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

Lecture – 26
Two level Fractional Factorial Design – IV

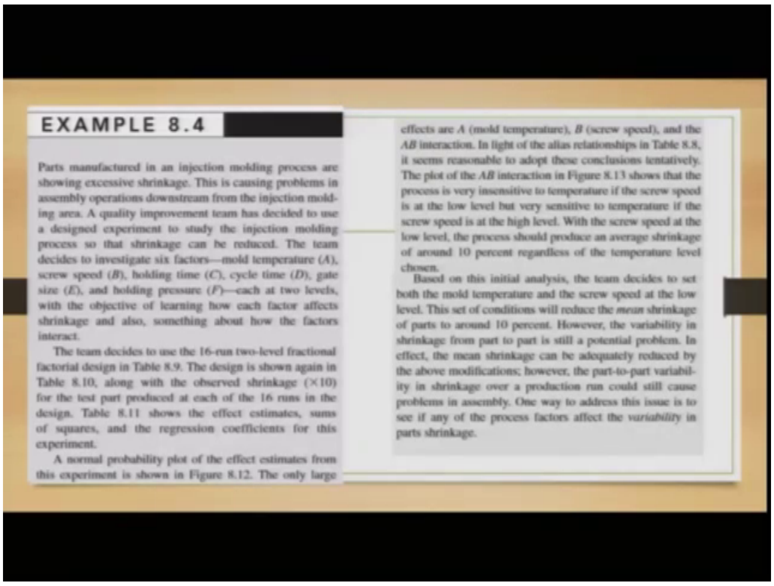
Very good morning good afternoon and good evening my dear friends welcome to this to come to lecture or course under the NPTEL MOOC series and I am Raghunandan Sengupta from the IME department IIT Kanpur and this is the 26 lecture. So, as you know I would just take your permission again summarize that this is a 20 hour lecture which would basic will be 40 hours, but considering their number of lecture each of half an hour it will 14 such lectures spread over 8 weeks and we are in the 26 learn lecture.

Now, in the last class which is 25th one I was summarizing that if you have a factional factor models to the; where the concept there are k number of factors and you are basically subsumed the effects by using the concept of k to the power 2 to the power k minus 2 which means that you subsume to generator subgroups and they were mentioned as p q. And now we see the third one will be p in p into q such that we are able to differentiate the overall effects of all the factors accordingly so, there k number of factors.

Now, many of you may be thinking in the last class, what if we had k to the power minus 2 to the power k minus m; where there were k factors and m what the. So, called subgroups it could have been done in this week, but the complications would be much more intensive, but; obviously, it would be a nice way of handing the problem, but will still stick in our case of fractional factorization model which is given by 2 to the power k minus 2.

So, considering that we will proceed with the discussion and this is the example and we basically go in the same way analyze the problem give the main table depending on the factors and the effect and then come to the answers and answers if you know that I would rather play stress on the factors and their effects rather than going to the (Refer Time: 02:26) of the calculation. So, I will just mention the values not going to how the values I have been calculated, because you have already dell that many times in the initial lecture starting the 7 after the 7th one because if you remember in till the 7th one we are basically completing few of the nomenclature and the basic structure of different type of distributions which are applicable.

(Refer Slide Time: 02:52)



So, let me read the example. So, parts manufacturing in an injection molding process as showing excessive shrinkage. This is causing problem in the assembly operations and assembly lines downstream from the injection molding plant or the area. A quality improvement team has decided to use the designed of experimental study the injection molding problem and find out what is the effect the team has found out there are six factors we will mention them as A, B, C, D, E, F.

So, what are those A is mold temperature, B is screw speed, C is holding time, D is cycle time, E is gate size, and F is the holding pressure each are at two level. So, that is why. So, if you if you have listen to me carefully this is the whole crux of the problem which will give you, what is the fractional factorization problem each are at two design so; obviously, it will be 2. Now what is the power would basically mean that how many factors are there which is 6. So, actually would have basically solved 2 to the power 6 way of combinations of problems.

Now, in this example we will consider 2 to the power 6 minus 2, because we will be basically doing the generator concept is use ling is utilizing that for a lower level of combinations of problems. Now another point which I will missed and excuse me for that is that we are always talking of the fractional factorization of 2 to the power something. So, we are laying more emphasis on those the powers which would basically mean, what is the how many factors are there and what is level of fractional factorizations? We are doing to what level of significance again the what significance in a very general English term not from the statistic point of view, but in all these discussions we have intuitively assume that the level of factors are 2. So, hence it is 2 to the power something. Now in case if the level of such just significance for any factors was more than 2; obviously, the overall concept of trying to analyze the problem would change.

So, which is very intuitive. So, I will come to that later on very flitting the, but let me basically discuss the problems as we proceed. So, as I read. So, there are factors A to F each at two levels with the objective of learning that how each factor affects the shrinkage and also something about how the factors interact with himself in order to affect the shrinkage the team which was doing all these analysis decides to use 16 run two level factorization problem. So, they will basically take 16 runs and informations about that which is given in table 8.9 the design in shown. Again in table 8.10 along with the observed shrinkages which are for the test part produced in each of the 16 runs in this example and table 8.11 shows the effects estimate sums of squares and the regression coefficient.

So, basically table 8.9 and 8.10 are actually the information for the all the 16 runs and as you remember the table where we do all the calculations, where in the first column again I am repeating it first column are the factors then their overall effects whether they are plus and minus that would be given in inside the. So, called matrix then is what are the eff effects which is coming out then you have basically the degrees of freedom and before that; obviously, degrees of freedom you have we have the scroll.

The sum of squares and after the degrees of freedom we have the mean square errors and after the mean square the F statistic based on which will pass judgment about the analysis of variance we will also do a normality plot of [vocalized-noise] of the effect of the estimates and also about the residuals and; obviously, try to find out whether the assumptions of normality is two whether the assumptions of the mean value of the error is 0 or whether the assumptions of the various of the error is sigma square which is constant with respect to time the they are true.

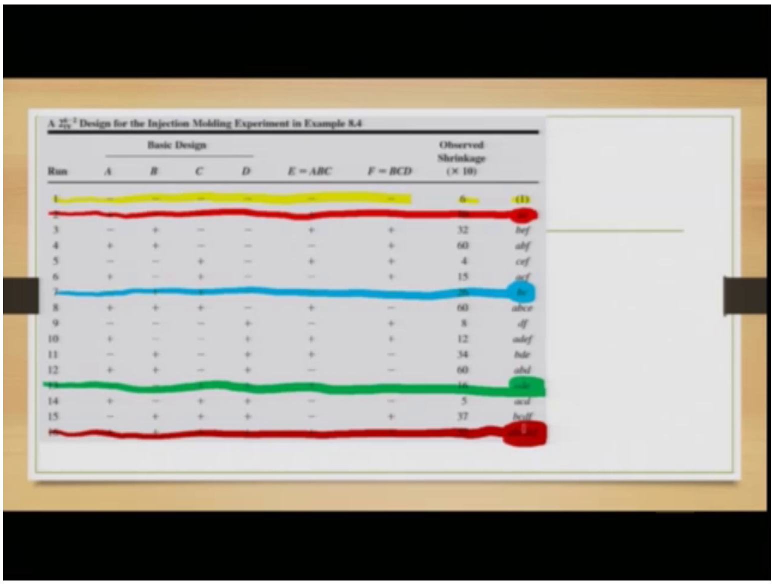
So, these are other three very important characteristics or assumptions which we intrinsically assume for all the examples we are doing the only large effects are basically A which is mold temperature and B, which is screw speed and the interaction of A B which is happening in light of this alias relationship between all these six factors it seems reasonable to adopt this conclusions tentatively and pass judgments accordingly the plot of A B introduction in figure 8.13 which will see later on.

So, note down the figures which I am mentioning it is 8, 9, 8.10, 8 11 so, on and so, forth. So, the plot of A B interaction which means we are taking the effects combined together for factor A and B which are the of the highest level highest level is not they are affecting the most they show the process is very sensitive to the temperature and if the screw speed is at low tempera low level, but very sensitive to temperature it is very insensitive to temperature sorry for that very insensitive to temperature with their screw speed is at low level, but very sensitive to temperature.

If the screw speed is high level with the screw speed at the low level the process should produce an average shrinkage of around about 10 percentage of the around regardless of the temperature level chosen based on this initial analysis the team decides to set both the mold temperature and the screw speed at the low level the set of conditions will reduce the mean shrinkage of parts to about 10 percent; however, the variability in the shrinkage from part to part is still a potential problem we will see that later on.

So, I am just reading the problem then come to the an analysis for that at the data and then the analysis using the charts as we have been doing for all the examples. So, in a affect them mean shrinkage can be adequately reduce by the above modifications; however, the part to part variability in shrinkage over production run could still cause problems in the assembly line one way to address this issue is to see if any of the process factors affect the variability of the part shrinkage. So, accordingly so that let us go to the diagram the tables.

(Refer Slide Time: 09:12)



So, now I will again repeat few things would be repetition please bear with me, but it will make things much clear to you. So, this is a fractional factorization problem of 2 to the power 6 minus t2wo because there 6 factors A, B, C, D, E, F, minus 2 because you are trying to basically find of the generators. Accordingly, they were sixteen runs if you notice the first column on this on this slide. So, that there are run numbers from 1 to 16. .

Now the affects which you are taking for the factors is like this will consider the effects of A, B, C, D and E and F would be a combination of the those four factors which you have considered as A, B, C, D and which is affecting positively and which is affecting negatively, then we will basically have the all of observed sign significance and then give the effects which are coming in terms of the factors by itself. So, let me go 1 by 1 it will become clear what I just did mentioned.

So, consider run one I will just highlight it using a colour blue yellow for run one the effect of A is negative B is negative C, D, E, F are all negative so; obviously, the observed shrinkage is 6 and it will basically the effects for which none of them are significant. Now, if I go to let me use a different color if I go to row 2 which is run 2 I have effect of A as positive E as positive.

So, if I see the overall factors which are affecting and E and E where I am just basically highlighting now if I go to I will use another different color if I go to see for example, the run number seven effect of B C are positive others are negative. So, you see the in the fractional factorization model the effects are coming from B and C if I use say for example, let me use another color see if I use the pay attention to the 13th run.

So, it is positive for C and D positive for E. So, the overall affect is C and D where I am just highlighting and lastly if I see the effect coming out let me use another color where blue let me use the dark one it may be difficult for you to read, but please bear with me if you see the 16th run it is positive for all and the observed shrinkage while is which we already know have been calculated. So, the effects are basically for A, B, C, D, E, F combined together.

(Refer Slide Time: 12:08)

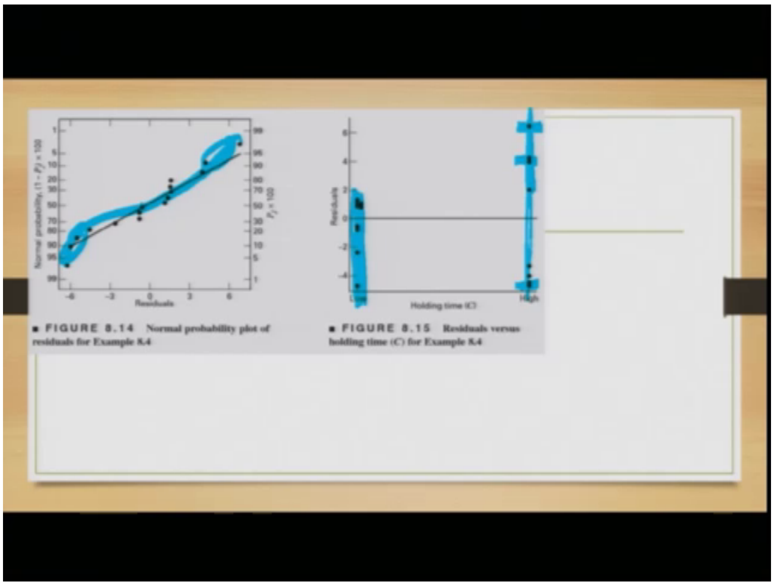
Now, each are at two levels. So, these are the factors that explanation in qualitative sense the levels of minus one and plus one. So, the variables are I will use the same colour the highlight continuing is A, B, C, D, E, F their mold temperature screw speed which I have said again repeat it holding time cycle time gate size hold pressure the minus 1 and plus 1 levels are given. So, far all of them if you remember it is 2 to the power something.

So, that that 2 basically de denotes the 2 levels for all these factors. Now, if I go to the overall averages. So, the overall averages now what I have is basically for A, B, C, D, E, F taken 2 at 1 at a time then the combinations goes of A, B and C, E then A, C and B E; that means, taken 2 at a time then A, D and E, F and then we go to the combinations of taking 2 at a time, but three combinations which is A, E, B, C and D, F. So, all the combinations are given here the regressions coefficients for the runs are given here the estimated effects and the sum of the squares are given here.

So, what is important is the sum of squares which I will basically the highlighted with a different colour. So, these are the sums of the squares which will be important for us and I will just highlight it this is important. Now once I basically have the effect estimates and if you pay attention to 8 po figure 8.12 and 8.13 this is basically like this we are trying doing the normality plots with respect to the effect estimates in 8.12 and we are doing the shrinkage plots with effect mold temperature.

So, shrinkage process basically the effect which are which you want to find out with respect to one of the factors which is mold temperature which is factor a. So, both in both this cases for say for example, for 8.13 they are linear because there are two levels. So, in this case when I am taking the effects of mold temperature which shrinkage at the lower level it is almost stagnant for values of higher levels of shrinkage it is basically increasing linearly, but very interestingly the effects of temperature will normal probability plots or not normal because if you pay attention here this always (Refer Time: 14:45) here and if you pay attention there is huge amount of outliers which are there technically they are not normal, but will stay continue with the example and continue with the discussion.

(Refer Slide Time: 15:02)



If you do the normality plots or residuals for ex for this example so, normality plot with the residuals here it is much better even though at the extremes there are outliers. So, it is basically like this. So, still it is much better and if you do the residual plots along the holding time the overall dispersion is very skewed. So, in low temperature it is as you remember the effect is almost NIL.

So, it is very on the negative side and for the high temperature is more on the positive side, because the effects are much more because if you see this is an effect six there are two force, but here everything is basically between the bandwidth of 0 to minus 4.

(Refer Slide Time: 15:52)

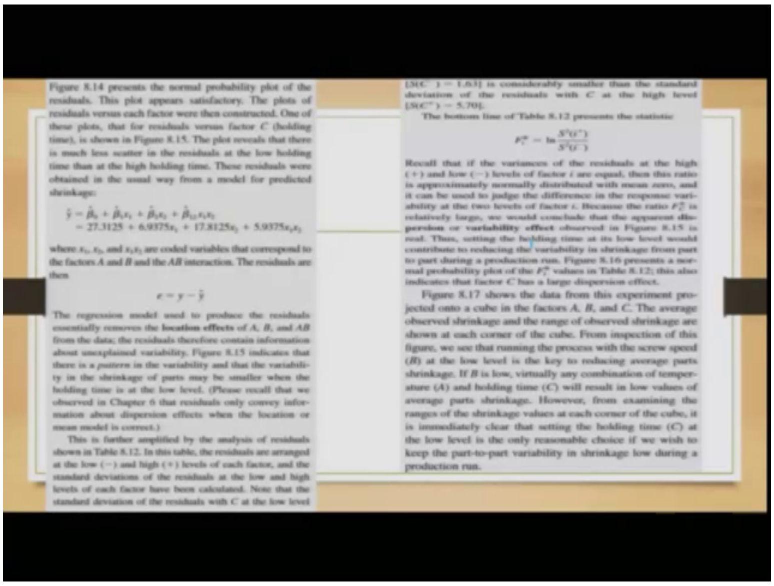


Figure 8.14 presents the normal probability plot of the residuals. This plot appears satisfactory. The plots of residuals versus each factor were then constructed. One of these plots, that for residuals versus factor C (holding time), is shown in Figure 8.15. The plot reveals that there is much less scatter in the residuals at the low holding time than at the high holding time. These residuals were obtained in the usual way from a model for predicted shrinkage:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_{12} x_1 x_2$$
$$= 27.3125 + 6.9375 x_1 + 17.8125 x_2 + 5.9375 x_1 x_2$$

where x_1, x_2, and $x_1 x_2$ are coded variables that correspond to the factors A and B and the AB interaction. The residuals are then

$$e = y - \hat{y}$$

The regression model used to produce the residuals essentially removes the **location effects** of A, B, and AB from the data; the residuals therefore contain information about unexplained variability. Figure 8.15 indicates that there is a *pattern* in the variability and that the variability in the shrinkage of parts may be smaller when the holding time is at the low level. (Please recall that we observed in Chapter 6 that residuals only convey information about dispersion effects when the location or mean model is correct.)

This is further amplified by the analysis of residuals shown in Table 8.12. In this table, the residuals are arranged at the low (−) and high (+) levels of each factor, and the standard deviations of the residuals at the low and high levels of each factor have been calculated. Note that the standard deviation of the residuals with C at the low level $[S(C^-) = 1.63]$ is considerably smaller than the standard deviation of the residuals with C at the high level $[S(C^+) = 5.70]$.

The bottom line of Table 8.12 presents the statistic

$$F_i^* = \ln \frac{S^2(i^+)}{S^2(i^-)}$$

Recall that if the variances of the residuals at the high (+) and low (−) levels of factor i are equal, then this ratio is approximately normally distributed with mean zero, and it can be used to judge the difference in the response variability at the two levels of factor i. Because the ratio F_C^* is relatively large, we would conclude that the apparent **dispersion or variability effect** observed in Figure 8.15 is real. Thus, setting the holding time at its low level would contribute to reducing the variability in shrinkage from part to part during a production run. Figure 8.16 presents a normal probability plot of the F_i^* values in Table 8.12; this also indicates that factor C has a large dispersion effect.

Figure 8.17 shows the data from this experiment projected onto a cube in the factors A, B, and C. The average observed shrinkage and the range of observed shrinkage are shown at each corner of the cube. From inspection of this figure, we see that running the process with the screw speed (B) at the low level is the key to reducing average parts shrinkage. If B is low, virtually any combination of temperature (A) and holding time (C) will result in low values of average parts shrinkage. However, from examining the ranges of the shrinkage values at each corner of the cube, it is immediately clear that setting the holding time (C) at the low level is the only reasonable choice if we wish to keep the part-to-part variability in shrinkage low during a production run.

So, figure 8.14 represent the normality pro normal probability plot of the residuals this plot appear satisfactory the plots of the residual versus each factor. When there when there are constructed would; obviously, give us the information as needed one of this plots that is for residuals versus factor C that is holding time is shown in figure 8.15, the plot reveals that there is much less scattered in the residuals at the lower holding temperature than at the higher temperature high holding time these residuals were obtain in the usual way from the model for the predicted shrinkage.

So, the shrinkage models what we have is basically y is equal to beta naught plus beta 1, x 1 plus beta 2 x 2 plus beta 12 in x 1 x 2, because you have now basically two factors and; obviously, there would be error term. So, the best fit areas are the best fit estimate which we find out for beta naught which is beta naught hat is 27.3125 for beta 1 hat is

6.9375 for beta 12 hat is basically comes for sorry sorry my mistake for be beta naught it comes out to 27.3125 for beta 1 hat it comes out to be 6.9375 for beta 2 hat it comes out to be 17.8125 and finally, for beta hat 2 suffix 1 to it comes out to be 5.9375.

So, these are x 1, x 2, x and x 1, x 2, x 1, x 2 separately x 1, x 2 and x 1, x 2 combine are the effects which is happening for A for B and for A B the residual adores or errors as we know is basically y minus 1 y hat which is the actual value is the predicted value. So, the regression model used to produce the residual essentially removes the location effects basically we are trying to find out and remember that figure 8.15 which we just discuss gives us the pattern or the type of relationship which already exist. So, will be continue reading it figure 8.15 indicates the these the pattern in the variability and that the variability in the shrinkage of the parts may be smaller when the holding time is at the lower level than at a higher level.

So, we will basically continue reading it with basically means the analysis of the residuals which are shown in 8.12 in this table the residuals are arranged at the lower level and higher level for each factors and the standard deviation for the residuals of the low and high level of each factors at r to be calculated have been calculated note that the standard deviation the residuals with c at the low level is considering smaller than the standard deviation of the residuals of the higher level based on that if you do the calculation we can find out what is the statistic which is used which is basically f I star where F i star would basically will be calculator it using the logarithmic of the ratios of the standard level squares at positive level and standard level squares with at the negative level.

That means at the higher level and the lower level [vocalized-noise] recall that if the variance of the residuals at the high and low levels of factors are equal, then this residual approximation normally distributed with a mean 0 and it can be used to judge the difference of the response variability all this, because the ratio of F star which we just found out and which we just discussed is relatively high we would conclude that happen and dispersion of the variable effect observed in this fig in figure 5.8.15 is d n.

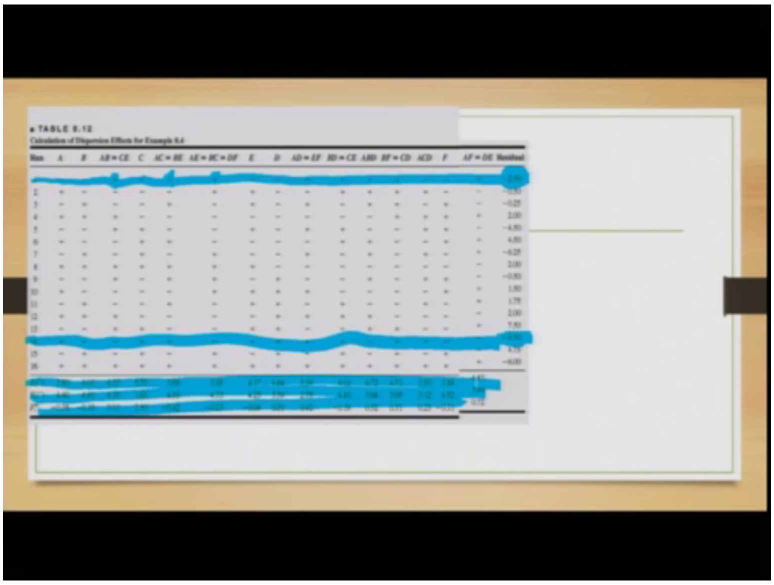
The setting the holding time at a low level would contribute to reducing the variability in the shrinkage from part to part during the production runs hence figure 8.16 represents the normality normal probability plot of F i star values in 8 table 8.12 this also indicates that factors C has a large dispersion effect with respect to the other factors figure 8.17 shows the data from the experiment pro projected onto a cube in the factors of A, B, C which are the factors which are significant the aver average observed shrinkage and the range of observed shrinkages are shown and each corner points of this cube.

So; obviously, they would be combination based on which at the cube points which will give the effects of these factors either this A, B or C or whatever combination you want depending on the outcome of the experiment. So, it is not arbitrary basically depends on the way that the problem has been analyzed and what information we are getting from the problem from inspection of this figure we see that running the process with the screw speed B at a low level is the key to reducing the average for shrinkage.

If B is low virtually any combination of temperature A and holding C will result in low values of average parts shrinkages highway from examine the range is of shrinkage values at each corner of this cube it is immediate to clear that setting the holding time c at the low level is the only reasonable choice if we wish to keep the part to part variability in shrinkage low during the production process.

So, all this things reading which have done I will just go to through do through these tables once.

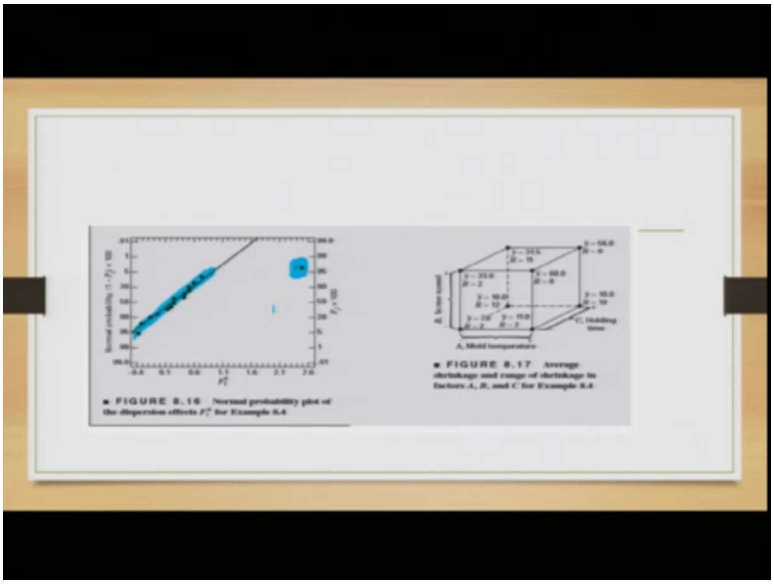
(Refer Slide Time: 21:07)



More now if you see the table which is 8.12 you have the runs on the leftmost column which is 1 to 16 and the combinations which I have are basically the factors affects which is coming out. So, if you basically pay attention say for example, to the first run you have the effect of A and B as minus then C as minus E and D as minus if we can find the overall effect the positive value is would give you the positive effects which we have based on which we find out the residuals at 2.5.

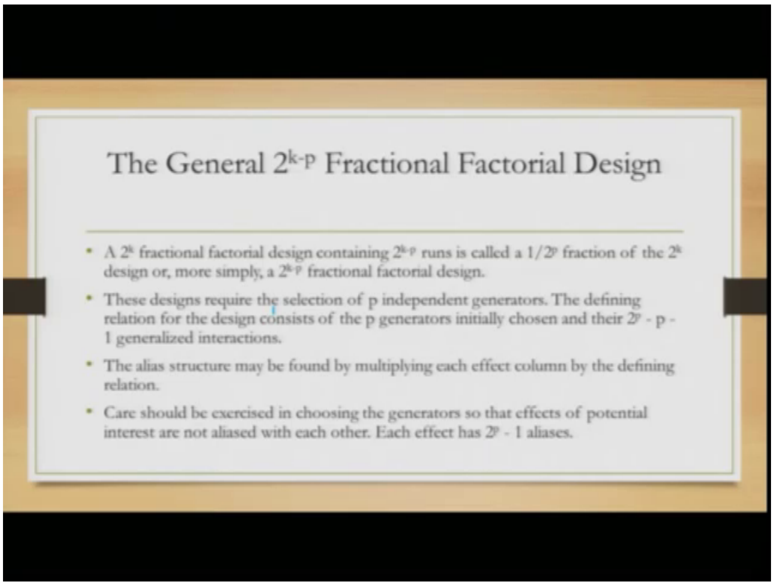
We consider arbitrarily again the 14th row the effects are positive for the values which I am which I am man mention for a I will mention some of them for several example c on so, and so, forth the value comes out to be minus 5.50. So, based on that I have the positive standard deviation the negative standard deviation and based on that I find out the F statistic as given by the ratio of logarithm values of s square positive divided by s square negative negative and positive means for the positive dispersion movement and the negative movement when we consider the normality plot or the dispersion effects of F star.

(Refer Slide Time: 22:22)



So, they are decently normal apart from one outlier. So, this is almost normal with one outlier here. So, I am basically plotting the normality plot with respect to when in the x axis here F i star if I do and consider figure 8.17 which is the average shrinkage and range of shrinkage in factor A, B, C com combine together so; obviously, you will have positive effect 2 levels for factor A 2 levels for factor B 2 levels for facer factor C. So, hence the cube and at this point are the cube what the values which we have which will give us the combinations of A, B, C which are important for us to consider in the example.

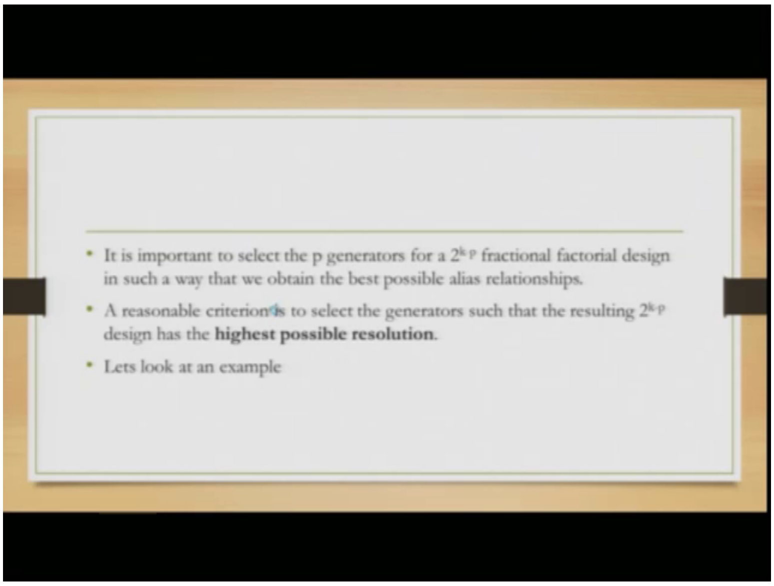
(Refer Slide Time: 23:15)



Now, if we remember that I did mention that we basically generis to the factors of 2 to the power k minus m. So, here m is basically p . So, general 2 to the power k minus p fractional factorial design problems and will consider the basic notions about that our 2 to the power k factorial fractional factorial design consists of 2 to the power k minus p runs is called 1 by 2 to the power p fractionals of this 2 k model, because we have key factors trying to reduce them by the power of 2 to the power p such that we are able to subsume the overall significance on all the factors to the to them maximum efficiency considering the word efficiency where we are able to subsume the effects of least number factors give us giving us the maximum amount of information.

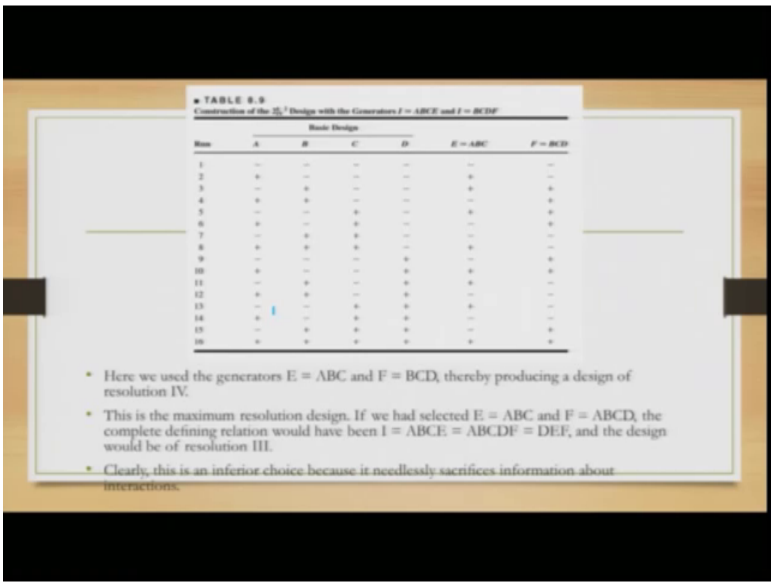
So, or simply will basically mentioned it 2 to the power k minus p fractional factorial design these designs requires selection of p independent generators from this k the defining relationship with the design consist of p generators initially chosen from this set off of generally introduction which will see the alias structure may be found by multiplying each effect to column by defining the relationship k should be taken to exercised in choosing the generators. So, those effects of potential interest or not aliased with each other's hence will basically considered 2 to the power k minus p depending on the alias structure and the generator concept.

(Refer Slide Time: 24:46)



It is important to select the p generators form for a 2 to the power k minus p fractional factorial design in such a way that we obtain the best possible alias relationship reasonable criterion is to select the generators such that the 2 to the power k minus p design has the highest possible resolution.

(Refer Slide Time: 25:09)

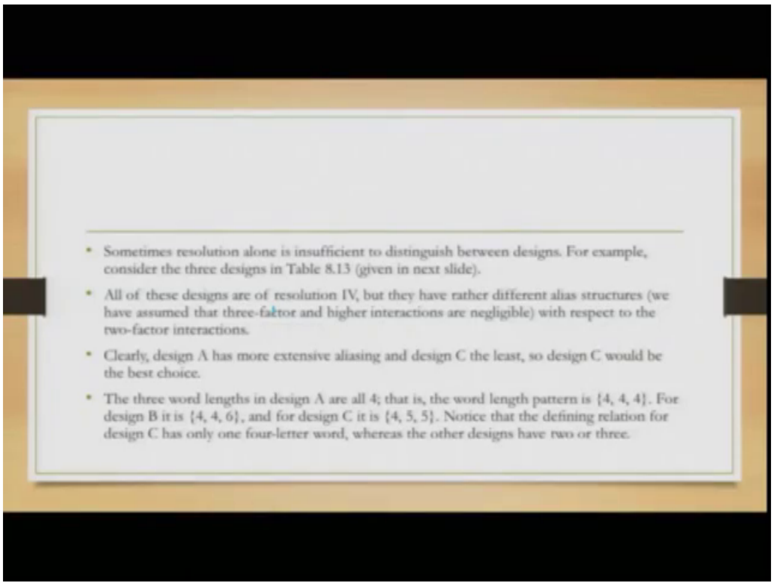


So, let us look at an example accordingly. So, here we have the constructions are 2 to the power k minus 2 ah. So, k minus p so, k was basically 6 p was basically 2. So, this is the 2 to the power 6 minus 2 suffix 4 design with generators I which was ABCE and I which would BCDE so; that means, we considering two generators the runs for 16 in number basic design factors are given for the column for the row, which is ABCD and the generator which is E and F here we used generators E and generators F there by producing a design of resolution 7.

So obviously, of 4 so; obviously, the resolution would depend on number of generators this is the maximum resolution design which we can basically get if we select E E as ABCF at ABCD the complete defining resolutions would be basically have three generators which is ABCE ABCDE and DEF and it would be a resolution 3 3.

So, depending on how many such generators we have and what is resolution you can basically design the generator generators according clearly this is an inferior choice because it and needlessly sacrifices information about interaction, because you are trying to basically have more generators and sacrificing the effects of individual factors in the analysis of the problem.

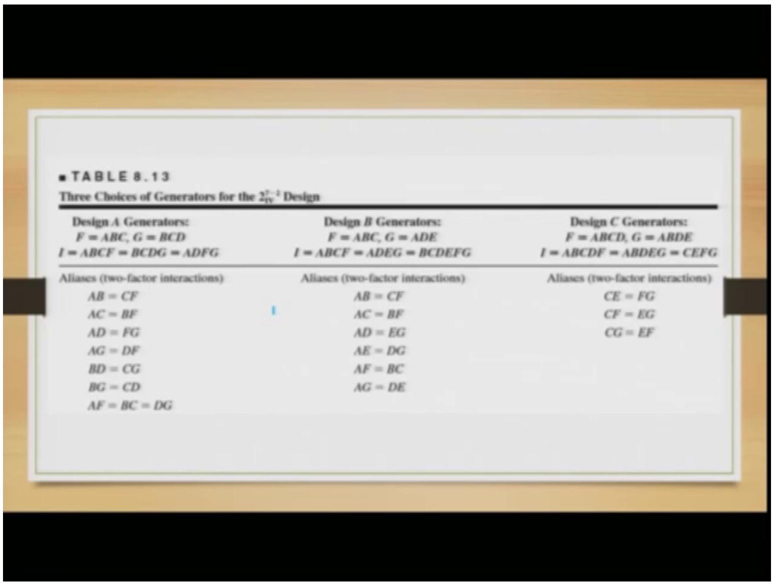
(Refer Slide Time: 26:34)



Sometimes resolution alone is insufficient to distinguish between designs for example, consider the 3 design designs in a problem as discussed in table 8.13 all of these designs are resolution 4, but they have rather different alias structures and as we are assumed that three factors and high international are negligible with respect to the two factor interactions clearly design a has more existing extensive aliasing and design C at the least.

So, design C would be the best choice they am based on the on the assumptions which have the three word length in design A are basically four then the word length would basically be 444 depending on how you are trying to analyze for B to be 446 and C to be 455 depending on the overall structure we have analyzed for the problem. So, notice that def defining relationship for design C has only 14 letter word where as the other designs have 2 or 3 depending on the problem formulation which have done so; obviously, A, B, C can be interchanged. So, the output would be designed accordingly. So, it a would not be 444 anymore it will be depending on how the design has been done.

(Refer Slide Time: 27:49)


■ TABLE 8.13
Three Choices of Generators for the 2^{7-2}_{IV} Design

Design A Generators: $F = ABC$, $G = BCD$ $I = ABCF = BCDG = ADFG$	Design B Generators: $F = ABC$, $G = ADE$ $I = ABCF = ADEG = BCDEFG$	Design C Generators: $F = ABCD$, $G = ABDE$ $I = ABCDF = ABDEG = CEFG$
Aliases (two-factor interactions)	Aliases (two-factor interactions)	Aliases (two-factor interactions)
$AB = CF$	$AB = CF$	$CE = FG$
$AC = BF$	$AC = BF$	$CF = EG$
$AD = FG$	$AD = EG$	$CG = EF$
$AG = DF$	$AE = DG$	
$BD = CG$	$AF = BC$	
$BG = CD$	$AG = DE$	
$AF = BC = DG$		

So, three choices of generation generators for 27 minus 2 of resolution 4 would be designing A generators. So, the generators are basically F G and I is basically ABCF generator 1; BCDG generator 2; ADFG which is generator 3.

So, now there are 7 factors which is basically ABCDEFG based on that as I mentioned the generators are ABCF, ABCDG and ADFG in the design of design B generators we have basically the generators given as I is ABCF ADEG and BCDE F G. So, based on the combinations of how we have assumed FNG. So, come if we considering the first problem F was ABCG was BCD in the second example of design factor based on B generations you will basically have F as ABC and G as ADADG and if I can come to the design of c generators f is ABCD and G is ABDE. So, hence generators would be ABCDEF, ABDEG and CEFG. So, we will consider these examples in motive details in that in the 27th class.

And with this, I will end this lecture. Have a nice day.

Thank you very much.