Statistics for Experimentalists
Prof. Kannan. A
Department of Chemical Engineering
Indian Institute of Technology - Madras

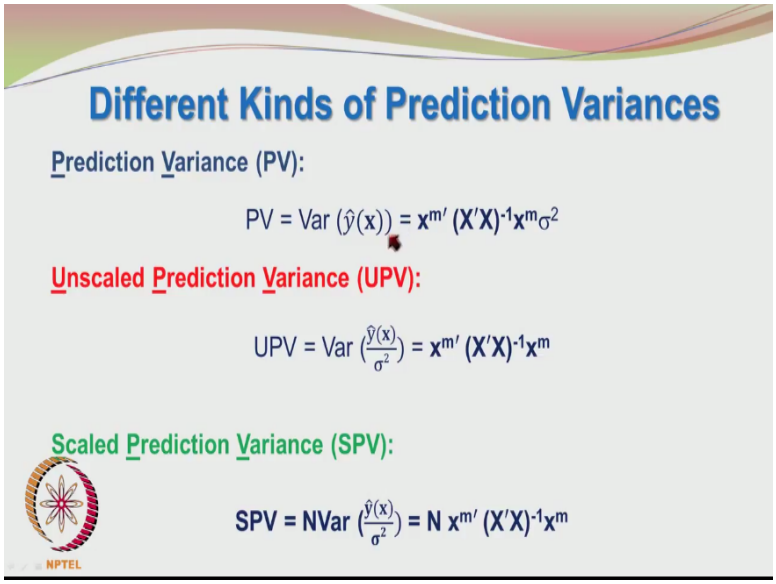
Lecture - 49
Response Surface Methodology - A

Hello welcome back. Today, we will be looking at the prediction variances, the different versions of the prediction variances. We have seen that after we develop the regression model or the Fx model for our chosen experimental design, we have to look at a few things. Of course, R squared will give the quality of the fit, but also we have to look at the prediction variances.

The concept is quite straightforward we want to ensure that the variances of the predictions are kept under control in the entire domain. We do not want certain regions in the domain where the prediction variances are very, very high. Then the quality of our predictions is not uniform in the domain and hence it is not reliable. You do not want too much of a variance in your predicted values.

So let us look at the measures proposed by Montgomery et al in the estimation of prediction variances.

(Refer Slide Time: 01:34)



So the variance y hat of x, here y hat refers to the prediction by the model, x refers to the point in the experimental space and that is given by xm prime X prime X inverse xm*sigma

squared. Here xm prime is the coordinate x expanded into the model space. So this is as per the model. We have seen this in the previous class. X prime X inverse is the usual matrix based on the factors we have considered.

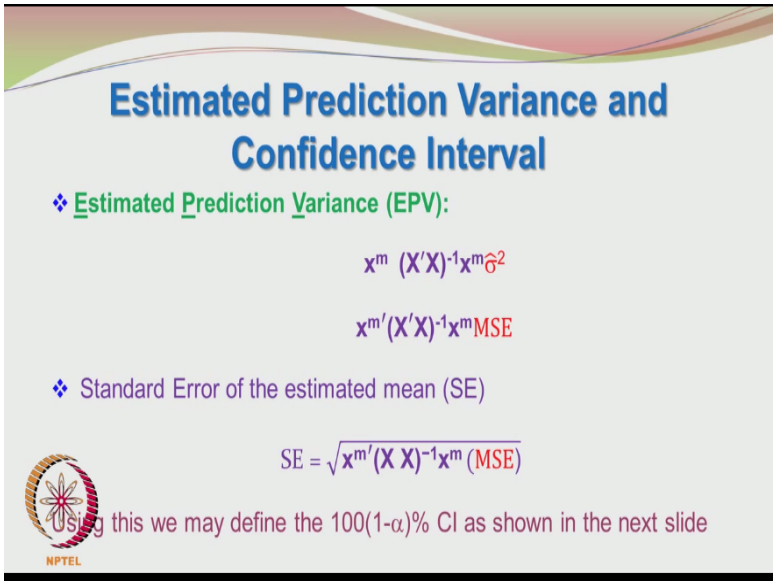
This is a very crucial component in the analysis and sigma squared is the unknown error variance. So when we divide this expression by sigma squared, we make it independent of sigma squared and so we have the unscaled prediction variance as variance of y hat of x/sigma squared, which is xm prime X prime X inverse xm and the scaled prediction variance SPV=N times variance of y hat of x/sigma squared=N xm prime X prime X inverse xm.

Why do we have to carry out all these things? Most importantly, please note that in these expressions we do not have Y which is the vector of experimental responses anywhere that is number 1. Number 2, dividing by sigma squared these unscaled prediction variances as well as scaled prediction variance become independent of sigma squared, which is not known and another good thing here is we are multiplying by N in order to scale the prediction variance depending upon the size of the run.

For example, if the run size is very high, then the prediction variance will become low. So when we compare design involving less number of runs with another design with more number of runs even though the second design involving more number of runs is less efficient. It may show a smaller prediction variance. In such cases, there will be some ambiguity.

To avoid this kind of number of run specific designs, we multiply by the total run size. So when we multiply by N, these issues are taken into account. That is why we scale the prediction variance and then it is called as SPV. Some times in our discussions, I will be referring to it as SPV that means scaled prediction variance.

(Refer Slide Time: 04:46)

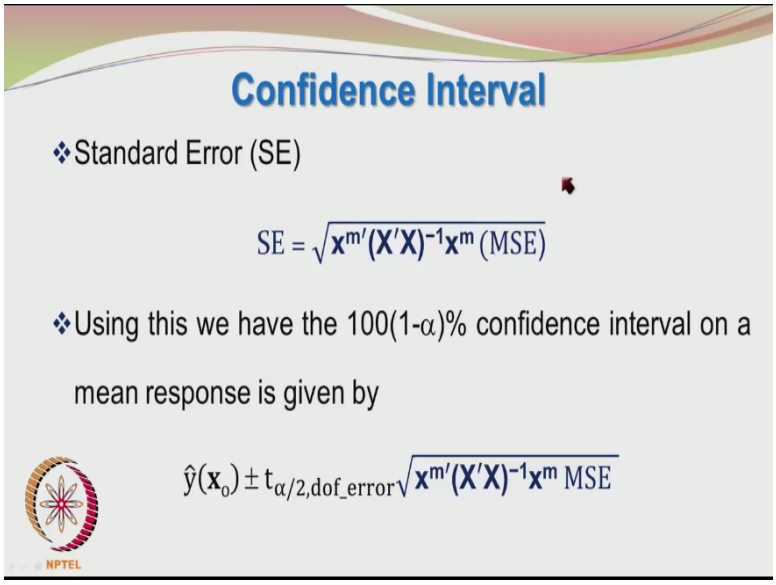


You also have the estimated prediction variance. Since we do not know the value of sigma squared we use the estimated value of the sigma squared and this comes from the residual mean square. We have the total sum of squares and then we have the regression sum of squares. The difference between the two will give you the residual sum of squares and that is divided by the degrees of freedom for the residual sum of squares.

So we get the residual mean square and that is used as a surrogate for the unknown error variance. So when you plug that in we have an estimated prediction variance and that is why I have written here as MSE, which is the mean square error and when you take the square root of that we have the standard error of the estimated mean. We will simply call it as the standard error.

So when we take the square root of the estimated prediction variance, we have the standard error. So the prime was missing and so we now have square root of xm prime X prime X inverse xm*MSE. Now that we have the standard error, we can very easily define the 100*1-alpha% confidence interval as will be shown next.

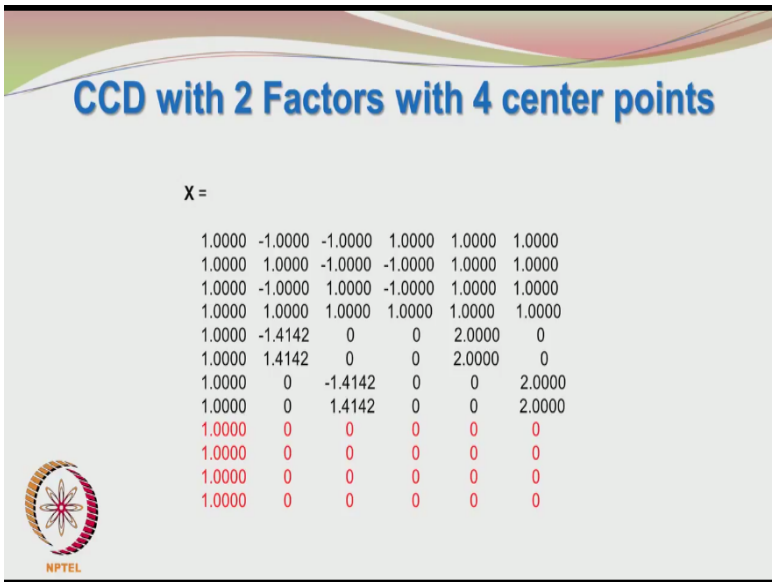
(Refer Slide Time: 06:13)



So we have the standard error as square root of xm prime*X prime X inverse xm*MSE. So the 100*1-alpha% confidence interval on the mean response is given by y hat of x0+t alpha/2 the degrees of freedom for the residual error*square root of this expression xm prime X prime X inverse xm*MSE. So this is the response at any particular point we are interested in and then you have the t distribution, which is based on a certain alpha level of significance.

We are doing a two-tailed test and so you have alpha/2 here. The degrees of freedom corresponding to the residual error and then we have the estimated prediction variance as given here and then we take the square root of that and again to repeat this refers to the prediction made at a particular point x0. It can be any value x in the domain.

(Refer Slide Time: 07:30)

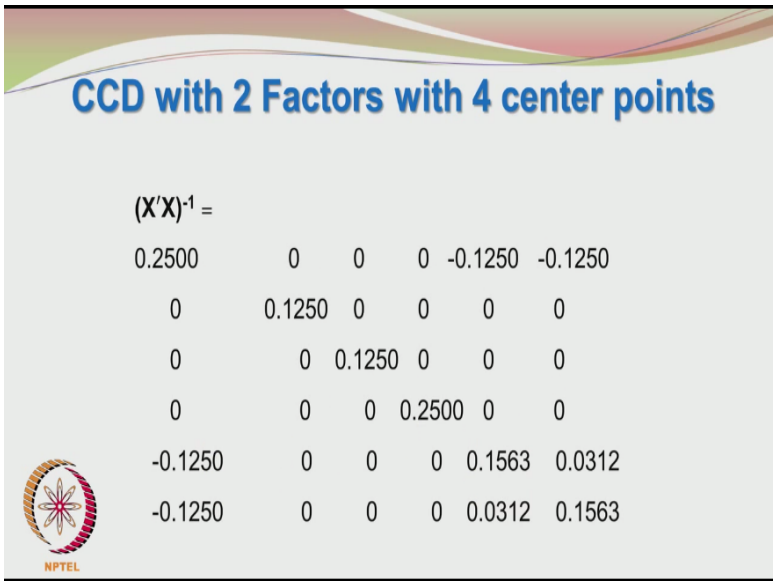


So now let us look at central composite design with 4 center points. You can see that this is the column vector comprising of ones in the X matrix and then you have the columns corresponding to factor A and this is the column corresponding to factor B and then this is the column corresponding to factor AB, the binary interaction.

Since we are considering second order model in the form of a central composite design, we also have the quadratic terms x1s or A squared and B squared so we are dealing with two factors. This is a central composite design for two factors A and B. So the column here is the column vector of ones, this column refers to the factor A, this column refers to factor B, this column refers to the interaction between factors A and B.

And then this would be factor A squared and this would be factor B squared. These values here represent the axial points for factor A and axial points for factor B and then when you square the root 2 you get 2 here and last but not the least we have 4 repeats here and the repeats are carried out at the center points and hence you have the 0 values here 0 0 0 0 0 0 0 0. And then when you take the binary interaction between A and B, again it is 0. When you square A, it is 0. When you square B, it is also 0 at the center point under consideration.

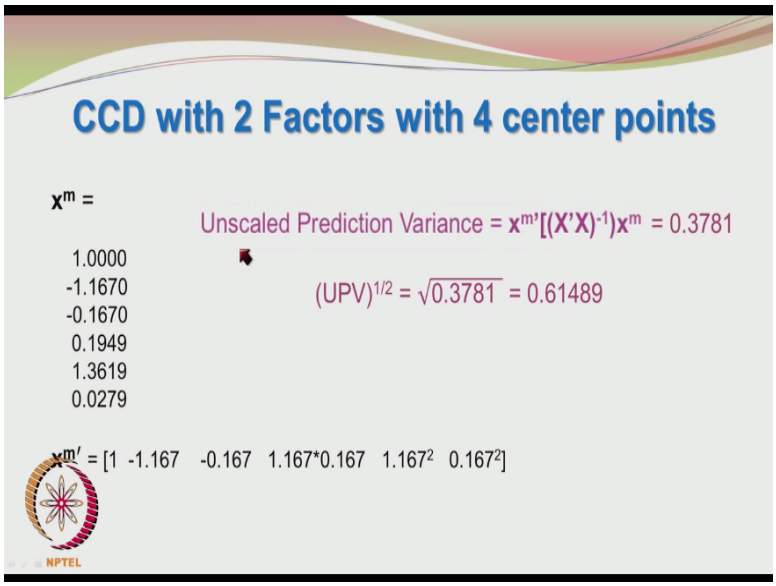
(Refer Slide Time: 09:26)



Now when you take X prime X inverse it is a sparse matrix, still lot of 0s can be seen and you do not have only the diagonal elements, there are also some off-diagonal elements and that is because of the central composite design structure. The experimental points are not located only at the extremes. They are also located at other locations or positions as well including the one at the center.

And hence it is not a variance optimal design or the design so constructed to reduce the variance.

(Refer Slide Time: 10:15)

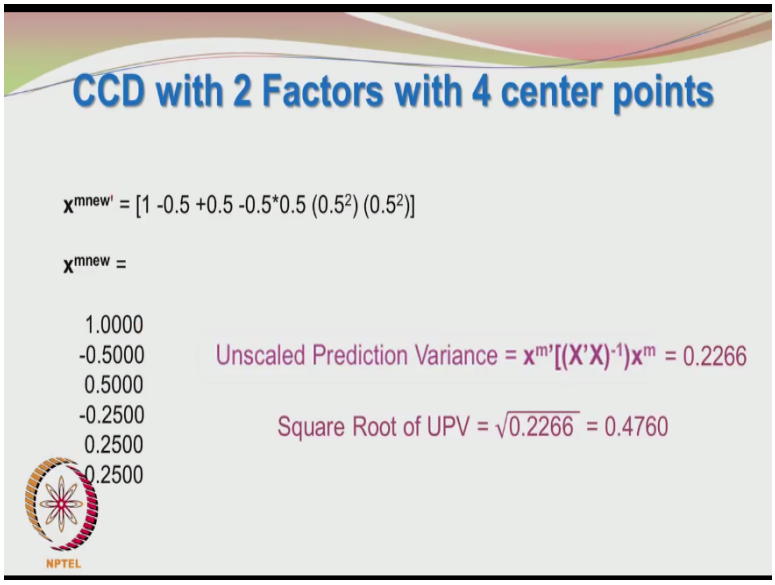


So first we will calculate the unscaled prediction variance, which is defined as xm prime inverse of X prime X*xm. So we have the xm vector, which is the coordinate point expanded to model space as shown here. So this would be the point we are interested in -1.167, -0.167. This corresponds to the column vector of ones and this represents the product of the A and B, so here you have -1.167*0.167 and that comes to about 0.1949.

And then this would be 1.167 squared and this would be 0.167 squared and since you have xm, you may also have xm prime and in the previous slide, we had calculated X prime X inverse so we have everything and the unscaled prediction variance is independent of sigma squared and so this is the value we get 0.3781 and we take the square root we get 0.615. So I have also shown how to calculate xm prime here.

This will be the transpose of xm and you can see that it is one corresponding to the vector of ones, factor A, factor B, factor AB, A squared and B squared. So we have calculated the unscaled prediction variance at a particular coordinate -1.167, -0.167.

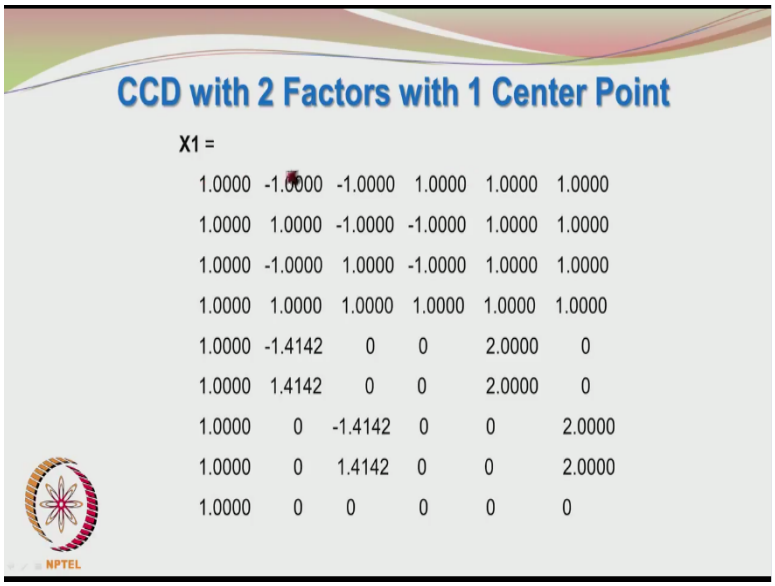
(Refer Slide Time: 12:08)



Now let us take another point xm new prime and that would be at the locations -0.5 and +0.5 and again you have, it is easier to compute now the interaction between A and B that would be -0.5*+0.5 that would be -0.25 that is what you have here and then you have 0.5 squared and 0.5 squared, which is 0.25, 0.25 and unscaled prediction variance here comes to be 0.2266 and square root of that comes to be 0.4760.

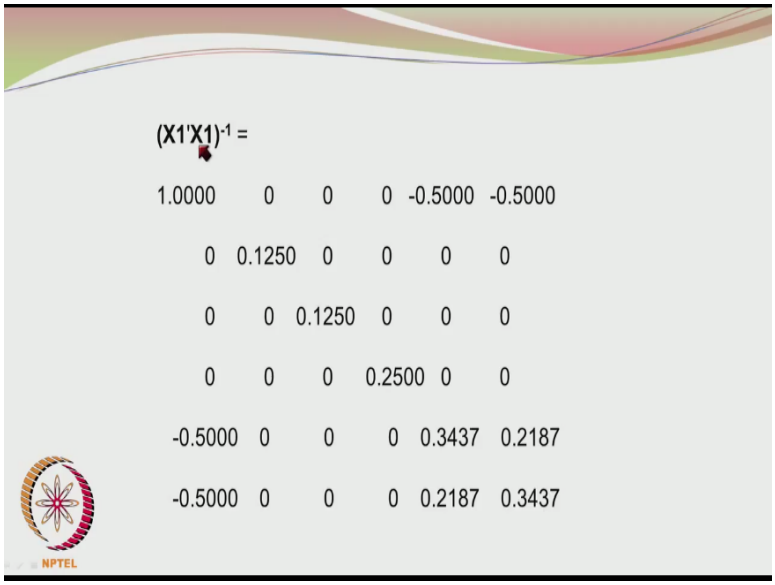
Please note that we are having 4 center points here in this design. Now when you have only one center point, what happens to the unscaled prediction variance? We have seen that in order to stabilize the prediction variance, we need more number of center points. Now we are having a central composite design with only one center point and let us see its impact on the prediction variance.

(Refer Slide Time: 13:12)



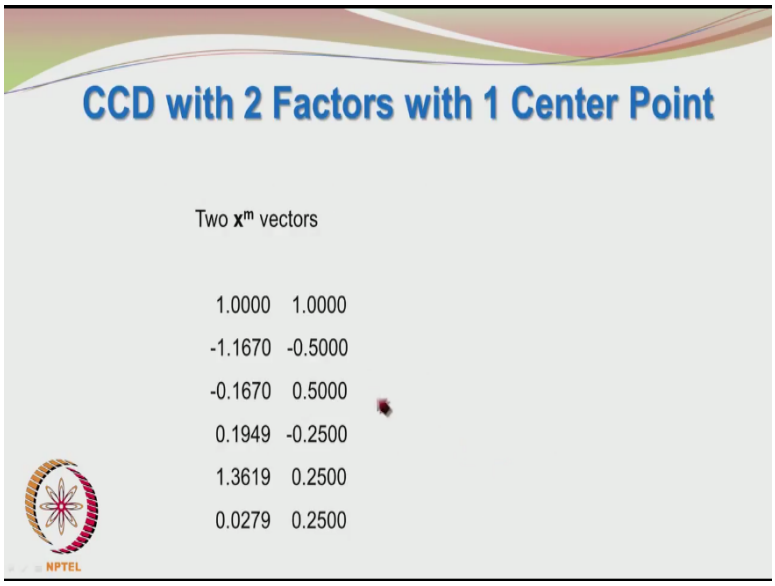
If you look at the design you can see that as usual we have the factors A and B and then the axial points. These are the factorial points, these are the axial points corresponding to factor A and corresponding to factor B and then you also have AB here and then you have A squared and then you have B squared but the most important thing to note here is when compared to the previous design or the example we saw, we are having only one repeat at the center point and we are having a new matrix X1.

(Refer Slide Time: 13:43)



So what we can do is we can take the inverse X1 prime X1 inverse and this is what we get. Again we do have non-diagonal terms also right.

(Refer Slide Time: 14:00)

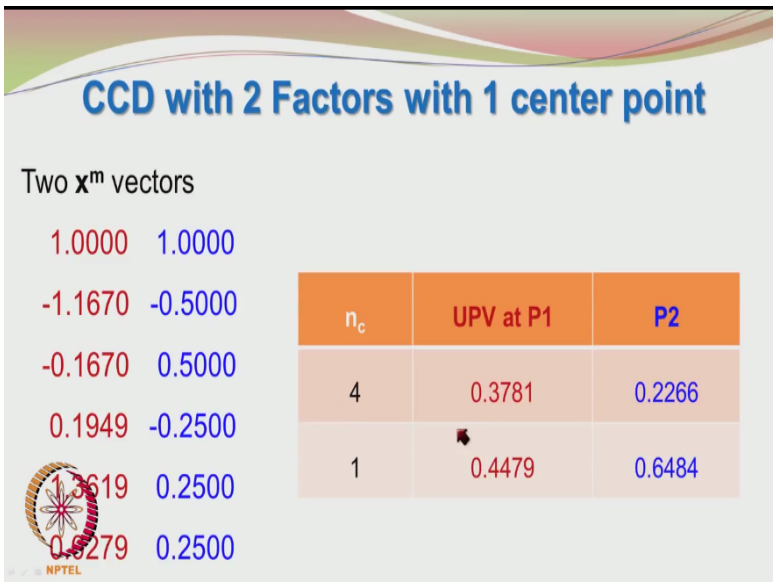


And let us now look at the two xm vectors, the same vectors which we consider in the previous case where we had 4 center points. So we are locating the xm corresponding to the

coordinates -1.167 and -0.167 and -0.5 and 0.5. So this is one coordinate and this is another coordinate and of course here again we are multiplying A*B and we get 0.194 and -0.25 and then we are taking square of -1.167 we get 1.3619.

Square of -0.5 is of course 0.25 and -0.167 squared gives you 0.0279, square of 0.5 gives you 0.25. So we have two different xm vectors. Let us see the scaled or unscaled prediction variance here.

(Refer Slide Time: 14:53)



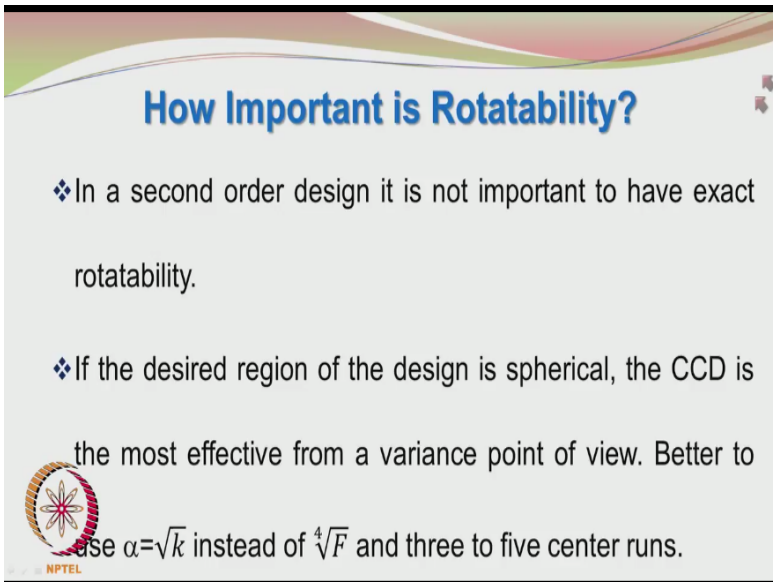
And so this is point 1 color coded red and this is point 2 color coded blue. So when you have 4 center points, the unscaled prediction variance at 0.1 was 0.3781 and it is 0.2266 at 0.2 okay. Let us just check that once again. So 0.3781 and 0.2266 corresponding to points 1 and 2 with 4 center points. So 0.3781 and 0.2266 at two different points. This is corresponding to 4 center points.

When you reduce the number of center points from 4 to 1, very surprisingly we see that the unscaled prediction variance at P1 are shot up from 0.38 to about 0.45 and at point 2 with a shot up from 0.23 to about 0.65. So this shows that the prediction variance increases dramatically when you reduce the number of center points. This also tells you why we should have center points in experimental design not only the center points help you to find the estimate of the pure error, it also minimizes or economizes the design strategy.

Because we do not now have to repeat all the experiments at the corner points and the axial points. We have to repeat the experiments only at the center points and it also as I said earlier

gives you an idea about the pure error. It also tells you whether curvature effects are significant or not and more the number of center points the scaled prediction variance comes down.

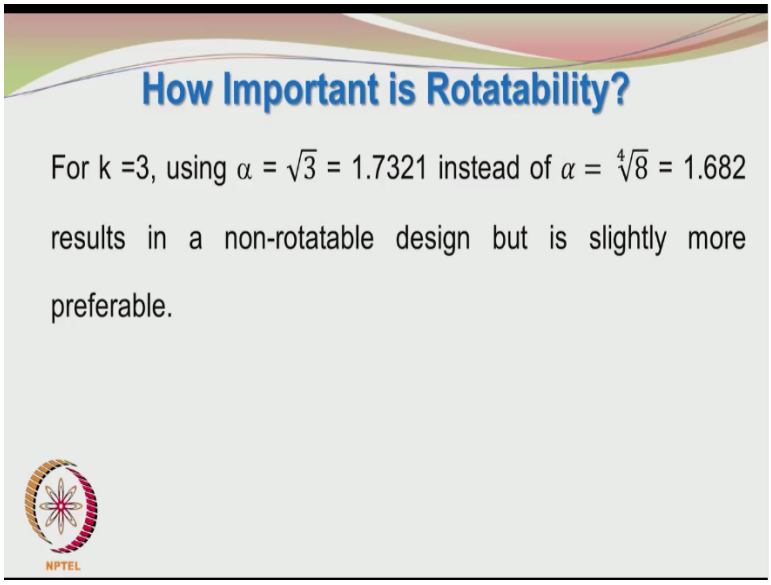
(Refer Slide Time: 16:59)



So you can see that the center points play a crucial role in experimental design strategies and how important is rotatability. In second order design, it is not important to have exact rotatability. If the desired region of the design is spherical, the CCD is most effective from a variance point of view. So what this slide recommends is more than rotatability it is the spherical nature of the design which is of importance.

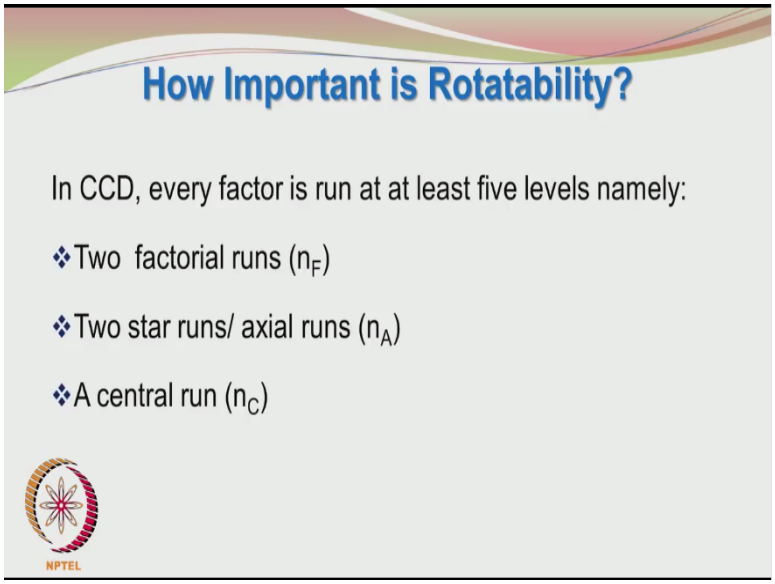
So for a spherical design, you have the axial points at root k instead of fourth root of F where k is the number of factors and F is the number of factorial points and 3 to 5 center runs.

(Refer Slide Time: 17:44)



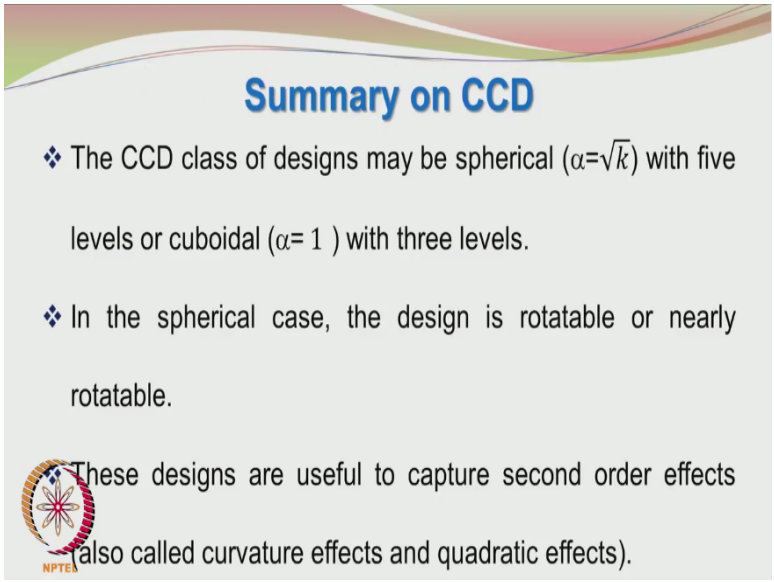
So for k=3 that means 3 factor design use alpha=root k or root 3 which is 1.7321 instead of alpha=fourth root of 2 power 3 which is 8 which is 1.682. So it results in a non-rotatable design but it is preferred.

(Refer Slide Time: 18:06)



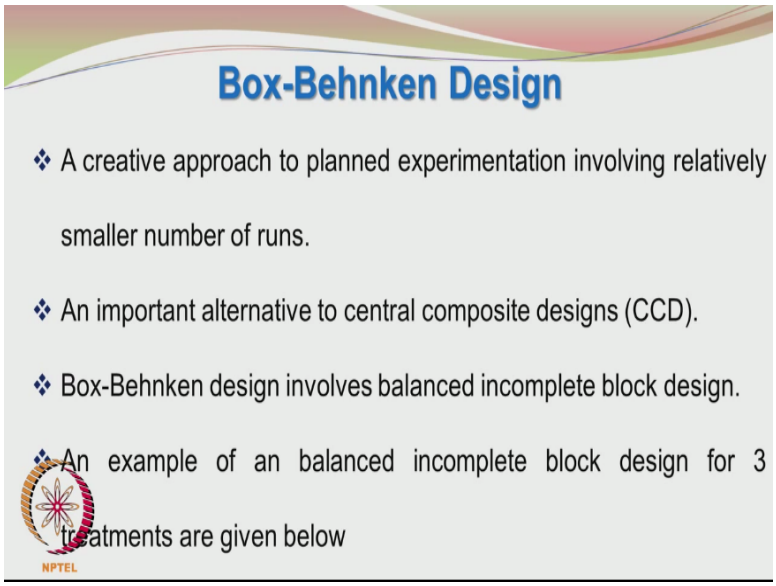
So when you look at the central composite design, you are running it at 5 levels, 2 factorial runs, 2 axial runs and then a center run.

(Refer Slide Time: 18:18)



So let us now summarize the central composite design. The central composite designs may be spherical, alpha=root k with 5 levels or cuboidal with alpha=1 with 3 levels. In the spherical case, the design is rotatable or nearly rotatable and these designs are useful to capture the second order effects also termed as the curvature effects or quadratic effects.

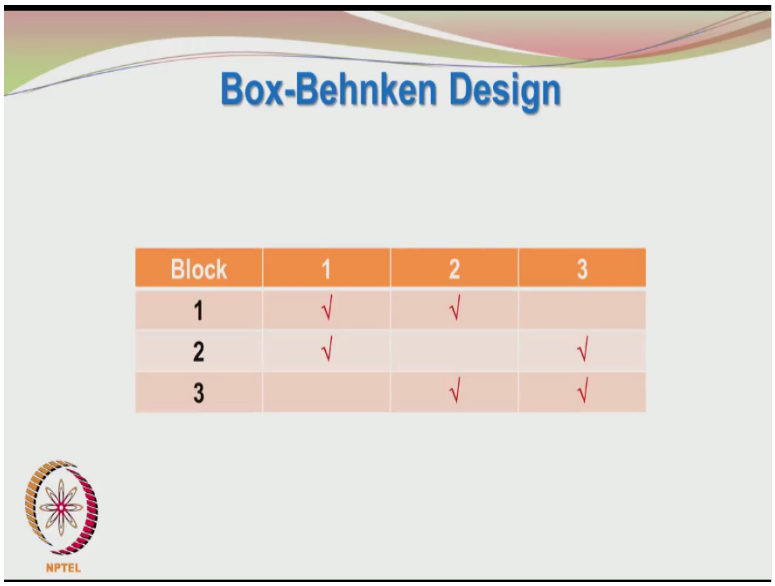
(Refer Slide Time: 18:46)



Now let us move on to the next design strategy after a small break. Now we are going to look at a new experimental design strategy, it is called as the Box-Behnken design. You might have come across this kind of design in research papers. Central composite design and the Box-Behnken design are most commonly encountered when second order models are getting discussed.

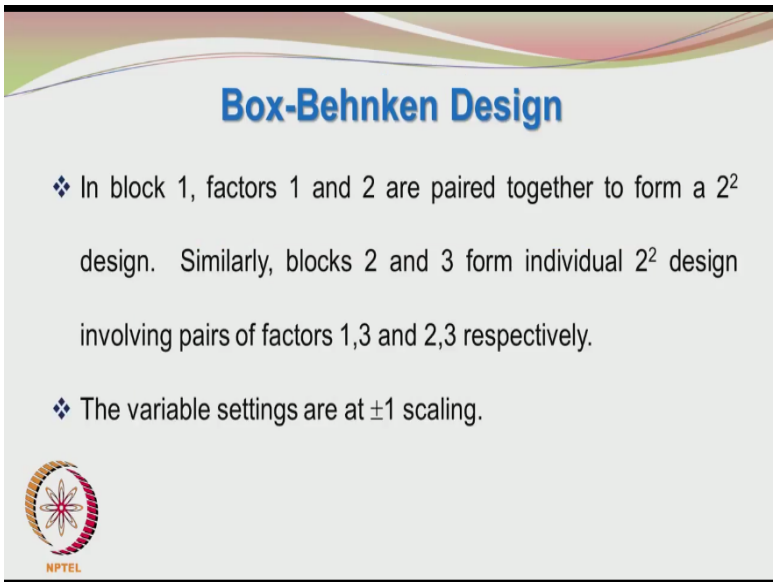
So let us look at the features of the Box-Behnken design. It is a creative approach to planned experimentation involving relatively smaller number of runs. It is an important alternative to central composite design and Box-Behnken design involves balanced incomplete block design. What this means I will tell you shortly and let us look at an example of a balanced incomplete block design for 3 treatments.

(Refer Slide Time: 19:53)



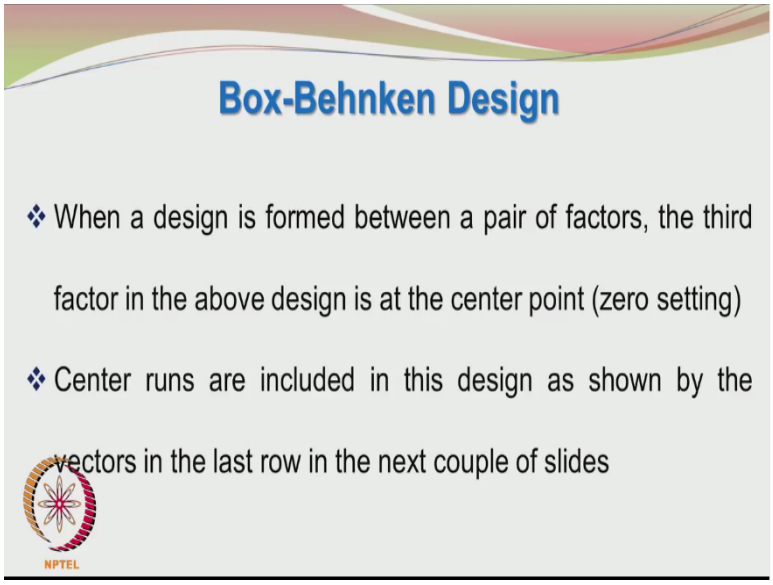
So in the Box-Behnken design, we are having 3 blocks and in the first block you are considering only factors 1 and 2 and in the second block we are paying attention to factors 1 and 3 and in the third block we are paying attention to only 2 and 3. So when you are looking at this particular design, you can see that each block has importance attached only to 2 out of the 3 factors, which are being considered.

(Refer Slide Time: 20:34)



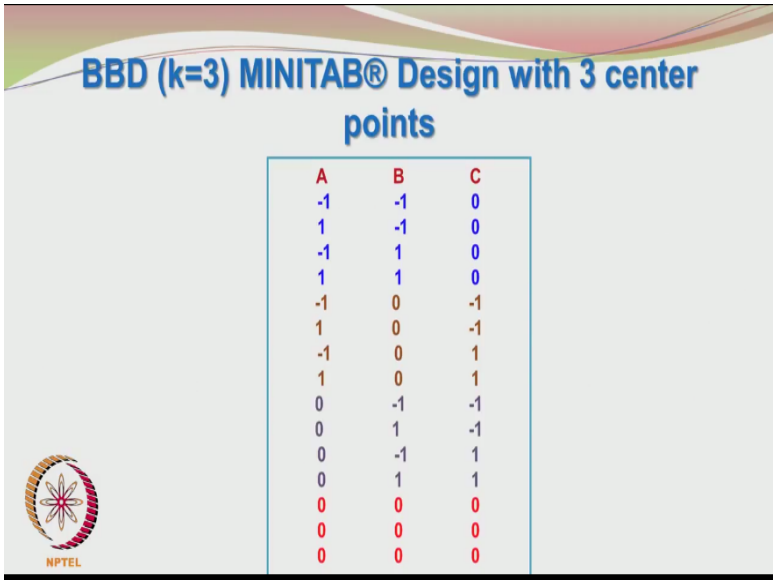
And in block 1, factors 1 and 2 are paired together to form a 2 power 2 design. So when you have importance attached to only the first and second factor you construct a 2 power 2 design out of them. Similarly, blocks 2 and 3 form individual 2 power 2 design involving pairs of factors 1 and 3 and 2 and 3 respectively and the variables are kept at + or -1 coded settings.

(Refer Slide Time: 21:12)



And what is happening to the third factor? When a design is formed between a pair of factors, the third factor in the above design is at the center point or 0 setting and the center runs are included in this design in the last row.

(Refer Slide Time: 21:32)

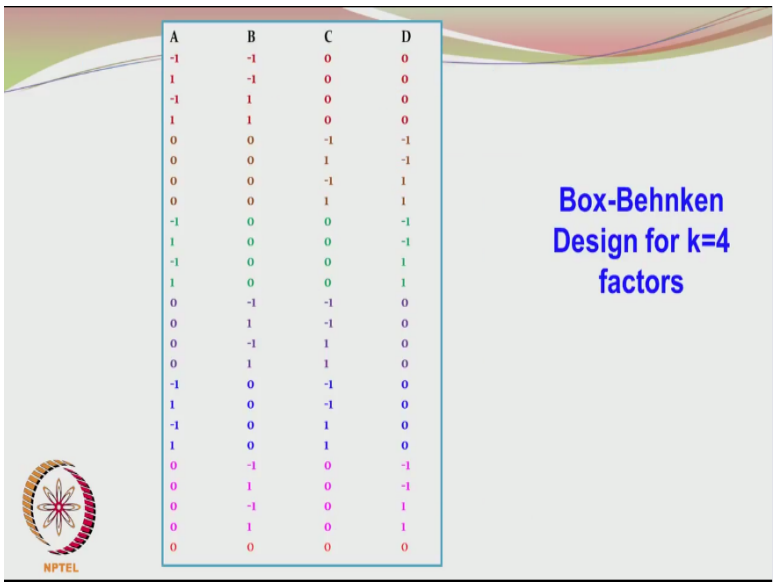


So when you look at the design you have A, B, C the 3 factors, you have -1, -1, 1,-1, -1, 1, 1 1 which corresponds to a regular 2 power 2 design involving factors A and B and then factor C is kept at 0 setting and since the values that may be taken by either A, B or C or -1 and +1 0

would mean a center setting and then when you look at the second block as shown in the brown color, we have factors (()) (22:09) B is kept at the center level or the 0 level.

And you can see that 2 power 2 design is formed between factors A and C. Next, we have the third block involving a 2 power 2 design between factors B and C while A is kept at the 0 level. In addition to the above, we also have center runs at the end where all the factors are kept at their mid values or 0 coded values.

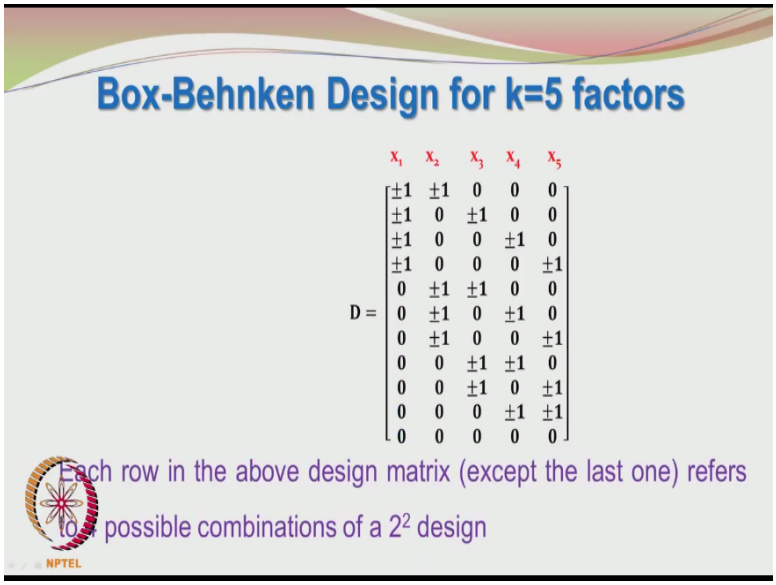
(Refer Slide Time: 22:42)



And then when you have 4 factors, the concept is pretty much the same. Here you have factors A and B forming a 2 power 2 design and the next you have a 2 power 2 design involving factors C and D and then you have 2 power 2 design involving factors A and B, then B and C, then A and C and B and D okay. So you are considering 4 variables 2 at a time and that would be 6 combinations.

Let us see 1, 2, 3, 4, 5 and 6 so we have the 6 combinations listed out in this table and then you also have the center point at the very end.

(Refer Slide Time: 23:35)

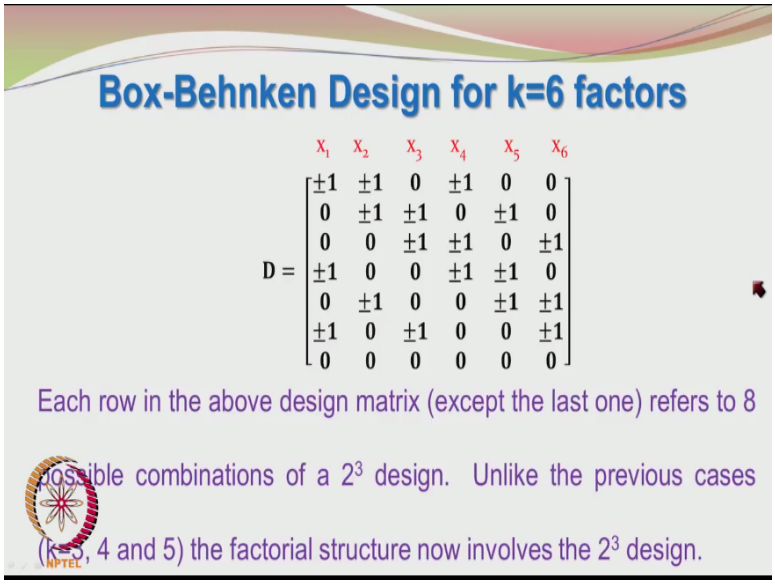


When you go for 5 factors, well we are increasing the number of factors and hence the number of experiments would also increase and this is shown in a very compact form. So in the first design you are considering factors A and B or X1 and X2 and that is why we are having + or -1 +or -1 here. So when you have 5 factors taken 2 at a time that would be 10 combinations, 5*4 by 2 that would be 10 combinations.

It is very large to show all the 10 combinations and hence we show it in a condensed form. This represents we are considering factors 1 and 2 X1 and X2 and so we form a 2 power 2 design whereas X3, X4 and X5 the third, fourth and fifth factors are kept at 0 level and in the next combination we take factors 1 and 3, third combination we take factors 1 and 4, next combination between 1 and 5 and then 2 and 3, 2 and 4 and 2 and 5, 3 and 4, 3 and 5 and 4 and 5.

So we have 1, 2, 3, 4, 5, 6, 7, 8. 9. 10 combinations and the last row represents the center runs maybe more than 1 in number. This is a vector notation and so you may have more than 1 center run at the end.

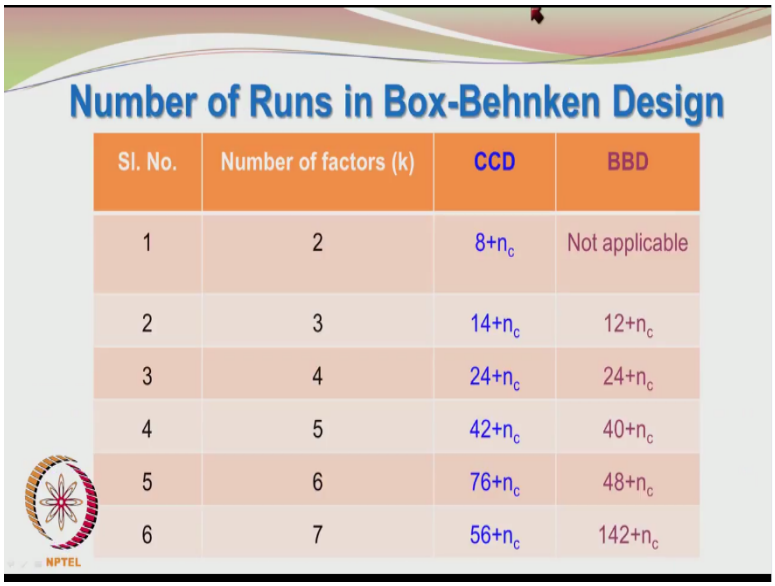
(Refer Slide Time: 25:21)



And then we go for 6 factors rather than taking 2 factors at the time to create 2 power 2 factorial design we actually take 3 factors at a time and create a 2 power 3 design. So this is shown in the condensed form. We are now looking at the Box-Behnken design for k=6 factors and instead of taking 2 factors at a time, we take 3 factors and create a 2 power 3 design.

And when you do that, we get a design where 3 factors are taken at a time and here this represents X1, X2 and X4 forming a 2 power 3 design whereas X3, X5 and X6 are kept at the center values or 0 setting. Next, we go for X2, X3 and X5 at 2 power 3 design with other factors kept at the center values. Then we go for X3, X4 and X6 which are kept at 2 power 3 design settings. Others are kept at center values and finally we have the center runs.

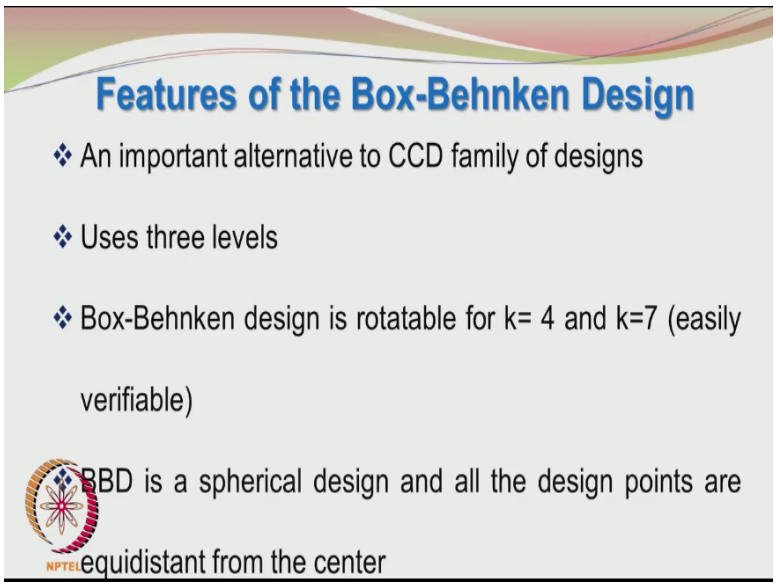
(Refer Slide Time: 26:41)



So when you compare the runs in a Box-Behnken design and central composite design, you cannot construct a Box-Behnken design with 2 factors whereas you can construct one with central composite design and here when you have 3 factors, you can see that CCD involves 14+nc and this is 12+nc and the number of runs are equal for k=4 factors 24+nc, 24+nc. For 5 factors, 42+nc, 40+nc.

And when you have 6 factors, the Box-Behnken design comes to only 48+nc whereas this is going to as many as 76+nc and for 7 factors you have 56 and 142 for the central composite and Box-Behnken designs respectively.

(Refer Slide Time: 27:37)



So the important features of the Box-Behnken design are they represent an interesting and practical alternative to the central composite design because for some combinations of factors you are going to have less number of runs and it uses 3 levels and it is a rotatable design for k=4 and k=7 you may want to verify this, this is very straightforward after you construct the relevant matrix.

And BBD is a spherical design and all the design points are equidistant from the center. So this is good from a variance prediction point of view. So this completes our discussion on design alternatives available to experimentalists. I have discussed the most common designs. There are of course many more designs, but I am sure that with this background you should be able to pick them up from standard text books and understand their implications.

For example, you can have the face centered design or the cubical design, other many designs exist but the Box-Behnken design and the central composite designs are the most popular ones and they are frequently encountered in many research papers. Rather than just implementing the BBD or CCD directly, it is important to understand the different implications of such designs.

So you have to think about the number of center points you may want to use and whether you want to have emphasis on rotatability or on the spherical nature where all the design points are equidistant from the center except the center points. So that is another important factor. You may also want to look at the scaled prediction variance properties.

So an important thing here is scaled prediction variance depends upon only the X prime X inverse matrix and then the coordinate location in the design space. It does not depend upon the experimental observation values Y. So even before you start your experimental work you may easily estimate the distribution of the scaled prediction variance in your design space and you may choose a design suitably based upon the distribution of the scaled prediction variance.

To emphasize this does not depend upon the experimental observations Y and another important thing to summarize here is the center points. The center points are used for getting pure error estimates. The center points are helpful in enabling you to identify the curvature and they are also helpful to stabilize the prediction variance.

The axial points are located in the central composite design for enabling you to identify the quadratic terms or the squared terms. So each and every point in the design space has its own role to play. The factorial points of course are useful to find the effect of main factors and the interaction between the main factors. It is very important for you to decide upon the number of factors and the interactions you want to consider in your model.

Look at the number of experimental data points then you decide upon the size of the model. If you go for a model with too many parameters very ambitious or a greedy model, then there would be the risk of aliasing and the X prime X matrix may have the danger of being undefined when the inverse is taken. So all these issues we must be aware of. So I request

you to go through the portions covered so far and get a clear picture of the different experimental designs. Thank you for your attention.