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

Lecture - 35
Fractional Factorial design

Welcome to the course on Biostatistics and Design of Experiments. We have started talking about Factorial Designs and then, Fractional Factorial Designs. Factorial designs are extremely powerful. We are looking at 3, 4, 5, 6 factors at two levels. Then, they are also called a 2 power n design, where this 2 indicates the levels; n indicates number of the factors or parameters. So, we can have 3 power n designs and so on actually. Then, I started introducing the terminology of Fractional Factorial Design. That means you do a fraction of those factorial design. So, what are the disadvantages of Full Fractional Design?

(Refer Slide Time: 00:49)

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

Suppose I take 2 factors a, b. Like temperature and pH. So, I can study the main effects of temperature and p h. I will be doing 4 experiments. Like yesterday, I talked about a at lower level, then a at higher level, b at lower level, b at higher levels. So, there are 4 experiments. So, we can study the effects, main effects. We can also study interaction a b effect that is called interaction, that is called a b effect.

Now, let us look at these 3 factors, that is temperature, pH

and carbon. That is a 3 factor. So 2^3 design, 2^3 designs that is 2 * 2 * 2, that is 8 experiments. So, that will, we can study the main effects that is temperature, pH and carbon. Then we can study 3, two way interactions, AB, BC, AC that is 3. In addition, we can also study a 3 way interactions a, b, c. Yesterday mentioned generally in systems you will always see a 2 way interaction, like the drug interacting with the rays or drug interacting with the gender or sometimes temperature interacting with pH, but 3 way interactions are very rare. 3 way interactions are very rare. Now, if we take 4 factors that is temperature, pH, carbon and nitrogen, we can study each one of these main effects, that is temperature effect of temperature alone, effect of pH alone, effect of carbon alone, effect of nitrogen alone. In addition, we can also study the 2-way interaction AB, AC, AD, BC, BD, CD and so on. So, there are 6 in number, and then there will be a 3 way interaction ABC, BCD, ACD, like that. BCD that is 4 and then, we will, there will be a 4-way interaction a, b, c, d like that it goes.

If I am going to have a 5 factors, 2^5, we are talking about in terms of 32 experiments and in addition, we are also talking about lot of 3 way and 4 way and 5 way interactions which are redundant actually. So, Box said there tends to be a redundancy in full factorial design, because we have excess number of interactions and generally interactions like 3 way, 4 way, 5 way, 6 way are very rare. 2 way yes, but others are rare. So we seem to be doing too many experiments from that point of view and also, if we do too many experiments it is going to be time consuming, resource consuming. So, why not go into Fractional Factorial Design. That means you do a fraction of the factorial design. You take the factorial and then, make a fraction. It could be half of that, it could be one fourth of that and so on actually. So how do you go about doing it?

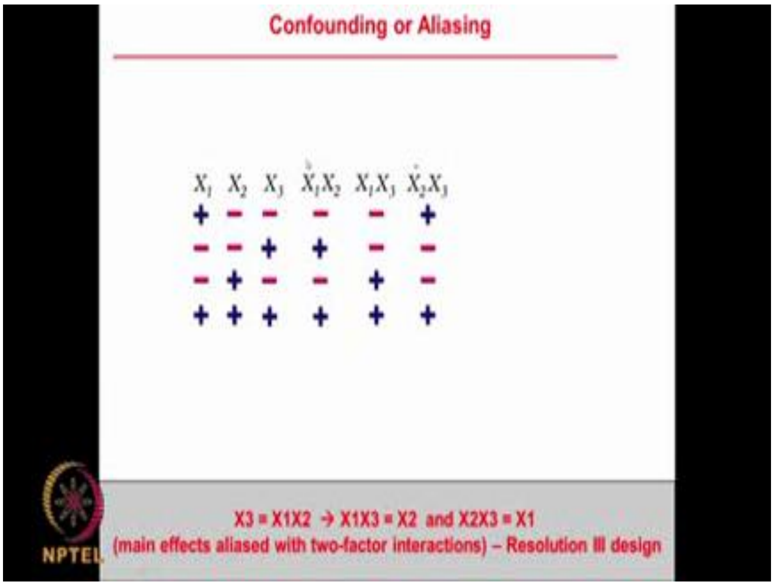
(Refer Slide Time: 03:58)

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

Yesterday I talked about it. Suppose, I have 3 factor A, B, C. I have to do 2^3. That is 8 experiments. So A is given like this, B is given like this, C is given like this. I also talked about the two things, the orthoginality and symmetry. That means you have 4 -, means you have 4 +. If we have say 4 - means, you will have 4 +. If we have -, you have 4 +. Even the interactions will be symmetry and that is why it is called orthogonal. It is called orthogonal. This is very symmetric design, Now, even the A, B, C will be symmetric.

Now, the question is we have 4 parameters or 4 factors, ABCD, ABCD. So, I do not want do a 2^4 experiment. 2^4 is 16 experiments, but can I do it in 8 experiments? That is a half factorial of 2^4, half of 2^4; that is 8 experiments. So, where do I put my d? So, I put my d here. This is the highest order interaction and as I have been telling that 3 way interactions are very rare. So, can I put d? That means this will correspond to how you change D. Then, what happens just like AB, BC, AC, we can also AD, BD, CD and if we do that interestingly, they all become like this. They all become like this. So, each main effect is interacting with the 3 way interaction. So, A is interacting with BCD, B is interacting with ACD, C is interacting with ABD and these variables are also interacting with each other, ok. CD interacting with AB, BD interacting with AC, AD interacting with BC and so on actually. Actually as I said this is a half of 2^4design or $\mathbf{2^{4-1}}$ fractional factorial. This tells you number of parameters. This tells you whether it is half or one fourth and so on. If it is one fourth, I would have put 2 here.

(Refer Slide Time: 06:14)



Let us go back to another problem. Suppose I have X_1, X_2, X_3, then 2^3 will be 8 experiments, but I do not want to do 8 experiments. So, I am doing only 4 experiments. That is this how do you write this, 2^{3-1} design because it is half of 2^{3-1}. Instead of 8 experiments, I am doing only 4 experiment. So, first we will look at this symmetry 2 +, 2 - the, the orthogonality; even the first order interactions of 2 +, 2 -. So, this is a good design. So, we have taken this half of 2^3 or 2^{3-1} design. Instead of doing 8 experiments for A 3 variables, namely 2^3, we are doing only 4 experiments.

If we look at this and this that is a confounding between X_3 and X_1, X_2, they are exactly the same. Similarly, if we look at X_2 and X_1, X_3 they are exactly the same. So there is a confounding between X_2 and X_1, X_3. Similarly, if we look here and there plus plus, minus minus, minus minus, plus plus, so there is a confounding between X_1 and X_2, X_3. So, main effects that is X_1, X_2, X_3 are called main effects are aliased with 2 factor interactions. Main effects are aliased with 2 factor interaction. So, I am using a new word alias. Main effects are aliased with, so there is a confounding of main effect with 2 factor interactions main effects (Refer Time: 08:05). So, this is called Resolution III design. This is called Resolution III design. Main effects are confounding with 2 factor interactions, main effects are confounding with 2 factor interactions as you can see here and this is called a Resolution III design. How do you decide? Look at the number of terms here. So, 3 terms, so we can call this as Resolution III design. Main effects are confounding with 2 factor interaction. So it is if we can do this experiment, but if there are any changes, for example, you cannot differentiate whether it is because of a main effect or the 2 way interactions. Do you understand?

For example, if I take these and this. Suppose, this is temperature, this is pH, this is a carbon amount. So, if I see a change because of carbon amount, I see a change in the yield. I cannot say whether it is because of the carbon amount or because of the 2 factor interaction that is temperature versus p h, because they are aliased. There is a confounding. Do you understand? So, I will not be able to tell whether it is because of the change in carbon. I am seeing some changes in my yield or is it because of the confounding of X_3 with X_1, X_2 which is temperature p * pH. So, a temperature into pH may also give the same changes. I will not be able to really differentiate. So because they are confounded, there is aliased and this is called a Resolution III design. Generally resolutions III designs are not liked because the main effects are of course are very useful. So, we do not want, but if the main effects are aliased with 3 factor interactions, then it is ok. Like in our previous problem, we have the main effects aliased with 3 factor interactions and then, but there is an interaction between the 2 factors, 2 factor interaction.

Now, this design is called a Resolution IV design because here ABC is equal to D. So, this is called Resolution IV design. This is called Resolution III design because main effects are interacting with 2 factor interaction. Here main effect is confounded with 3 factor, 3 factor interactions, whereas in this case the main effects are confounded with 2 factor interactions. So, this is Resolution III design, this is a Resolution IV design. Resolution IV designs are, but Resolution III designs are not good at all. So, in Resolution IV design, what happens? Main effects are confounded with three way interactions, I mean 3 factor interactions. So, generally 3 factor interactions are very rare scenario. So, if there will be change, we can say it is because of the main effects, but then of course there are interactions, there are confounding between the two-way interactions confounding between the two-way interactions. So, I will not be able to say whether something is because of AB or because of CD, something is because of AC or because of BD. So, they are aliasing. So, there is a confounding between these two. So, in a Resolution IV design 2, 2 factor interactions are confounded, main effects are confounded with the three-way interactions, 3 factor interactions, whereas in Resolution

III design, the main effects are confounded with 2 factor interaction and Resolution III design is not a good idea to derive, ok.

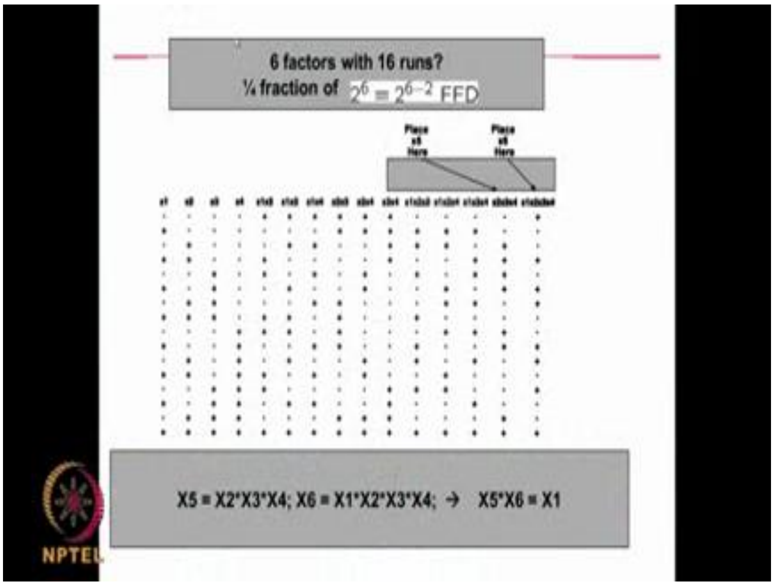
(Refer Slide Time: 11:45)

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

Let us look at same problem. We looked at this A, B, C problem. So, it is A 2 bar 3, so 8 experiments. So, what I did was, I put D here. So, I made it into A 2^{4-1} design. Now, I want a put E also, and then it becomes one fourth of 2^5 that is 2^{5-2} here. So, if I put E that is 5 factors, ideally I should be doing 32 experiments, but I want to do only 8 experiments. Instead of doing 32 experiments, I want do only 8 experiments. So, that is one fourth of 2^5 or 2^{5-2}, this is what it is called 2^{5-2} design. So, I want to put D. I have put D in A under A, B, C but if I put E under this, the main effect E is confounding with the two way interactions, and this is a Resolution III design as I said before, and this is not generally liked. So, if I have 5 factors, it is not a good idea to do a 1/4 2^5 design because I end up with some two way interactions here. This two way factor interaction, interacting confounded with the main effect. So it is confounded with the main effects. So, this is called a Resolution III design, whereas the corresponding if you look here, this is Resolution IV design. So, do you understand this picture, and picture, this picture and this picture? So instead of 32 experiments, I am thinking of doing one fourth of 32, that is 2^{5-2}. So this is a Resolution III design because in the main effects are confounded with the 2 factor interactions, whereas here instead of doing with the 4 parameter, instead of doing 2^4 16 experiments, I am doing half of 16 that is 2^{4-1}. Here the main effects are confounded with the 3 factors interaction and this is a Resolution IV design. Generally

Resolution IV designs are ok. Resolution III designs are definitely not ok, ok. Let us go forward further.

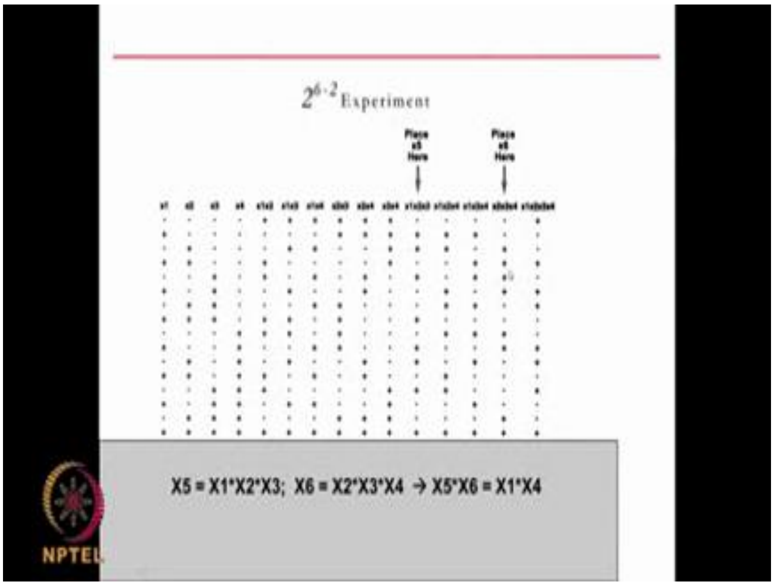
(Refer Slide Time: 14:01)



Let us look at 6 factors. 6 factors is what? 2^6. That means, 2 * 2 * 2 * 2 * 2, 64 experiments, ok, but I want to do only one fourth of that 16 experiments. So, it is 2^{6-2}. So, quarter fraction of 2^6 is 2^{6-2} fractional factorial design. Now, we have X_1, X_2, X_3, X_4. So, you start with the 2 full factorial for 2^4 because that is 16 experiments. 2^4 is 2 * 2 * 2. So, you put X_1, X_2, X_3 and X_4 and then, write the all the interactions. You are going to have two way interactions; you are also going to have three way interactions. X_1, X_2, X_3. X_1, X_2, X_4. X_1, X_3, X_4. X_2, X_3, X_4. X_1, X_2, X_4. So, that is how you first make. I want to do only 16 experiments. So, what gives me 16 experiments? 2^4, that is 4 parameters. So, I write the full.

Now, I have to introduce 2 more variables or parameters. I have to introduce 2 more effects or variables or parameters in order to make it 6. So, where do I introduce? Do I introduce here? Do I introduce here? Do I introduce here? Here? Here? So, you have a lot of choices. Suppose, I introduce X_6 here and X_5 here, so X_5 is equal to X_2, X_3, X_4 and $X_6 = X_1 X_2 X_3 X_4$. So, if I multiply X_5 by X_6, what will happen? I will be left with X_1. X_2, X_2 will become 1. X_3, X_3 will become 1. X_4, X_4 will become 1. So X_5, X_6 has become X_1. This is not desired. As I have been telling always a main effect should not be confounded with two-way interaction. So, the selection of putting here and here is not a good idea.

(Refer Slide Time: 16:02)



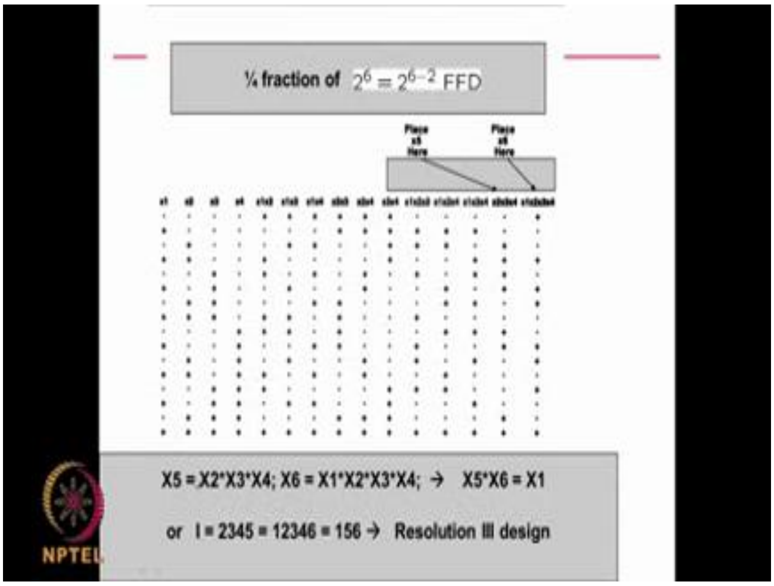
So, I will look at some other place. I am looking at this place instead of this, and this I am looking at this place. So, here X_1, X_2, X_3 is X_5. X_2, X_3, X_4 is X_6. So, when I multiply X_5 and X_6, what happens? I have X_1, X_2 into X_2 is 1. X_3 into X_3 is 1. X_4 comes here. So, X_5, X_6 is X_1. X_4 that is a two way interaction or 2 factor interaction, it is confounded with another 2 factor interactions. So, no problem; this is allowed. This is a Resolution IV design. No problem, whereas this becomes a Resolution III design. So, it is a good idea to put your X_5 here, it is a good idea to put your X 6 here. You understand? Do you understand how to go about creating such a design? I want to have 6 parameters, but I want to do only 16 experiments. So, 16 experiments means it is 2^4 for a full factorial. 2^4 is 2 * 2 * 2 * 2 will give you 16, 4 parameters X_1, X_2, X_3, X_4. So, I write down the whole design. Whole design means I will have a column for X_1, column for X_2, column for X_3, column for X_4 and then I will say -, -, -, -. Then, I will say +, -, -, -. Then, I will put -, +, -, -. Like that I will build up. Then how do I generate the two way interaction? It is simple. $X_1 * X_2$ means -, - is +. +, - is -. -, + is -. +, + is +. -, - is +. Like that I build X_1, X_3. I will multiply X_1, * X_3 -, - is +. Like that I build this. Like that I build all the two way interactions.

Then, I go to three way X_1, X_2, X_3. So -, -, - will give you -. Like that I build up all the three way. Then, finally I get the three way interaction. So this is the entire design table with the interactions for a 4 parameter. Now, I want to introduce 2 more parameters, X_5 and X_6 because I want to do a one fourth design of 2^6. So, how do I represent 2^{6-2} or one fourth fraction of 6^6? The question is where do I introduce these 2 variables? You know X_5 here and X_6 here, or I have so many choices here, here because a main effect can be always confounded with three way interactions, 3 parameter interactions, 3 factor interaction. So, when I do this, $X_5 = X_2, X_3, X_4$ and $X_6 = X_1, X_2, X_3, X_4$. What happens when I multiply X_5, X_6. I will have $X_2 * X_2$ is 1. $X_3 * X_3$ is 1. $X_4 * X_4$ is 1. So, I will be left with X 1. So, X_5, X_6 is confounded with the X_1. This is a Resolution III design which is not desired at all because the main effect is confounded with 2 factor interaction. So, I need to look at some other place, where to put my X_5 and X_6. So, I think of this place and this place. So, when I do this, when I do that, $X_5 = X_1, X_2, X_3$. X_6 is $= X_2, X_3, X_4$. So when I multiply X_5, X_6 you will get X_1, X_2, X_2, 1. X_3, X_3 will become 1. So, X_1 and X_4. So, you have a 2 factors confounded with each other which is this is a Resolution IV design. This is allowed. So, I can put X_5 here, I can put X_6 here. What about these two places? Can I put X_5, X_6? Why do not you take this as your homework and try to work out now because you have different places where you can put here, X_5, X_6 here, on here or here and here. So, I want you to try and see whether you end up with Resolution III or Resolution IV design, whether it is to allow, do you understand how to go over doing this type of approach we talked about. We introduced some terminologies like aliased X, Y is aliasing with X_1, X_2, X_3 or X_5 is confounded with the X_1, X_2, X_3. So, this is confounded with this and so aliasing. So, you understand the aliasing relationship, so 1, 2, 3 and 5.

(Refer Slide Time: 20:38)

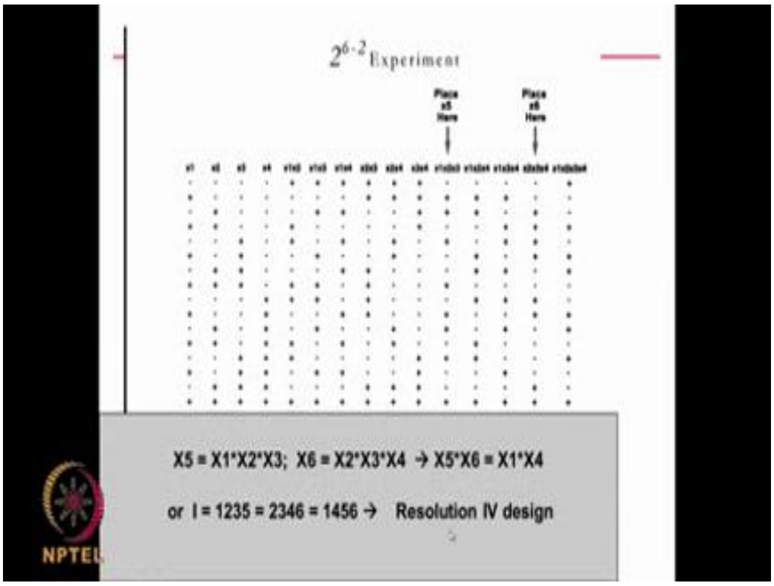
So, this is called the Design Generator 1235 aliased with 2346 or we can say 2346 or 1456. Sorry, yeah1456. So, the main effect 1 is confounded with 2351 is confounded with 235 which is confounded with 456, which is confounded with 2346; 2 is confounded with 135, 2 is confounded with 346 also confounded with 1456; 3 is confounded with 125 confounded with 246 is confounded with 1456. Like that we can build up, right. 4 is confounded with 12, 35 or 236 or 156, 5 is confounded with 123 or 2346 or 146 and these are the 2 factor interactions which may be confounded. 12 is confounded with 35, 13 is confounded with 25, 23 is confounded with 15 and so on actually, right. So, this a very useful relation to build up initially and then, with that we can build up everything. Do you understand? This is a IV, Resolution of design of IV because these you have 3 plus 1, right, the minimum. This is how do the aliasing relationship. Do you understand? So, we take up 1235 as I12346 as I 1456 as equal to I right here and then, we build up how they are going to be related. 1 will be remaining is 2352 will be 346 or 135 like that. Do you understand and then, how to build up the 2 factor interactions also is given here.

(Refer Slide Time: 22:39)



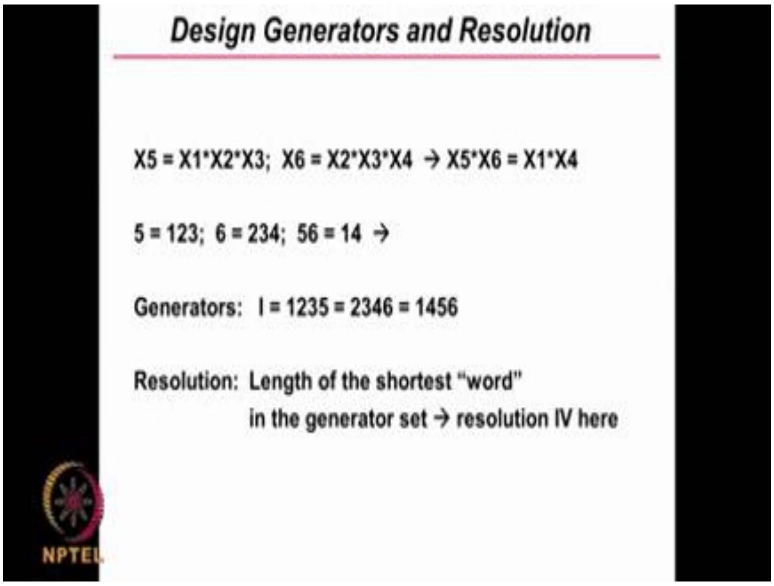
So, as originally we said in the previous picture when we put here, this becomes a Resolution III design because we have this type of relation. Now, what is the aliasing relation? I = 2345 which = 12346 which = 156 just like in this previous case, it is like this. So, we can build up all the relationship here. No problem, right. So, 2 = 345 or 2 = 1345, 1346 I like that. So, here comes 1 = 56 or 1 = 2346, 1 = 2345 and so on actually, and 2 =3452 = 1346 or 156. Like that we can build up the I relationship. Also that is the aliasing relationship. This is the Resolution III design because we have 3 minimum comes out to be 3 of that, ok.

(Refer Slide Time: 23:50)



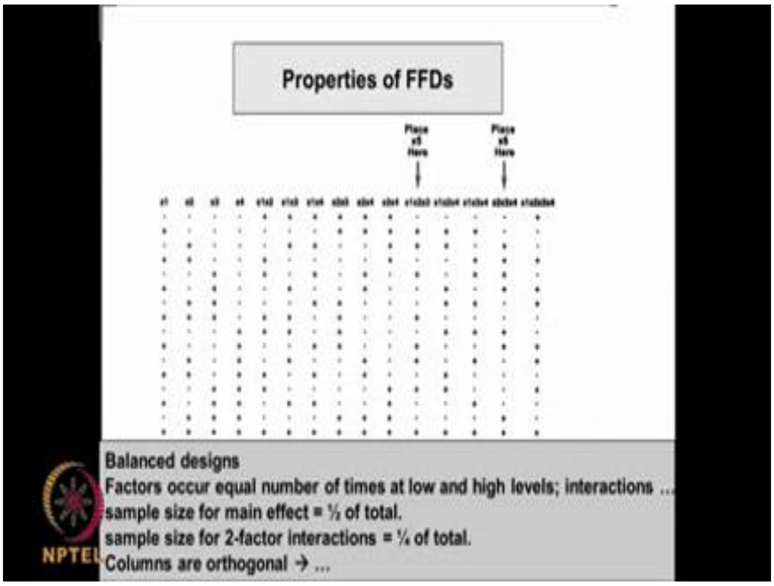
Now, the other one where when we selected this and this, $X_5 = X_1, X_2, X_3, X_6$ = to X_2, X_2, X_4, so your I = 1235 or = 2346 or = 1456. Do you understand? So, 1 will be confounded with 235 or 2346 or 456. This is of Resolution IV design because we have relationship involving minimum comes out to be 4 parameters. So, you understand the concept of how to introduce a new parameter if you are doing a fractional factorial design, how to identify the resolution of that design in the factorial and how to look at the aliasing relationship, ok.

(Refer Slide Time: 24:55)



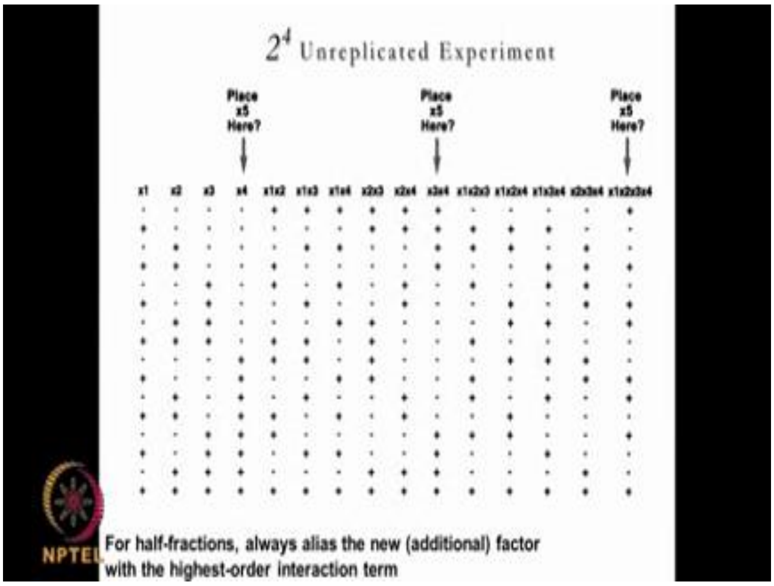
So, this X_5, X_1, X_2, this one this is called a Design Generator because we can look at all the combinations, right. I = 1235, 2346, 1456 and so on actually, right. So, 5 equal to 1235 =2346 = 1456 like that you know. So, this is called a Design Generator, and as I said resolution is length of the shortest "word" in the generator set, this is resolution IV whereas if you look at this I; you have X_5, X_6, = 1. So, it is the shortest now. 1 = 56, so obviously resolution III. Do you understand how to build up these resolutions in this, ok?

(Refer Slide Time: 25:23)



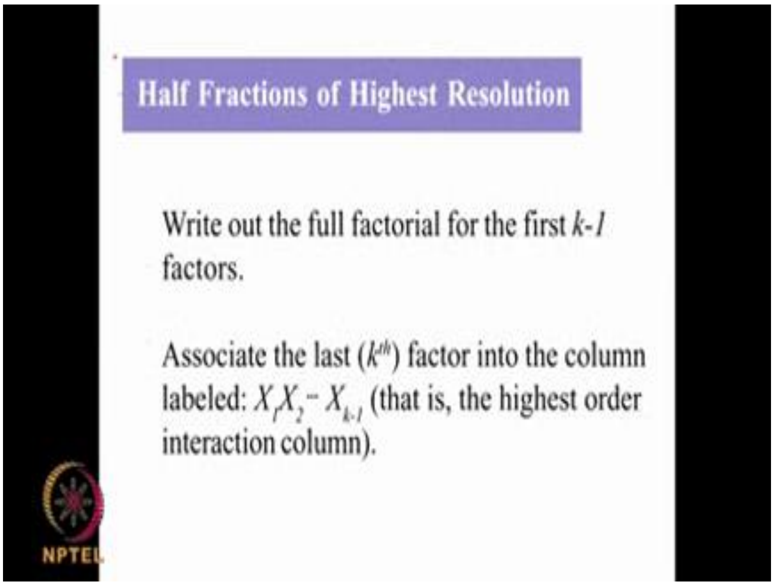
The properties of Fractional Factorial Design just like your Full Factorial Design or Fractional, it has to be balanced, ok. All the plus in the main effect should be = all the minuses and it should have orthogonal. 1 is balance orthogonal. All the interactions, also the pluses = minus. So, factors occurs equal number of times at low and high levels. Do you interactions also same thing happen, otherwise if you do not have like that, if you have more pluses than minus, then it is biased. That means you would like to do more experiments at the higher level than lower level. So, it is not a really balanced design. That is not very so sample size for main effect. We will take half of total sampled size for 2 factor interactions, one fourth. So, these are properties. Once you make your design, check cover in case you made any mistake, whether the design is balanced, whether it is orthogonal, whether the resolution is more than 3 and so on. Actually we need to understand these properties.

(Refer Slide Time: 26:32)



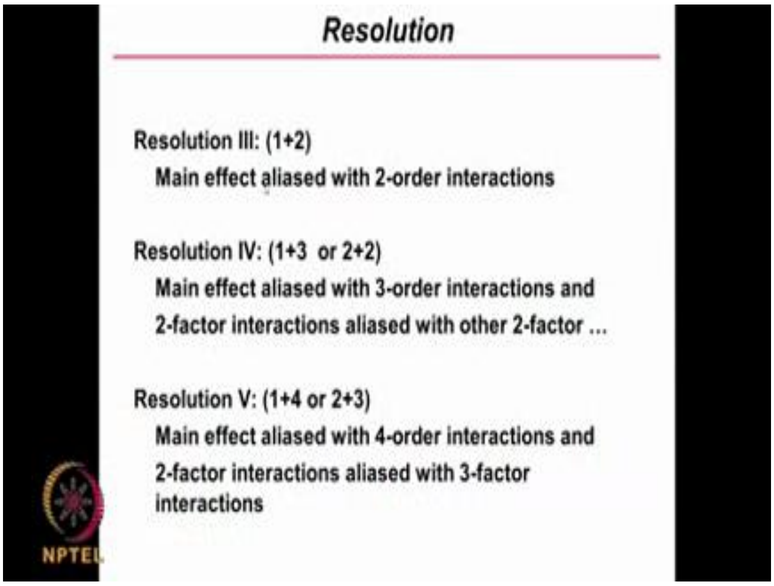
Another thing is for half fractions that always alias the new factor with the highest order interaction terms. So, the question is, suppose I have a 2^4 design. That means 16 experiments 2^4, 16 experiments. Now, I want to place 5, the 5th factor that is that means half of 2^5; that is ½ of 32, that is it is called 2^{5-1}. Do you understand where should I place? The best place of course is here. You cannot place it here because it will become a Resolution III design. You can try it placing, but then you may have problem. So, best place is to always alias the new factor with the highest order interaction term. Here you are safe, you are very safe. So, x1, x2, x3, x4 = x5. That is how your equation will look like, the I equation will look like, ok. Aliasing equation will look like.

(Refer Slide Time: 27:38)



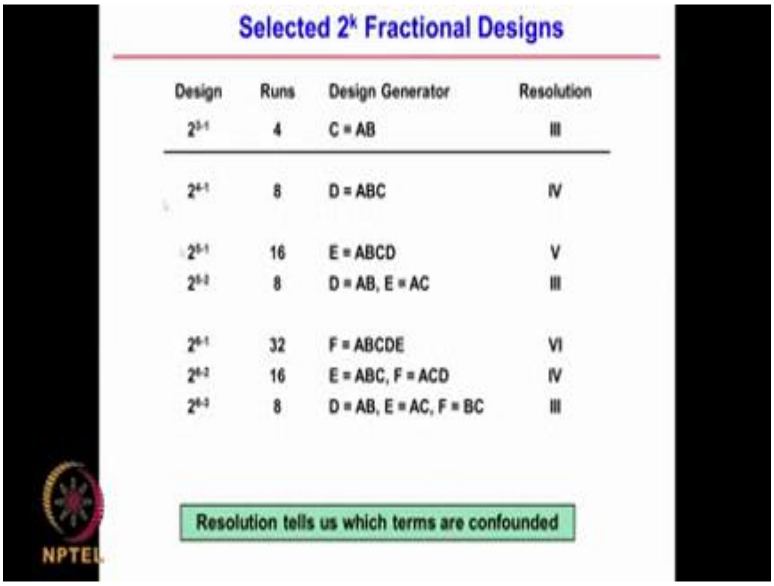
Do you understand? So, write out the full factorial for the first k - 1 factors for a ½ fraction. So, if I am doing a 2^4 1/2 of 2^4, I will write 2^3 full factorial design. If I am writing half, if I am interested in half of 2^5, then I will write the 2^4 design fully because ½ of 2^5 is 2^4. So, I write the 2^4 fully. So, first I will write the full design for the k - 1 factor and then, associate the last k the factor into the column that is the highest order interaction and put the new variable there. That is always good. So, half fractional is very easy to do. It is not very difficult. One fourth fraction as I showed in the example especially if I want to add two variables, where to put it is little bit tricky, but when we are doing half fraction, it is very simple. What you do is, you take a design which is number of factors minus 1 and then, you put your new variable in the highest order interaction. That is how you do have actually. So, now you understood what is how to estimate resolution and how to look at the concept of the design generator, the aliasing equation?

(Refer Slide Time: 29:03)



So, Resolution III design as is said has main effect interacting aliased with 2 order interactions Resolution IV design, main effect is aliased or confounded with only 3 order and 3 order interaction, and 2 factor interactions are aliased with each 2 factor interactions. That means AB = CD and so on. Resolution V design main effects are aliased with only 4th order interaction, that is 4 factor interaction and 2 factors are aliased with 3 factor interactions. This is very good resolution V at the number of experiments go up. So, generally this is Resolution IV design is, but this is generally not ok. This is how we go about thinking process.

(Refer Slide Time: 29:48)



So, let us again come back. Suppose I have $\mathbf{2^{3-1}}$ that means there are 4 runs. Instead of doing 8 runs for a 3 factor, instead of doing 8 runs, I want to do only 4 run; that is ½ of $\mathbf{2^3}$. Design generated is C is equal to AB. So, I have A, B, then I have AB. Under that AB column, I will put it as C, this III resolution design. Now, let us go to the next one. I have 4 parameters, 4 factors. So, I want to make a half of that. That means I want to make ½ of 16 experiments that is $\mathbf{2^{4-1}}$. So, I first take that 2^3 design, 2^3 design is 3, 2 * 2 * 2 that is 8 and then, I introduce a new variable D under the column of ABC. This is my Design Generator. I is equal to ABCD. This is Resolution IV design. Design generator I is equal to ABCD and this is a 4 design because I have 4 terms here. This is ok.

Now, let us go the next one. 2^5, that means 5 parameters. So, I have 32 experiments, but I want do only 1/2 of that. So, 2^{5-1}. So, I am doing only 16 experiments. So, I introduce the new variable E under ABCD column. So, this is the Resolution V design. Understand Resolution V design, but if you want to do one fourth of 32 experiment, that is 2^{5-2} that is 8 runs, then I end up having D equal to AB, E equal to AC. Then, that becomes Resolution III design. This is not good. If you go for 6 parameter, 2^6 maximum will be 2 * 2 * 2 * 2 * 2. That means 64 experiments. If I want to do half of that 2^{6-1} design, that is 32 experiments, so how do I introduce that 32? That is 2^5. I introduce F, the 6th parameter under this. So, this is a resolution 6th design because you have 6 terms here. If I do one fourth of that 32 experiments, that is 16 experiments 2^{6-2}, what do I do? I put E as ABC, F as ACD. That is the Resolution IV design because 4 terms, but if I make instead of that, instead of doing 64 experiments, I am doing 8 experiments. 1/8 2^{6-3}, then I end up D = AB, E = AC. F = BC. That is a Resolution III design. So, Design Generator here is I = ABD = ACE = BCF. This is a Resolution III design because we have these types of interactions. The main effects are confounded or aliased with 2 factor interactions. That is a resolution. Do you understand how it looks like for a 2 level designs? So, it will, we will continue further in the next class on a different types designs.

Thank you very much.

Key words - Fractional Factorial design, Full factoraial, Confounding, Aliasing, Aliasing Relationships, Design Generators, Resolution, Balanced designs, Selected 2^k Fractional Designs