothesis testing. So, ho is alpha into beta subscript ij equals to 0 or all i and alternative hypothesis at least one alpha beta subscript ij is different from 0. So, how do you compute the F statistics that is if a be equal to mean square ab divided by mean square error

So, if the calculated value of F AB is greater than F alpha alpha is the level of significance a minus 1 into b minus 1 that is degrees of freedom and again a b into n minus 1 that is the degrees of freedom for the error the null hypothesis is reacted. So, you please remember this condition and the interactions between factors are assumed to be significant at the chosen level of significance that is alpha, is it ok.

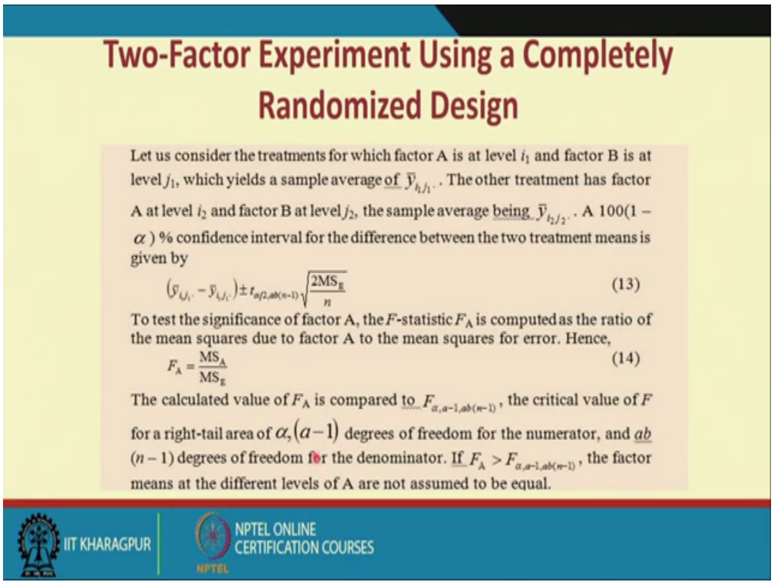
(Refer Slide Time: 29:27)



So, this is the this typical approach we have and then what you try to do; that means, you go for the estimation of say the confidence interval; that means, 100 into 1 minus alpha percent is constraint interval for a treatment mean when factor A A A is at the level i and A factor B as of the level j is given by this one, is it ok. So, this is the t statistics you have and t alpha by two comma abn minus 1 the degrees of freedom is a t value for a right tail area of area of alpha by two and a b into n minus 1 degrees of freedom.

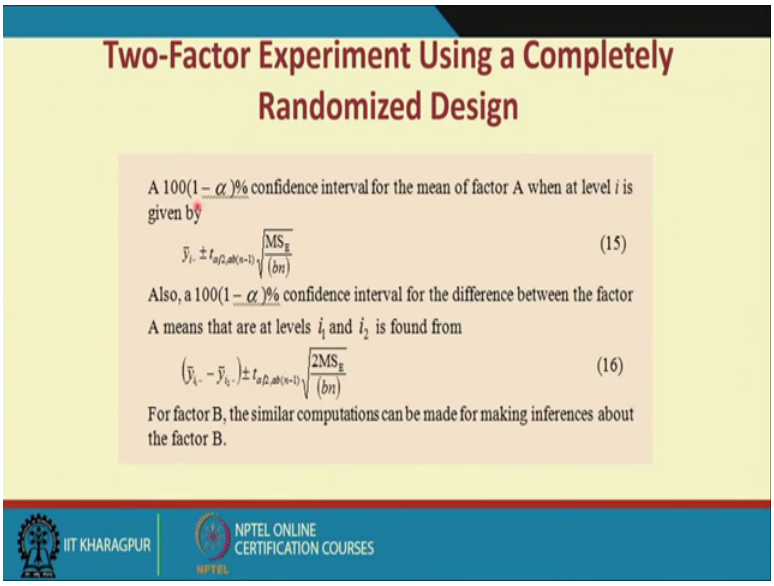
The standard deviation of the difference d between 2 treatment means are given by this one ok. So, you follow these steps.

(Refer Slide Time: 30:12)



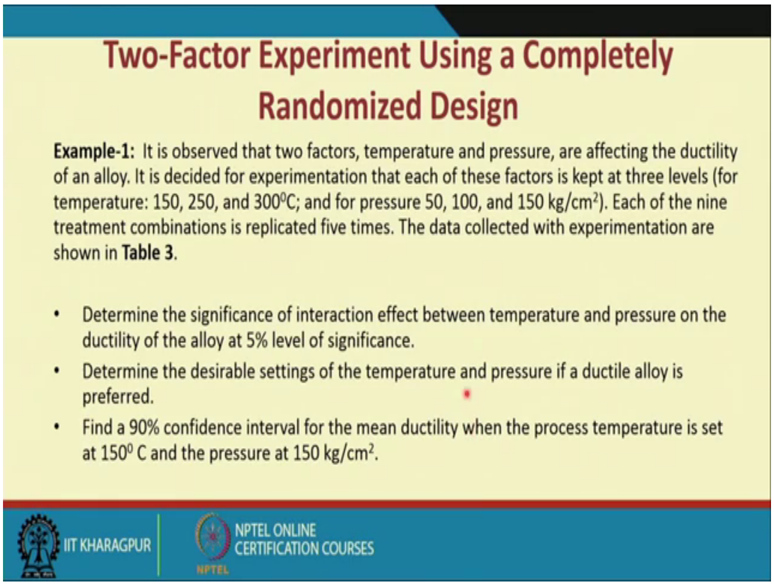
And then you determine the say the difference between the two treatment means is given by this, is it ok. So, this is you know the confidence interval and the test significance of factor A, they have statistics of FA is computed that is this one MSA by MSE I have already explained this part.

(Refer Slide Time: 30:35)



And similarly hundred into 1 minus alpha percent confidence interval for the mean of factor A when at level, i is given by this and for the factor A means that at intervals i I one and I 2 is found from this one, is it ok. So, for factor B the similar computations can be made for making inferences about the factor B.

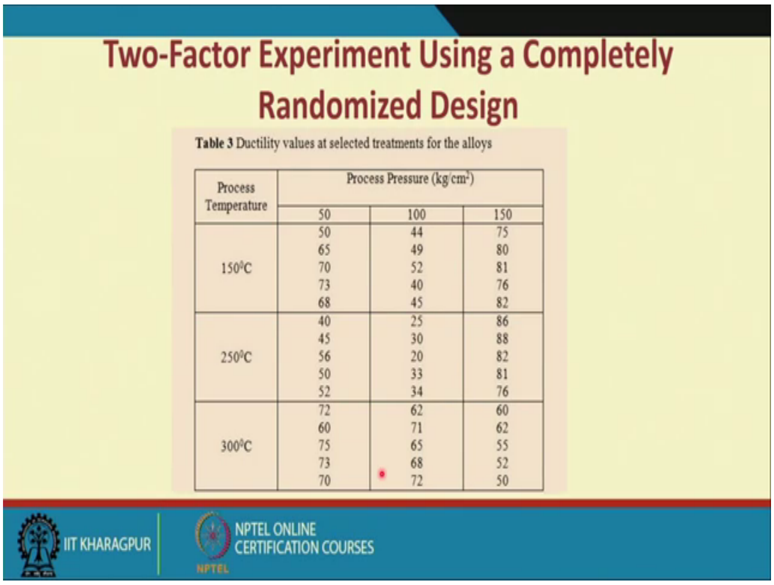
(Refer Slide Time: 30:58)



So, here is one example. So, I suggest you just go through these examples and whatever you know we have studied till now, right and as far as 2 factor the factual experience is consider, please go through this only thing before I close this session I will just highlight that this sort of you know the questions you may have to answer like determine the significance of interaction effect between the factors here the factor is temperature and pressure at the 5 percent level of significance determine the desirable settings of the temperatures and pressure if a ductile alloy is preferred.

That means, it is an exercise for determining the process settings and find a 95 percent confidence interval for the mean ductility when the process temperature is set at this. So, what is this ductility that is the quality and this is the temperature at this temperature and the pressure of 150 kg percentage centimeter square.

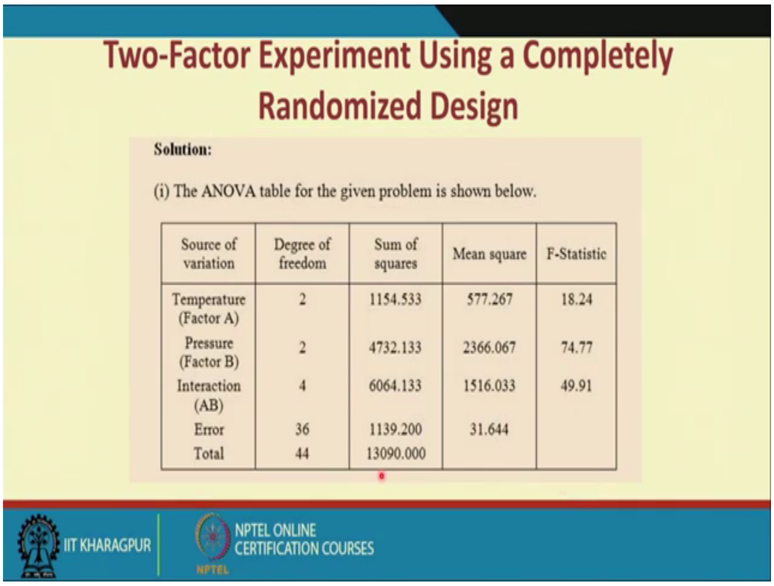
(Refer Slide Time: 32:08)



Process Temperature	Process Pressure (kg/cm^2)		
	50	100	150
150^{0}C	50 65 70 73 68	44 49 52 40 45	75 80 81 76 82
250^{0}C	40 45 56 50 52	25 30 20 33 34	86 88 82 81 76
300^{0}C	72 60 75 73 70	62 71 65 68 72	60 62 55 52 50

So, the specific values; so, this sort of the questions are asked and you need to answer these questions.

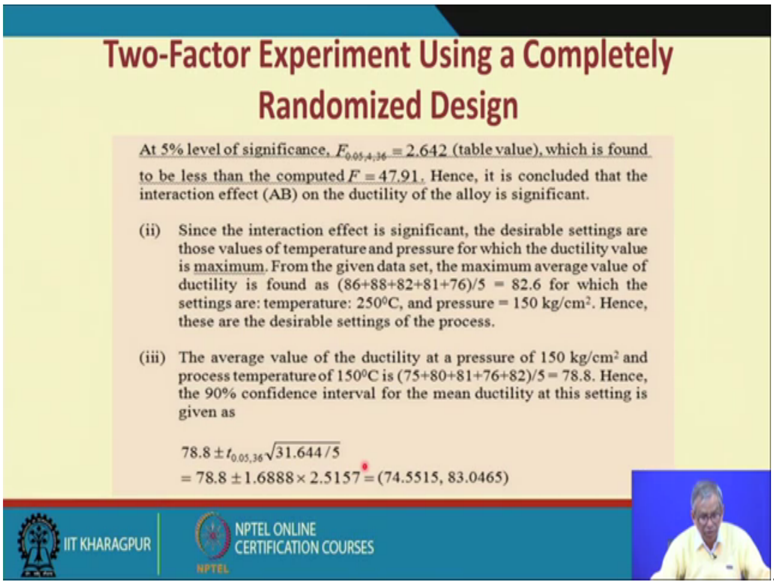
(Refer Slide Time: 32:13)



Source of variation	Degree of freedom	Sum of squares	Mean square	F-Statistic
Temperature (Factor A)	2	1154.533	577.267	18.24
Pressure (Factor B)	2	4732.133	2366.067	74.77
Interaction (AB)	4	6064.133	1516.033	49.91
Error	36	1139.200	31.644	
Total	44	13090.000		

So, please go through these particular typical examples and you will come to know how to create this ANOVA table for the factorial experiment ok.

(Refer Slide Time: 32:19)



All these details are given over here ok.

(Refer Slide Time: 32:24)

So, you refer to this particular the textbook. So, I close this session.