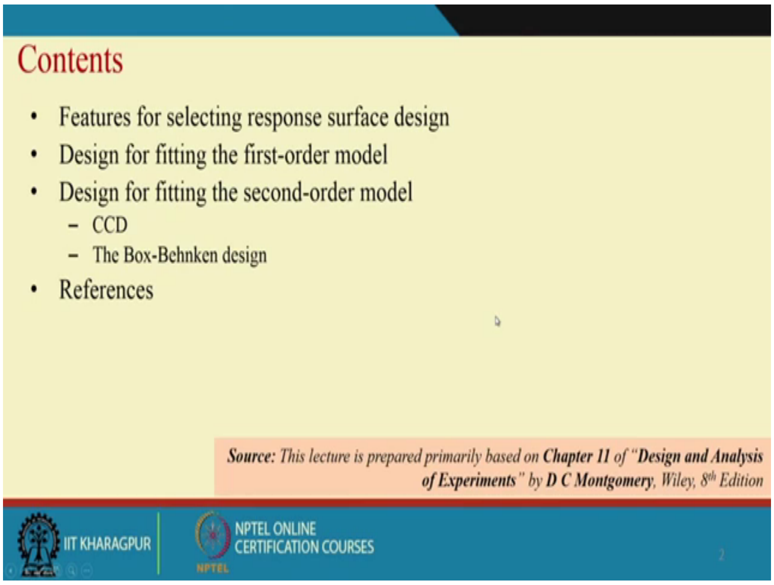
Design and Analysis of Experiments
Prof. Jhareswar Maiti
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture - 54
Experimental Design for Fitting Response Surfaces

Hello, welcome today's topic is Experimental Designs for Fitting Response Surfaces Experimental Designs for Fitting Response Surfaces.

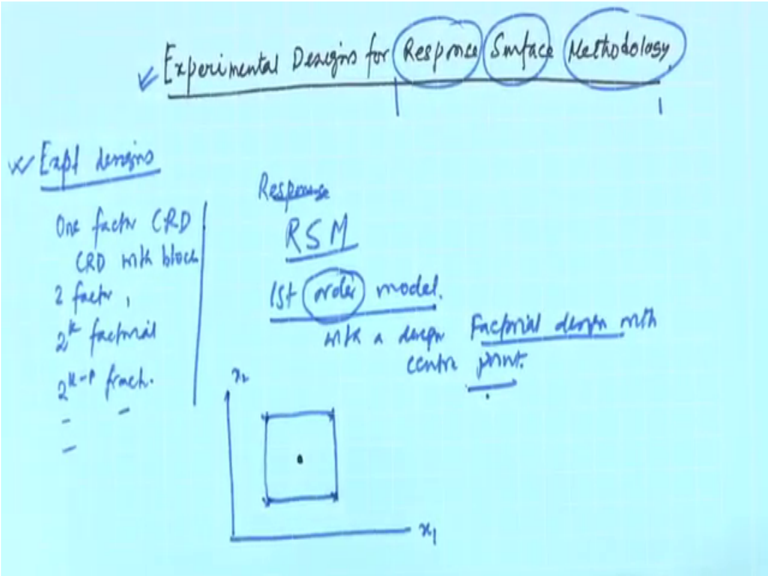
(Refer Slide Time: 00:29)



So, the contents of today's representation are features for selecting response surface design for fitting; first-order model designs for fitting, second-order model under this. We will discuss center composite design and Box-Behnken design and there are some references primary source of shell material is chapter 11 of the book entitled agenda analysis of experiments by Montgomery which is published by Wiley, 8th Edition.

So, let us first look back little bit. So, when we talk about experimental design designs in the.

(Refer Slide Time: 01:10)



First, few classes I have explained the names and then thereafter; what happened? We have elaborated all those designs and that experimental data based on those designs did need to be analyzed and the analysis schemes were also given to you. For example, we have started with one factor randomized design complete randomized design. Then complete randomized design with blocking, then we have gone for two factor design three factor multi factor design then special class 2 to the per k factorial design, then 2 to the per k minus p fractional factorial design and so, many things.

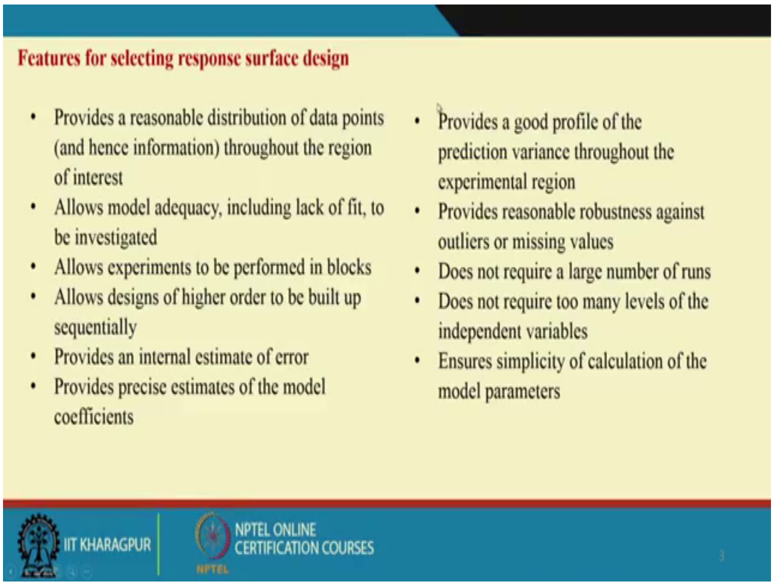
So, thereafter we have started the; in last lecture we have started with response surface metrology response surface response surface methodology RSM response surface methodology. And, I have explained the first-order model first-order model with a design called factorial design with central point factorial design with central center point.

So, into today's lecture, we will discuss those popular designs experimental designs which are huge to feed response surfaces. So, that is, why? I have written that design for fitting response surfaces or experimental design for response surface methodology. We will start with the first factor model first-order model, and then we will go for second-order model. What are the different kinds of designer needed? So, if you look back the first-order model, what first-order response surface?

There we have used factorial design with centre point these are the factorial points and this is central point and in code coded term x 1, x 2; obviously, with reference to two factors. So, there are will be more factors, and then it will be basically factorial design with center points. So, at the factorial points usually one run each experiment is conducted, but at the same point several experimental runs are conducted and it is a very important point, because at this point the data will give us to estimate the m s e mean square mean square error which is basically also a estimate of sigma square ok.

So, we will go by this manner that first we will see that what are the features?

(Refer Slide Time: 04:51)



That is important for selecting response surface design. So, the design should provide a reasonable distribution of data points throughout the region of interest. So, there is the operating region operating zone of interest or re region of interest for the behavior of y; which is response variable and which is known to the process engineer of the experts. So, you your design should give valuable information on the behavior of the response at the at any point on the design surface or design space. So, design allows model adequacy including lack of fit to be investigated.

So, that mean the better what you collect that master represent the population and must be sufficient in size. So, that the adequacy of the model used for analyzing the data will be found out and also at the same time the lack of feet; I think if you can recall the regression lectures, where we have described lack of fit. So, lack of fit also to be estimated, then the design would be such that it allows experiments to be performed in blocks very important, because you know randomization blocking and replications are the three basic principles and designs the any experiments they are dependent or useful resources including raw materials.

So, you require to do either intentionally blocking or you are forced to do blocking, but blocking is an important one and your design should allow you to do this designed of higher order to be built up sequentially; that is, what we have discussed in 2 to the power k factorial? 2 to per k minus p fractional factorial design so; that means, you design should be sequential. So, that the higher the order of interactions; if you require to estimate those orders of interactions it can be estimated the design provides an internal estimate of error.

So, this is that a part of this is basically the error part. So, estimation error is very important it is it is basically that between levels within levels all kinds of errors must be estimated provides precise estimate of the model coefficients. So, what does it mean? The beta value regression coefficient or the effect in ANOVA you use this should be this should be estimated in such a manner that the variance of those estimates should be the minimum.

So; that means, that your design should allow you to I will the minimum variance estimates design should provide a good profile of the prediction variance throughout the experimental region. So, when you pray using the analysis of the model you employ for analyzing the experimental data; that model ultimately will give you the response surface kind of thing and you want to predict or you will get regression model you want to predict the output or response value. Now, the out the predicted variable value that must be within having the variance part that also should be should be at the minimum level.

So, what do you want that the data what you get out of the designed experiment it will allow you to have or develop a good profile of the prediction variance and; obviously, throughout the experimental region or design space design provides reasonable robustness against out layers or the missing values this is another important feature, because many a times what happened you may encounter outliers or some available reasons which causes your experimental runs to be to be skewed to certain direction or you may miss some values to record properly under such situation the missing data handling and outlier removal.

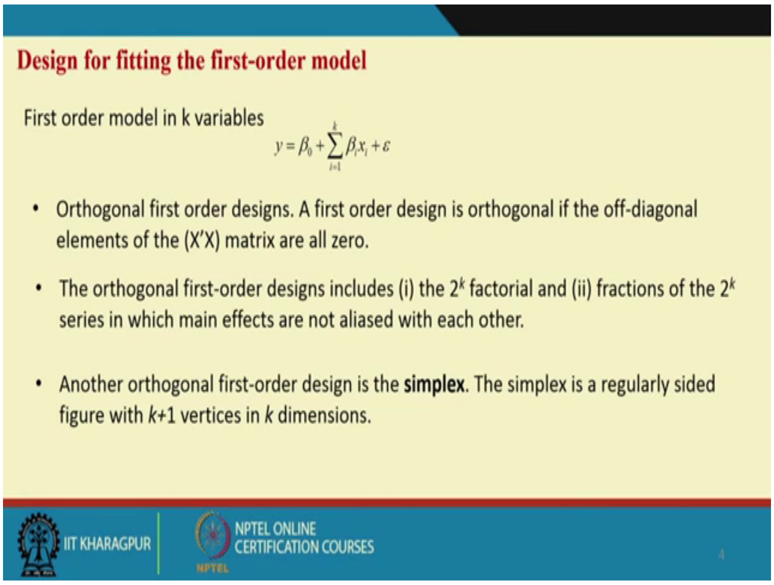
All those things are required and at the same time your even if all those things are done the out outcome of this will be. So, robust that you can use the experimental data for the objectives for which the design was done. So, design does not require a large number of rounds large number of runs means costly. So, it should be it should be the minimum possible runs; obviously, representative one. So, there that sample size determination we have I think three late to two lectures on some percent determinations.

So, you have seen that how the sample size is determined. So, what you want your sample size should be representative, but at the same time it is not be. So, high that it we it will there will be cost overrun design does not require too many levels of the independent variables suppose you have 5, 5 factors and if each factors have been 5 levels each. So, there will be 5 into 525 retreatment combinations.

So, this is probably little exhaustive we do not want such high level of levels or. So, many treatment combinations with minimum level the hot maximum information that can be achieved. So, that should we got from the; that we must get from the design you imply design ensure simplicity of calculation of the model parameters. For example, you will be able to use ANOVA, because ANOVA is a very simple one easy to understand very powerful effective.

So, it should be your data should be such that it will it will help you doing ANOVA; for example, sometimes if you consider continuous factors along with categorical factor at different levels. So, what will happen you may require analysis of covariance of some other kind of model, but it is always better if you can make it simple? So, that may be the simple ANOVA will work there. So, please keep in mind all those features while choosing a design.

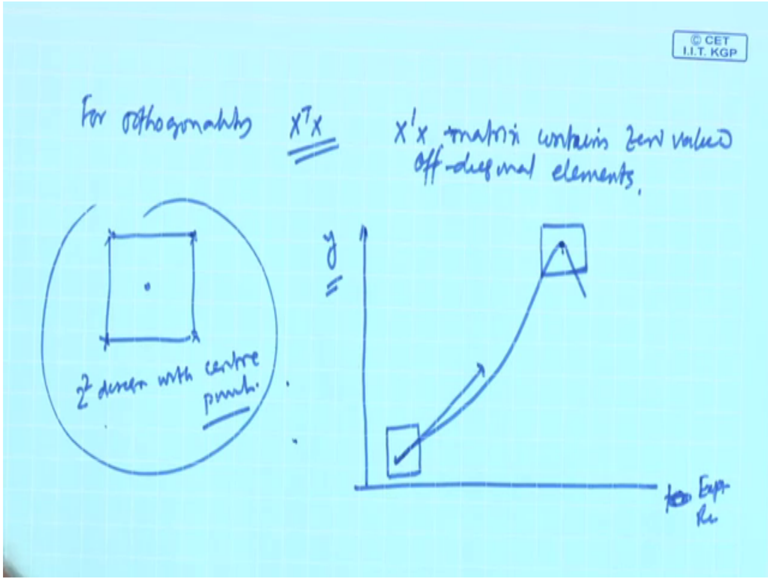
(Refer Slide Time: 12:12)



Now, let us discuss first-order model and this model you have seen earlier and here what happened? We say that we want to adopt orthogonal design. Now, I have given you X matrix, which is basically design matrix and we have shown you earlier, that if you know with coded variable that if you multiple if you take the dot product of two up column of the design matrix including the factors of interest you will find out that that product is 0, which means that orthogonality is obtained.

So, now, you can you can find out you by computing the X transpose X, if you if you your design matrix each X, then X transpose X when you compute if the design is orthogonal then, what will happen the off diagonal elements of X transpose X become 0.

(Refer Slide Time: 16:12)

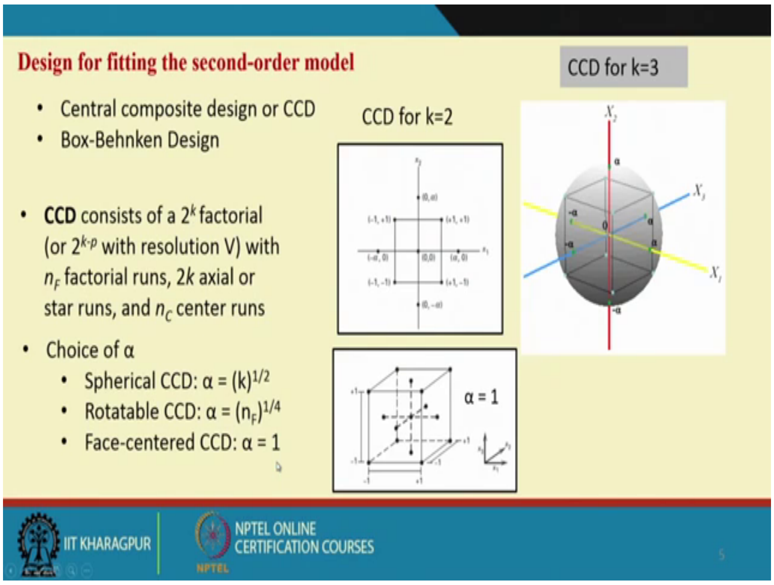


So, for orthogonality: for orthogonality X transpose X or we are defining like this X transpose X matrix contain zero off diagonal elements valued off diagonal elements.

So, for first-order model 2 to the power k factorial or 2 to the power k minus p fractional; factorial design is sufficient the reason is you are, what you require in first-order model? You require that the main effect should not be aliased with other main effects, because your model is first-order only main effects you have considered. So, 2 to the power k factorial or 2 to the power k minus p fractional factorial will work there like the factorial design with center points.

This is a very good design; that means, this 1 is 2 to the power 2 design with center point. And there is no replication at the center point at the factorial points and another design that is also used in first-order is known as simplex will not discuss this, but the simplex by definition is a regularly sided figure with k plus 1 vertices in k dimensions.

(Refer Slide Time: 15:02)



So, we will come to the second-order response surface designs; I have explained the first-order response surfaced in plus two classes. So, I told you that when you achieve the using first-order response surface model you achieve the point of optimum. So, we you if you start from here and you ultimately your direction will be given by the first-order model and then what happened your that is experimental runs here you start running.

So, the experiment and you get the y result and ultimately this is the zones where you require second-order model, because first-order will not be fit here and to understand that whether optimality is reached or not. So, that mean the in the second-order, we require second-order model here and second-order model. Ultimately, has more number of parameters and such design like these 2 to the power to case with center point or 2 to the power k factorial, 2 to the power k minus p partial factorial will not to be able to provide sufficient data. So, that the parameters of a second-order model can be estimated all the parameters second-order model can be estimated. So, you require more information you require special kind of design and those designer discussed here.

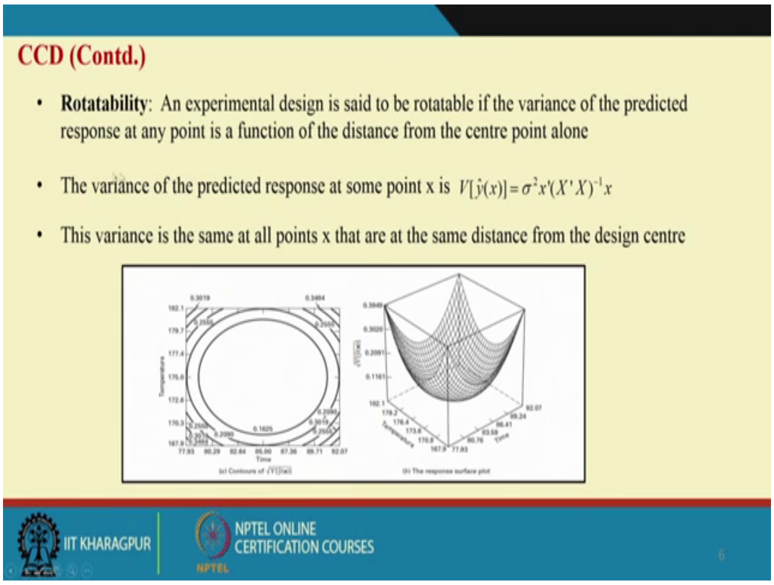
Now, see we are saying that central composite design and Box-Behnken design this one is central composite design; what is central composite design? So, central composite design has three kind types of three types of points I can say what is this one is factorial points which are the corner of this rectangle axial point and the center point and axial points. Please understand this with reference 2 to the power 2 design k equal to 2 with reference to 2 to the power to case here you have 4 factorial points, 1 center point and 4 axial points and the distance from center point to any of the axial point is alpha.

So, you required to find out the value of alpha; that is important this axial point is also known as star point. If it is 3 per factor case, 3 factor cases then you will be getting the cube and there will be 8 factorial points 1 center points and; obviously, how many faces are there 1, 2, 1, 1, 2, 3, 4, 5, 6. So, accordingly and the axial point soon or. So, will be chosen axial point will be chosen along the 3 dimensions both side and it will be alpha distance apart from the center point plus minus alpha distance effort from the solution there are when there are 2 factors, 2 axis; so 2 into 2, 4 axial points; 3 factor, 3 axis; 3 into 3 into 2; 6 axial points.

Now, issue is that what will be the alpha? Alpha will is this is chosen and depending on the value of alpha the design central composite design also classified as spherical central composite design. When alpha equal to k to the power half rotatable central composite design, when alpha equal to n F to the power 1 by 2, where n F is basically the number of factorial runs and face centered CCD when alpha equal to 1; that means, that mean the axial points are all on the middle of the side of the polygon ok.

So, CCD have three kinds of that is points where the or arcic are away I can say settings where the experiments need to be conducted 1 is factorial points, another one is center point, another one is axial and star points. Depending on the k value of number of factors the number of runs as will be chosen depending on the on the alpha value chosen it can be of spherical CCD it can be rotatable CCD it can be face enter CCD.

(Refer Slide Time: 20:13)

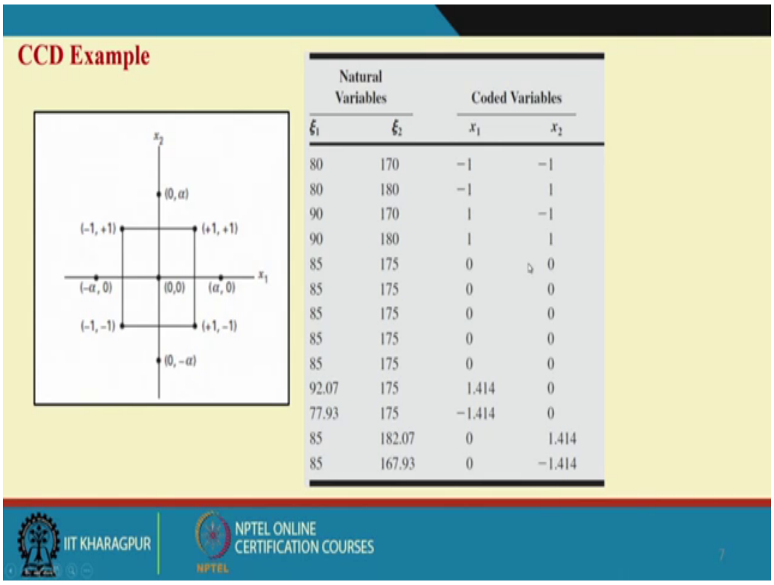


Now, let us understand what is the concept of rotatability, rotatability and experimental design is to be rotatable if the variance of the predicted response at any point is a function of the distance from the center point. For example, you consider this figure here temperature and time and is plotted along x and y and along z the response is variance parties response part is plotted. Now, what happened what is the center point? What is the center point here? You have already seen.

So, now, you take any distance any distance from the center point and then you measure what will happen you measure the variability of the predicted response at a point and all those points which are equidistant from the center then the variance will be same that is the rotatable one if from through corner the contour plot you can understand that if suppose this is the center point and then this is; what is the distance? So, you see that the very there what happened the; variance part along this contour line these are basically same.

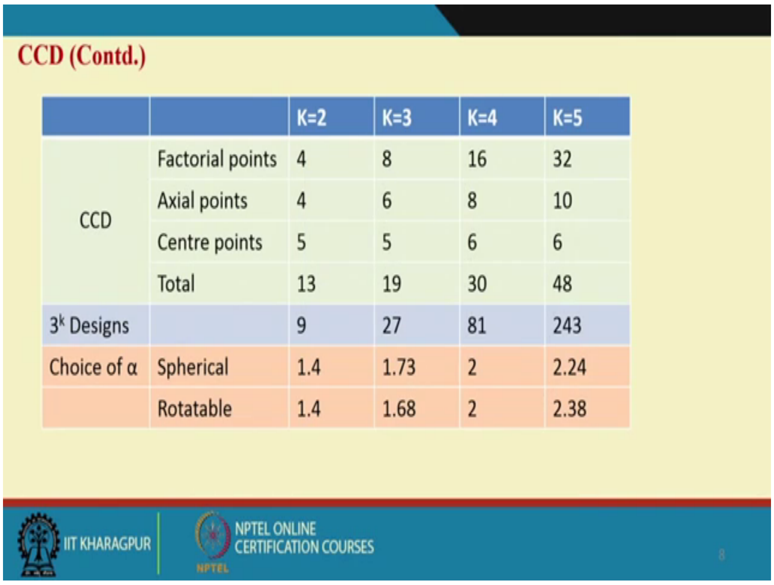
So, they are same distance apart this is one, now how do you compute the variance of predicted response at any point each sigma square X transpose X transpose X inverse X this is what we have discussed earlier and in regression time also we have discussed this variance is the same at all points that are at the same distance from the design center ok.

(Refer Slide Time: 22:08)



So, now this is one example of CCD. So, you see that this is a 2 to the power 2 factorial point. So, 4 factorial points that say minus 1 minus 1 minus 1 plus 1 like this and then there are center point is 85 and 175, runs at center points and then there are 4 axial points runs are there. So, if you conduct experiments based on all those experimental settings either through; that means, if you can very easily understand from this coded variable and this can, because this week must come from the natural variable and you conduct the experiments accordingly and you will get the y data that is the response data and this response data can be used use actually this design element response data can be will be used to feed the second-order response surface ok.

(Refer Slide Time: 23:18)



Now, here is a comparison for example, whether should we go for theta over K design or CCD? 3 to the power k means every factor at three levels. For example, what you are doing here, if you see that go along any di any dimension suppose x 1, what are the design points? 1, 2, 3 here also 1, 2, 3, in addition here these two are there this point this point is there.

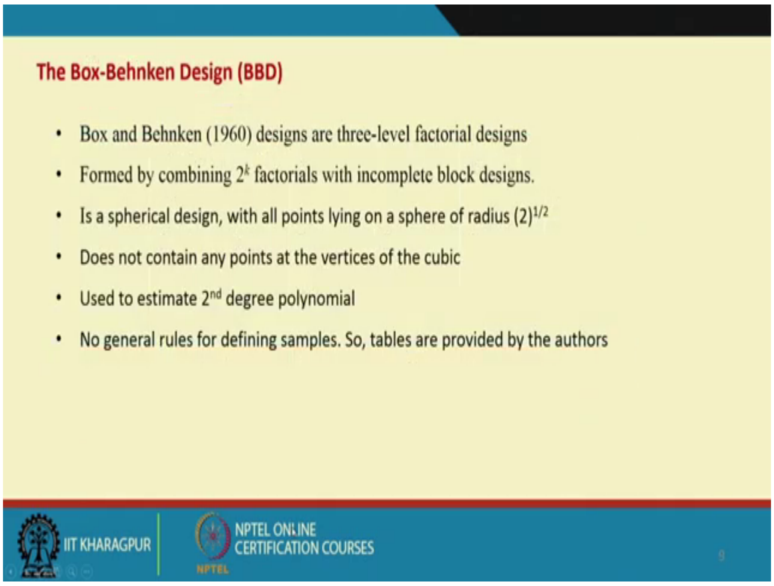
So, what we are saying that instead of this if we say x 1 or 3 levels and x 2 at 3 levels, then because then we will get this design 3 to the power k design. Now, here what we are comparing? We are comparing that number of experimental runs, when K equal to 2 to K equal to 5, if you go for CCD and if you go for 3 to the power k design, what is the difference?

So, from number of runs point of view if your number of factor is 2, then CCD requires 13 runs if your central point run is 5 in number whereas, 3 to the power k design requires 9 runs, but when you go for K equal to 3, 4, 5. Immediately, you see that the reverse picture the number of runs required for CCD is much less than the number of runs equal for 3 to the power k design. So, what we can say here from this that the we not to fill the second-order surface, CCD is an efficient and better one compared to this provided and the purpose for which the model is building sir means you will get the enough information on the on the on the intelligent space and that is what is possible also using CCD.

Now, another one important one is; if you choose alpha equal to that 2 to the power half, then actually k to the power half, then with spherical design when K equal to 2 and K equal to 3, K equal to 4. The values are given and if you say no at, we want the rotatable one rotatable one means the variance at the disto predicted or response at a distance from a equidistance from center point will be equal.

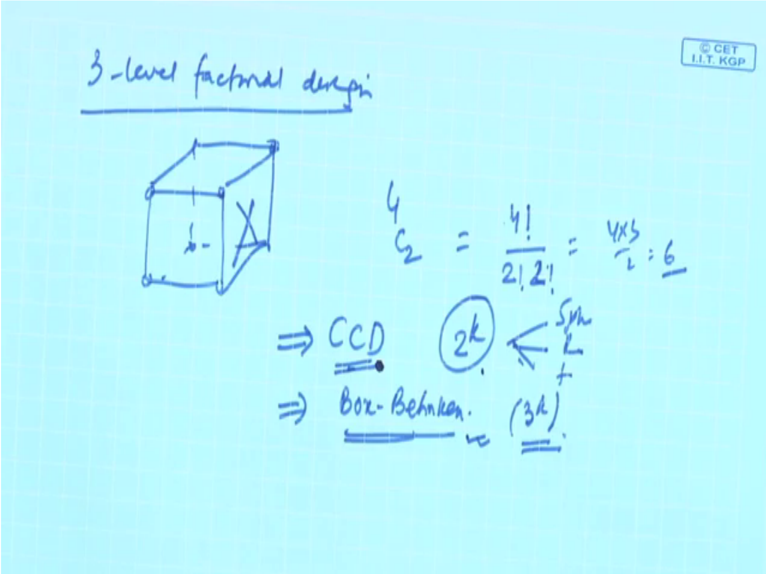
\So, then if you compare this what you are getting the alpha value they are more or less same similar so; that means, a rotatable design and a spherical design spherical design rotatable one and rotatable one is spherical on that sense. Now, with little and little marginal difference it is almost equal at least up to K equal to 5 that is what is given in this slide?

(Refer Slide Time: 26:17)



So, fine, but what will happen if you if you require 3 to the 3 level.

(Refer Slide Time: 26:28)

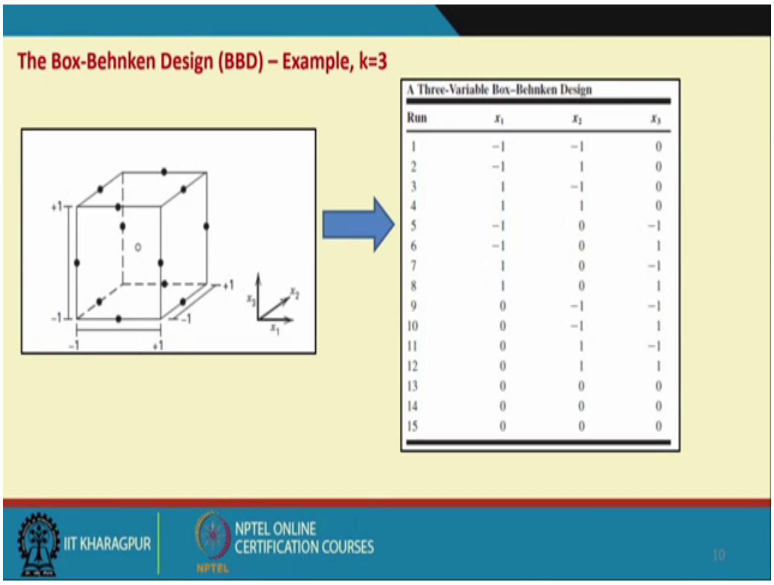


Factorial design 3- level factorial design we have seen that till if I go for 3- level factorial design instead of 2- level factorial design and if I employ, CCD and if the purpose is served then, what happened we are we are peppering CCD CCD way for to level factorial design, but the situation is such that 2-level will not give you the required information you require to go 3-level. So, under such situation, what happened? Your number of experiments will be large.

So, how can you make it efficient from number of experiment point of view and also at the same time you will get the required information, so that the purpose for which the data is collected will be solved. So, one such design is given one such approach is given by Box-Behnken and the designs adopted based on Box-Behnken approach are known as Box-Behnken design or b b d the historical perspective of Box-Behnken. Box-Behnken design is that Box-Behnken Box-Behnken designs are 3-level factorial design developed in 1960.

Actually, these are formed by combining to the power of k factorials with incomplete block designs. So, that mean you 2 to the power k factorials design and then do another one and finally, you combine in; that means, in the sequence you can combine them in a spherical design with all points lying on the fair it is a spherical design and with all points lying on the sphere of radius 2 to the power half does not contain any points at the vertices of the cube. So, know that vertices mean we have seen in factorial experiment that in the cube the vertices there are there these are the settings so, but here what happened does not contain any point on the vertices of the cube. So, this is not correct. So, then used to estimate second degree polynomial; obviously, a second-order response surface can be fit no general rules for defining samples. So, tables are provided by the authors; that mean, the sample size and other things. So, it is a group of tables are provided by Box-Behnken; Box-Behnken and the starting point. So, we can we can use those tables for experiments.

(Refer Slide Time: 29:19)



A Three-Variable Box-Behnken Design

Run	x_1	x_2	x_3
1	−1	−1	0
2	−1	1	0
3	1	−1	0
4	1	1	0
5	−1	0	−1
6	−1	0	1
7	1	0	−1
8	1	0	1
9	0	−1	−1
10	0	−1	1
11	0	1	−1
12	0	1	1
13	0	0	0
14	0	0	0
15	0	0	0

Here one example for Box-Behnken. So, we have 3 levels; so x 1, x 2, x 3, 3 variables.

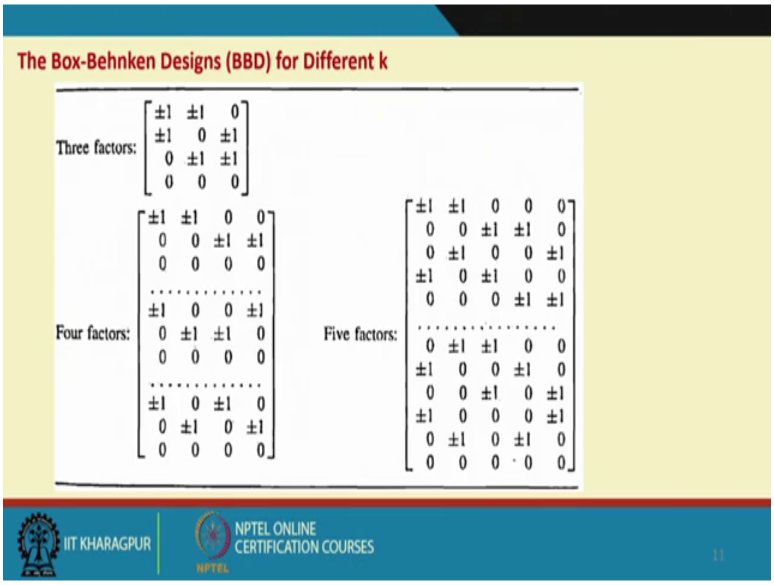
So, you see that ultimately this is, what is the design how you are doing it you are you are considering you see that there are three factors: x 1, x 2, and x 3 and you are you are basically at a time you are basically choosing the 2 to the power 2 factorial design for a run for example, run 1 x 1 at minus 1 x 2 at minus 1 and x 3 at 0. So, this is run 1, run 2. So, I means first what you do you find out 2 to the power 2 factorial design for the first two variables and the say third one you put equip and put it zero means it is at the center point.

So, then you go for the second's two variables and the third one will be at 0-level and in this manner you go. So, if you proceed the; in this manner what happened there are three variables; x 1, x 2, x 3 and first you consider x 1, and x 2 and you know that two to the power 2 means 4 factorial runs so; that means, you have 4 runs for first combination x 1 and x 2, where x 3 will be 0 second combination x 1 and x 3 x 2 will be 0; that is, another 4 and 3 combination, x 2, x 3 and 1 one is at zero level that is another combination another four.

So, altogether that mean you require 12 12 experiments to be conducted in this manner. In addition, what happened in addition what happened you require experiment at the center point ok? So, that we minus 1 low and plus 1 high, but 0 is basically the middle value the center point value. So, first well in this case for K equal to 3, first 12 experiments are done by considering 2 to the power two factorial design for three combinations x 1, x 2, x 1, x 3 and x 2, x 3 and the last here last three experiments are conducted conduct will be conducted using the center point.

So,. So, what I mean to say here? You require here 15 experimental runs. So, if you say no I do not conduct three experimental runs at the center point only one you want to condor, then at least 13 you require, but at the say it is customary that the central point you consider more number of runs and, because this will give you this will this will allow you to estimate error independently and your tests other things will become more robust and at the same time, what happened the center point is; the point which is the most familiar to the operators.

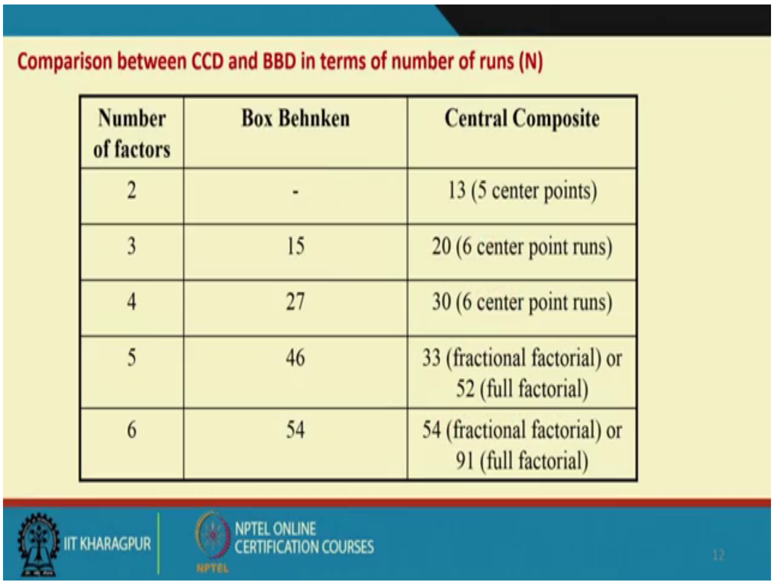
(Refer Slide Time: 32:34)



So, using this analogy what happened the three factor tables, four factor tables, five factor tables are provided by the author. Here you see there are three factor tables, what I have explained plus minus 1 plus minus 1. I mean this is the first one is the combination for 2 to the power 2 factorial design for x 1, x 2 and x 3 will be at 0; that means, 4 observations is run required. Similarly, second one, third one, 4 into 3, 12; 12 plus 1, 13 minimum, but at this level we require more maybe 3 to 5, then in case of 4 factors, what happened you see? What you have done first any 2 to the power 2 here, 2 to the power 2, then here 2 to the power 2 so; that means, 2 to the power 2 combinations you make.

So, how many you out of force factors you require two factors at the factorial case. So, that mean you have how many such combination 4 c 2 combination that mean factorial 4 by factorial 2 and factorial 2. So, they mean 4 into 3 by 2 6, 6 such combinations will be there ok. So, here one this will give you 4 this will give you 4, 4, 8, 8, 8, 24 and then again center points are there. So, if you go for 5 factors accordingly you see that 2, 2 combination 2 to the power 2 combination you use and find out all the design settings means the point the combination for the different factors where you will conduct the experiment ok.

(Refer Slide Time: 34:23)



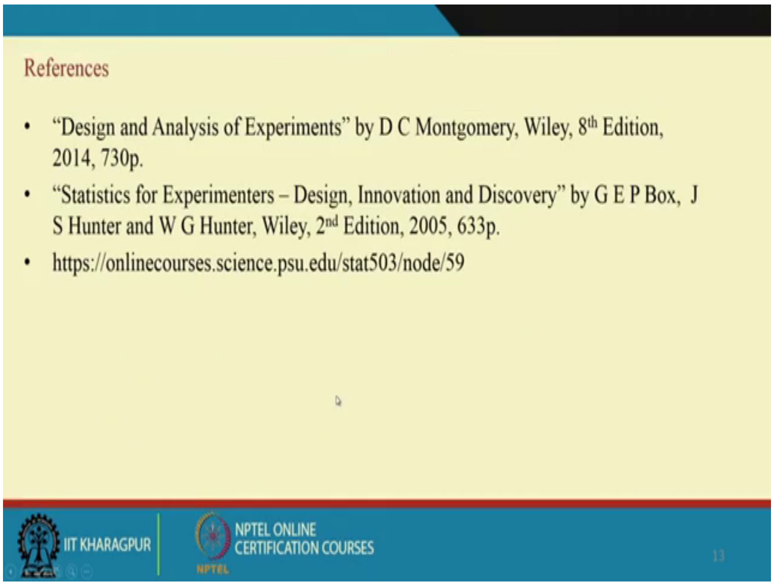
Number of factors	Box Behnken	Central Composite
2	-	13 (5 center points)
3	15	20 (6 center point runs)
4	27	30 (6 center point runs)
5	46	33 (fractional factorial) or 52 (full factorial)
6	54	54 (fractional factorial) or 91 (full factorial)

So, here we are giving you some comparative study in terms of number of runs, when number of factor we need varies that Box-Behnken, but the central company how the number of runs varies. So, Box-Behnken started from fact number 3 or more factors. So, if it is 3 Box-Behnken request 15 runs central composite design require 20 including 6 enter points 20 ok. Then, when number of factors is 4 Box-Behnken 27 central composite

30 number of factors is 5, this is 46, then 33 fractional factorial or 52 runs for full factorial and 54, then here central company also 54 with fractional factorial, but 91 with full factorial.

So, what is happening here then? So, if the number of factors; when it is increases so, if you go for 3 to the power Box-Behnken is 3-level design and central composite design each 2 to the power k-level design and with center and axial points here the number of runs are required more when you go for full factorials, but Box-Behnken design matches with the fractional part already I am giving the giving the efficient results so, but if you go by 3 to the power 6 k design here. So, 3 to the power 6 k it is a huge number of runs experiments required, but using Box-Behnken this is much less it is 54 only.

(Refer Slide Time: 36:06)



So, with this I conclude that that in order to estimate the second-order regression model or response surface you require a different kind of design and that designs which are most popular one are CCD, when it is it is 2 to the power k effect level case 2 to the power k means each factors are two levels and another one is Box-Behnken design when it is 3 to the power k type of case.

In this in the first case; there will be spherical CCD spherical CCD there is your rotatable one there is face centered CCD. So, CCD requires three kinds of observations one at factorial points another said at axial points and another set at central points and it has very useful property and heavily used in fitting in doing experiments for fitting second-order response surface Box-Behnken design employs the console concept of 2 to the power k, but it is basically a 3 to the power k design where 2, 2 factors at a time consider another factorial points.

And the other factors will keep at the at the at the center or level and all the combinations are used and that will that will give you an efficient one compared to 3 to the power k full factorial design ok. So, next class we will discuss the second-order model second-order model means second-order response surface model. So, hope you have understood the CCD Fully.

And in next class I will I will use the CCD in experimentation and the results through CCD; whatever we will get that will be analyzed and that will be used for fitting the second-order response surface.

Thank you very much.