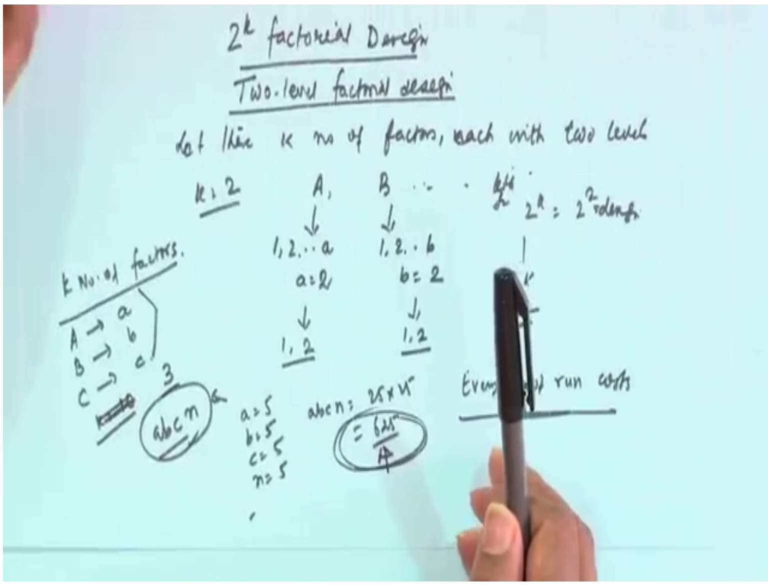
Design and Analysis of Experiments
Prof. Jhareswar Maiti
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture – 36
Two Level Factorial Experiment

This lecture we will discuss, 2 to the power k factorial design.

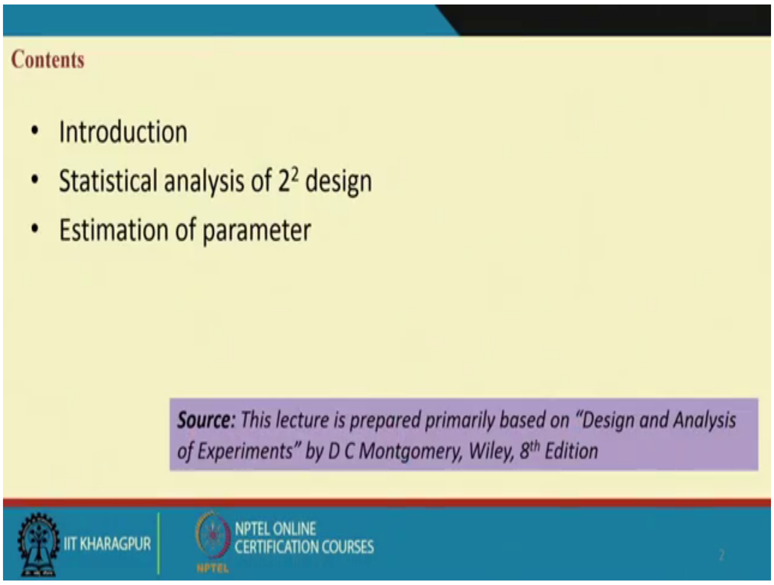
(Refer Slide Time: 00:26)



Actually, we are saying that factorial designs with two levels or two - level factorial design also sometimes we say like this. What does it mean? Suppose, let there are k number of factors each with two levels. Suppose, if I consider the k equal to 2, then number of factors 2 and A and b and now instead of A having a levels B having b levels. We are saying that this small a equals to 2 and small b equal to 2. So, as a result a having 1, 2 levels as well as B also having 2 levels, then the design will be 2 to the power k equal to 2 to the power 2 design.

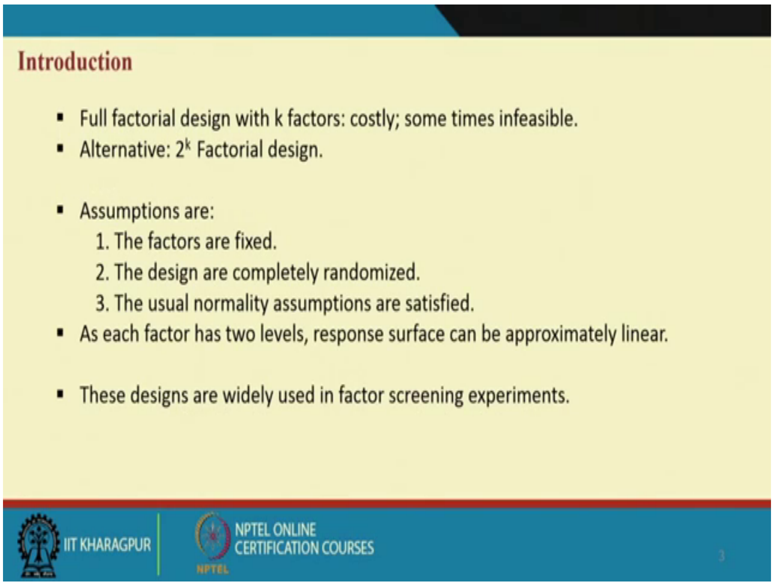
So, as such there are if there are k levels k-th factor then the design will be 2 to the power k. This is what we will be introducing in this lecture and later will continue few lectures on 2 to the power k designs.

(Refer Slide Time: 01:56)



So, in today we will discuss that why this kind of design is needed and then what could be the statistical analysis and how the parameters will be estimated.

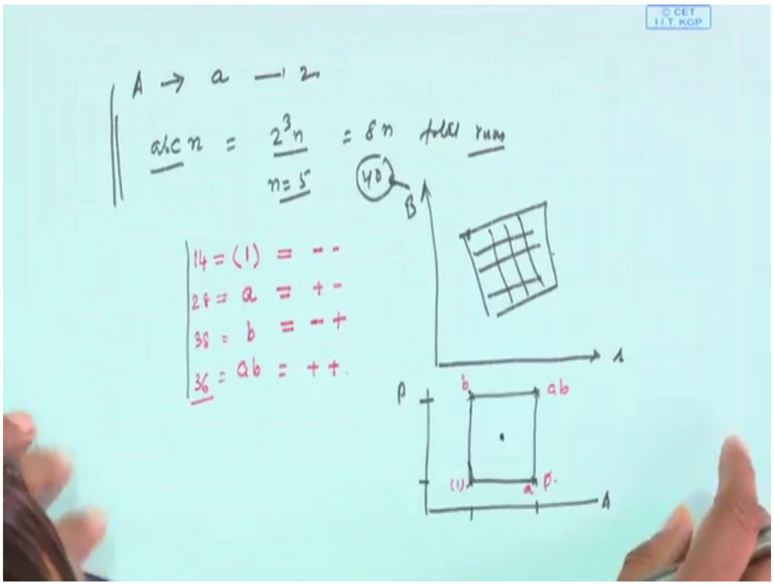
(Refer Slide Time: 02:11)



So, you just consider a case with full factorial design. Suppose, I have k number of factors and each having different levels A having a number of levels, B having b number of levels, C having c number of levels like these. If my number of factor is 10 or let us number of factors is 3 only and then the number of independent runs will be a b c and if I have n replication, this is what the total number of experiments to be conducted.

Suppose, if I say a equal to 5, b equal to 5, c equal to 5 and n equal to 5. So, that means, my abcn this will be 25 into 25. So, this will be 625. So, that number of experimental run will be very large and every run every experimental run costs. For example; it require raw materials, it requires operator, it requires time, it requires process to be used. So, it may be infeasible if we require this much of experiments to be conducted, may be the process owner will not allow us to do this.

(Refer Slide Time: 04:29)



Under such situation can we do some kind of adjustment where that instead of going for A level for factor with a different levels rather make it 2 levels. So, in that case what will happen abcn when each of the factor at 2 level there may be 2 to the power 3 into n. So, you require 8n total number of runs for the factors. Now, if I consider n equal to 5 then you require 40 experimental runs to be completed. So, possible it will be less costly, but if you want to use this kind of things you have certain assumptions that is to be satisfied and definitely the purpose that will decide by analyzing the experimental data will not be the exhaustive one because it has reduced levels. So, it will be used for specific purposes.

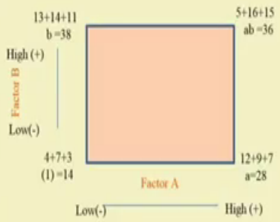
So, what are the assumptions the factors are fixed. The designs are completely randomized. The usual normality assumptions are satisfied as each factor has two levels response surface can be approximately linear. If it is not then what will happen in between the non-linearity you will not be able to estimate. For example, you think of a design where 2 levels so, low and high. So, this is low and this is high, this may be low and high then this is A and B. Now, what will happen, suppose, this A and B if I want to go for the response surface, you may be getting a surface something like this or approximately linear, but if the surface having some kind of suppose, now linearity at this point some kind of quadratic part. So, you cannot capture this because your experimental data are coming from these four factorial points. So, if it assumed it is assumed that or it is quite practical that this surface is linear a parallel or approximately parallel then this kind of 2 to the power k design will help you.

So, as a result also the purpose is limited I told you these designs are widely used in factors screening experiments. Factors screening experiment mean suppose, if there are k number of factors, how many factors are significantly contributing on the response and how many are not. So, your sole purpose is to eliminate those insignificance factors and to reduce the number of factors from higher level from a high level to a lower level and you can do further experimental run with small factors maybe complete only full factorial design you will employ there.

(Refer Slide Time: 08:06)

Data Representations of 2^2 design

Factor		Treatments	Replicate			
A	B	Combination	I	II	III	TOTAL
-	-	A low, B low	4	7	3	14
+	-	A high, B low	12	9	7	28
-	+	A low, B high	13	14	11	38
+	+	A high, B high	5	16	15	36



IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

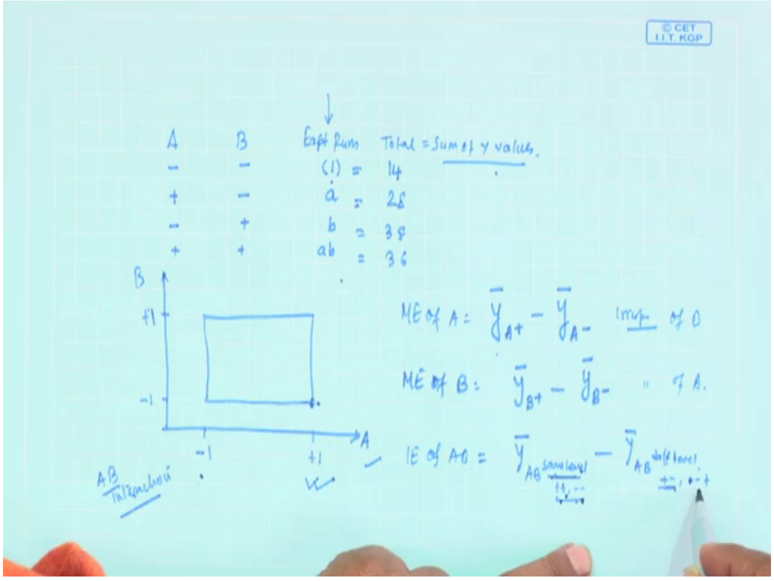
So, let us now understand that the effect parameters calculations with reference to a 2 to the power 2 factorial design which is the most simple case of 2 to the power k factorial design.

So, we will do this kind of tabulation. There are 2 factors A and B. Now, each has 2 levels minus and plus. So, we are denoting like this A minus plus minus plus low high low high B when A low B low A high B low. So, that low high and the combination is a low B low A high B high A low B high B high A high B high. So, these are the points low, high, low, high and suppose, in each this experimental settings you have done 3 experimental replications giving a 3 experimental runs conducted; obviously, these are randomly experiment and noised one. So, then suppose the response value is 4 here, 7 here, 3 here and here 12, 9, 7 like this. This is the data you got.

So, what happens here then you have 3 runs in each statement settings. So, where you want to first know the that is the total here this total will be 4 plus 7 plus 3, 14 here when low, low. Similarly, at this point is 28, similarly at this setting is 38 similarly at this setting is 36, these are known as totals. These totals are given a special notation in 2 to the power k factorial design the reason is it helps us defining many things also in computing the effect as well as the sum squares. What is the notation? The notation will be 1, when both are at low level that is minus, minus. Then we are saying a if a changes to plus when a goes to high factorial a goes to high factorial at low level we are denoting b the situation when a at low b at high and we are denoting by ab when both at are high. So, this is the notation from that I mean this is my 1, this is a, this is b and this is ab.

So, I can say these are the experimental combinations, but in addition from total point of view we give some Newman number to all those things these are nothing, but the total values of y response at 1, at this point, at this point, at this point, at this point also. So, this is the situation, but we also give that what is the total here. So, what is the total here and here? So, with reference to this for example, then this will be 14, a will be 28, b will be 38 and ab will be 36.

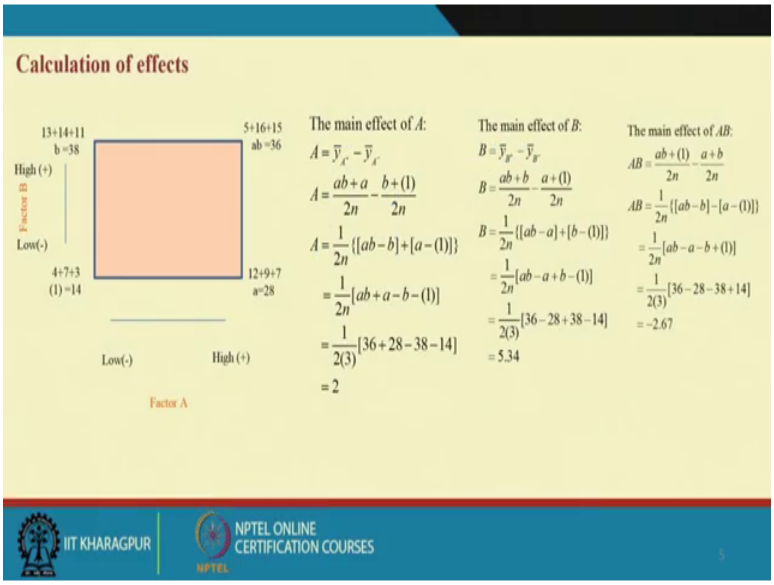
(Refer Slide Time: 11:57)



So, what I mean to say here that you have seen the situation like these situation is that when A low, B low and suppose experimental runs low low this setting is denoted by 1, then high low a, that mean low high b and high high ab and we have calculated total here with reference of this 1 is representing 14, 2 is representing the total value that is a is 28, b total value is 38, ab total value is 36.

So, this will be the experimental setting one end and another end Newman number means suppose, if there is one replication only then this value is the one y value. So, sum of y values.

(Refer Slide Time: 13:05)

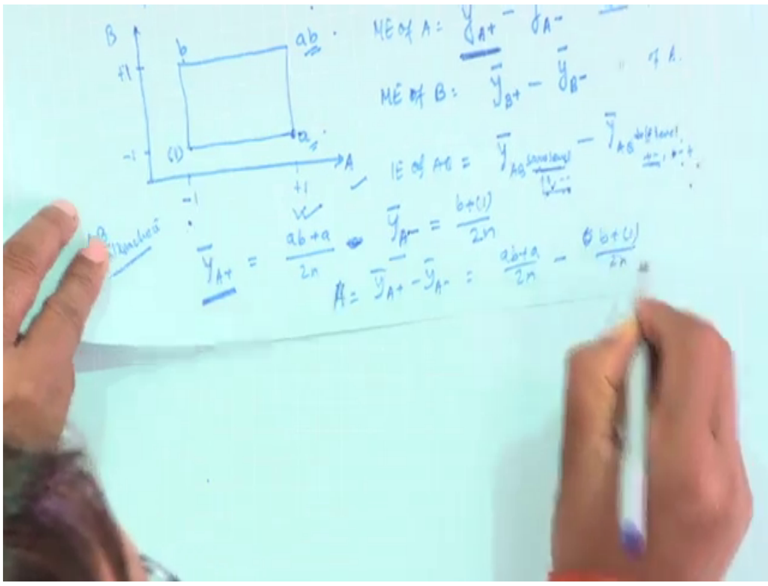


So, then we want to estimate the effects. What are those effects? What is the A effect? Suppose, what is the b effect and what is the interaction effect AB interaction effect. So, A is changing from minus to plus or minus 1 to plus 1, we can write minus 2 plus or minus 1 to plus 1, B from minus 1 to plus 1. So, that mean now A effect will be the mean values here, total you get here mean values of y when A at high level minus mean values of y when A it at low level irrespective of level of B.

So, that is the main effect of A equal to y A at positive that is y bar, A at positive minus A at negative, that mean when A at high value high what is the average of y and A at low what is the average of this? Similarly, main effect of B will be same thing y bar B high minus y bar B low this plus minus for high and low. Here, irrespective of B and here irrespective of A, then and what will be the interaction effect AB, that will be when that mean the y bar AB same level minus y bar AB different level. What does it mean? That mean either same level means A at positive – positive or negative – negative. Here, either positive negative or negative or negative positive, means A at high - high or low – low, that is the same level and A at high B at high A at low B at low that is same level. Here, what is happening A high B low or A low B high that is the different level. So, what is the average of this and then minus the average of this.

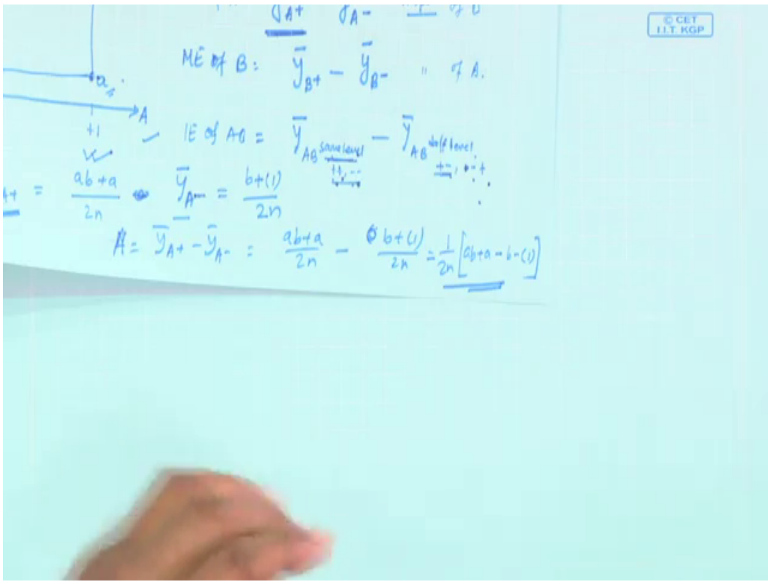
So, now what will be the average of this the y total at this two points. So, this is if there are n replicates here also n replicates then here n plus n, 2n will be the observations. So, 1 by 2n and the total, that many 2n will come out. What you will get? you will get an equation like this.

(Refer Slide Time: 16:33)



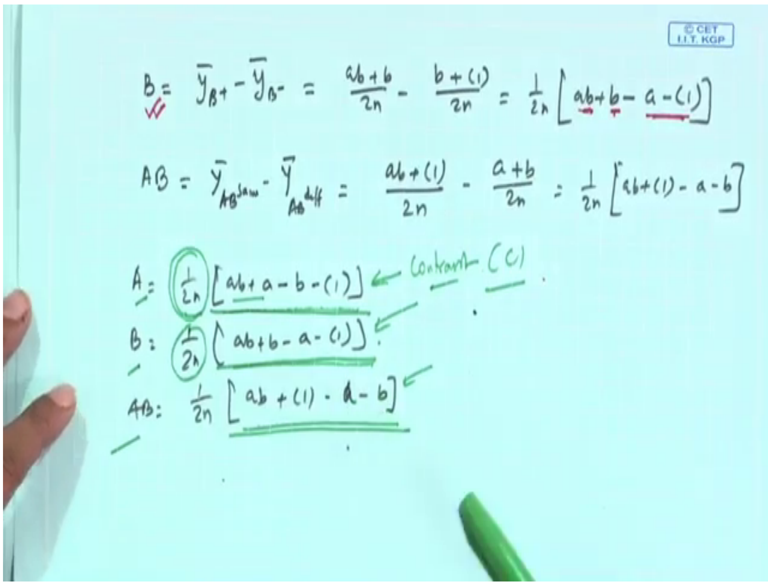
Why? The equation will be something like this suppose. I want to know what is y A positive bar. So, we know the total here is 1 within bracket, here is a here is b and here is a b. So, y A bar is when A at high what is the total then ab plus a, what is the number of observations here n, here n, 2 n; then, minus as this is similarly y A minus bar this will be this and this average b plus 1 by 2n. then any that effect A will be the difference between this two, y A plus bar minus y A minus bar then this will be ab plus a by 2n minus b plus 1 by 2n then the resultant is 1 by 2n ab plus a minus b minus 1.

(Refer Slide Time: 17:40)



So, see this situation this interesting one, when we are talking about A effect you see that this small a is coming wherever small a is there that to those are basically is summed up and wherever there is no a related terms are there they are subtracted.

(Refer Slide Time: 18:22)



So, using these if I want to tell you what is the effect of B, so, B effect it is y B plus bar minus y B minus bar. So, this will be basically when B plus ab plus b divided by 2n when 1 minus a plus 1 divided by 2n. So, that mean ab plus b by 2n minus b plus 1 by 2n then this will be 1 by 2n into ab plus b minus a minus 1. Interestingly, you are getting the same one, that we are interested in B, the small b is related to B. So, here b term is there ab term is there this is plus, wherever there is no b term this is minus.

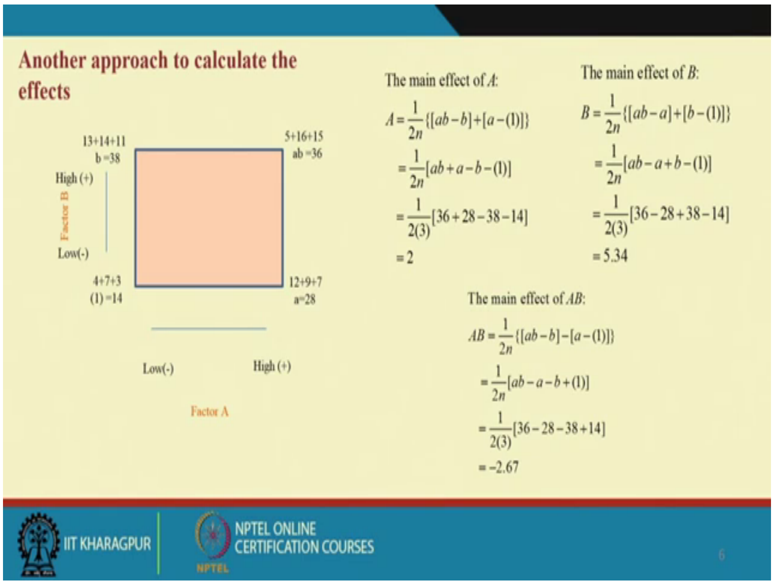
Now, what will be the AB? AB effect is when both are at same level. So, that mean either high – high or low – low and both are at different level, so, y bar AB same minus y bar AB different, that is what I said; same means ab plus 1 divided by 2n minus a plus b by 2n. So, this will be 1 by 2n a b plus 1 minus a minus b. So, here when if we see that A is 1 by 2n ab plus a minus b minus 1, B if you see that is 1 by 2n ab plus b minus a minus 1, AB if you see this is 1 by 2n ab plus 1 minus a minus b.

So, you have interestingly within third bracket a quantity third bracket everywhere a quantity you have this is known as contrast; suppose, we say this contrast c. What we see here? that means, when we keep in a effect contrast if I keep if I change from low to high what if the change? So, that means, with reference to low, what is the high position? Similarly, for B and for AB that will be same level to different level; so, contrast. This value is giving us contrast value.

Now, we have two part; one part is the what will be the division factor 2n, 4n whatever maybe that you have to identify and you have to know contrast. If you know contrast you are in a position contrast for A contrast for B contrast for AB, as such if you have many more factors and their interactions you have to know the contrast then you will be able to find out the effects.

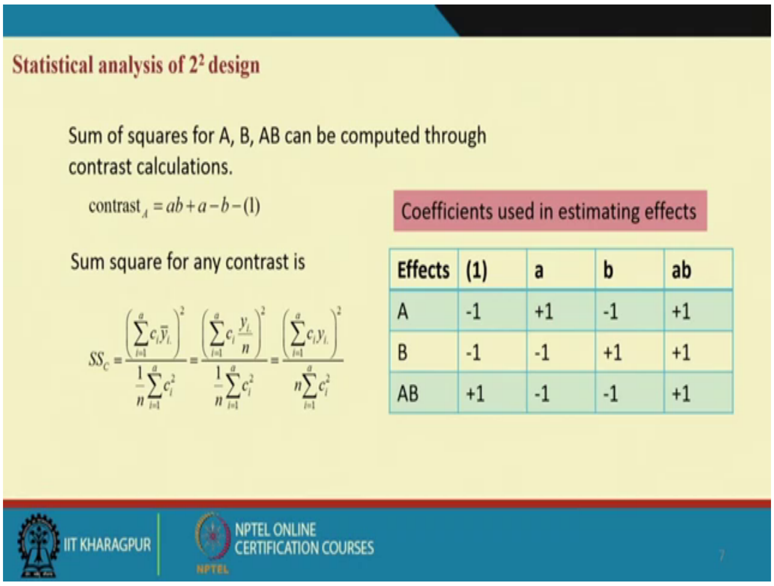
So, with this description and giving this data we found out that A equal to 2, what does it mean? If you change a from low to high or minus to plus then the y change will be 2 units, then the main effect of b is 5.34. If you change b unit from minus 1 to plus 1; that means, when two units change in B, in this coded scheme minus 1 to plus 1 will change y by 5.3 per units and the interaction effect is minus 2.64. So, that means if change from the same level to the different levels, there will be y change will be in the y will be decreasing by minus 2.67 units. So, this is from average point of view.

(Refer Slide Time: 22:38)



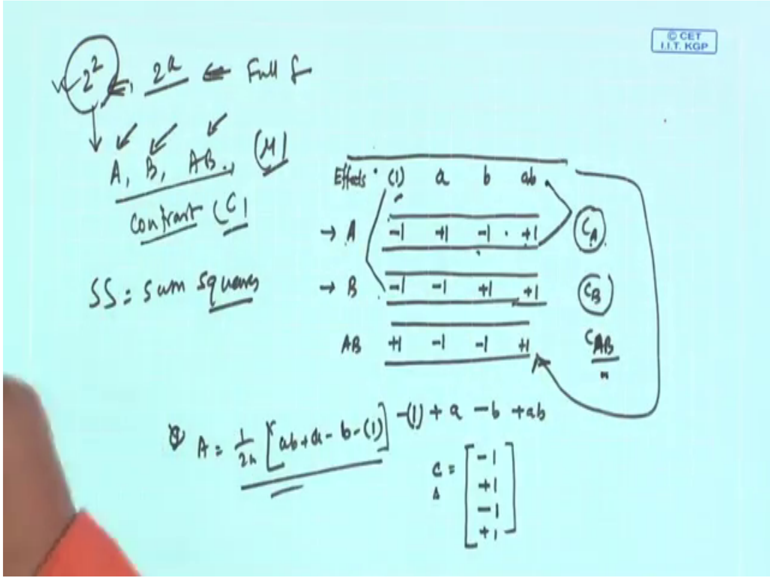
Now, as I told you the contrast point of view if you see use the contrast the same equation you will come and same formulas you will see.

(Refer Slide Time: 22:49)



So, what we have discussed so far then, we have discussed that 2 to the power 2 design or 2 to the power 2 design is the special case of 2 to the power k design and 2 to the power k design is the special case of full factorial design that we have discussed.

(Refer Slide Time: 22:59)



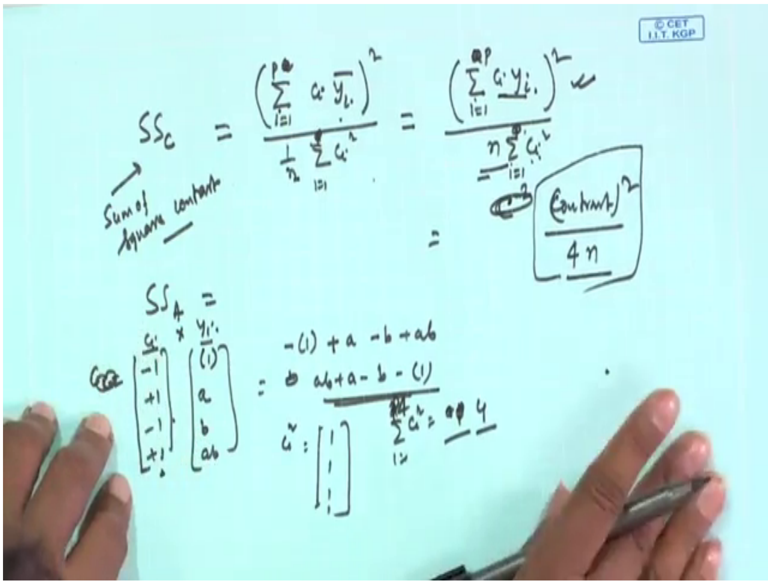
Then, we have showed with reference 2 to the power 2, I have told you that how to compute the effect parameters A, B and AB we have seen another important thing called contrast, which will be denoting by c. So, now you know how to compute this A, B and AB also you will be I will tell you how to compute mu later on that grand mean and now, we will see that with reference to 2 to the power 2 designs, how to compute the sum squares how to compute the sum squares. So, if I go back to estimate of effect of A and if I say this is the contrast and then how this contrast is coming from the data, suppose I say that we are denoting this A, B and AB what we have seen that A at low, B at low, so, minus 1 minus 1 these we have denoted by 1, then when it is plus 1 this is minus 1 we denote it by a, when this is minus 1 this is plus 1 denoted by b, when this is plus 1 this also plus 1 denoted by ab.

So, these are basically effects of on y values. Now, then if I want AB you multiplied these row with this row. If you multiplied the 2 minus into minus this will be plus 1 this will be minus 1 and this will be your minus 1 and these will be your plus 1. So, you write here 1, here 1, here 1. So, this is these row these row represent contrast for A this row represent contrast for B this row represent contrast for AB, this is what is contrast. Now, when we say this basically gives you this fine basically that effect is there. So, if I go back again let it be like this, I am saying this portion is the contrast when multiplied by this. So, by this is basically the constant basically contrast constant for A, for B and for AB.

Now, if when you multiplied these with these, you will get this, contrast; when you multiplied these with the effect you will get the contrast and when you multiplied these with these you will be getting AB contrast. So, essentially what happened that if I say this into this then it is nothing, but minus 1 plus a minus b plus ab, first one and then if I see that what is the A effect 1 by 2n sum of ab plus a minus b minus 1, this is 1 is there. So, this way this is basically this is coming. So, this is the contrast.

So, we can say that this are the c. So, general one c for small c for A is like this, in this case minus 1 plus 1 minus 1 plus 1, this is for A, for B something like this. Now, this is what is represented here now how do you that mean then the then the contrast of A is as I told you this row multiplied by the effect row as row multiplied by the contrast constant for A and effect values this 2 multiplied you will get contrast A. If this and this multiplied you will get contrast B, if this and this multiplied you will get contrast c.

(Refer Slide Time: 22:24)



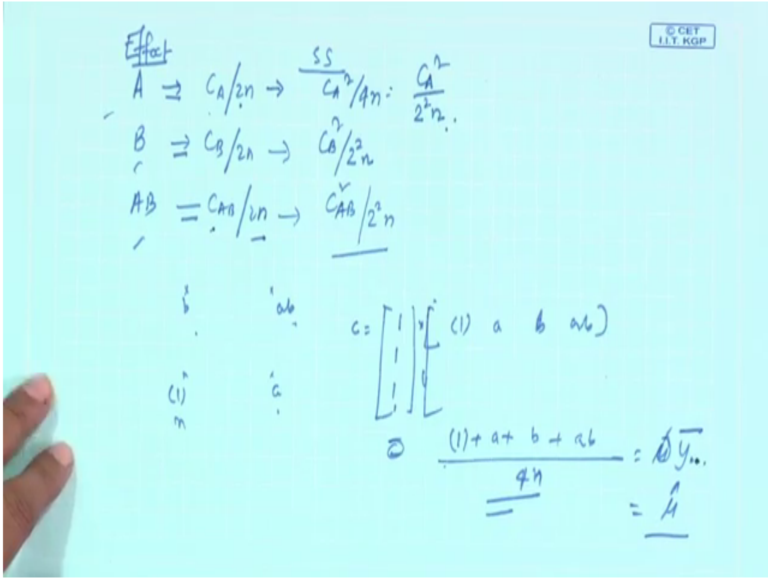
Now, how to compute the sum squares contrast? So, the sum square contrast if I say what is the SS C means sum square contrast; sum of squares contrast. So, these will be sum of i equal to 1 to now you have to see that if it is a you have to know where from changing this one suppose I am writing 1 to a, then this is c i and y i dot bar, the total square divided by 1 by n sum of c i square, i equal to 1 to a and ultimately this will give you this quantity. This will give you if I change into i equal to 1 to a and c i then y dot y i dot because this is the total and if I make the square then n square will out then n sum total of c i square i equal to 1 to a.

So, here again I have used the word a, i equal to 1 to a, rather let me tell you basically we are saying that these are the effects against each treatment combination. So, let me write out not write out a maybe I can write out that treatment combination. So, let this be my p if there are such treatments combinations are there, no problem. Whatever may be you understand that how many treatment combinations are there a minus 1 plus 1 minus one plus 1 for each level you will be able to do it, this is p, this is p.

So, now, if I want to know the contrast for SS A, suppose I want do this SS A then what I will write down, what is c a for me where c i values are minus 1 plus 1 minus 1 plus 1. So, I have to multiplied these by the total is here when it minus 1, it is 1; this is a, this is b, this is ab. So, that mean if I multiplied by c i and y i total, if I multiplied this two columns, what it is giving me, it is giving me minus 1 plus or I will write like this plus a minus b plus ab which is nothing, but ab plus a minus b minus 1. So, that mean this is my contrast square. So, I can say this is contrast if contrast is C then C square or otherwise I can write contrast square.

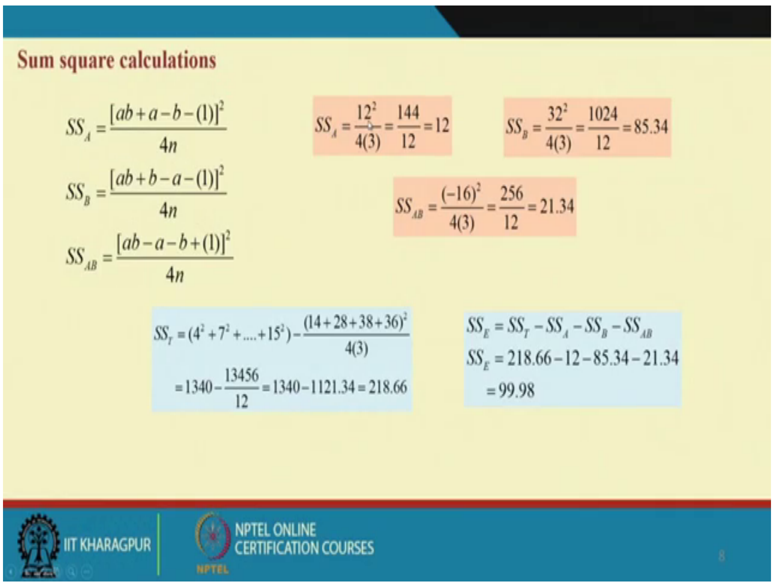
What is the lower value then? Lower value is basically you are multiply squaring c i square. So, c i means this multiplied by this will give you only all 1 minus 1 square this all. The sum of these will be give you i equal to 1 to a c i square will be ultimately what this will be this many suppose if I say p here, it will be p. So, now, in this case what is the settings four, settings sum of four, so, these will be 4. So, then this quantity is 4 into n. So, please remember this one, interesting one.

(Refer Slide Time: 32:50)



So, you want to compute any effect you first know what is the effect either A, B or AB as such for others later on we will only discuss, find out the contrast is C A, C B, C AB. If you want effect, so, effect of A then it is contrast divided by 2n, divided by 2n, divided by 2n, these are the effect, equal to these, equal to this, equal to this. If you want SS then what will happen? for A the SS will be contrast square by 4n means I can write contrast square by 2 to the power 2n, here 2n, 2 to the power 2n. This is C B square by 2 to the power 2 into n and this will be C AB square by 2 to the power 2n.

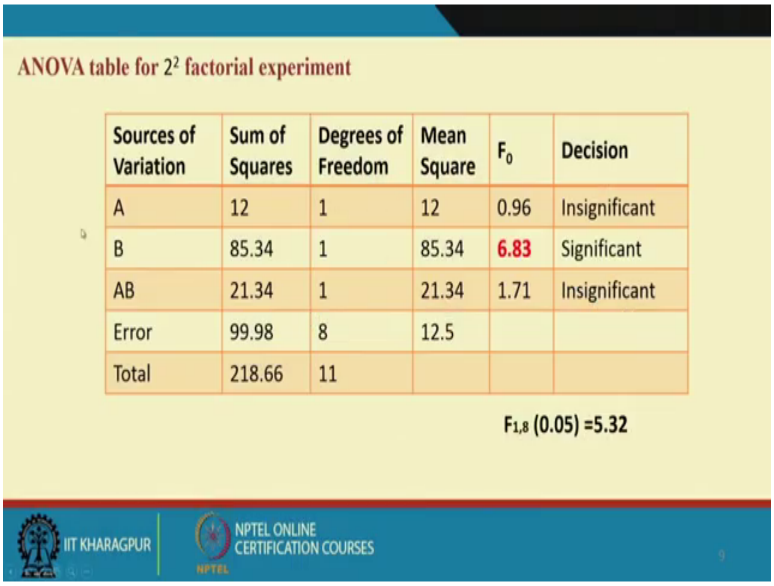
(Refer Slide Time: 34:20)



So, here things become very simple because this contrast is known for each of the effects where then the resultant quantity here what is happening SS A it is SS B this and you just you know the SS A you will be this quantity is 12. 12 square by 4, 3, 12 if it is 12 only SS AB is these, then SS T will be the square of all the observation then sum minus the correction factor like this.

So, how do you compute the mean value then? In case of mean value ultimately what is mean value? For all settings whatever values are there all total divided by number of observations. So, all total will be possible if it is 1, this is a, this is b, this is ab, that means, you require you create a contrast kind of thing for all 1 only and then multiplied by the effect that is 1, a, b and ab then you will be getting 1 plus a plus b plus ab; obviously, there will be how many observation are there? Here n, here n, here n, here n, then 4n observations only total observations this will give you the estimate of y bar y if it is double dot or I think i j k is there, this which is basically estimate of mu.

(Refer Slide Time: 36:00)

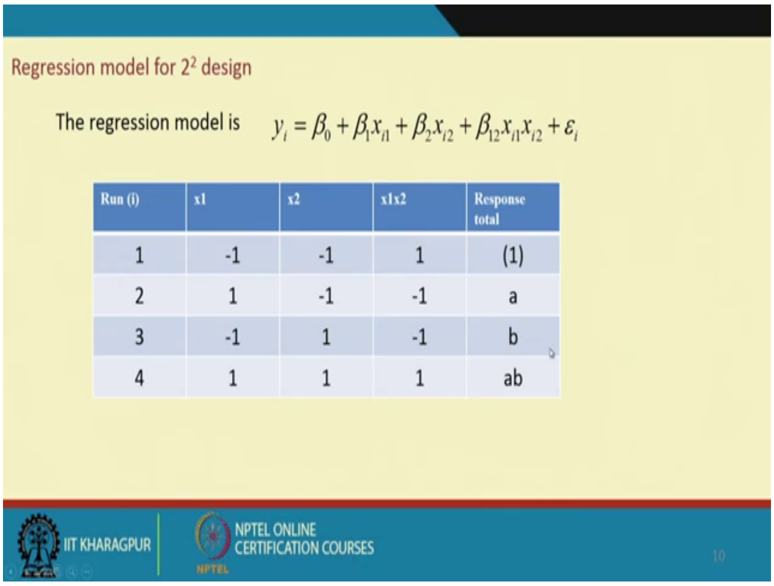

ANOVA table for 2^2 factorial experiment

Sources of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	Decision
A	12	1	12	0.96	Insignificant
B	85.34	1	85.34	6.83	Significant
AB	21.34	1	21.34	1.71	Insignificant
Error	99.98	8	12.5		
Total	218.66	11			

$F_{1,8}(0.05) = 5.32$

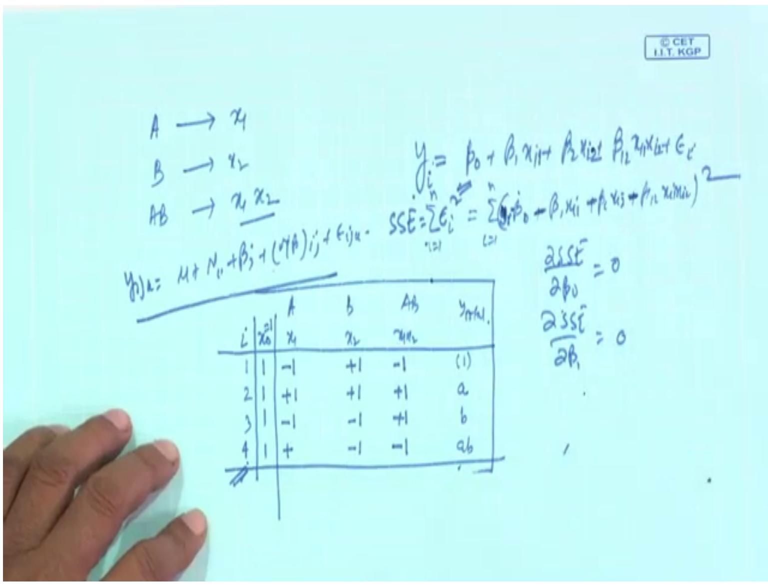
Now, once you know A, B all those things squares are known and you will know that it is significant or not and accordingly you decide interestingly here what will happen whether A, B, AB all effect these will be only one degree of freedom. So, as a result you required to see the error degree of freedom term and as all the effects will be one degree of freedom terms. So, F 1 and error degree of freedom terms 0.05 this value once you will get to take for this any value F value more than this value for this example will be significant.

(Refer Slide Time: 36:34)



Now, I will show you very quickly that how regression approach, that means, what we have seen.

(Refer Slide Time: 36:48)



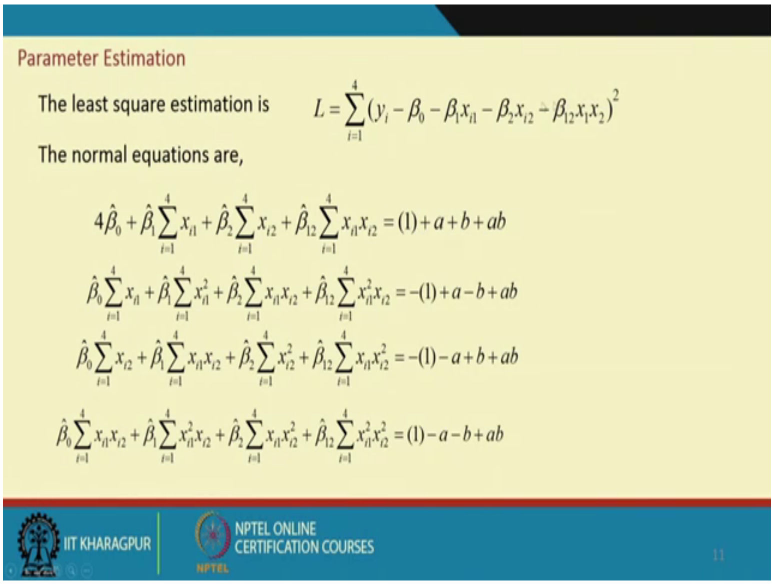
We have seen earlier that we in regression A we denoted by x 1, B we denoted by x 2 and then AB basically it will be x 1 times x 2.

So, then our your fixed effect model is y ijk equal to mu plus tau 1 plus beta j plus tau beta j plus tau beta ij plus epsilon ijk, our regression will be, we will write like this y equal to beta 0 plus beta 1 x 1 plus beta 2 x 2 plus beta 12 x 1 x 2 plus epsilon. So, now, what happen, we know that different settings those different settings we will denote it by observations. So, of settings are with reference to A means x 1, B means x 2 and AB means x 1 x 2 the setting. What will happen? First setting is A minus 1 this is plus 1 then this one is plus, 1 this one is plus 1, this is again minus 1, this is minus 1 then, this is plus 1 this is minus 1.

So, then in independent around 1, 2, 3, 4; then what will be x 1 x 2? This will be minus 1 plus 1 plus 1 minus 1. So, this is our and then what happen then you will get y total somewhere y total you will get y total for this it is 1, this is a, this is b, this is ab as if this is what is our data for regression. So, we have 4 independent settings there are x values, there are y values also. In addition what we require here we require a const call x 0 which will take all is one for beta 0 this will writing will be writing like this 1, 1, 1, 1 this manner we will write.

So, then you what you do you find out SS E. SS E will be epsilon square if I write this i then this is i 1, this will be i 2, this will be i 1 and this will be i 2 then this is epsilon i. So, that mean SS E will be i equal to 1 to n; here it is 4, fine epsilon i square this will be the sum of equal to 1 to n beta 0 plus beta 1 x i1 plus x beta 2 x i2 plus beta 12 x i1 x i.2 These y 0, this one y i minus this y i minus these minus these minus these minus these this square. So, what you do now del SS E by del beta 0 put to 0 del SS E by del beta 1 put to 0 in this manner you put and you solve all these things.

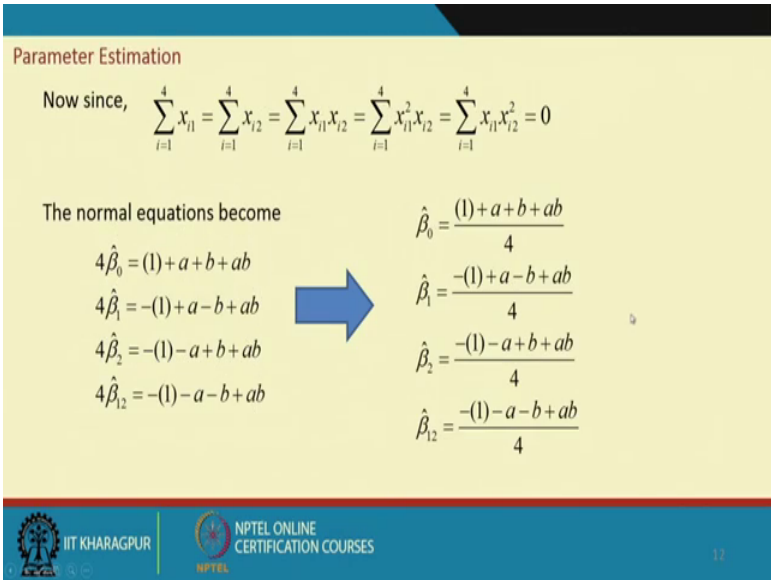
(Refer Slide Time: 40:28)



The resultant equations be a simple one will like this. So, this is my function that is to be minimized. This is basically the SS E. If you make L beta 0 is 0 and then put two 0 equal to 0 you will get this equation. If you do with beta 1, you will get second equation; beta 2 third equation, beta 1 2 fourth equation you will be getting and here you see this sum of x i 1 will be 0, because x i 1 means sum of this column minus 1 plus 1 minus 1 plus this will be 0. This column will be 0, this columns sum all sum will be 0 and if just take the square then it will be had 4 independences square and sum that will be 4.

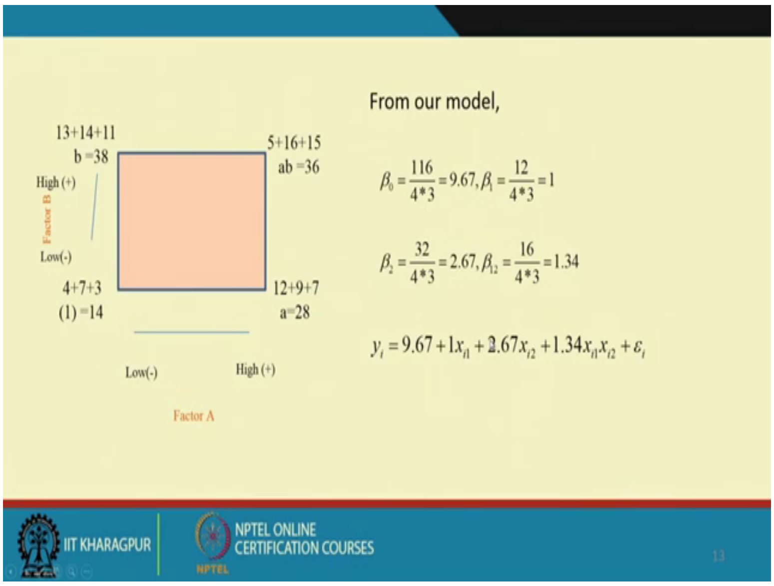
So, that is why, what happened; you are getting these equal to i equal to 1, 2 this y i this total will come here, this is the all sum and here with reference to this x i 1 is multiplied so that the minus 1 plus minus plus coming. Here, the x 2 will be multiplied, so, that is why what happened minus 1 minus plus will be multiplied with these and here what happen that same and different levels. So, same level plus, plus; minus, minus. So, this total is coming.

(Refer Slide Time: 42:01)



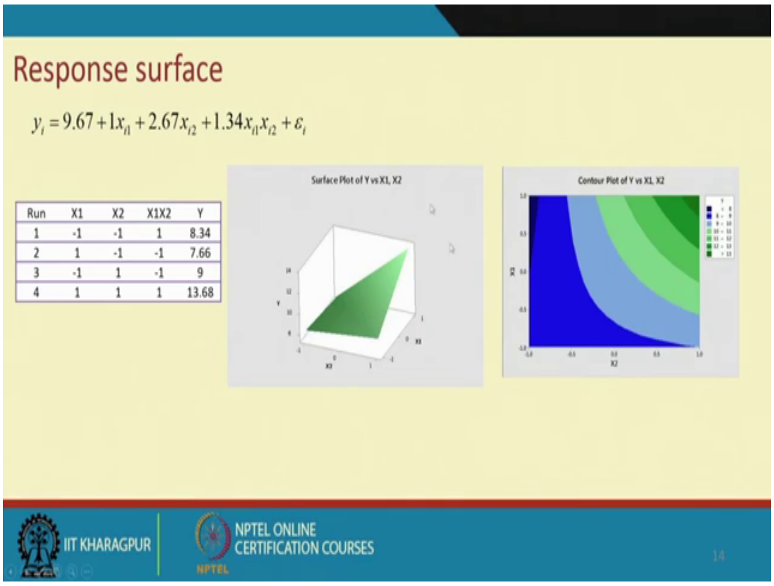
Now, putting the other condition this equal to 0 this equal to 0 sum of these, what you will get; if you put these values as 0, then from this equation 4 beta 0 plus 0 plus 0 plus this also will become 0. So, that will 4 beta 0 is equal to these, then beta 0 equal to 1 plus a plus b plus ab by 4. Similarly, beta 1, 4 beta 1 will be these, 4 beta 2 will be this and 4 beta 1 2 will be these, then you will be getting this beta 1, beta 2, beta like this, these are the estimates.

(Refer Slide Time: 42:33)



So, once you get this beta values, you will see that the beta values are nothing, but a by 2, beta 1 is a by 2, beta 2 is b by 2, beta a 1 2 will be ab by 2, the effect parameters what you have estimated in fixed effect model.

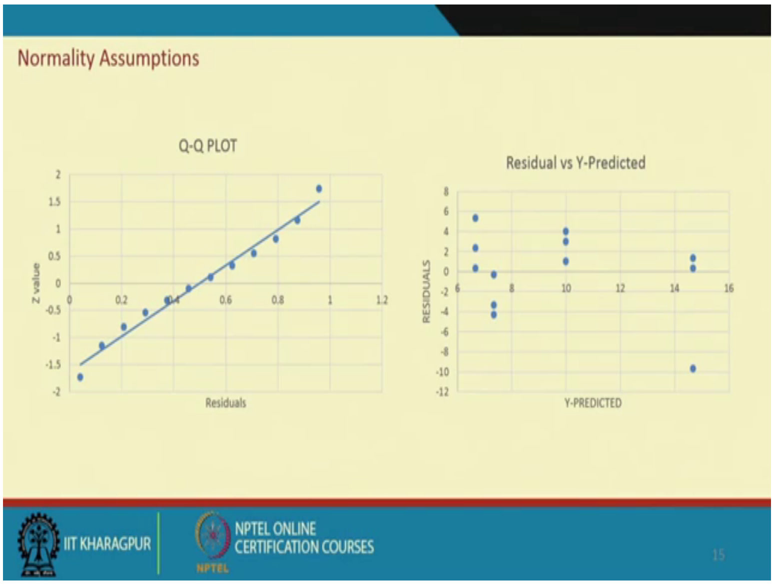
(Refer Slide Time: 42:54)



Now, the advantage of having this model is that we can go for response surface. So, this is our run, that is, data and this is the response surface plots for this particular equation. Here, x 1 x 2 and how y is changing you see that almost a plane, perhaps because of this what is the beta 1, 2 values you have seen that beta 2 effect is significant earlier, but beta 1, beta where is your beta 1 as well as this beta 1, 2 is not significant will be there. Anyhow, those things you will be tested through the regression model the way we have done that can be same way it will tested here.

Now, from here we have here we are giving a contour plot like for a different settings for x 1 and x 2, what will be the range of y values here. So, there a color combination suppose if it is a maximization problem, what you want y values should be maximum that is greater than 13 here when that is zone this is the zone of interest where you can set your process. Suppose, it is a minimization case then with reference to this is the zone where y values are we will be getting minimum y values. But, if it is some target values is there suppose 10 to 11, then this is the zone which gives you the y value 10 to within 10 to 11 and accordingly, you choose what will be your x 1 and x 2 values.

(Refer Slide Time: 44:30)



I hope that you have understood and you please remember that, there will be test on assumptions like normality homoscedasticity every time so, you have to do all those things. It is simple once you have residual; you do residual analysis and get it done. Thank you very much. See you again.