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

Lecture – 25
Two level Fractional Factorial Design – III

Welcome back my dear friends. A very good morning, good afternoon good evening to all of you and this is the TQM – II lecture series under NPTEL MOOC and we are going to start the twenty fifth lecture and as you know that this is 20 hours that is 40 lectures and that is spread over 8 weeks. And, before I start this 25th lecture because this being second of January, 2018, I would like to wish all of you a very happy new year and a prosperous year and I am sure everybody would do very well mentally, physically, spiritually for yourself and for the society with this words I will start this lecture twenty fifth.

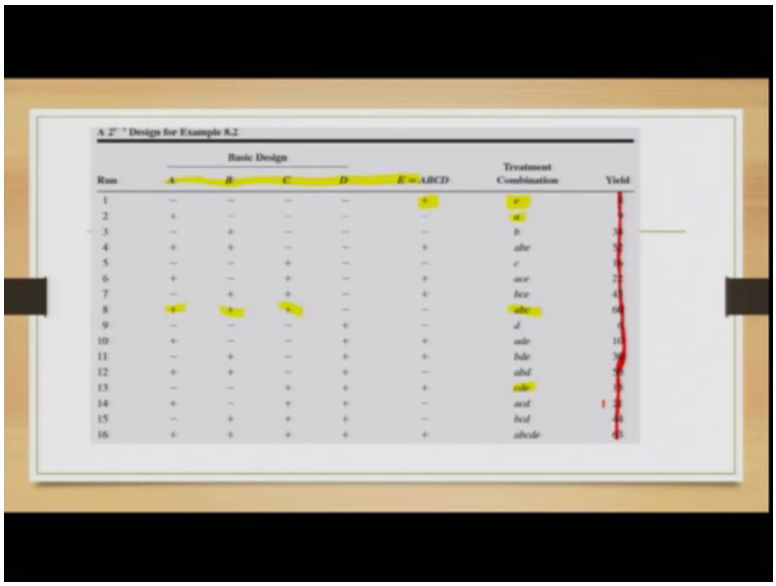
Now, if you remember in the last class which is the 24th one we had just started an example and in the in example there were five factors which were basically mentioned as etch time and the run down time and for a some circuit electronic circuit. So, the factors was ABCDE and for this example, obviously, it can change for this example for each factors there were two levels now also interestingly for this example you can consider the factors as to be quantitative and qualitative like say for example, quantitative can be; So, etch time can be in minutes or in seconds another one can be say for example, the amount of impurity or the amount of purity of some material or whatever quality factor which you are using may be very high or very low or say for example, humidity may be very high or very low temperature may be very high and very low, but temperature can also be done on a quantity scale.

So, with this we want to basically see that how ANOVA modeling can be utilized to basically find out the effects based on the fact that we have basically and have hypothesis to test and remember that having hypothesis would be either for the greater than type or less than type depending on the level of significance which you want to assign for this. And, obviously, for each and every hypothesis small hypothesis which you have I will I will come to that the level of significance to which it will affect the overall example of overall experiment would differs say for example, one example it can be 12 level of

percentage of one person in some case it can be 2.5 percent it can be 5 percent, 10 percent whatever it is.

So, you will see that and basically plan our discussion of the analysis of the problem accordingly.

(Refer Slide Time: 02:57)



So, if you consider the overall data. So, I will just go to the actual example as which you always do. So, there are basically five factors that is A B C D E and, obviously, it would be an fractional fact factorization model of 2 to the power 5 minus 1 and whatever the level of the roman concept of level of the problems we will discuss we will we will come to that also but, let me go through the examples slowly.

So, if you see the slide in the left most column you have the runs. So, basically they start from run 1 to run 16 and in the box or the cell values you have basically as usual as you can understand plus and minus significance plus means effecting positively and minus means effecting negatively and the top most row were basically the factors which we have denoted as I just mentioned few second back it is basically A B C D E. So, if you if you note down let me use the highlighter.

So, the factors are A B C D and E and obviously, we will consider that that is e would be basically be a combination or the effects were A B C D E's are not there. So, basically they would we will try to consider them on a standalone basis.

Now, when you consider the treatment combination so, say for example, for run 1. So, run 1 basically means the effect of a if you if you just look at the cell they are negative which means the effects of A B C D are not there only the effects E is there. Hence you see a plus here which means the treat treatment combination is coming out to be coming from factor A.

Now, remember if you are doing the actual mon ANOVA modeling and you are using the so called regression model where say for example, y what are the suffixes I am going to come to that later y is equal to mu plus say would be effects of I which would basically be for a factor A where I basically changing from 1 to A then they would be basically the con the factor only B this I am talking about standalone basis this is only the effect of be where the suffix would be j, j changing from 1 to small b because there is the total number of such effects coming from factors for factor B then they would be a factor C then would be a factor D then be a factor E.

So, it will be ijklm, so that means, ijklm for ABCDE and the last factor would basically be n where n would change technically here from 1 to n. So, ba consequently and very interestingly the suffix and the total number of observations for the number of samples is same. So, n would change from one to n and these ijklm would change in basically respectively to A to B to C to D to E.

Now, this I am repeating if we have done the modeling as per the norm where the number of combinations would have been in the next stage would be combination of two taken at a time from this 5 that is 5 C 2. So, one of the combinations would be AB then another would be AC then ad AE then we will go to BC BD so on and so forth. Then will go to the next there would be 3 combinations taken from the 5. So, it would be ABC ABD and so on and so forth. Then would be four combinations at a time taken from the 5 it will be ABCD ABCE and so on and so forth and, the last one would be basically be the combination of 5 taken from the 5 is ABCDE.

Now, for each of this combinations they would be a so called dispersion and as you remember they were denoted for I am for the factor A it was tau factor B it was beta so, basically we will denote as tau one suffix I tau 1 is basically for a tau 2 for b tau 3 for c tau 4 for d tau 5 for e and the suffixes along with this tau 1 tau 2 tau 3 tau 4 tau 5 would basically be i j k l m respectively denoting the number of such factors which are there for

each and obviously, then you would basically have in the next step when you are taking two combinations I will only repeat one it will be tau 1 tau 2 suffix i j, that means, combinations of ab taken together.

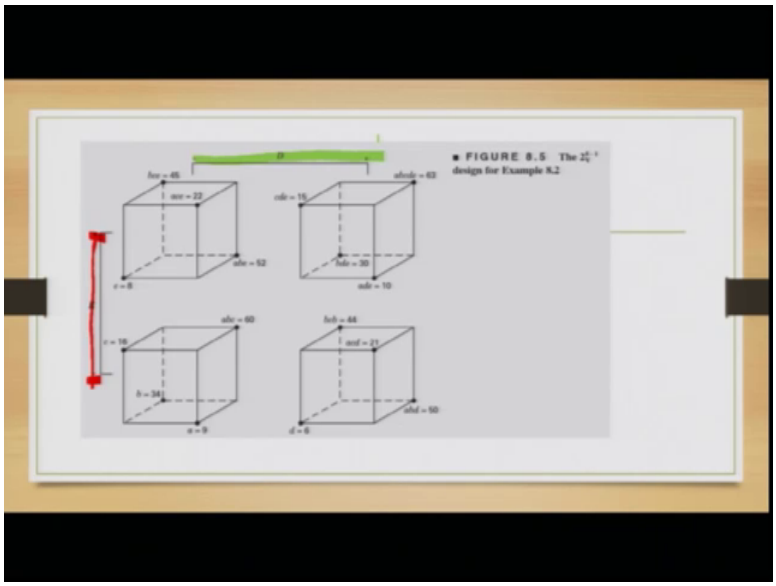
Then others combinations can be done accordingly then when we go to the 3 combinations it would be tau 1 tau 2 tau 3 suffix ijk when it is abc and so on and so forth for the for a combination of four and the last one would tau 1 tau 2 tau 3 tau 4 tau 5 suffix ijklm.

Now, in order to do away with that we consider the fractional factorization and hence we are discussing this. So, again coming back to slides the next row would basically be where the effect of a is there so hence the treatment combination would be only for a I will highlight that if I go to the 8. So, repeatedly if I go to the eighth row the combinations effects are coming positive from A from B from C which is here and there is no effect from D and E, so, obviously, the treatment combination are abc. If I go to the thirteenth row I am going I am taking one at a time I you can go from 1 2 3 4 5 6 till the sixteenth one, but I am taking arbitrarily few of them, but the analysis obviously, obviously remains the same.

So, if I basically take the thirteenth row so, obviously, and it is minus a minus b, but this plus for b for cde. So, hence the treatment combination would be first cde as mentioned here and when I come to the last row which is the sixteenth one the combinations for all of them are positive, hence the treatment combination would be abcde.

Now, if you see the yield that is coming from the data. So, the yields I will just change the color. So, the yields are basically if you consider they are for the first one it is 8, second one is 9 so on and so forth, till the second last one 44 for the fifteenth row which is the combination treatment combination for bcd and for the combination treatment combination of abcde the value is 63.

(Refer Slide Time: 09:49)



Now, when I go to the fractional factorization for the combinations of because why it is based to with 2 I would not use the word base 2. It is 2 to the power 5 minus one because for each of this factors there are two levels. There are 5 factors and we are com combi considering this fraction factorization to the level of 2 to the power minus one base of a roman subscript 5 denoting the level of significance which you want the word level of significance. I am using is not from the statistical parlance, but basically from the parlance of the discussion which we are having for the last few days for few lectures let me use the word lectures.

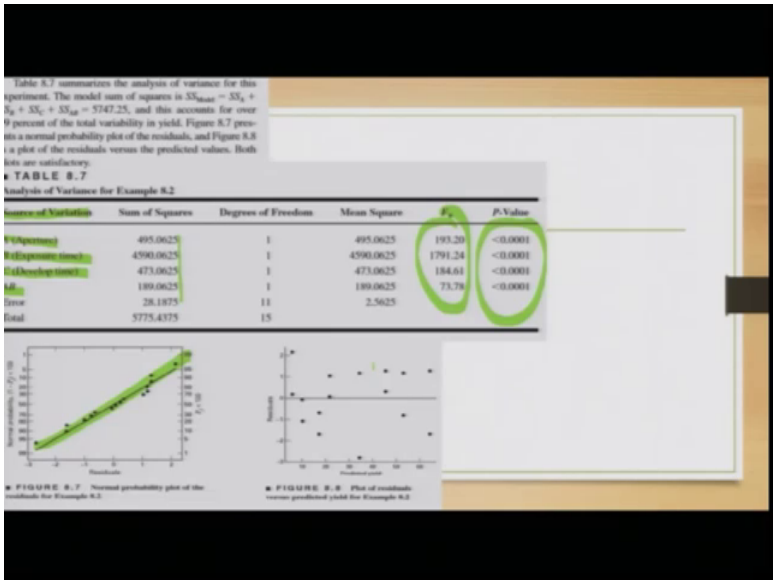
So, if I con considering combination of E, so, E would basically have the positive effect and a negative effect. Now, let us basically not dwell into the effects on positive level inside E what are the factors and then for the negative effect for inside E what are the factors. So, it is like this. So, these are positive and negative for E and these are the combinations for D which is positive and negative.

So, obviously, you will have now let us consider the combinations as this at a level of e is 8 or a combination of abe is 52 you have these this this 3 dimensional figures which you have already discussed which will give me the combination for abcde depending on the fact which is positive and which is negative such that we will be able to find out the fractional factorization diagram. So, which means the design of the experiment for 2 to the power 5 minus one depending on each factor as 2 levels they have 5 factors and the

level of significance is for given by roman subscript the 5, so, these combinations can be done.

Accordingly and remember that when you are basically going to the to the higher level the combinations and would be more complicated point of number one and it could also have been done where we could have taken the effects of bcdee and then consider a as an effect which basically subsumes all the effects which are not considered under bcde, that means, e would get replaced by a and the corresponding diagrams could be done accordingly and table could be done accordingly as just discussed in the previous slide.

(Refer Slide Time: 12:28)



So, the tables 8.7 summarizes the analysis of the variance for the experiment the model sum of the squares are one is basically for sum of squares for A, sum of squares for B, sum of squares for C and sum of squares of AB. So, if this comes out to be 5747.25 you can do me the equations accordingly remember that I am just mainly helping on the F x rather than going to the calculation, but if you remember the first example which we did for the one was the was for the elimination melting another was for the battery one I dint go into the details of the calculations how do you do it.

So, but subsequently I am only highlighting at the factors which are the factors which are effecting and how the factors are basically being utilized in order to find out the mean squared errors and then find out the F statistics. So, remember all these those analysis has been to draw a table where the first column would be the factors next column would be

the sum of squares then would the degrees of freedom then would be the mean squared errors then would be the F value for the F statistics and then the label level of significance. So, the overall game plan of trying to tackle any problem remains the same.

So, this 5747.25 and this accounts for overall 99 pa percent of the total variability in the yield. So, figure 8.7 which is there here ma gives the norm normality plot for the residuals because that is important to find out whether they are outlined or not and figure 8.8 is plot of the residuals versus the predicted values such that we can find out there is a one to one point correspondses with that. So, both this plots are satisfactory and let me come to the this analysis elements little bit more in qualitative sense.

So, the sources of variances are as usual that we are basically been able to reduce the fractional reducing the concert of fractional factorizations. So, it is A which is aperture, B is exposure time, C is developed on time and the other factor is AB which is aperture and exposure time. So, they so, those values comes out to be so, it is 495.0625 for a 4590.625 for B similarly 473, I am not talking about the decimal part for C and for AB which is aperture and develop exposure time it comes out to 189. So, if you add up these values and basically find out the overall percentage with the total which is 5775 that is comes out to be 99.

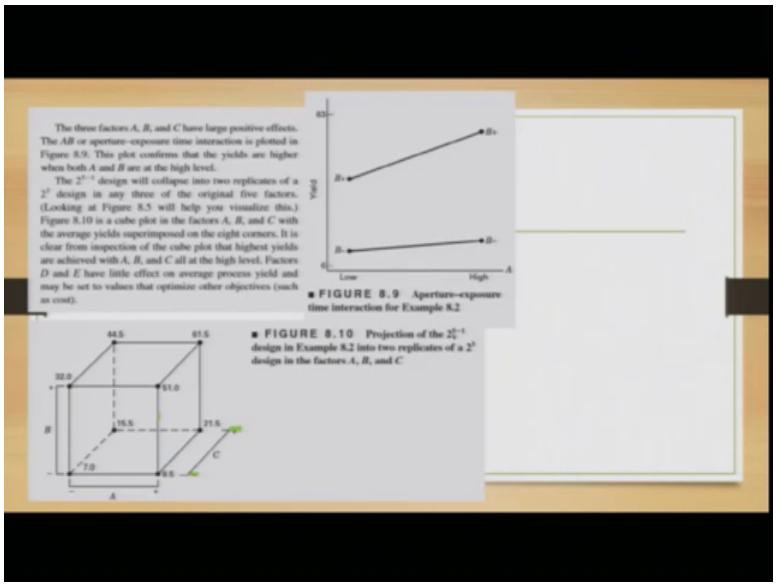
So, the total error is very insignificant with respect to the total sum of error, that means, the white noise is you have been able to design the overall experiment in such a way that you have been able to do a go away with the white noise and as far as possible and able to basically find out some possible reason that what are the effects which are basically effecting your total experiment. So, the degrees of freedom as I mentioned in just few minutes back would be the third column. So, the degrees of freedom are given as 1 1 1 1 and 11 and 11 is for the errors.

The mean square can be found out by again I am repeating please forgive me for that. So, for a it would be 495.062 divided by one for the error part it would be others would be always be divided 1. So, they would not change in the value of the mean square for the error it would be 28.1875 divided by 11 and the to and the F values again found out and given by finding out the ratios on the mean square and this level of significance are given based on which you can say that what is the p value based on which you can pass on some judicious judgment about this example.

So, if you see the normality plot, so, they are almost normal which it should be 45 degrees line if you do the or if you go through the basic fundamental concepts of the cube plot the plot of the residuals are done in such a way. So, there is almost even balance amount or dispersion both on the positive side and the negative side of the mean value.

So, if which will would be basic technically the average value of the errors is 0 and standard deviations would basically be sigma square, but you should be remembered that the sigma square value of the errors are independent of time which have been repeating time and again from day one when we started the concept of design of experiment.

(Refer Slide Time: 16:56)



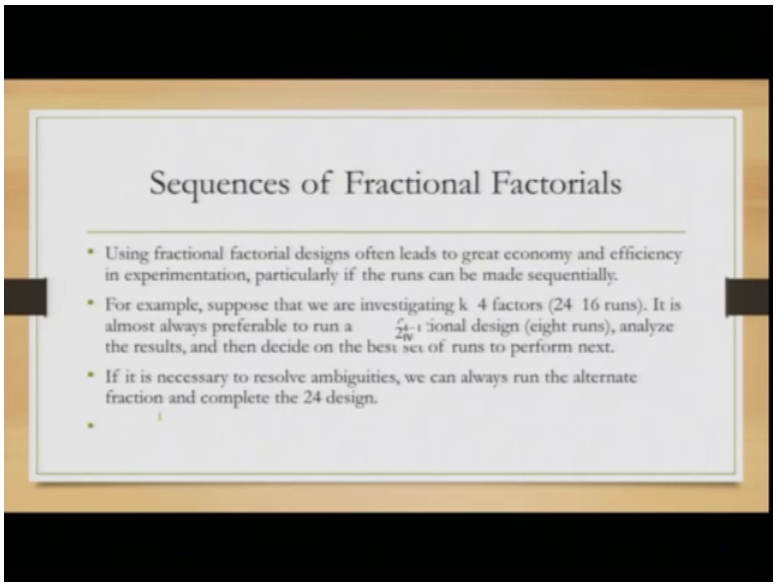
The three factors abcd have A, B, C have last positive effects. So, which we just saw so, the if the AB factor which is aperture and exposure time fraction is plotted in figure 8.9 which is this and we want to find out with respect to the different other factors which is the overall significance of them. This plot confirms that the yields are higher when both a and b are at the higher level. So, the 2 to the power 5 minus 1 design will collapse into 2 in case of 2 to the power 3, because now the here basically only three A, B, C as there design in any 3 of the original 5 factors.

So, looking at the figure 8.5, we it will basically have us to visualize because now we have been able to do is that try to find out the level of significance level of significance means again I am using this word very broadly, that means, the overall effect which is

coming out is basically from A, B, C and others have been subsumed in such a way that we are able to not ignore I would not use the word ignore, but we have been able to reduce the overall effect for the other factors least possible value. So, hence now it becomes actually it was 2 to the power 5 then we are basically considering 2 to the power 5 minus one, now it is basically being reduced to your 2 to the power 2 to the power 3 which means that we have been able to grasp our problem to using the simple concept of fractional factorization in such a way that the overall effect is basically being highlighted only 3 factors not 5.

So, figure 8.10 is basically the cube plot of this factors A, B, C which are important with the average yields superimposed on the eight corners. So, you have basically if you see the effects of A are positive and negative. So, I am taking F A is positive means effect is there negative is not no effect B again I am plotting along this vertical line along this the plane where I am basically showing you the slide and C is basically coming going from this point towards into the slide. So, these are the negative and positive value. So, if you find out all overall effect the values are the corner points gives you the overall average values based on which we will do the calculations.

(Refer Slide Time: 19:22)



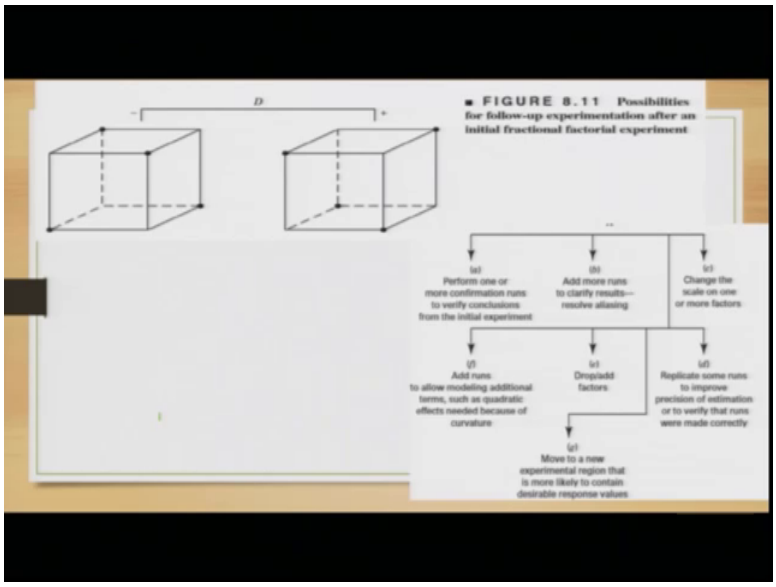
Now, we will consider the concept sequences of fractional fact factorials or factorizations. So, using fractional factorial design often leads to greater economy efficiency in experiment because we are able to use less number factors, but basically as

the higher level of efficiency such that overall effects is mimicked by the by least number of factors the maximum amount of amount of a level of efficiency let me use this word. So, ba they often lead to greater economy and efficiency in experimentation particularly if the runs can be made sequentially.

So, if you consider that we are taking the factors in such a way the sequentially we are able to consider the effects in such a way that we are able to eliminate maximum of the factors which are in significant. For example, suppose that we are investing four factors. So, if it is there it is always almost always preferable to run of a factor a model of 2 to the power 4 minus 1 suffix roman subscript 4 fractional design problem analyze the results and then decide on the best set of runs to from the experiment in next; that means, if we have basically four factors we take them in this way analyze and find out the effects of which factors are significant and which are not that significant.

Hence, they can be not eliminated in not be considered to be that details such that the amount of calculation which you want to do would be reduced to the maximum possible extent. If it is necessary to resolve ambiguities we can always run the alternate fractional models and complete total set of designs which is 24; that means, rather than immediately going to the for the combination of 2 to the power 4 we can consider them 2 to the power 4 minus one and find out if the efficiency is good. So, we need not consider greater number of calculations as required.

(Refer Slide Time: 21:24)



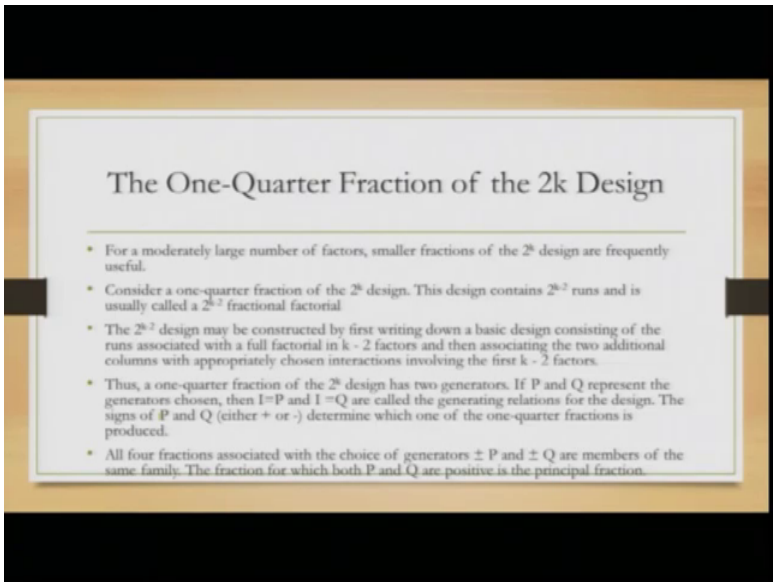
So, the possibilities for the follow up example could have been that we perform one or more conf conversion confirmation runs to verify conclusion from the initial experiment or we can add more runs to clarify the results and resolve the analyzing problem the concept so that we are able to take the factors which have significance and the sig the steps we could be basically scale of one or more factors. So, hence which would mean the combinations to or the levels to which we are going to consider for this all this factors are on one to one parity and they can be basically considered in the simplest form.

So, you add runs, so, you if you basically go in to the details the next level of significance would be replicate some runs to improve precision of estimation or to verify that runs were made correctly. The other way it would be can be that we add and drop as we mention that if we considering 5 factors rather than considering 2 to the power 5 we can consider them 2 to the power 5 minus 1 and basically con consider only those factors which are significant others may not be considered for the for the combination.

Because, if you saw that in the last experiment initially we had 5 factors, but the at the end of the day we only took two three which is A, B, C the f level of analysis or decisions could be add runs to a modeling additional terms which quadratic effects needed because of the curvature of the level of relationship which is there.

So, it is not always linear it can be quadratic also or polynomial I am using the word polynomial is very general sense such that we are able to replicate the effect maximum degree and the last way of handling this would be which is g is basically move to new experimental region that is more likely to contain desirable response values and do our experiment accordingly.

(Refer Slide Time: 23:25)



We will now consider the one quarter fraction of 2 to the power k design or 2k design for a moderately large number of factors smaller fractions of 2 to the power k design are frequently useful because if you are able to find out the effects of the factors by by considering the effect of the overall factors how many factors are there by using the least number of factors then, obviously, our calculations are must less cumbersome and they are they can be utilized such that we are able to attain the answer in the least possible manner or least possible calculations so, least possible time whatever it is.

So, we consider the one quarter fraction of 2 to the power k design this design contains 2 to the power k minus runs and is called 2 to the power k minus one factorial fractions, that means, initially we have 2 to the power k minus 1 or it is 2 to the power k minus 2 and what is that we will discuss that in details. The 2 to the power k minus 1 design may be constructed by first writing down a basic design consisting of the runs associated with full fractional of k minus k 2 factors and then associating the 2 additional columns depending.

So, the 2 if you remember the table initially you had the runs then in the column which I mentioned here of if you consider the last experiment it was a b c d and in e if we consider the effects of a b c d in a way; that means, e subsumes all the effects in very simple parlance which are not being subsumed under a b c d combined together so, whatever may be the case.

Now, what we are going to consider is that consider our second level of deviation such that we are able to break the factors in such a way that the effects can come out more easily, but remember that trying to be basically be simplistic in a overall calculation would be nice from the computation part, but may not be able to give us the exact answer which you are talking, but if you are able to design and find out the factors as far as possible then obviously, this job is done very carefully and nicely such that we still get the same level of robustness in our answer by or a calculation doing less amount of calculations.

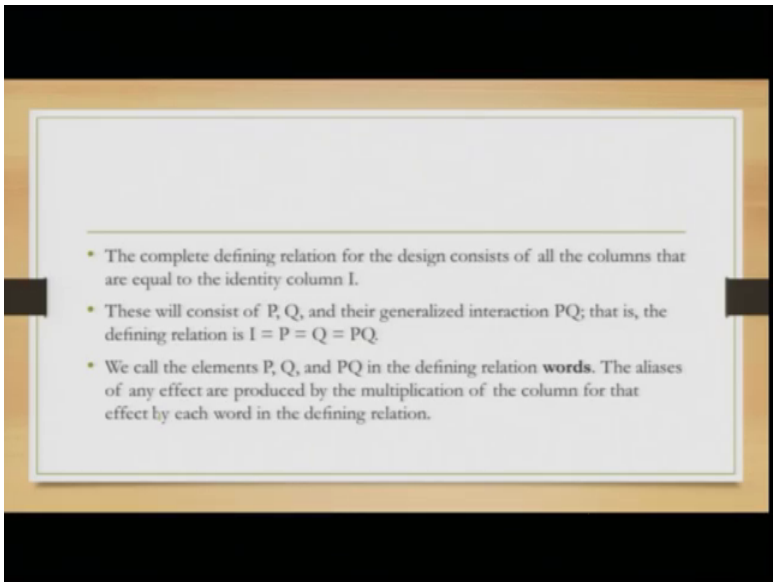
Thus is thus let me continue reading fourth point thus a one quarter fraction of 2 to the power k design had 2 generators if. So, initially if you remember the generator was I, so, if P and Q represents the generators chosen. So, obviously, we would have P and Q and are called the generating relations for the design; that means, P and Q would become similarly the generating design based on which we will do our calculations. The signs of P and Q either they are plus and minus determine which one of the quarter fraction is produced and which would basically be utilized for a calculation.

All four fractions are associated with the choice of generators for either plus minus P or plus minus Q are and would be the members of the same family which means that if you are considering plus and minus P on plus minus Q the overall set of the factors from where the generators are being formed would be the same. So, the level of significance for a may be positive under P the level of significance for factor a means at the factor may be negative under the generator Q.

So, those are answers would be coming out and the moment we are basically able to differentiate the generators it will give us the level of significance of the effects which is coming out from the factors ABCDEF whatever it is, that means, I mentioned six factors it can be more or less depending on how you are being able to design.

So, this k value 2 to the power k minus one k would basically signify the number of factors which you are initiating. So, if it is ABCDEF, that means, we have basically 6 factors, so, k would be 6.

(Refer Slide Time: 27:31)



The complete defining relation for the design consists of all the columns that are equal to the identity column I, these will basically consists of P and Q. So, we are basically going for this factor of k minus 2; hence it is P and Q generators and let me continue reading and their generalized interaction is basically given as P and Q that is the relationship would be I is basically P generations Q gen generators and they would be combinations of P and Q.

We call the elements P Q and PQ in defining the relations of the words such that P Q and PQ basically have the effect which is coming out from all the k number of factors which are there in the example which you are going to consider. Then aliases of effect are produced by the multiplication of the columns for that effect for each word in defining this relationship.

So, with this I will end this 25th lecture and continue twenty sixth about the discussion over the factor of fractional fraction of 2 to the power k minus 1 and consider examples accordingly. Have a nice day.

Thank you very much.