Total Quality Management - II
Prof. Raghunandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

Lecture – 28
Two level Fractional Factorial Design - VI

A very good morning, good evening, good afternoon to all my dear friends and participants; welcome back to this TQM-II lecture on the NPTEL MOOC and we are continuing the discussion of the regression model and the factorial design for that and this is the 28th lecture so that means we have another 12 lectures to go and this is the penultimate week almost going on because in each week if you finish 5, you can calculate accordingly.

Now coming back to the regression model if you remember which I will repeat and please excuse me if I am just going through the same concepts again, but I am sure hearing that twice would not do you any harm, you will understand it much clearly. So, when you are trying to find out the beta hat, beta hat why it is a hat, because you are trying to estimate because you cannot find the exact value, if you found out the exact value; obviously putting that and multiplying beta 1 into X1, beta 2 into X2 would give you the exact value why that you are able to predict beta Y exactly, so obviously the error should be zero.

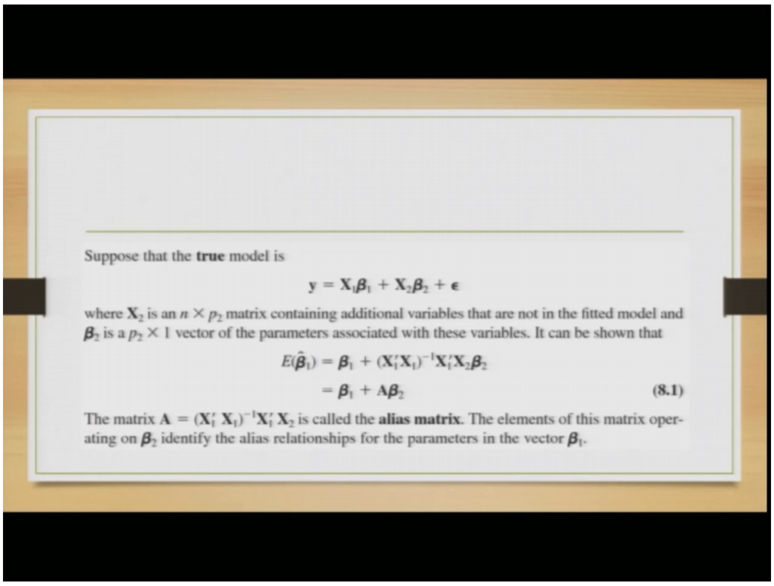
So, where it means technically the average of the error should is zero, but they would be in each and every day reading there would be some plus minus value. So, if you find out the average is zero, but in each reading it cannot be zero. So, hence it basically is a predicted value or the estimated value is beta hat. So, what you try to do technically is that try to minimize the error. So, now, the error is basically the difference between the actual value of Y and the predicted value of Y which is also equal to the epsilon error which you have said.

Now if you want to basically minimize the error, what you will try to do is conceptually you will try to minimize the variance. So, the variance of the errors would basically that you are trying to find out the best fit lines as the sum of the squares of the differences of the errors, so that means you find out the error in each reading square them up and add up the whole sum, so you are trying to minimize that.

When you do that what is actually you are doing is like this, when you put the Y values along with the predicted value of Y, which we consider as Y hat as we are doing it for beta hat, those beta terms would be there which are unknown, you have to find it out. So, if you want to find out the beta values, individually you have to basically partially differentiate the error square with respect to the each betas, put it to zero, find out the betas that gives you the beta hat and once you get the beta hat that formula which came out was basically X transpose into X that multiply them you will find out the inverse of that, multiply it by X transpose into the Y.

So, this transpose or Ys which are all talking about are all vectors and matrices, not scalars. So, let us continue with the discussion so true value is there are two values X1 and X2 which are independent one and the error is epsilon.

(Refer Slide Time: 03:30)

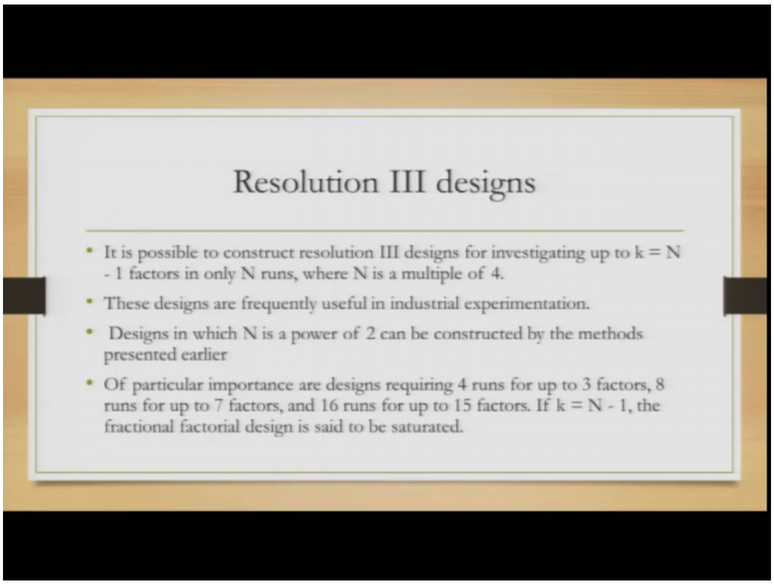


Now for the easier explanation of these models, we will consider the number of factors XP by the suffix. So, initially we did mention it was K then we said it was P and that suffix was L depending on the number of variables which are there. Now as there are only 2 independent variables, it will be P suffix 2. So, that suffix would denote the number of Xs which are there. So, let me continue with the discussion further.

So, where X suffix 2 is a matrix of size X into P2. So, P2 matrix contains additional variables that are not in the fitted model and the beta 2 is basically P2 cross 1 depending on what you think is basically the coefficients we want to find out. So, it can be shown that once you find out the expected value of beta hat, so obviously the expected value in beta hat in the long run in the asymptotic sense should be exactly equal to beta. So, when you try to find out the expected value of beta hat, it comes out to be beta 1 into A into beta 2, so these beta 2 are basically the external effects which are there.

So, the matrix A is called in this model the alias matrix and obviously A is given as X transpose multiplied by X1, I am not repeating the suffix, that when you multiply you find out the inverse multiplied by X transpose is X2, so it is called the alias matrix. The elements of the matrix operating on beta 2 identify the alias relationship which is existing for the parameters in the vector beta suffix 1.

(Refer Slide Time: 05:21)

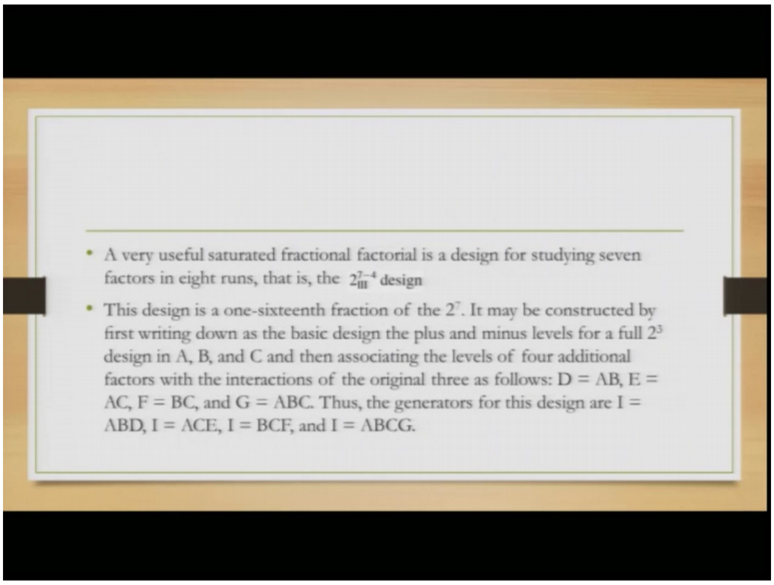


So, a resolution of 3 design models or consonant to the level of 3, it is possible to construct the resolution 3 design for investigating up to K which is N and minus 1 factors in only N runs, so obviously it will may be N minus 2, N minus 3 depending on the runs you are considering. These designs are frequently used in the industrial experiments when you want to predict or forecast. So, design in which N is a power of 2, so obviously, we are taking the 2 factor models only, so 2 to the power N or 2 to the power of N minus 1 so on and so forth can be constructed by the method present earlier.

So, what you have done is K minus P, so K would basically be combinations of the number of factors. Initially, we are basically taking A B C D then we combine to find out E F G, so which combination is you want to take to find out the best so called effects from the factors would be dictated accordingly. So, of particular importance of the designs requiring 4 runs for you each up to 3 factors are there, 8 runs for up to 7 factors, 16 runs for up to 15 factors so on and so forth.

So, if K is N minus one, so you remember that it was 2 to the power K minus P, so if K is N minus 1 the fractional factorial design is said to be saturated and because anymore addition of the number of factors would not give you any additional benefit in trying to predict or find out the so called effect of the models.

(Refer Slide Time: 06:52)



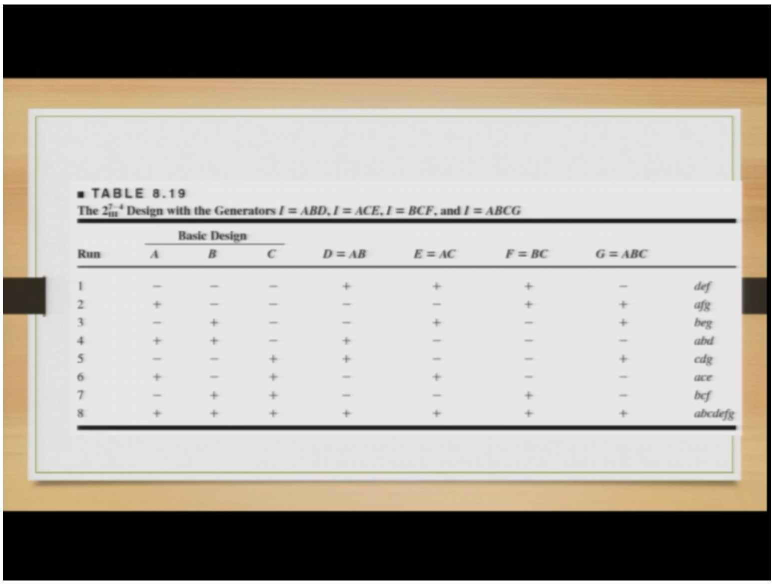
A very useful saturated model of the factorial design is the design of studying 7 factors in 8 runs which is basically given by 2 to the power 7 minus 4, so this design is about 1-16th factors which basically becomes 2 to the power 7 because you are going to have 7 factors, it may be constructed by first writing down the basic ones which are considerably 2 to the power 3, so they are 3 factors A, B, and C and then associating the levels of this 4 additional based on how you design the problem so these are as follows as D as is a combination of AB, E as is a combination of AC, F is a combination of BC, and G is a combination ABC which means that if you have 3,.

So first what you do is that take 2 of them at a time, so if it is 3C2, so obviously the combinations would be C you have taken then you have taken the AB values are there which is D you are basically saying that and the last combination is basically BC, and the next stage what you will do is you will take the factors combination of 3C3 so obviously it will be combination of ABC. Now, if you remember it did mention and that was not very explicitly given in the slides, I did mention that if you increase the number of factors that if you take, say for example, 4 of them and you are trying to combine the factors, it will be 4C2 first combinations, so it will be AB, AC, AD then it will go into the combination of BC, BD, then it will go into the combination is CD.

Then, when you combine the three at a time from this 4, it will be 4C3 then it will be 4C4. So, if you basically keep increasing the factors so if you make it 5, the combinations would be 5C2, 5C3, 5C4, 5C5. So, depending on the combinations which you take, if you have K factors it will be KCK would be the individual one then obviously will come to KC2, KC3, KC4, till KCK and the combinations would obviously give you the effects of the combined factors.

So, here you are considering 3, so again I will repeat, so D is AB, E is AC, F is BC, and G is combination of 3, thus the generators for the designs are ABD, ACE, BCF, and ABCG so depending on the combined effect which you are taking.

(Refer Slide Time: 09:11)


■ TABLE 8.19

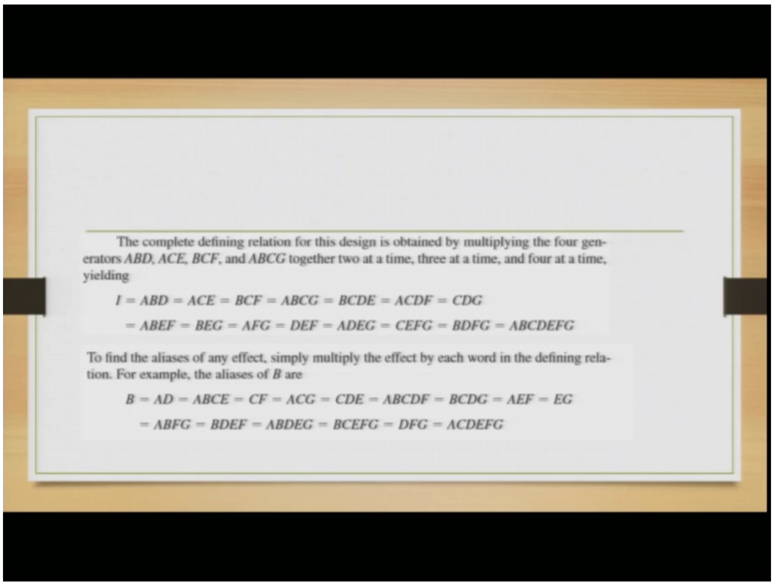
The 2_{III}^{7-4} Design with the Generators $I = ABD, I = ACE, I = BCF$, and $I = ABCG$

	Basic Design							
Run	A	B	C	$D = AB$	$E = AC$	$F = BC$	$G = ABC$	
1	−	−	−	+	+	+	−	*def*
2	+	−	−	−	−	+	+	*afg*
3	−	+	−	−	+	−	+	*beg*
4	+	+	−	+	−	−	−	*abd*
5	−	−	+	+	−	−	+	*cdg*
6	+	−	+	−	+	−	−	*ace*
7	−	+	+	−	−	+	−	*bcf*
8	+	+	+	+	+	+	+	*abcdefg*

So, as usual first we define the factors, define the combinations, define the overall effect and then draw the table so this is the general plan which you are following. So, again we will come to the table 8.19 which will have all the informations about the combinations of the factors. In the first column, you have the run, which is the run number starting from 1 to 8 then the basic design factors as we mentioned that we have taken K is equal to 3 which was ABC then the combination which we took, if you remember we took 2 at a time then 3 at a time, so 2 at a time was basically AB, AC, and BC, so they are mentioned as D which is AB, E which is AC, F which is BC and the combinations of 3 of them taken all at the same time was G which was ABC.

So, if you take the combinations of the effects of A being minus, B being minus, C being minus and if we continue that effects, so this table will give you the positive and the negative effects that is higher effect, lower effect and from that you will basically have the combined effect as coming as if you read the last column of this table which is DEF, EFG till the last one which is A, B, C, D, E, F, G. So; that means, you are combining them at different levels.

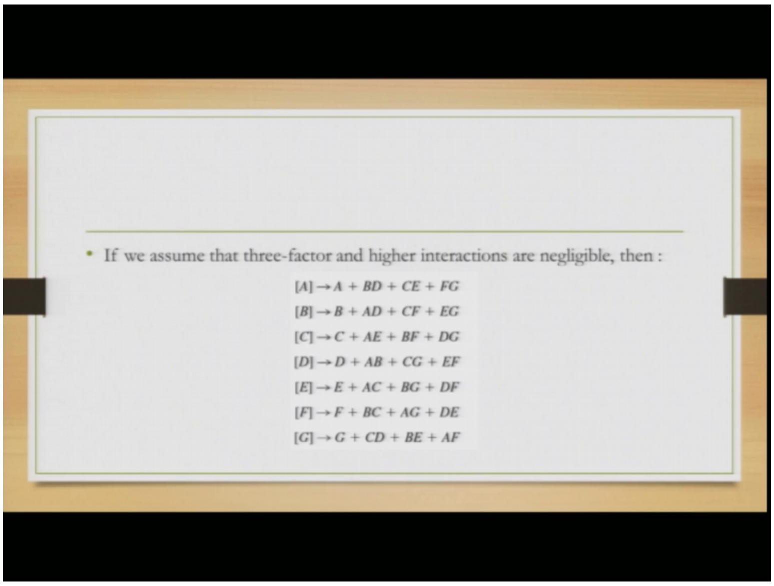
(Refer Slide Time: 10:28)



The complete defining relationship for the design is obtained by multiplying the 4 generators which we have seen which is ABD, ACE, BCF and ABCG into 2 at a time then 3 at a time, 4 at a time and can go on depending on the number of such combinations which you have. So, the I which is the so called combined factor the indicators are and the combined mentioning of the factors are the first one is ABD, ACE, BCF and it goes on till the last one which is A, B, C, D, E, F, G.

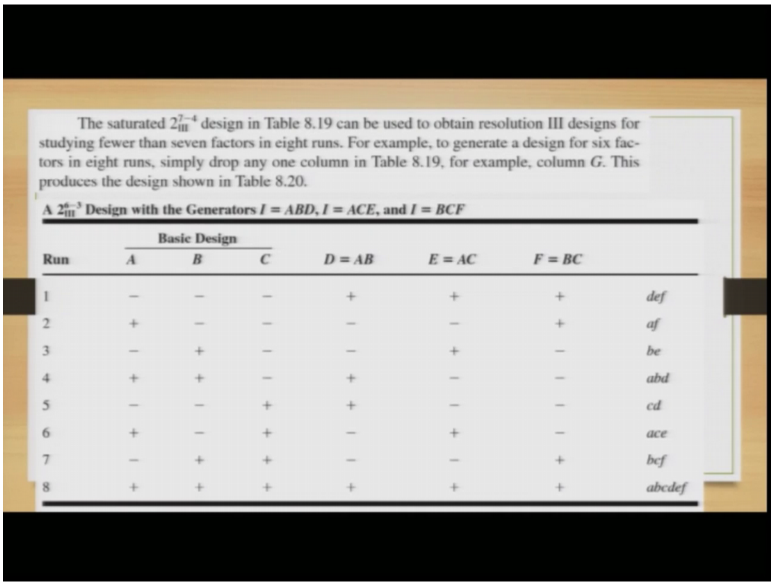
To find the aliases or the effects of this combined factors what we do is that we simply multiply the effect by each one in the defining relationship, for example, the aliases of B is would basically B as we have seen would be the combination of, I am reading the last line only, ABFG, BDFA, ABDEG, BCEFG, DFG and ACDEFG, so based on that you can find out for A for C and as you increase K they would be different combinations.

(Refer Slide Time: 11:45)



So, if you assume the 3 factors and a higher interactions then the effects are A would be in the level of A, BD, CE, FG similarly for B it will be B, AD, CF, and EG and combined to the last effect for G it will be G, CD, BE, AF. So, I am not going to the details how you do it, but once you make a calculation of the number of factors and how you going to combine them, you will come up with these tables and fractional factorials very easily.

(Refer Slide Time: 12:20)



The saturated 2_{III}^{7-4} design in Table 8.19 can be used to obtain resolution III designs for studying fewer than seven factors in eight runs. For example, to generate a design for six factors in eight runs, simply drop any one column in Table 8.19, for example, column G. This produces the design shown in Table 8.20.

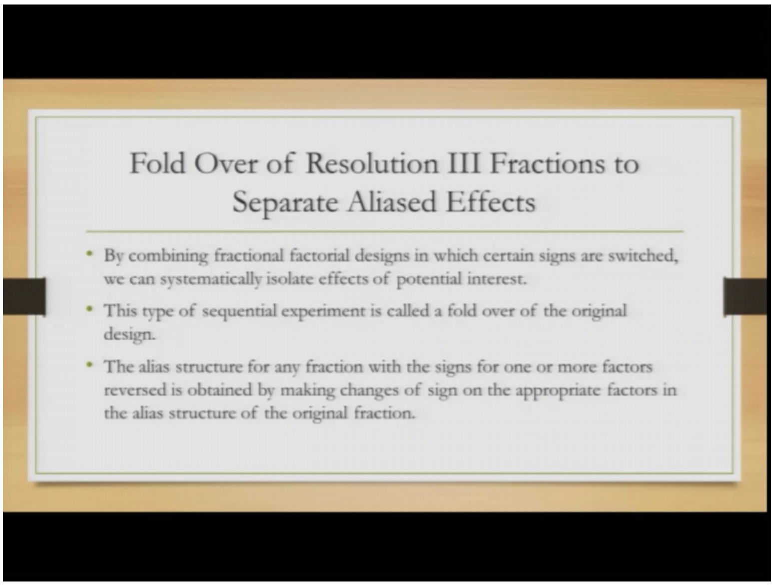
A 2_{III}^{6-3} Design with the Generators $I = ABD$, $I = ACE$, and $I = BCF$

	Basic Design						
Run	A	B	C	$D = AB$	$E = AC$	$F = BC$	
1	−	−	−	+	+	+	*def*
2	+	−	−	−	−	+	*af*
3	−	+	−	−	+	−	*be*
4	+	+	−	+	−	−	*abd*
5	−	−	+	+	−	−	*cd*
6	+	−	+	−	+	−	*ace*
7	−	+	+	−	−	+	*bcf*
8	+	+	+	+	+	+	*abcdef*

So, the saturated model is 2 to the power 7 and suffix 3, so design which we have already considered and table 8.19 which we have already discussed can be used to obtain the 3 design models for studying the fewer of these 7 factors, for example, to generate a design for 6 factors we may simply drop any 1 column as given is 8.19, so let us consider the design with the generator, so the generators Is are ABD, ACE, and BCF so again you have the same fundamental principle of layout of the table.

The first column is basically the runs, the basic design factors which you have consider is ABC, it could be increased or decreased if you remember I have been keep repeating this and the other combined factors were D, E, so similarly we have seen that the by finding on the effects we have found out and considering the design of the generator being as ABD, ACS, and BCF we have the combined effect coming as DEF, AF, and the second last one being BCF, ABC and DEF.

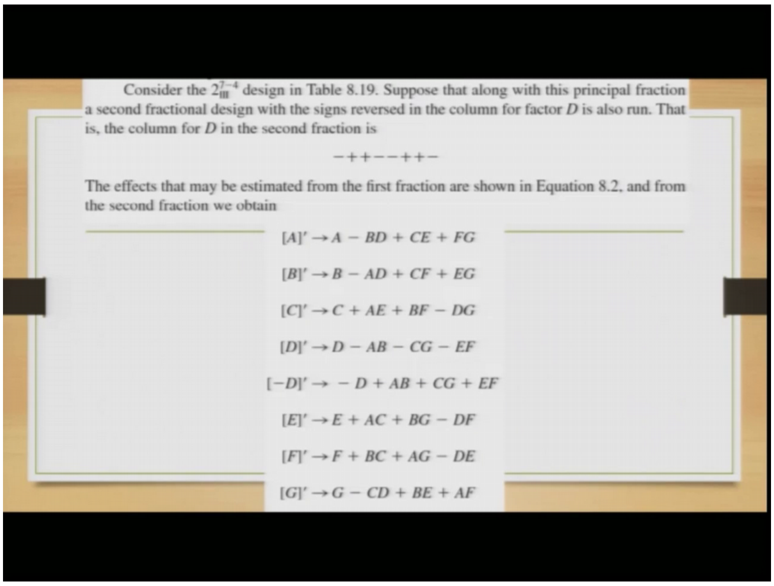
(Refer Slide Time: 13:30)



So, continuing the discussion, fold over or resolution free factors to separate aliases effects and separate them in such a way that we are able to find out the individual as the combined effects also. So, by combining fractional factorial design in which certain signs are switched depending on plus and minus or the effects changing we can symmetrically isolated effects of the potential interest or the factors and make a judgmental decision about that such that the activity shows of your combinations or the effects of the factors is maximum.

This type of sequential experiment is called the fold over. So, they are basically fold over or take the information set in such a way that you could keep repeating step by step and go forward and combine them. The alias structure for any fraction with the signs for 1 or more factors reversed is obtained by making the changes in the signs and appropriate factors and in the appropriate way such that the alias structure of the original fractions can be obtained accordingly still maintaining the accuracy and the predictive power.

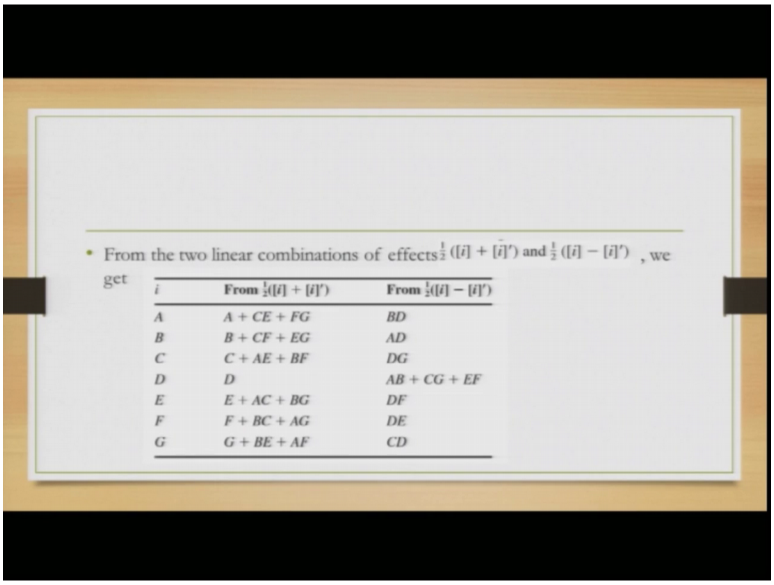
(Refer Slide Time: 14:36)



So, consider the 2 to the power 7 minus 4 designs, suppose that along the principal factors a second factorial design with the sign reversed is given and in the column. So, thus this column of D would be in these combinations minus, plus, plus, minus, minus, plus, plus, minus, minus so the different combinations can be done. So, the effects may be estimated for the first factors as shown here.

So, consider we are mentioning as A transpose even though that is a symbolic notation, you have A then you have the effect of BD on the negative sides; that means lower then CE and FG are positive. Similarly, if we go to EE prime, again nomenclature way I am saying that the effects of all the first 3, which is AE, AC and BG are positive and DF is negative. Similarly, you can basically find out the effects of all the so called factors where the change of the signs has been incorporated considering the fold effort is considered.

(Refer Slide Time: 15:40)



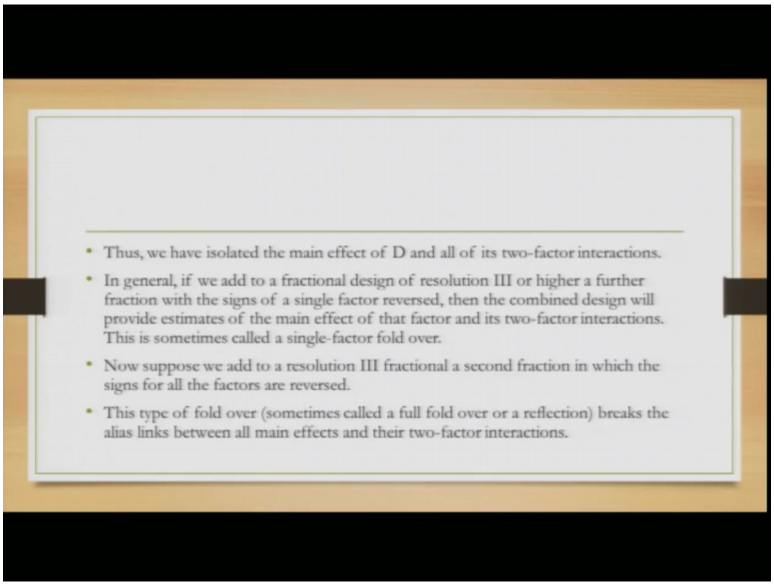
From the two linear combinations of the effects, so you are not basically trying to take without the fold and with the fold, but in this case you are going to consider the change in the sign effect. So, it can be done in different ways, if you take 2 different combinations of the change effects of the signs it will basically be trying to find out the average value. If you take 3 such sign effect changes in combinations, it will be one-third. So, you are putting equal weightage of the signs in order to find out the overall effect of the fold and the fractional factorial design.

Now remember that these probabilities which you are giving half or one-third depending on 3 different combinations sign change or one-fourth for 4 different sign changes, we are giving equal probabilities to the sign change, it may not be true in actual practical examples, so obviously it will change depending on how the problem and how the overall conceptual framework of the problem formulation has been done. So, in this case if you consider again the Is, Is means the overall combined effects of the basic factors and their actual combined factors which you have.

So, you have basically ABC and then if you remember you have combined them as AB, AC and BC then you consider the combined effect of ABC. So, which we mentioned as DEF for the first 1 which D was basically for AB then E was for AC, F was for BC, and G was for ABC. So, this was the first column gives you the basic factors along with their combined factors then from the combination of this full effects you are considering 1 sign change, so original plus the 1 sign change you found out the average, so averages values which are given in the second column are for 1 case is basically for A it would be A plus C plus FG. Similarly, for B it will B plus CF plus EG.

Similarly if you go to the last one which is G it will be G plus BE plus AF. So, this word which are mentioning on the combined effect of the factors. So, from a combination if it is a negative sign, so the combination basically gives me. So, the third column BD for A, AD for B, DG for C and till the second last one is for F which was initially for the positive sign change was for the fold effects was F plus BC plus AG now it is only DE which is where I am hovering my finger and the last one for G which was for the positive fold effects, I am using the word positive fold effects in order to make you understand, it is not an actual technical word, it was G plus BE plus AF and for the other combined for the negative fold effects was basically CD.

(Refer Slide Time: 18:42)



Thus, we have isolated the main factor of D and all of the 2 factors interactions and made the best possible combinations, in general if we add to a fractional design of resolution 3 or higher, so it can be 3, 4, 5, 6 whatever a further fraction with the sign of a single factor is reverse then the combined design will provide the estimates of the main effect of that factor and it is 2 factor interactions correspondingly. So, consider these combinations which I did mention in the last slide, I will again repeat it.

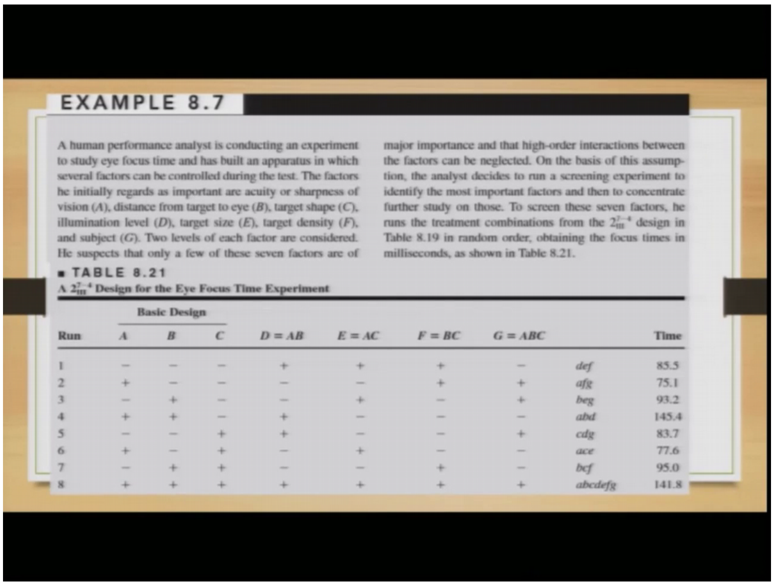
So, for the 3 factors, 3 means 3 suffix or the 4 or the 5 when you are considering the fractional factors. So, the factors may be 3, 4, 5, 6 but the level of accuracy which you want 3 or 4 or 5 would basically be dictated by the combinations which you are going to take, if it is three; that means you are going to take the combinations of 2 at a time, three at a time. If you are going to consider the 4 so called factors being considered to a 2-level, so you will only consider 2 at a time, 4C2.

If you are considering, say for example, K number factors in P such combinations so obviously it would be go as KC2, KC3, till KCP and combinations can be done accordingly depending on the accuracy we want, but for all of these if you consider the fold effect; obviously, this change of signs would increase correspondingly. So, if you want to basically make an accurate calculations. Obviously, the question would be asked that how many different such combinations of folds, the different combinations of factors, different combinations of combining these factors would be taken into consideration, so which you have to decide depending on the problem which you are going to solve.

If I say that and repeat it and obviously calculations would entail a lot of time, so whether you want to ask your question, whether it really suits your time and effort and energy to do such a detailed calculation you get a result, where the effect may not be that substantial, so obviously you will lessen the amount of calculation to get the best effect in the least possible time because that entails energy, time, money so and forth. So, let me repeat it so in general if we add a fractional design of resolution 3 or higher a further fractional with the signs of single factor is reversed then the combined design will provide estimates on the main effect of this factor and its 2 factor interactions accordingly, this is sometimes called a single factor fold over model.

So, it can be single, double, triple depending on number of such folds you are doing different combinations. Now suppose we add as a resolution 3 fractional second fractions in which the signs of all the factors are reversed, so this reverse or combination of the factors signs can be done depending on the problem lay out, this type of fold over sometimes called a full fold over or a reflection fold or depending on what combinations of folds you are taking, it breaks the alias links between all the main effects and there are 2 factor interactions such that even if you are trying to do a detailed calculation what your actual aiming at is to basically break up this effects and try to find the individual factors or individual combination of factors such that are accuracy increases in order to predict the overall effect of the fractional factor models.

(Refer Slide Time: 22:17)



EXAMPLE 8.7

A human performance analyst is conducting an experiment to study eye focus time and has built an apparatus in which several factors can be controlled during the test. The factors he initially regards as important are acuity or sharpness of vision (A), distance from target to eye (B), target shape (C), illumination level (D), target size (E), target density (F), and subject (G). Two levels of each factor are considered. He suspects that only a few of these seven factors are of major importance and that high-order interactions between the factors can be neglected. On the basis of this assumption, the analyst decides to run a screening experiment to identify the most important factors and then to concentrate further study on those. To screen these seven factors, he runs the treatment combinations from the 2_{III}^{7-4} design in Table 8.19 in random order, obtaining the focus times in milliseconds, as shown in Table 8.21.

■ TABLE 8.21
A 2_{III}^{7-4} Design for the Eye Focus Time Experiment

	Basic Design								
Run	A	B	C	$D = AB$	$E = AC$	$F = BC$	$G = ABC$		Time
1	−	−	−	+	+	+	−	*def*	85.5
2	+	−	−	−	−	+	+	*afg*	75.1
3	−	+	−	−	+	−	+	*beg*	93.2
4	+	+	−	+	−	−	−	*abd*	145.4
5	−	−	+	+	−	−	+	*cdg*	83.7
6	+	−	+	−	+	−	−	*ace*	77.6
7	−	+	+	−	−	+	−	*bcf*	95.0
8	+	+	+	+	+	+	+	*abcdefg*	141.8

So, let us consider an example, I will read it and then come to the one. So, human performance analyst is conducting an experiment to study eye focus time and has build an apparatus in which the several factors can be controlled during the test. The factors initially guards as important or acuity or sharpness of the vision which we mentioned as factor A, distance from the target to the eye, so you are basically checking the power of the eyes and doing all the tests which is distance from the target which the eye want to see or you want the patient to see or the experimenter wants the persons on whom they are doing the experiment to see, that is B.

What is the target shape depending on depend on how this model is being developed, so smaller the size, larger the size; obviously, you will have an effect on your eye. So, what is the illumination level in lumens where the room is bright or dark then you take this target size, so apart from the target shape which is how it is designed the size is also important then the target density how is considering you are seeing a picture and if it is very dense colour, the number of pixels are very large; obviously, it is much better visible if it is very light; obviously, it becomes difficult.

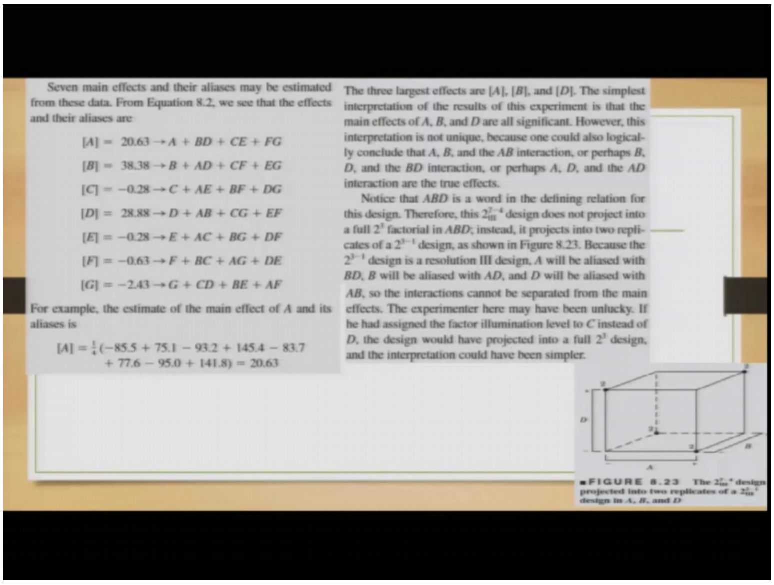
So, if we have a screen, computer screen far off or a tv screen, so the density of pixels or density of number of such dots per square inches per square centimeters would basically entail you in much better visibility of that screen whatever the pictures, whatever the objects which are there on the screen and the other factor is basically the subject. So, you have basically A, B, C, D, E, F, G such individual factors So, 2 levels of each factors are considered. So, the experimenter suspects that only a few of these 7 factors which you have considered A to G are of major importance for the study and that higher order interaction between the factors can be neglected.

So, say for example, you combined AB, AC or if you combine ADE whatever the combination so; obviously, you can combine 2 at a time, 3 at a time, 4 at a time and continued till the highest level. On the basis of this assumption, the analyst decides to run a screening experiment to identify the most important factors and then to concentrate further and do the analysis in more details. To screen these 7 factors, he runs the treatment here, he runs the treatment combinations from 2 to the power 7 minus 4 or the 3rd level design in as given in table 8.19 in random order in order to obtain the focus time on a level of milliseconds and that is given here which is table 8.21.

So, what you have is 7 factors and you are considering level P value of 4 and through the third level of so called accuracy. So, this design for the eye focus time experiment which say for example, ophthalmologist or the optometrician or the doctor is doing, so the combinations are given. So, you have basic factor designs, technically they were combined as A, B, C. So, the other factors are D, E, F, G which we consider as the combinations of AB which is for DE is basically AC, B is BC and G is ABC.

So, if you do the runs which are given in the first column run 1to 8 and if you combine them, the combinations are given as say for example, for the first row you have the runs as the values of ABC are minus and the values are given and then you find out the combinations of DEFG as positive, positive, positive, minus respectively and the time is given in the last column; that means, you are trying to find out the time of the focuses and all these things in milliseconds. Similarly, for the second run, you take a different combinations where effect of A is positive, BC is negative and the combinations of D to G can be seen from this table. Similarly for 3rd, 4th, 5th till the 8th when you are considering all the effects are positive and the time in milliseconds are given in the last column.

(Refer Slide Time: 27:05)



The seven main effects of these aliases may estimated it from this data, so they are as mentioned. So, the equations can be found out and we see the effects and the analysis as follows. So, for A you have the value of 20.6, those are in the milliseconds, and the effects are for AB, BD, C, and FG. I am talking about the so called the combined effects. So, it is B, AD, C, and EG. Similarly for 3rd, 4th, 5th and 7th which are the combined effect of C D E F G respectively.

The effects are the values are given as minus 0.28, 28.88, minus 0.28, minus 0.63, minus 2.43 and the values are individual are given, for example, if I want to find out, so what I will do is that I will try to find out the effects of for A combine would be A, BD, CE, and FG, so if you put those values of the milliseconds which you found out, so how many such combinations are there, there are 4, what are those four, A, BD, CE, and FG, hence you would basically give a weightages of one fourth based on that when you calculate you find out the value of 20.6. So you can do the calculations accordingly.

The three large effects which are basically the combined effect of A, combined effect of B and combined effect of D are found out. The simplest interpretation of these results of the experiment is that the main effects of A, B, D are all significant and we will consider them accordingly. However, the interpretation is not unique because one can also logically conclude that different combinations of AB and AB interpretation or perhaps interaction of BC and D interactions and different combinations can give you the same effect. So, notice that the ABD is in defining the relationship for the two factors 7 minus 4 to the level three accuracy can be done accordingly and we can basically do the calculations and find out the so called models and do it.

So, with this we will end the 28th lecture and continue more discussions about the factorial models, the four models, so on and so forth in the in the next remaining 12 lectures. Have a nice day.

Thank you very much.