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

Lecture - 36
Other Designs

Welcome to the course on Biostatistics and Design of Experiments. We will look at more designs and how one goes about performing these design strategies.

(Refer Slide Time: 00:22)

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

So, we talked about Full Factorial design, and as you know, Full Factorial is represented by 2^n, n is the number of factors; 2 here denotes the number of levels. Of course, you can have 3^n, 4^n, and so on actually. Initially, when you are doing a screening design, that means, when you are looking at large number of factors, we generally do it at 2 levels only. So, a full factorial design may have, if you have 3 parameters, 2^3, that means, 2 * 2 * 2, that is 8 experiments, and so on actually; 2^4, 4 factors, that will be 16 experiments, and so on. So, the advantage is that we can look at the interactions; like, if you have say three factors, 2^3 design, 8 experiments, we can look at interaction AB, AC, BC. In addition, we can also look at the three-way interaction; that is A, B, C. But, that is also a disadvantage of this full factorial design, because, suppose you look at a 2^4 design; not only look at 4 factors, the main effects, you are looking at 6 interactions; that is, AB, AC, AD, BC, BD, and CD - 6 interactions; but you will look at three-way interactions like, ABC, ACD, BCD and so on. And then, you will also look at four-way; all these are very superfluous, because I did mention that, these three-ways and four-way interactions are very, very rare.

You generally have two-way only. So came into being, something called the fractional factorial designs, which are like, ½ of 2^n, that is 2^{n-1}, or 1/4 of 2^n, that is 2^{n-2}, and so on. See, these fractional factorial designs reduce the number of experiments dramatically. And, we can also get interactions, but there are some disadvantages also in fractional factorial, because, unless you choose the number of experiments and where to put in your new factor, you may end up having something called confounding; I mean, sometimes, a two-way interaction may be confounding, and sometimes a one-way interaction may be confounding with three-way interactions and so on. So, the fractional factorial design…

(Refer Slide Time: 02:43).

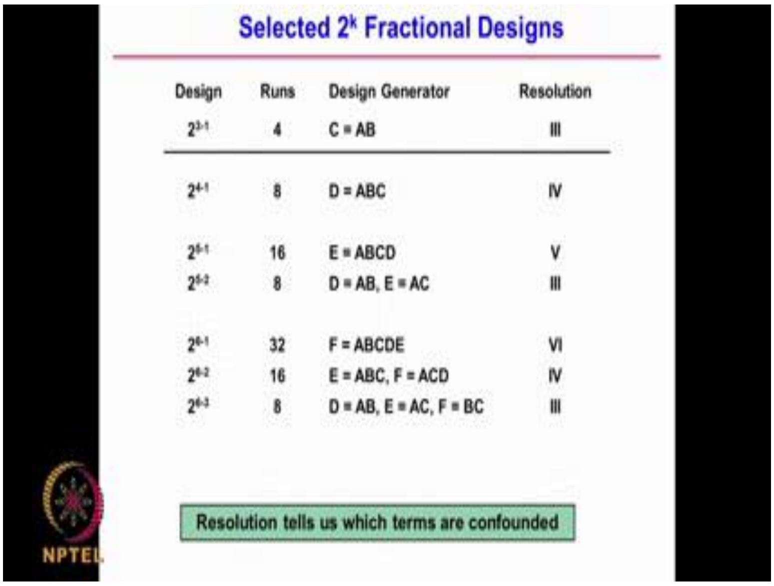
Run	BCD	ACD	ABD	CD	BD	AD	D
	A	B	C	AB	AC	BC	ABC
1	-1	-1	-1	+1	+1	+1	-1
2	-1	-1	+1	+1	-1	-1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	-1	+1	+1	-1	-1	+1	-1
5	+1	-1	-1	-1	-1	+1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	+1	+1	-1	+1	-1	-1	-1
8	+1	+1	+1	+1	+1	+1	+1

Let us look at this, I showed you yesterday also. Imagine I have 3 variables, A, B, C; so, 2^3, 8 experiments; everything is ok. Now, I want introduce another variable D, with these 8 experiments; so, it is a ½ of 2^3, or 2^3, 2^{4-1}; instead of ½ of 2^3, it is actually 2^{4-1} design.

So, you are doing only 8 experiments; so, I put D here, ABC. So, what happens in this, there is an interaction, confounding of A with the BCD, B with ACD, C with ABD; confounding, we have to use the correct word. The two-way interactions are also confounded; AB in confounding with CD, AC confounding with BD, BC confounding with AD, and D confounding with ABC. And, this confounding, and this is called a Resolution IV design, because, we are having D confounded with ABC. And, so, the main effects are confounded only with three-way interactions, and two-way interactions are confounded.

Now, you want to introduce one more variable, fifth variable; that means, 2^5 design; that is 32; but you want to do only 8 experiments, that means, 1/4 of 32, which is 2^{5-2}. So, the question is, we put E here, what happens, we have main effects confounded with two-way interactions; main effect confounded with this, which is called a Resolution III design, because E = CD, right, or E = AB; that is called a Resolution III design, which is generally not liked; Resolution IV designs are ok, but Resolution III designs are not ok, because we find lot of confounding.

(Refer Slide Time: 04:46).



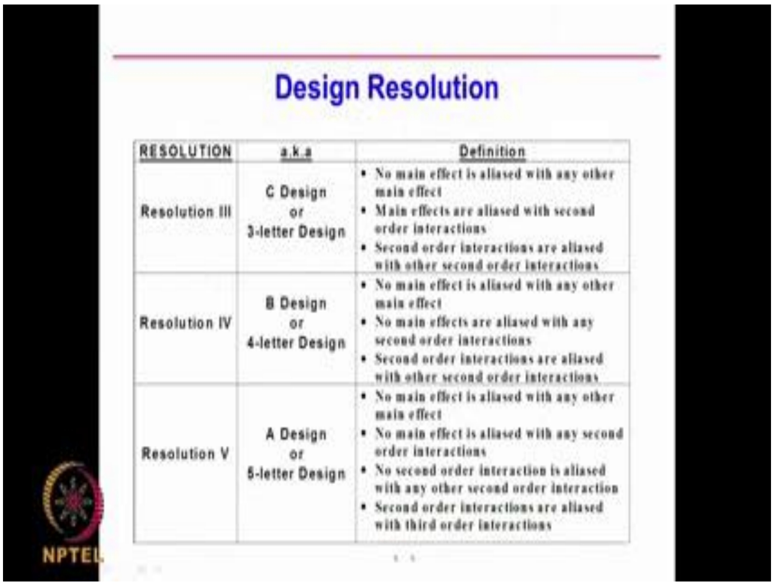
So, this table gives you nice idea. So, suppose I have 3 parameters or 3 factors, so you have, a full factorial will be 2^3. I want to do ½ of this; so, 2^{3-1}, that is, 4 experiments.

So, 4 experiments, we have the AB; instead of AB you put it as C. So, this is a Resolution III design; understand, how to get, understand the resolution which is not liked very much. Now, let us go to, the 4 parameter problem, ABCD. I want to do a half fractional factorial design. So, what I do is, I take the 2^3 model, I mean, sorry, 2^3 design table, and then, I will introduce ABC with D. So, such a design is called 2^{4-1}; 2^4 is 2 * 2 * 2 * 2, that is 16 experiments. So, the - 1 gives you 1/2 of that, 8 experiments. So here, the design generator is D = ABC. So, this is a Resolution IV design, because you have 4 variables; 1, 2, 3, 4; this is ok.

Let us look at another, one more variable, variable E; so, that is 2^5, that is 32 experiments. Suppose, if I want to do only half of that, 16 experiments, what will I do? E equal to ABCD; this is a Resolution V design, very nice. But, if I want to do one-fourth of 32, that is 8 experiments, 2^{5-2}, then, D will become AB, or E will become AC. So, this is a Resolution III design. The main effects are confounded with the two-way interactions which is not desired, and so on. So, now you know how to do a design generation.

And then, you can also calculate how to do a resolution calculation. So, many softwares will automatically do this. So, it is a good idea for you to understand it from fundas, so that, you can yourself generate the designs. You can even crosscheck the softwares. Most important thing is, you need to have balance and orthogonality into the term, and when you are talking about a fractional factorial design, you should have at least Resolution IV design, not Resolution III design. So, you do not want a main effect confounding with the two-way interactions; main effect confounding with three-way is ok, But, main effect confounding with two-way interactions, is a Resolution III design, and generally, that is not desired. So, this is, this is the entire logic of this. These are factorial designs. But then, there are many types of designs. There are designs like Taguchi designs. There are designs like Plackett-Burman design. All these are 2 level designs, and we call it as screening design. Let us look at more of them.

(Refer Slide Time: 07:54)



Same thing, I showed you yesterday, Resolution III design; main effects are aliased with the… main effects are not aliased with main effects; but main effects are aliased with second order interactions. So, that is bad; and second order interactions are aliased with other second order interaction. Resolution IV design, main effects are not aliased with main effects; main effects are aliased with the three-way interaction; and two-way interactions are aliased with two-way interaction. Second order interactions are aliased with second order interaction. Resolution V design, no main effect is aliased with any main effect, and no second order interactions are aliased; second order interactions are generally aliased with much higher, third order. So, Resolution V is great, and then, of course you need to do more experiments. So, generally, a Resolution IV is ok.

(Refer Slide Time: 08:51)

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

I = BCD = ABCE = ADE Design is of Resolution III.

So, as you can see here, we have the A, B, C, D, E type of situation. So, 5 parameters; 2^5, that is, 32, but we are doing only 8 experiments. So, this is a 2^{5-2} design, ¼ of 32 experiment. So, what happened? We put E as ABC, but we put D as BC; so, this is a Resolution III design; not liked; not good. So, ideally, if you want to do a 5 parameter, you must have at least 2^{5-1} type of experiment; that means, 16 experiments.

(Refer Slide Time: 09:32)

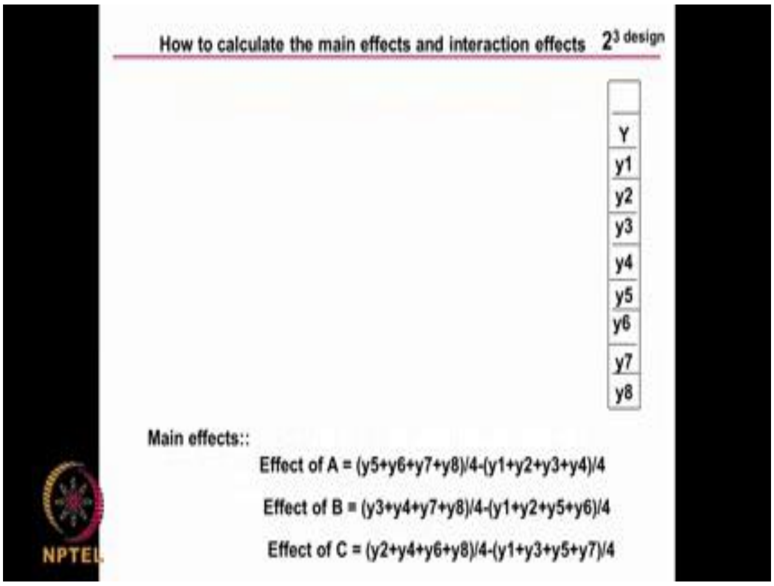
Run	A	B	C	AB	D = AC	E = BC	F = ABC
1	-	-	-	+	+	+	-
2	-	-	+	+	-	-	+
3	-	+	-	-	+	-	+
4	-	+	+	-	-	+	-
5	+	-	-	-	-	+	+
6	+	-	+	-	+	-	-
7	+	+	-	+	-	-	-
8	+	+	+	+	+	+	+

I = ACD = BCE = ABCF = ABDE = BDF = AEF = CDEF Design is of Resolution III.

.

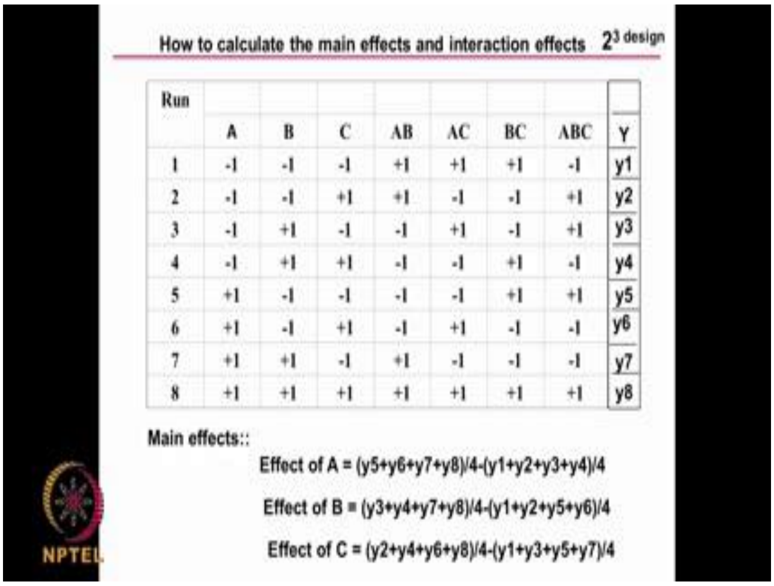
Same thing here. So, suppose I want to have more variables, say D, E, F, and I want to do only 8 experiments, this is a Resolution III design, as you can see; D = AC; E = BC; F = ABC. So, the I, that is the design generator, = ACD, BCE, ABCF, and so on, actually. Again, this is a Resolution III design. So, main effects will be confounded with two-way interaction. So, we will not know whether it is two-way, or it is a main effect.

(Refer Slide Time: 10:11)



.

(Refer Slide Time: 10:17)



Having done your experiments, imagine you have done, say 8 experiments, with 3 variables A, B, C, and you got some results. So, this could be, say temperature, this could be pH, this could be carbon amount; we have changed it at lower level, higher level; - 1 means lower level; + 1 means higher level. Imagine, I have done 2^3, 8 experiments; a full factorial, 3, 3 level, 3, 2 level, 3 factor, so 8 experiments and I got some results. Say, it is a biomass. Now, how do you identify the main effects? What is the effect of A? What is the effect of B? How do you do that? So, what you do is, these, these 4 experiments are done with higher level of A; these 4 experiments are done with lower level of A. So, what you do is, you add up these 4 outputs, ÷ 4, -, + these 4 outputs, ÷ 4. So, that tells you, effect of, main effect A, understand. Because, these 4 are +, these 4 are -. So, you add up these 4, to that. Now, if I want to know effect of B, I look at these pluses, y7, y8, y4, y 3; add up, ÷4, minus; minuses are y1, y2, y5 and y6; ÷4. This will tell you effect of B. Effect of C. So, you have 4 places where you have pluses, here, here, here, here. So, we take, y2, y4, y6, y8; ÷4, minus, 4 minuses here, here, here, here; so, y1, y3, y5, y7, ÷4.

So, this tells you the main effect. What is the effect of, say, temperature? What is the effect of pH? What is the effect of carbon? Now, how do we get the interactions? We can do that also. Now, ok. So, look… For example, if A is your temperature, B is your pH, AB gives you temperature into pH, and I told you how to calculate this, right? You know how to calculate this; multiply - 1 -1 which is + 1. So, - * - is +; + * - is +; - * + is +and + * + is +. So, - - is +; - - is +; - + is -; - + is -; + - is +; + - is -; + + is +; + + is +.

So, how do you calculate AB? You have 4 places plus; that is, y7, y8, y2, y1. So, y1 + y2 + y7 + y8 ÷4; and you have 4 places you have minus. So, y3, y4, y5, say y6, ÷ by 4. So, this result will give you the interaction effect AB; very simple. So, when you do a factorial plan, well planned out factorial design, the results, analysis also becomes very very simple. First, we have the results. Then, we can perform an anova. We can find out what are the main effects that are important. Then, if there is any interaction also we can calculate, using this; so, it is very nice.

Now, if I want to know AC interaction, for example, temperature and carbon amount, you know how to get this AC. You just have to multiply these 2, and get this; these 2, this; these 2, this; like that. So, for AC, there are 4 places you have plus; here, here, here; so,

Effect of AC = (y1+y3+y6+y8)/4-(y2+y4+y5+y7)/4

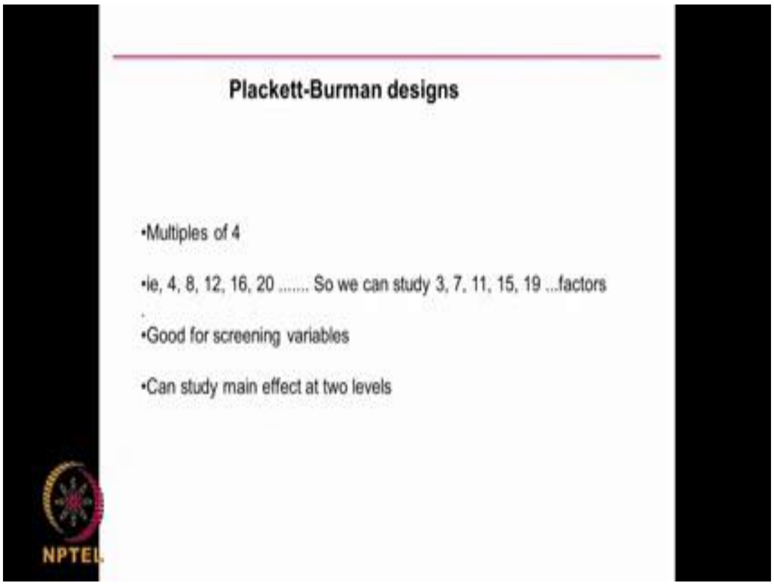
Same thing with BC; 4 places you have +, 4 places you have -. So, you pick up those pluses,

Effect of BC = (y1+y4+y5+y8)/4-(y2+y3+y6+y7)/4

. So, this gives you the interaction effect.

So, you can even get the ABC effect, because ABC also has, you know, the plus is here, the minus is here. Suppose, if you had made this as a D, that means, if you had done a 2 power 4 minus 1 design, that is, half of 16, and like I said, we can call this as D, all, although it is a Resolution III design, suppose, if you have done that D here, then you can get the main effect for D, by adding this, this, this and this, dividing by 4, and subtracting this, this, this, this and divide by 4. So, you get that, got the effect of D also. So, you see, the factorial tables are extremely useful. It tells you how to plan the experiments; it tells you how to do the analysis later, and do the various calculations. So, that is the main advantage of these type of factorial designs; so, it is very nice.

(Refer Slide Time: 15:48)

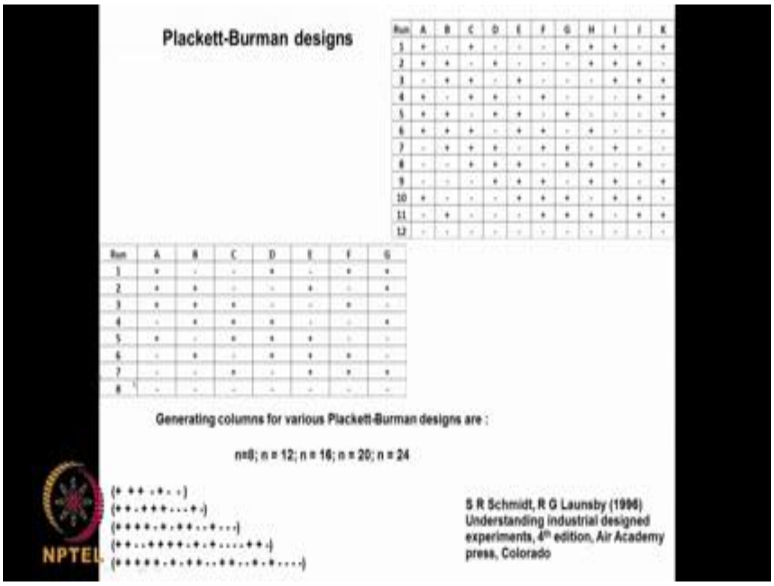


There is another design which is also a screening design, which is useful if you have lot of parameters, and you want to screen them quickly, these parameters, shortlist only a few important ones, and that is called a Plackett-Burman design, Plackett-Burman design. So, generally, it is a multiple of 4, 4, 8, 12,16; that means, you have to do either 4 experiments, or 8 experiments, 12 experiments, 16, like that. So, 3 factors can be studied with only 4 experiments. It will give you the main effects. 7 factors can be studied with the 8 experiments; it will give you the main effects. 11 factors can be studied with 12 experiments. So, it is extremely good; quickly we can screen out factors; and ideally, if we have say 3, 7, 11, 15, 19, all you have to do is 4 experiments, 8 experiments, 12 experiments, 16 experiments, and 20 experiments. So, it is extremely good for screening variables; large number of variables, we can study. So, we will study the main effects at 2 levels; minus 1 and plus 1.

There are, again, you have something called design generator, here also, in Plackett-

Burman. Let us look at those design generators.

(Refer Slide Time: 17:06)



This is the design generator; have a look at it. So, for 8 experiments, design generator is +, +, +, -, +, -, -. So, it gives you only 7, please note. So, what do you do? For variable A, we put +, +, +, -, +, -, -; and the 8th row is all -. Do you understand? The last row will always be minus. So, what do you do, for variable A, we put +, +, +, -, +, -, -. Now, how do we get these columns? What you do is, you take this and put it on top; and then, push everything down. So, take this, put it on top, and then, + will become + here, + will become +, + will become +, - will become -, + will become +, - will become -. And then, for C, take this and put it on top, and push everything one down; so this will become here, this will go here, this will go here, this will go here, this will go here, this will go here; of course, do not forget that 8th row, will always be all -.

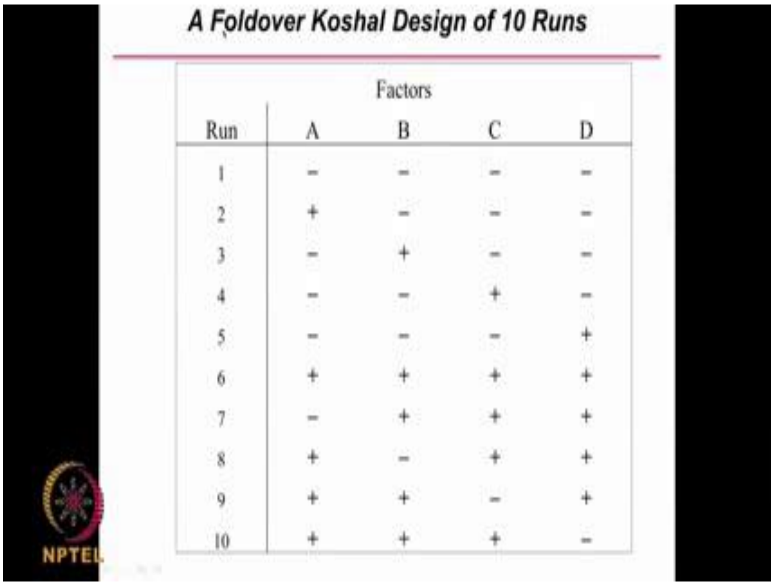
How do you get D? Take this and put it there, and then, push everything down 1; - -, - -, + +, + +, + +, -. E, push it; this + will come down, this, this, this; like that you build it up. So, you have a design generator for 8 parameters. It gives you 7 numbers, here; the last row will be all minuses. So, you put them here, in this column, and then, take the last one, 7th one; this is not, not the last one; this is ignored generally, always. So, you take this, and then, put it there, and then push everything down 1; take this, put it there; push everything down 1. So, this will come like this; take this, put it there; push everything down 1; take this, put it there; push everything down 1. As you can see here, so, you get this; and the 8th row will be all minuses.

Again, you can, you have to crosscheck, see, it is also a symmetry, balanced; 4 pluses, 4 minuses, 4 pluses, 4 minuses, 4 pluses, 4 minuses; very nice. So, I can screen 7 variables with 8 runs. It is very very good, minimum number of variable. What happened if I had 6 variables? What do I do? I can use the same design; I will call one of them as dummy variable. So, I will use the same, so, you can have one dummy variables. Sometimes, if you want 5 parameters, and I want to do 8 experiments, 5 parameters and 8 experiments, for example, factorial design is to be 2^5, that is 32 experiments, half factorial will be 16 experiment, but I want to do only 8 experiments. So, I take this table, and I will put two variables here dummy. so, I can do this. So, you see, Plackett-Burman is extremely powerful, in that sense. But, you can study only main effects, and you cannot study any interaction effects; and ideally, because it is in multiples of 4, ideally, if your parameters are closer to 4, then, it is very good; if not, we can always put dummy variables and like that.

Now, imagine, I have, I want to do 12 experiments; that means, 11 variables; 3, 4, 5, 6, 7, 8, 9, 10, 11 variables. So, I take this design generator, +, +, -, +, +, +, +, -, -, -, +, -. This last row, please remember, will be all -. So, how do I get B? I take this, put it here, and push all of them 1. So, this plus will come here, this plus will come here, this minus will come here, plus, plus, plus, minus, minus, minus, plus. Do not put anything on this 12th row, and so on. So, if I have 11 variables, all I have to do is a 12 experiment; whereas, if I want to do factorial design, 11 variables is 2^{11}, which is a huge number; even if I can of course look at one-fourth or one-eighth fractional factorial designs also, but, this gives you the minimum number of experimental runs, the Plackett-Burman designs, because, for 11 variables, I can just do 12 experiments, quickly screen, and eliminate many of them which I think is of no use at all. Then, I can go for a detailed design, later on; that is a beauty of Plackett-Burman design. And, these tables, I took it from this particular reference; this is a good reference to look at. It is quite simple, and I have been looking at it, actually.

So, you have different design generators, for, the n = 8 experiments, or 12 experiments, or 16 experiments, 20 experiments, 24 experiments, remember, that last row will, will be minus always. So, you take the last sign and put it in the next top of the next variable, and push everything down, and just insert it, that is it; that is how you make the, that is the design generator for the Plackett-Burman design. And, this design is extremely useful, if you are looking at screening designs; and, if it is multiples of 4, nothing like it, use a Plackett-Burman design; do not go for factorial or fractional factorial design; because, even if you use a fractional factorial design, you have some confounding. So, quickly do a Plackett-Burman, look at the main effects, remove some of the main effects, take only a limited number, and then, you do a detail study. That is, this is to do with Plackett- Burman design.

(Refer Slide Time: 23:13)

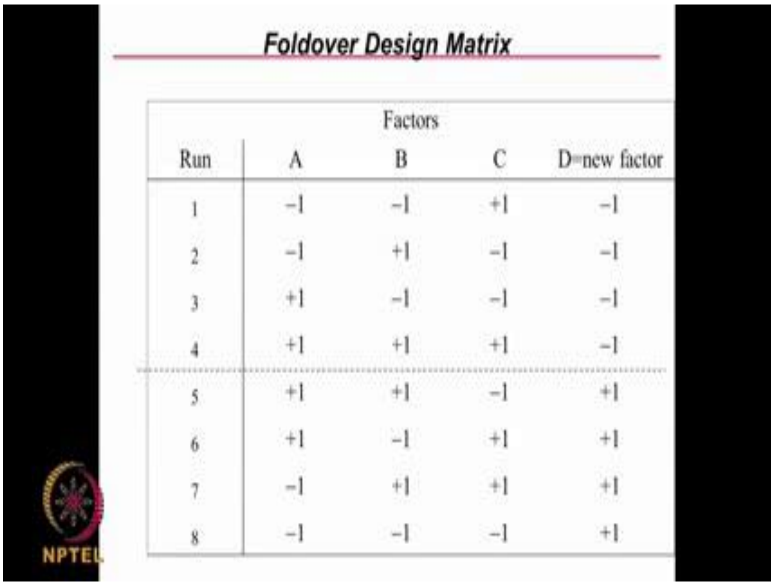


Run	Factors A	B	C	D
1	-	-	-	-
2	+	-	-	-
3	-	+	-	-
4	-	-	+	-
5	-	-	-	+
6	+	+	+	+
7	-	+	+	+
8	+	-	+	+
9	+	+	-	+
10	+	+	+	-

There is something called Foldover Koshal Design of 10 Runs. You have 4 variables, 4 variables, Plackett-Burman design, if I have to do, I can do a 4 experiments also. I can even do half of 2^4 factorial design, that is, 2^{4-1}, that is, 8 experiments, also we can do. This is called a Koshal Design. These are generated by different statisticians who do, who look at the importance of certain parameters, and then, they come up with the design. So, always remember, in any design, you will always have a balance. See, look at this, 5 pluses, 5 minuses, 5 pluses, 5 minuses; that is very important.

Any design you do, you are doing it by writing down by hand, always crosscheck whether balance is maintained, orthogonality is maintained. That is very important. It is called a Koshal Design. Let us not go too much into these type of designs. Plackett-Burman is very commonly used in softwares. The factorials are used. Then, of course, the Taguchi designs are also used.

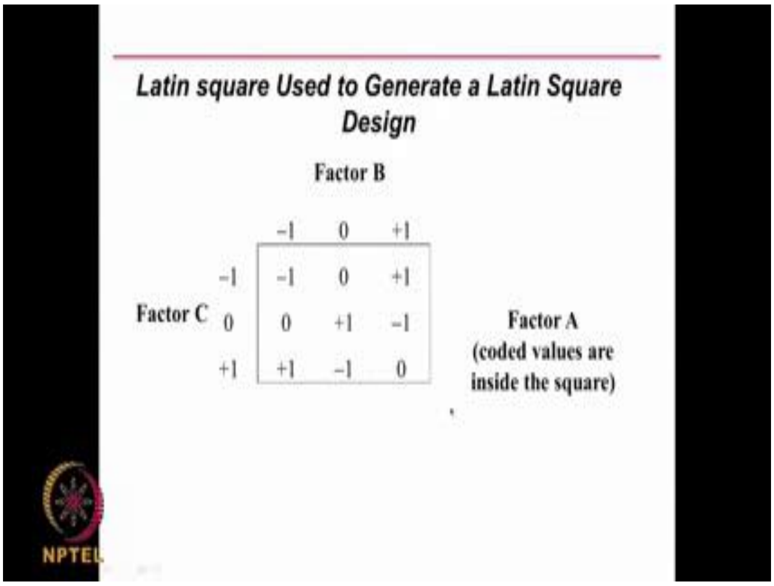
(Refer Slide Time: 24:36)



Run	A	B	C	D=new factor
1	−1	−1	+1	−1
2	−1	+1	−1	−1
3	+1	−1	−1	−1
4	+1	+1	+1	−1
5	+1	+1	−1	+1
6	+1	−1	+1	+1
7	−1	+1	+1	+1
8	−1	−1	−1	+1

See, you have Foldover designs for, if you are doing 8 experiments, previous one is 10 experiments, 4 factors. Here, you have 8 experiments, 4 factors. So, this is called a Foldover design; this is called a Foldover. Why it is called a Foldover? You can see, you can fold it, it will be exactly like a mirror image; you know this -, this -, this -, this -, +, this +, this +. So, that is why, it is called a Foldover design. So, ABC, you have, and then, D variable alone is a new variable you are adding, actually. So, if you look at only the ABC portion, it is exactly foldable; you know, this line, if you fold it, it is like a, the mirror image, right? That is why it is called a Foldover design. Then, you are adding a new parameter called D. So, if you have a ABCD Foldover design, then, we can add a new, new parameter E in this fashion. These are called Foldover design. Again, of course, do not forget the balance, balance is maintained, right; that is very very important.

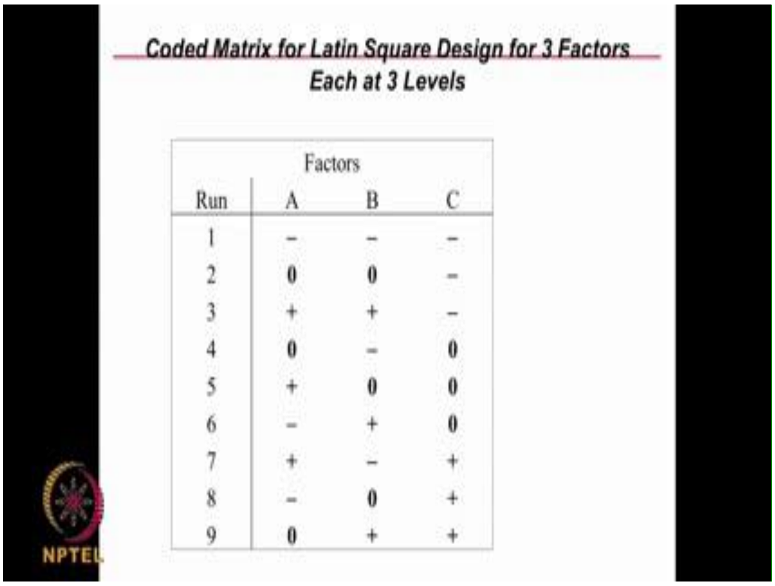
(Refer Slide Time: 25:43)



Then, we have the Latin Square Design. I talked about Latin Squares before also. Suppose, I have a Latin Square design, and we are looking at, say 3 levels, - 1, 0, 1; just like 2 levels, we call - 1 and + 1; if it is 3 level, generally, the methodology they suggest is, use 0, as the middle level. So, this is like, I am doing 9 experiments; 1, 2, 3, 4, 5, 6, 7, 8, 9, and this is Factor B, this is Factor C, and this is Factor A. 9 experiments you are doing; 3 * 3, 3 * 3, 9 experiments; look at it, nice symmetry. So, the first experiments will be all the 3, at the lower level. Second experiment will be C, will be at lower level; A and B in the middle level. The last and third experiment will be C, will be at a lowest level, and A and B will be the highest level, like that, you know. So, you do 9 experiments, and the results also we can write down here, and we can do some anovas. We did some problems on Latin Square Designs, anovas, with 3 factors also, if you remember, long time back. So, this is called Latin Square Design.

So, we have 3 parameters, 3 levels, we are doing only 9 experiments. You can see; otherwise, if you are doing a factorial design, 3^3, that is 27 experiments. If I want to do a one-third of that 3^3, that will be 9. So, it will be 3^{3-1}, type of experiment. So, this called a Latin Square, because, we have a square here, and then, we put in the factors, the levels of each of the factor, and in the middle also, we can put in some factors. If you remember, we did a problem, carbon, nitrogen and organism, for the production of biopolymers, 7 by 7 Latin Squares.

(Refer Slide Time: 27:51)



So, again, Latin Square for ABC, 9 experiments here, right. So, like I said, -, -, -, that is first experiment; the second experiment will be 0, 0, - 1; third experiment will be +, +, - 1, like that. This is how you do the experiment, 3 levels. Important to notice, each of the factors, there is a balance. 3 times you are doing at - 1 level, 3 times you are doing at 0 level, 3 times you are doing at plus level. So, it is very well balanced, as you can see here. Again, B also, we have 3 times at negative, 3 times at 0, 3 time at plus, plus. So, this is called the Latin Square and this is, we are writing down in the matrix form, it will be that.

(Refer Time: 28:55).

Foldover Design Matrix

Run	A	B	C	D=new factor
	Factors			
1	−1	−1	+1	−1
2	−1	+1	−1	−1
3	+1	−1	−1	−1
4	+1	+1	+1	−1
5	+1	+1	−1	+1
6	+1	−1	+1	+1
7	−1	+1	+1	+1
8	−1	−1	−1	+1

So, you do this experiment, and you get the results. So, if I want to know the effect of A, here, and like plus and minus, we also have 0. So, that means, we have effect of A at 3 places, +, 0, -. So, we can generate second order type of mathematical relation; whereas, if it is 2 levels, we will be able to generate only linear type of mathematical relationship. So, we will continue in the next class.

Thank you very much.

Key words - Other designs, Full factorials, 2-Level Designs, Selected 2^k Fractional Designs, Design Resolution, Plackett-Burman designs, A Foldover Koshal Design