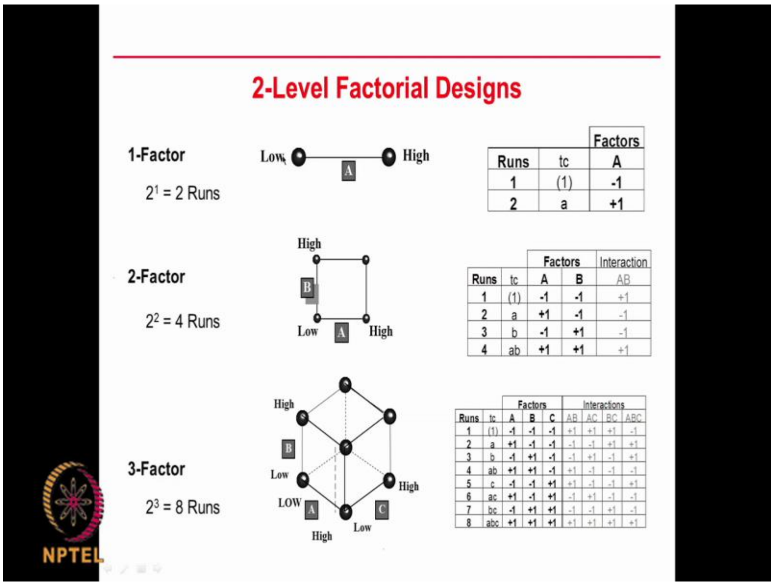
Biostatistics and Design of Experiments
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology, Madras

Lecture - 34
Full Factorial design

Welcome to the course on Biostatistics and Design of experiments. We will look at Full Factorial Design. In the previous class I talked about what is the factor and what is the level.

(Refer Slide Time: 00:24)



Let us look at, for example this particular number tells you the level and the exponent tells you the number of factors. Suppose you have 2^1 that is only 1 factor, you will have 2 experiments here, 1 at low, 1 at high right, the factor a will be done with the low level and with high level. You will have 2 experiments done. Suppose you have 2 factors like a and b at 2 levels, 2 levels 2 factors 2^2 that is 4 experiments. We need to do 4 experiments run 1, 2, 3, 4 and this is one way of representing them factor a, when you put - 1 that is at low level, when you put it at + 1 you call it high level. Factor a at low level factor b at low level, factor a at high level factor b at low level, factor a at low level, factor b at high level, factor a at high, b at high. So this is one way of representing - 1 - 1, + 1 - 1, - 1 + 1,

+ 1, + 1. Another way of representing these 4 experiments is like this actually, we call it 1 that is the base level -1, - 1 and then we call it small a, that means factor a at higher level, when you call it factor b, factor b is at higher level, when we call a b, factor a as well as factor b at higher level. This is one way of representing; this is another way of representing. If you look at only 1 factor now, we will call it the 1, that is lower level and when we say small a, that means factor a is at higher level.

Generally, we like this one because we can do lot of things with this. Now if I want to look at interaction A, B that means temperature pH example I am coming back so how do I study interaction A, B? All I have to do is multiply these 2 numbers to get this number - 1 * - 1 is + 1, + 1 * - 1 is - 1, - 1 * + 1 is -1, + 1 * + 1 = 1. Do you understand?

Interaction AB is given like this. This is the column for interaction AB and this is the main factor a and b this is the interaction A, B. So here you are doing four runs and we can get information about the effect of A and the information about effect of B and we can also get effect of interaction A, B. How do we get that? In the previous time, I mentioned to you right? That if I am having a results output, four experimental outputs. If you want to look at the effect of A, I add up the results from here ÷ 2 subtract it from adding up this and this / 2 because this is at higher level this at lower level. If you want to look at effect of B, I add up the results from these 2 experiments ÷ 2 subtract it from these 2 experiment ÷2 that is average. If I want to look at interaction AB, I will take the results of this experiment.

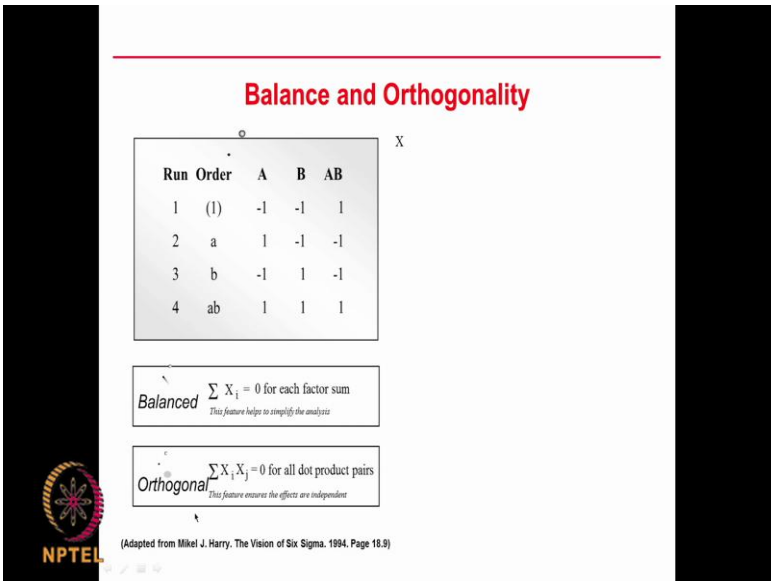
This experiment add up ÷ 2 subtract it from the results of this experiment plus this experiment by two. It is so nice, we can look at effect of interaction very nicely from this type of factorial experiment that is the beauty of factorial experiment and another interesting thing you can see is the number of pluses in this column will be equal to number of minuses. That way you will have a very symmetric 2 + , 2 -, 2 +, 2 -, So there is the type of experiments we do also will be symmetric, that means we will do in equal number of higher level experiments with the equal number of lower level experiments. Even for A, B if you see you will have 2 +, 2 -.

Now, let us go to the three factor experiment that means I am looking at temperature, pH and dissolved oxygen amount on a product yield. For example, A, B, C what will be the design? 2^3 the 2 means 2 level, 3 means 3 parameters. Of course instead of 2 levels you can also go for 3 level x, 3 levels and so on. We will talk about it later, but the most of these factorials experiments we go with the 2 levels.

2^3 experiments that is 2 * 2 * 2 is 8 runs. It is like a cube a will be at low level and high level, b will be at low level and high level, c will be at low level and high level 8 runs. How do you generate the factor experiment? Look - 1 - 1 -1, then we put a + 1 - 1 - 1, then for b - 1 + 1 - 1, then AB + 1 + 1 - 1. Now we keep changing C - 1 - 1 + 1, + 1 - 1 + 1, - 1 + 1 + 1 then all of them are pluses. This is one representation. If you want to represent the other way you have 1 A, A means level A factor A is in higher level, B means only factor B is in higher level, AB means both factor A and B are in higher level, C means factor C is at high level, A,C means factor A and C are at high level, B,C means factor B and C are at high level, ABC means factor A B C all 3 are at higher level. So we have 8 experiments, if you look at the column you will have 4 pluses 1, 2, 3, 4 you will have 4 minuses, look at column b you will have four pluses you will have 4 minuses, look at column c you will have 4 pluses 4 minuses. So it is very well balanced this called a balanced design.

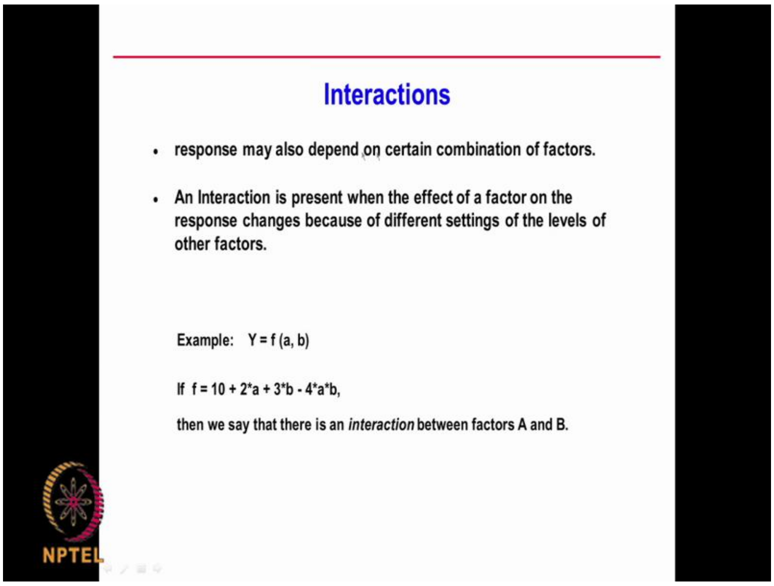
Now how do we get the interactions? AB interaction you multiply A and B column - * - is +, + * - is -, - * + is -, + * + is +, - * - is +, + * - is -, - * + is -, + * + is +. So BC, AC column, A into C, BC is B into C, ABC is all three together multiplied. So -, -, - will give you -; +, -, - will give you +; -, +, - will give you +; +, +, - will give you -; -, -, + will give you +s; +, -, + will give you -; -, +, + will give you -; +, +, + will give you +. Look at the interaction columns also so beautiful you have 4 +, you have 4 minuses. It is so balanced so the factorial designs are always balanced. They are called Balanced because if you look at the effect 2 +s 2 -s, B effect 2 +s - and they are also called Orthogonal because these interactions are also balanced 2 minuses, 2 pluses if you go to the previous slide look at these BC is got 4 +s, 4 -s, look at AC is got 4 +s, 4 -s. So, balanced and as well as orthogonal.

(Refer Slide Time: 08:04)



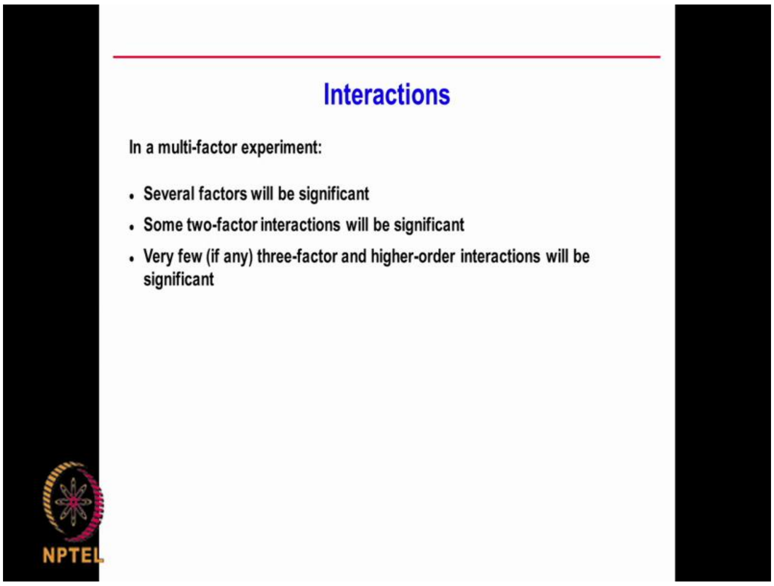
So factorial designs are always like that and it you have to be very careful when you select it. You should always have a balanced and an orthogonal designs, please remember that. If you do not have a balanced design, that means if you do not have enough plus equal number of pluses and minuses. Suppose you pair a design where you have one more =s than -, that means you are doing biased design where you are trying to look at the performance at higher level, which is wrong because any design should be unbiased. You should always have a equal number of +s and - that is one thing. Other thing is your interactions also are balanced. That way you can study the interactions also in a equal unbiased manner. All these designs, especially the factorial design are always balanced and orthogonal. When you select designs later on you should always look whether your designs are balanced, if it is not then you have a biased that means you are focusing more on one particular a level of factor rather than looking at all the factor all the levels in uniform way that is very important when you plan your design. The balance and orthogonality are very, very important when you perform a strategy of design. You can look re interaction as I said combination of factors. So we can look at ab, when you develop your regression relation you will not only have equation terms like this.

(Refer Slide Time: 10:03)



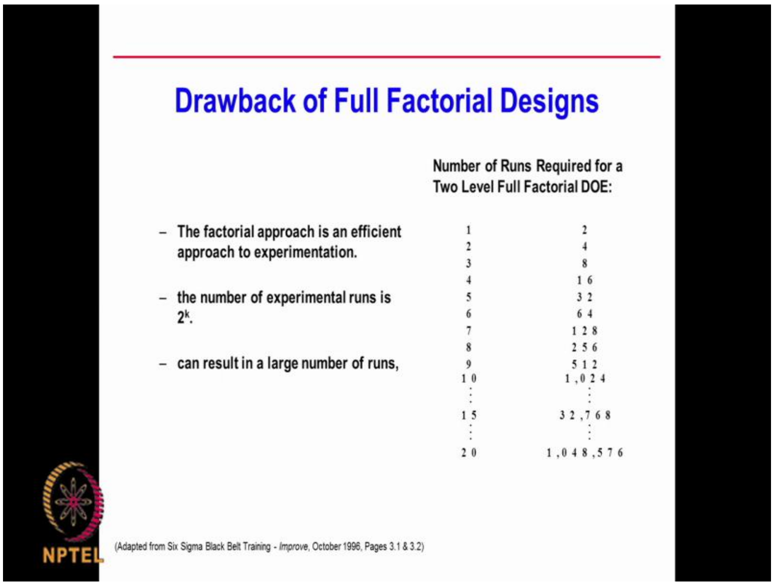
We will also have terms like a and b we looked at it long time back temperature and pH and so on actually.

(Refer Slide Time: 10:20)



Several factors will become significant, sometimes 2 factor interactions may be significant but 3 factor interactions might not be significant at all very rare. We will not have a situation where temperature pH and agitator r p m interacting with each other. But you can have situation where temperature and pH may be interacting, pH and agitator RPM may be interacting. For example, when I increase agitator RPM may be I am dispersing the material better; pH is more uniform and so on. pH so 2 factor interactions are generally possible. But 3 factors and 4 factors are very, very rare. You need to keep that point in mind. It does not happen every day, it does not happen at all in most of the situation. Main factors will be significant 2 factors interaction will be significant.

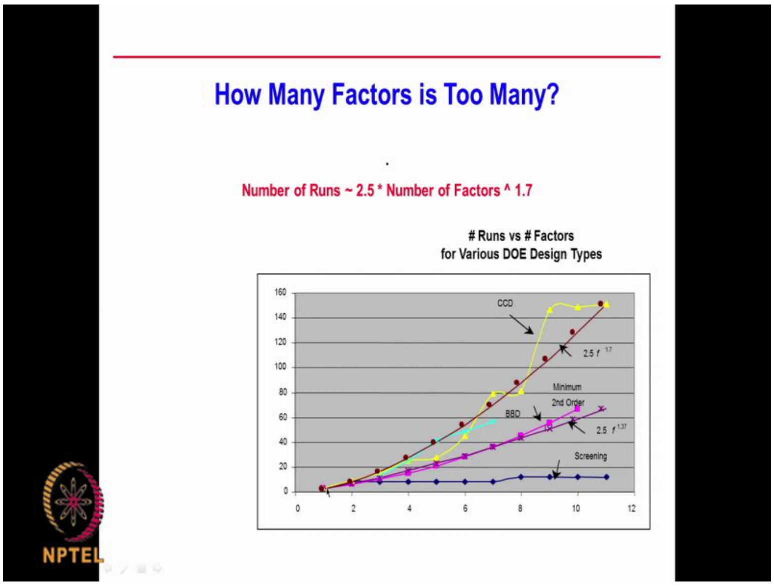
(Refer Slide Time: 11:12)



But the there are drawbacks of these full factorial design like a it is 2^n right, so n is 1 means 2 experiments, 2 means 4 experiments, 3 means 8 experiments, 4 means 2 into 2 into 2 into 2, 4, 16, 5 means 32 like that it keeps on increasing actually you know it is a huge. imagine if I am doing a 15 parameters factors 2^{15} that is a huge number of experiments now that won't do at all. There is something called fractional factorial design.

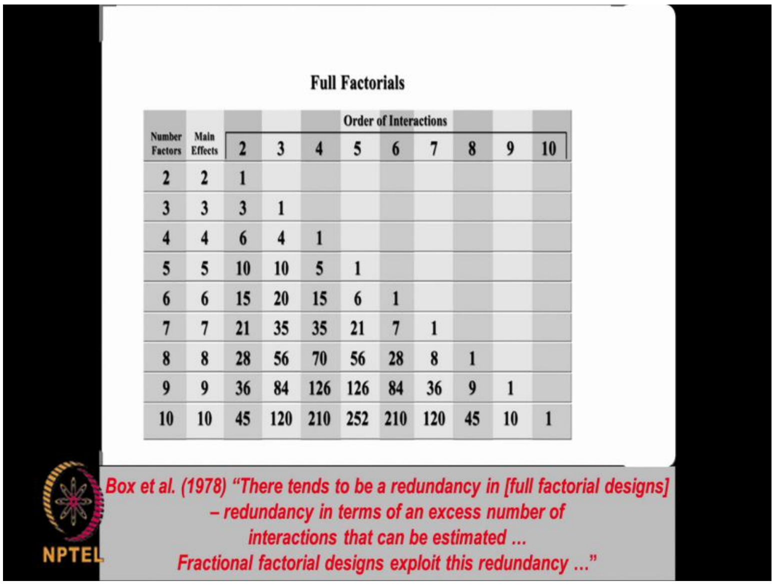
We will look at fractional factorial design in more detail as we go long. But fractional factorial designs are fractions of this main full factorial design. You can have half fractional factorial that means you will do half of the full factorial, you can have one-fourth of full factorial, you can even have one-eighth of full factorial, so you will do much less number of experiments like half 1, 4, 1, 8 but, you lose out on some other parameter we will talk about it later on. But the number of experiments goes down, if you are doing half or one-fourth of the full factorial experiment.

(Refer Slide Time: 12:26)



This figure tells you as you keep on increasing the factors it keeps on going up, that is not a very good idea to do actually.

(Refer Slide Time: 12:33)

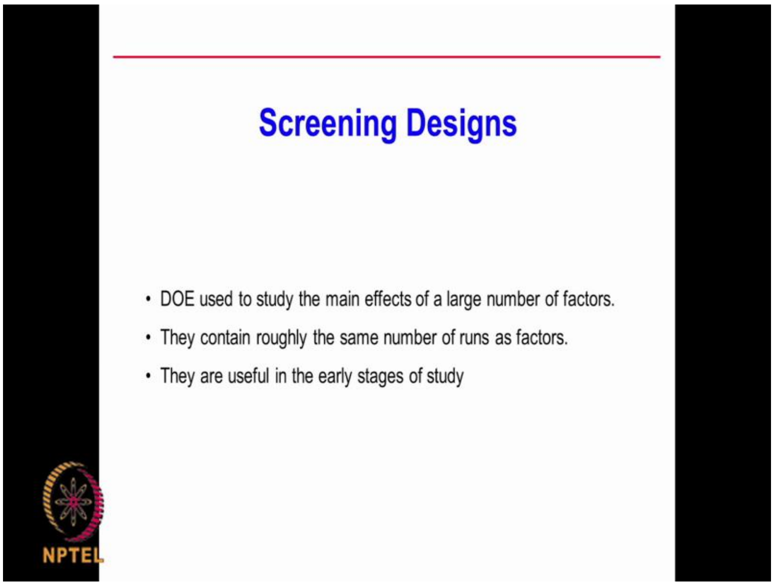


Number Factors	Main Effects	Order of Interactions								
		2	3	4	5	6	7	8	9	10
2	2	1								
3	3	3	1							
4	4	6	4	1						
5	5	10	10	5	1					
6	6	15	20	15	6	1				
7	7	21	35	35	21	7	1			
8	8	28	56	70	56	28	8	1		
9	9	36	84	126	126	84	36	9	1	
10	10	45	120	210	252	210	120	45	10	1

If you do a full factorial for example, if you take 2 factors like a and b you are doing 4 experiments. So you are looking at main effects of a, main effect of b that is 2 main effects and we can also study the interaction AB, that is called a second order interaction. Suppose we have three factors a b c there are 3 main effects. Then you can have 3 second order interaction ab, ac, ac and you can also have a third order interaction a, b, c. If you look at 4 variables or parameters or factors A, B, C, D you will have main effects A, B, C, D then you will have 6 two way interactions like AB, AC, AD, BC, BD, CD that is 6, you can have 4 three way interactions ABC, ABD, BCD, ACD and so on, you will have 4 and there will be 1 four way A, B, C, D like that it goes right. You, in many situations you will end up having very large number of interactions and as I said it is very, very rare to have a three way interaction, so it is very rare that means this portion is actually useless. We are doing too many experiments and we are not going to anyway look at three level or higher level interactions, higher order interactions that mean we are doing too many experiments. Look at box he said there are tends to be redundancy in full factorial design because there are redundancies, because 3 level or three way interactions are very rare. That is why it is always good to do fractional factorial designs to exploit this redundancy. Like I said fractional factorial if you do a half fractional factorial we will be doing half of a full factorial or if you doing one-fourth full factorial that means that could be a one-fourth fractional factorial design because nobody is interested in three way or four way or five way interaction that is very, very rare. All this portion is all redundance.

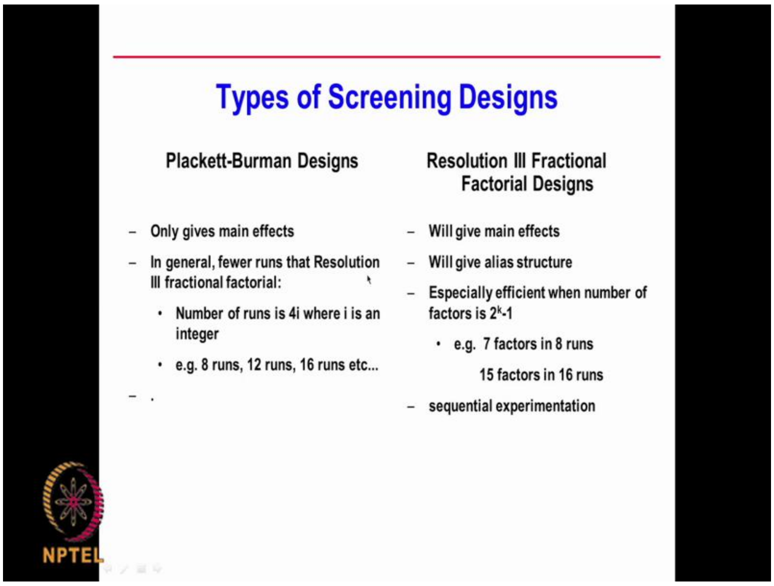
All these experiments are waste you know that is why it is better to always do a fractional factorial design and as you know it is always everybody likes to do less number of experiments because of time constraint, because of resource constraint but with less number of experiments we should be able to get as much information especially, the main effects and the main interactions. Two factor interaction AB, AC, AD, BC, BD these are the things we are interested in, if it is A, B, C, D we are interested in main effects like A, B, C, D and then we are interested in AB, AC, AD, BC, BD, CD we are not interested A, B, C that is very rarely going to happen, so such is not generally absorbed in many situations. That is why one goes about doing something called the fractional factorial studies. How do you go about selecting a fractional factorial? Let us look at it in very systematic way.

(Refer Slide Time: 15:57).



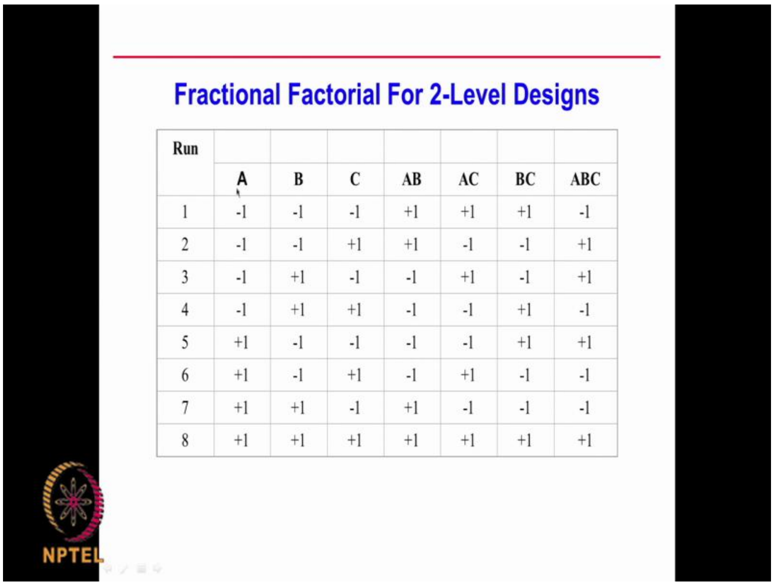
If you go to screening designs, like I said initially you may select large number of parameters and you perform some experiments. In that screening design you are interested to identify which are the critical parameters, which are noise parameters eliminate the noise take only the critical and spend more time in screening designs we take large number of factors. We are more interested in only main effect we are not even interested in interactions. In screening design you will perform only few experiments get the main effects and then take only few important ones and then go and do a detailed design that is called Screening Design.

(Refer Slide Time: 16:40)



There are many types of screening designs one is called the Plackett-Burman design, the other one is called Resolution three Fractional Factorial design. I will talk about what are these actually you know these are some screening designs which are used to just screen large number of parameters or factors eliminate quickly, which you think is not necessary by doing minimum experiments and select only those which are only critical and then do a detailed study. That is what is screening design is all about actually, we will spend more time on this Plackett-Burman design and resolution three designs later on. Generally the screening design involves lot of fractional factorial designs. Let us look at a simple picture.

(Refer Slide Time: 17:27)



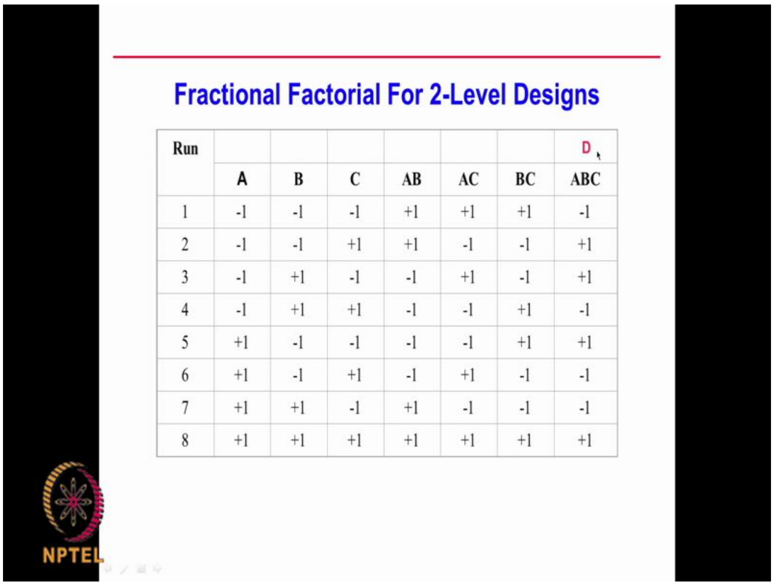
Fractional Factorial For 2-Level Designs

Run	A	B	C	AB	AC	BC	ABC
1	-1	-1	-1	+1	+1	+1	-1
2	-1	-1	+1	+1	-1	-1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	-1	+1	+1	-1	-1	+1	-1
5	+1	-1	-1	-1	-1	+1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	+1	+1	-1	+1	-1	-1	-1
8	+1	+1	+1	+1	+1	+1	+1

Imagine I have three parameters or factors A, B, C so 2^3 design. That means 8 experiments. So how do I change? - 1, - 1, - 1 so each one I am changing to - 1 to + 1 so, the last experiment will be +, +, +. Do you understand? How do you get AB? This into this -, -, +; -, -, +; -, +, -; +, +, -; +, -, -; +, -, -; +, +, +; +, +, +. So look at here you need to always cross check whether the balance is there number of negatives equal to number of positives and the north orthogonality the interactions also there are 4 positives, 4 negatives, 4 positives, 4 negatives like that.

Now I want to go into a fractional factorial design. Imagine I want to introduce one more factor A B C was say the pH temperature and carbon but actually I want to study pH temperature carbon and nitrogen. So 2^4 will be 2 * 2 * 2 * 2. 16 experiments, I do not want to do 16 experiments I want to do only 8 experiments and get lack lot of data. What do I do? I take this design matrix and do a fractional factorial to involve d also. Now, where do I put d here that is the question, Do you understand? I have 4 factors or parameters A, B, C, D temperature pH, carbon, nitrogen. If I have to do a full factorial design it will become 2^4 that is 16 experiments but I do not want to do 16 experiments I want to do only 8 experiment that is half of 2^4 or 2^{4-1}. So 1 of them I want to make it d generally, you take the highest order interaction that means A, B, C as I said before people generally see these first order interaction but, never second order interaction. So you call this D.

(Refer Slide Time: 19:48)



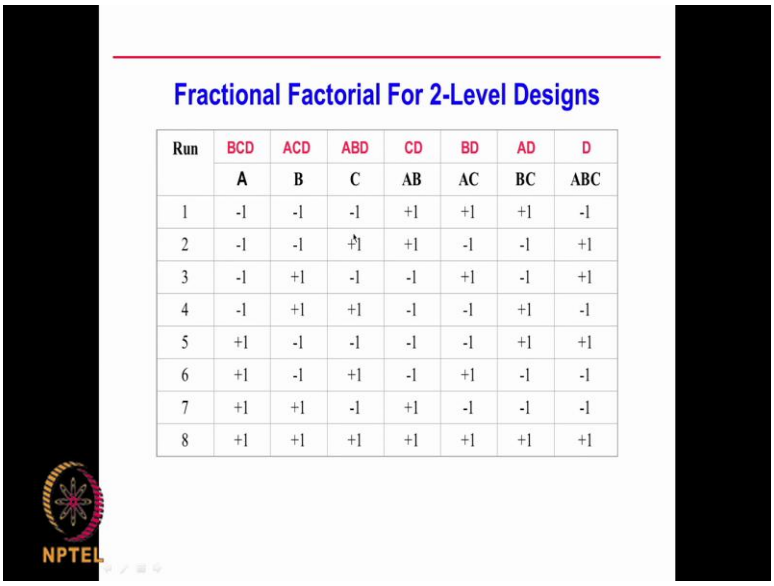
Fractional Factorial For 2-Level Designs

Run	A	B	C	AB	AC	BC	D ABC
1	-1	-1	-1	+1	+1	+1	-1
2	-1	-1	+1	+1	-1	-1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	-1	+1	+1	-1	-1	+1	-1
5	+1	-1	-1	-1	-1	+1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	+1	+1	-1	+1	-1	-1	-1
8	+1	+1	+1	+1	+1	+1	+1

Your design will be only 8 experiments, 4 factors a will be like this, a is modified like this, b is modified like this, C is modified like this and d is modified like this. If I have temperature pH, carbon, nitrogen, 4 factors instead of doing 16 experiments I am doing half of 16 so I will have only 8 experiments so I take the design matrix for 8 experiments A, B, C and then I will put the A, B, C A * B into C as D, so I will change D in this fashion.

Now the question is there is a catch? What will happen? You have already have certain interactions AB, AC, BC. What will happen to D involved? What will happen to AD? What will happen to BD? What will happen to CD?

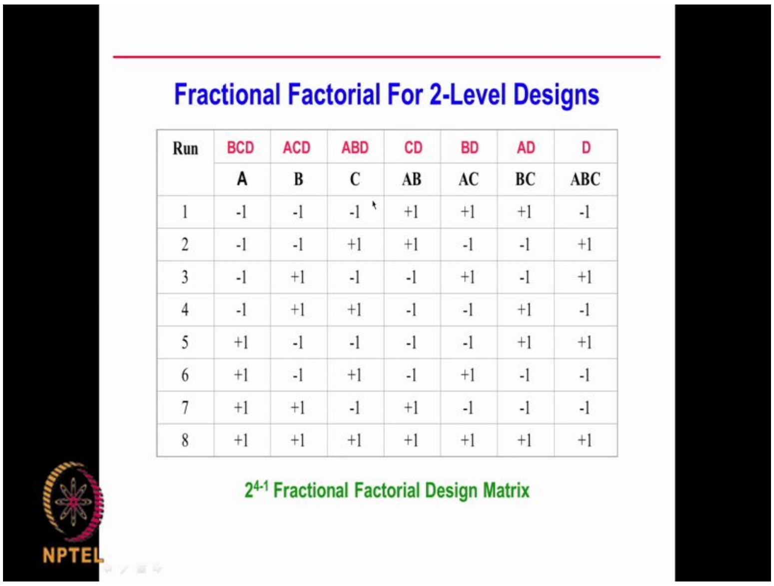
(Refer Slide Time: 20:45)



Run	BCD	ACD	ABD	CD	BD	AD	D
	A	B	C	AB	AC	BC	ABC
1	-1	-1	-1	+1	+1	+1	-1
2	-1	-1	+1	+1	-1	-1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	-1	+1	+1	-1	-1	+1	-1
5	+1	-1	-1	-1	-1	+1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	+1	+1	-1	+1	-1	-1	-1
8	+1	+1	+1	+1	+1	+1	+1

When you do that you find AD same as CD, AB I said -, -, +; -, -, +; -, +, -; -, +, -; + * -, -; +, -; so +, -, -; +, +, +; +, +, +. Now, if you look at CD, -, -, +; +, +, +; -, +, -; +s, -, -; -, +, -; +, -, -; -, -, +; +, +, +. AB looks exactly like CD same thing I can show that BD looks exactly like AC, same thing I can show AD looks exactly like BC and also a looks exactly like BCD just like D looks exactly like ABC, A looks exactly like BCD, B looks exactly like ACD, C looks exactly like ABD.

(Refer Slide Time: 21:51)



Run	BCD	ACD	ABD	CD	BD	AD	D
	A	B	C	AB	AC	BC	ABC
1	-1	-1	-1	+1	+1	+1	-1
2	-1	-1	+1	+1	-1	-1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	-1	+1	+1	-1	-1	+1	-1
5	+1	-1	-1	-1	-1	+1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	+1	+1	-1	+1	-1	-1	-1
8	+1	+1	+1	+1	+1	+1	+1

And this design is of course 2^{4-1} because I have A, B, C, D as 4 parameters I am doing only a instead of 16 experiments 8 experiments, I am doing 2^{4-1}.

What has happened here? I am able to do away with instead of 16 experiments only 8 experiments, but some of these factors are confounded, new term I am using some of the factors are confounded AB will look like CD, AC will look like BD, BC will look like AD. I will not able to differentiate between AB and CD, I will not be able to AC differentiate AC and BD I will not be able to differentiate BC and AD and of course I will not be able to differentiate d against A, B, C so I will not be able to differentiate a against BCDB, against ACDC, against ABD. So, these have been confounded. These 2 level interactions have been confounded and the principle effects or main effects are confounded way these 3 parameter, 3 factor interaction. Do you understand? So these 2 factors interactions are confounded with 2 factor interactions and the main effects or main factors are confounded with these 3 factor interactions.

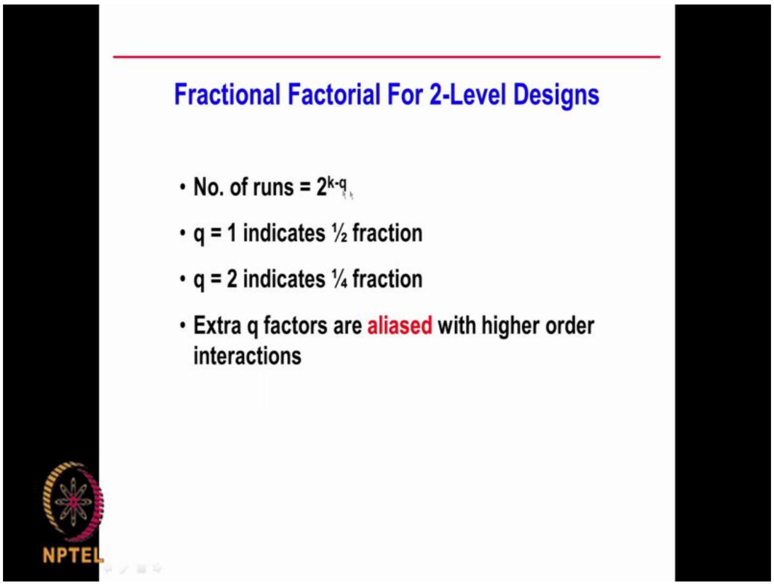
Some information is lost but, what has happened? I have reduce the number of experiments instead of doing 16 experiments that is 2^4 that is 4 factors A, B, C, D I have reduced it by half so this called A , 2^{4-1} designed, 2^{4-1} is 2^3 which is 8 experiments. But in the bargain what has happened? Some interactions are confounded AB interaction is confounded with CD, AC interaction is confounded BD, BC interaction is confounded with ad that means I will not be able to differentiate, whether if some change is there is it because of AD or BC, whether it is because of AC or BD, because of AB or CD. Same thing for the main effects also they are confounded with 3 factor interaction and that is not a big problem because generally we do not observe any 3 factor interacting. Whatever change we see because of main effect we are sure that is because of the main effect only. Do you understand?

We have lost some information, which is not of very big significance unless there is some problem here, AC the 2 level interactions are confounded unless you think there is going to be some important ah knowledge lost. Otherwise this is very good sort of a screening design for 4 parameters or 4 factors instead of doing 16 experiments I am doing only 8 experiments and this is called a 2^{4-1} fractional factorial design. So number of runs =

- 2^{k-q}

.

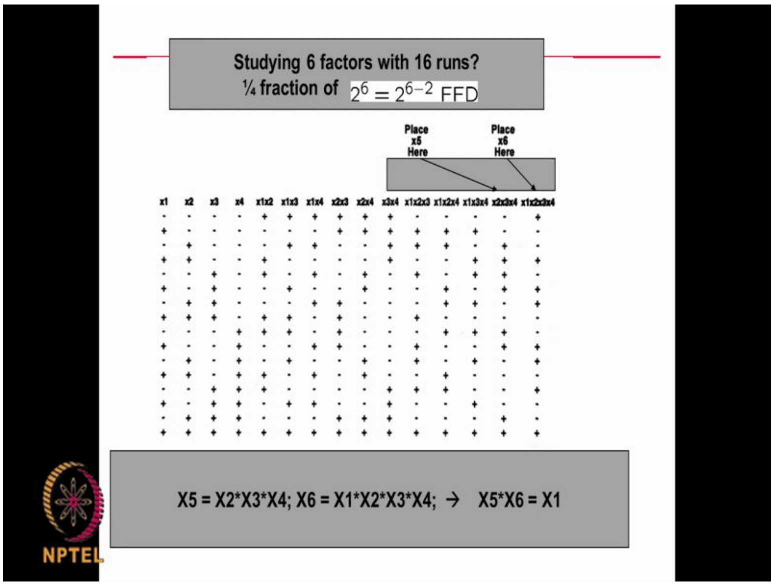
(Refer Slide Time: 25:16)



- This 2 is the number of levels your k is the number of factors, q is 1, then say in the previous case 2, 4 - 1 that is a half instead of 16 experiments I am doing 8 experiments if q is 2 indicates quarter. Extra q factors are aliased with higher order will take about it later. So this is called a 2^{k-q}

or 2 4 - 1 design this is a half fractional factorial design for 4 parameters so instead of 16 experiments I am doing only 8 experiments in this particular case.

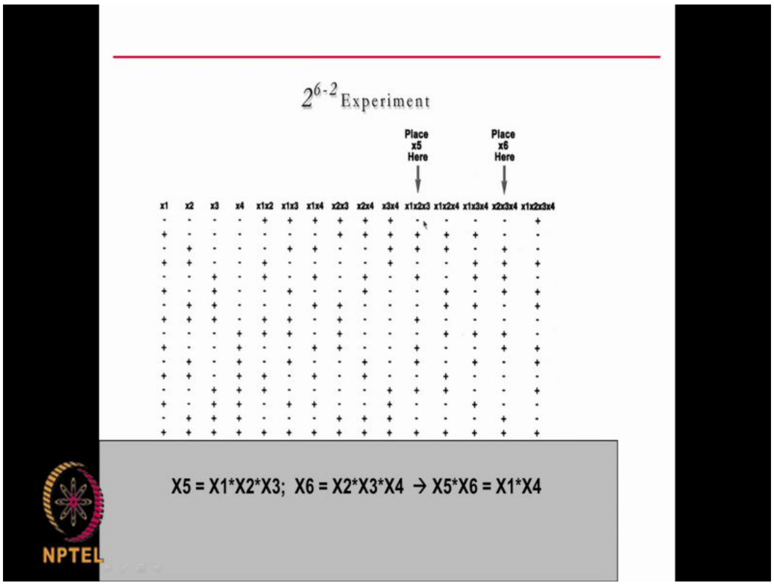
(Refer Slide Time: 26:04)



Now let us look at six factors x 1, x 2, x 3, x 4, x 5, x 6. Now imagine first I start with the design matrix for x 1, x 2, x 3, x 4 that is 4 parameters 2^4 is 2 * 2 * 2 *2 16 experiments, understand. So I write down this sixteen experiments and then I can write down the interactions x 1 into x 2, x 1 into x 3, x 1 into x 4, x 2 into x 3, x 2 into x 4 then x 3 into 4 then 3 level x 1, x 2, x 3 like that like that it goes.

Now I want to introduce 2 more factors x 5 and x 6. Now where do I introduce, the problem is I could introduce in different places right, I could Introduce in different places the x 5 and x 6. This if I introduce this x 5 and x 6 that is 2^{6-2}that is a quarter fraction design right instead of 6 2 into 2 into 2 6 time 4, 8, 16, 32, 64 experiment I am doing only 16 experiments that is a quarter fraction of 64 experiments 2^{6-2} fractional factorial design, do you understand? So do I introduce x 5 here, that is x 5 = x 2, x 3, x 4 and x 6 = x 1, x 2, x 3, x 4. Then if I multiply x 5 * x 6 what will happen that will become x 1, this is very dangerous what it means is the principle effect x 1 is confounded with their 2 way interaction x 5 and x 6. I daily we always said that principle effects should interact have an interaction with more than 2 factors at least 3 factors or more then that is a good design, whereas, here it is interacting with 2 factors that is not a very good design at all. What do we do? We need to look at some other columns. So that the principle effects will have confounding only with 3 factors.

(Refer Slide Time: 28:29).



Now if we look here x 1, x 2, x 3, put it as x 5 and x 2, x 3, x 4 put it as x 6 then x 5 into s x 6 will become x 1 * x 4. There is only a 2 factor interactions unlike the previous case where x 1 is interacting with 2 factor x 1 is a principle effect right whereas, here 2 factors are interacting with 2 factors so this design looks good. You need to think about how to select the column in which you want to put in your new parameter or new factor or new x or new variable. That you do not end up situation like this, where a two factors interacting with a principle effect or main factor, whereas two factor can interact with 2 factor like this. So, this happens in this type of situation where you can have the new variables and choice is either out of this 5 how to select 2 of them, when you selected these 2 they are not good. So what did you do you selected these 2 so that is how you get your design done actually.

Let us continue more on this which fractional factorial design I introduce the terminology called Confounding. We have these factors confounding in some situations, we have three factors confounding with the principle effect or main factor which is allowed but, you do not want a main effect or main factor confounding with 2 factor interaction but it can confound with 3 factor interaction.

So x 1 can confound with x 2 into 3 into x 4, but x 1 should not confound with x 2, x 3. But 2 factors can be confounding with each other. Like here x 5 into x 6 can confound with x 1 into x 4 or in the other situation BC is confounding with AD, BD is confounding with AC, AD is confounding with AB. in this problem there was no problem we could put D exactly here whereas, in this problem these 2 factors or 2 effect parameters could be put in any place because there is lot of choice here. Do you understand? And selection of the correct column will depend upon which factors get confounded. So we will continue on this concept of a fractional factorial designs and confounding in the 4th coming classes.

Thank you very much.

Key words Full Factorial Design, variables, Interactions, factorial experiment, two-factor interactions, Plackett-Burman Designs, Resolution III Fractional Factorial Designs