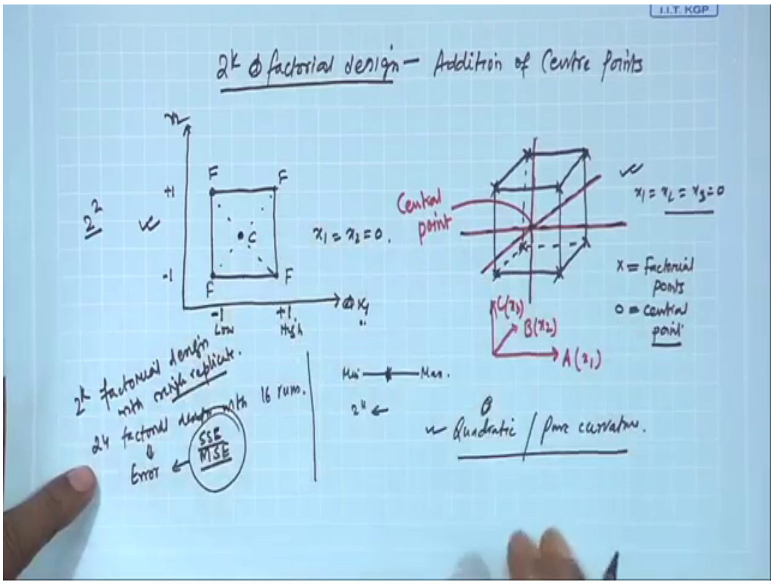
Design and Analysis of Experiments
Prof. Jhareswar Maiti
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture - 39
2 k Factorial Design Centre Points

Welcome, today we will discuss 2 to the power factorial k factorial designs, addition of centre points.

(Refer Slide Time: 00:27)



So, you have seen in last class that we discussed 2 to the power k factorial design with single replicates and we have given an example 2 to the power 4 factorial design, with 2 to the power 4 16 runs. And in analysis we have seen that with this single replicate you are not able to estimate the error part, error particularly we are talking about MSE or SSE not able to estimate this one because of no degrees of freedom available for SSE and that mean there are no independent observations available to help us to compute the error terms.

So, we have adopted a policy, there we say that estimate all the effects and their contribution and find out the percentage contributions, and then see which are the effects having low percentage contribution and discard them primarily, following the policy of sparsity of effect principles means the higher order interactions will have negligible effect.

So, the same problem this problem error calculation can problem can be sorted out with single replicate plus, having another experimental setting which is known as central point for example, if we discuss with reference to 2 to the power 2 factorial designs, then this is the diagram. These are the factorial points this four coordinates are factorial points, where we usually conduct experiments and at every factorial points we conduct more than one experiment to get replications. Here what is happening that in the earlier example, that we have only single replications and as a result we could not contour found out the SSE.

Now if we find define central point is basically the point where you that x 1 equal to x 2 equal to 0, what will be that point? You know the factorial points this one is minus 1,this is plus 1 low and high. Similarly for second factor this is minus 1, this is plus 1 at low and high. So, x 1 will be 0 at the mid of this point, x 2 will be 0 at the mid of this line means here. So that means, that mean if you intersect the diagonal lines of these rectangle you will be getting the centre points.

So, this is geometrically this is am saying centre point; in reality what is this? In reality it is the point; it is the process at which x 1 and x 2 the coded values are 0. So, that mean this is the plus point where the effect of x 1 and x 2 in the coded scheme is 0. So, if you conduct several experiments here what will happen, by using those experimental runs those that y values you can compute the error independently ok.

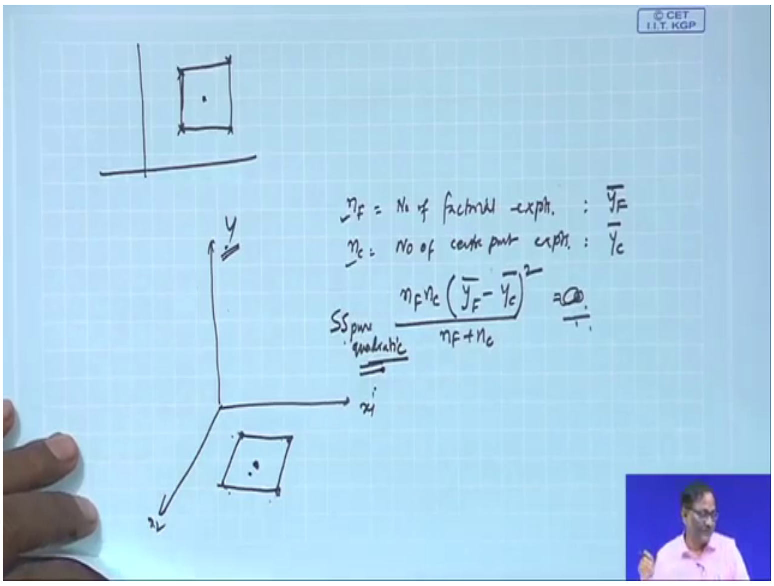
So, as we have seen earlier that at every point factorial point you conduct experiment to find out the find you conduct more ex more number of experiments to find out the error terms, error values here this is one. Suppose you go for the three factorial case in that case the central point is 1 where x 1, x 2, x 3, equal to 0 and the eight corner points are the factorial points.

So, centre points is one where the effect of the factors is 0 the way we have defined the scheme, another important thing is that this is the point where mostly the process is run and this is the point which is more familiar to the production system means actual operations. The reason is the reason is you usually we define the high low like minimum is low and maximum is high and most of the time we will be seeing that in between the two the middle point the operator will try to fix the processes or set the processes at this point which is that is why we will see the most the familiar condition in the production actually in during actual production.

So, this is one advantage of having centre point second one is that when we do 2 to the power k factorial design we assume that the response surface is linear. So, although higher order effects are higher order interaction effects are negligible plus, the effects are linear that is that is the best possible situation when you go for 2 to the power k factorial design it is basically used for screening purposes.

So, if the there is linear relationship then what will happen you have to test it beforehand otherwise, the linear model with interaction if you fit it may not be the correct one and now the concept here is suppose if the suppose if you develop the response surface, suppose this is x 1 and this is x 2 and suppose y is in side.

(Refer Slide Time: 07:26)



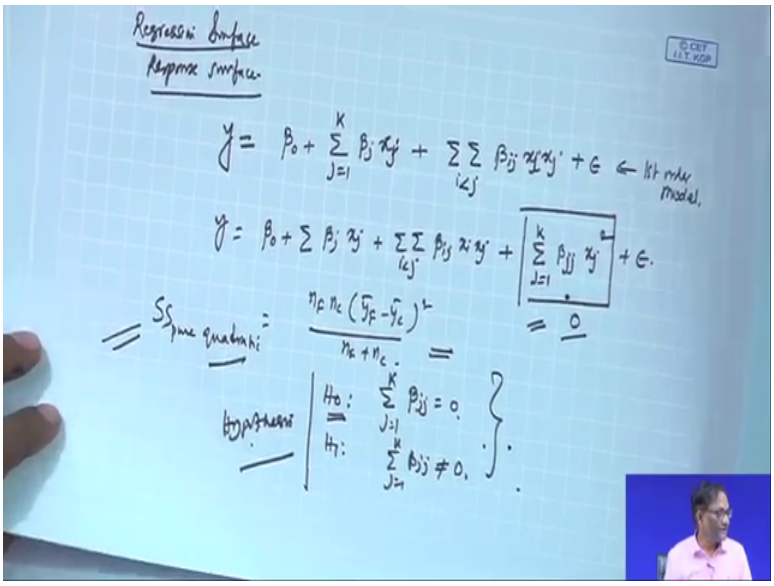
So, this is y response this is x 1 and this is x 2 so, what you assume that for any particular value whether it is a factorial points, suppose if I do like this whether its factorial points or centre points if the surface is linear.

So, what will happen, the average value at this points and average value at the centre point that should be almost equal. So, that mean if there is nonlinearity then, the y will not be equal at all the factorial points as well as the centre point; now there will be difference in average value of response taking observations of the factorial points and observation at the centre point. So, there the difference is significant then there is quadratic effect; second order that higher order effect not linear that first order will be interaction it second order may be and with interaction effects also.

So, we will be in a position to if we use centre points we will be in a position to find out whether there is quadratic or quadratic error is there or not quadratic error or pure curvature is there or not. So, this is another one, one end that at the factorial points if you have single replicate observations with centre points the estimation of error that is possible. Another one is if there is any quadratic error that can be tested with using the centre points centre point give you more data.

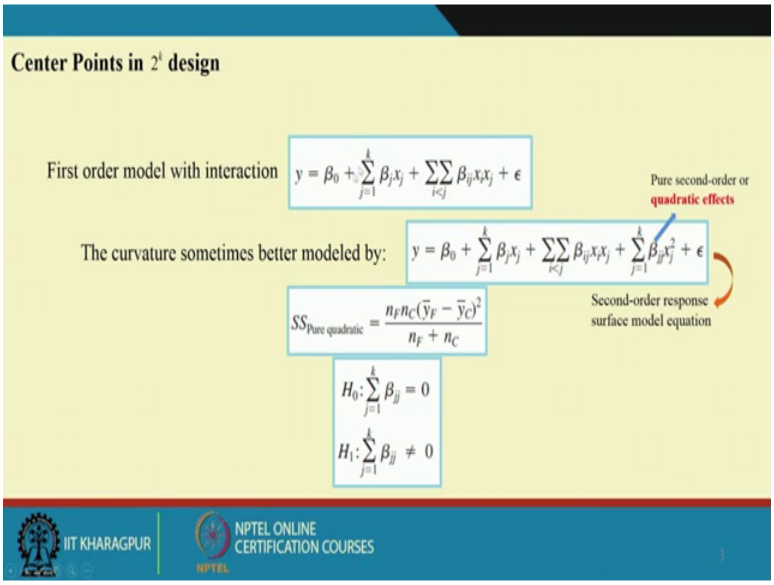
So, as a result what happen we have more information and we can do more kind of analysis. So, with this a background I will let me tell you that this preparation we have made using the Montgomery book, now we will show you the regression surface.

(Refer Slide Time: 10:06)



Regression surface is which is also known as Response surface so, let us see the slides.

(Refer Slide Time: 10:22)



If I use first order with interaction then this is our regression equation now if there is curvature what is known as we are talking about the quadratic effect is there or pure curvature second order then your model will include the second order interaction the quadratic term. So, suppose, you have when you do your factorial experiment there you will find out that you will not be able to estimate this other parameters, the quadratic part also.

So, here is what we are that is why we are assuming that so, as I have already discussed. So, if they you have to test whether there is curvature in that sense the quadratic effect is there or not you require to have a test. So, with reference to this example what you are doing when you conduct experiment or this example or higher order example. Suppose you have n F number of n F is the number of factorial experiments and let n c is the number of centre point experiment.

So, as I told that then if you compute here the average of responses is y F bar and here average of response is Y c bar. So, if there is no quadratic curvature effect what will happen ultimately, even that it is expected that y F bar minus Y c bar theoretically it should be 0, but it will be there will be some value you will not get exact value because of that randomness initial factors and all those things are there, but if the quadratic effect is there then let us square it and you have how many data points n F and n c. So, you multiplied this two and divided by n F plus n c, this is a measure of SS pure quadratic ok.

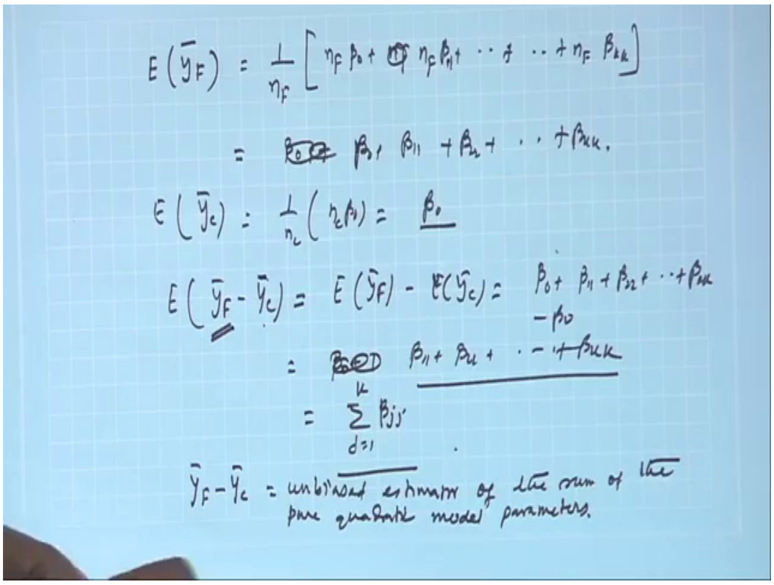
So, what I mean to say with reference to a 2 to the power k factorial experiment centre points I am saying that you conduct experiment all those points, take their average conduct experiment at the centre point, take the average and the difference in average can be used to compute the SS pure quadratic and which is this formula. Now, what you require to do you require to test whether this one is significant or not, if this is significant with reference to a threshold value then pure quadratic effect is there otherwise it is not there.

So, in other words in other words this can be represented like this so, my response surface is y equal to beta 0 plus sum total of J equal to 1 to K beta j x j that is your main effects, then you write down i less than j beta i j x j x i and x j that is the interaction term plus epsilon this is my first order model. Now, if there is the quadratic effect is present then your model will be this beta j x j plus i less than j beta i j x i x j plus, there will be another term J equal to 1 to K beta j j x j square plus epsilon.

So, we want to see that this part is negligible, this will be negligible when beta j j equal to 1 to K beta j j this will become 0 because x j square will be always 1, if we go for coded variance. So, when I say that SS pure quadratic we say this equal to n F n c y F bar minus Y c bar square by n F plus n c this one will follow certain distribution and then with using this we will basically test, that the two hypothesis is H 0 is sum total of J equal to 1 to K beta j j equal to 0, alternate hypothesis is sum total of J equal to 1 to K beta j j not equal to 0.

So, this is the test this is what is the statistics this is the hypothesis using this statistics you will test this hypothesis. So, if this statis hypothesis is say null hypothesis is satisfied this becomes 0 so, this is a first order model with interaction. How do you basically arrived at that relation between this SS pure quadratic vis a vis this J equal to 1 to K beta j j equal to 0, that some explanation is like this ok.

(Refer Slide Time: 17:14)

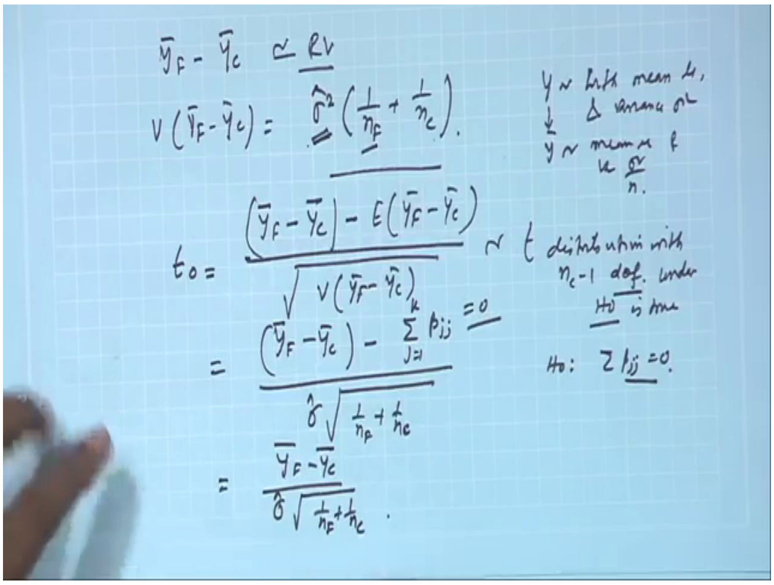


So, expected value of y bar at factorial points this will be your 1 by n F into n F beta 0 n F beta 1 like this you go up to n F beta k k. So, n F beta 0 n F beta 1 1 something like this, this, this, this, this n F 1 1 and if you follow this ultimately your result will be this will be beta 0 beta sorry beta 0 beta 1 1 beta 2 2.

So, like this beta k k and expected value for y bar c this will be using the same way that 1 by n c, n c beta 0 and this will give you beta 0 and ultimately that expected value of y bar F minus y bar c this will be the difference between that mean expected value of y bar F minus expected value of y bar c. So, this will be beta 0 plus beta 1 1 plus beta 2 2. So, like this beta k k minus beta 0 so, this is basically beta 0 sorry beta 1 1 beta 2 2 so, like this beta k k this is nothing, but J equal to 1 to K beta j j.

So, we can say that y bar minus y c unbiased estimator, unbiased estimator of the sum of the pure quadratic model parameters, that is sum of pure quadratic model parameters. What is y F bar? That is average of responses of all the experimental runs conducted in factorial points, y bar c is the average of responses of all the experimental runs conducted at the centre point.

(Refer Slide Time: 20:07)



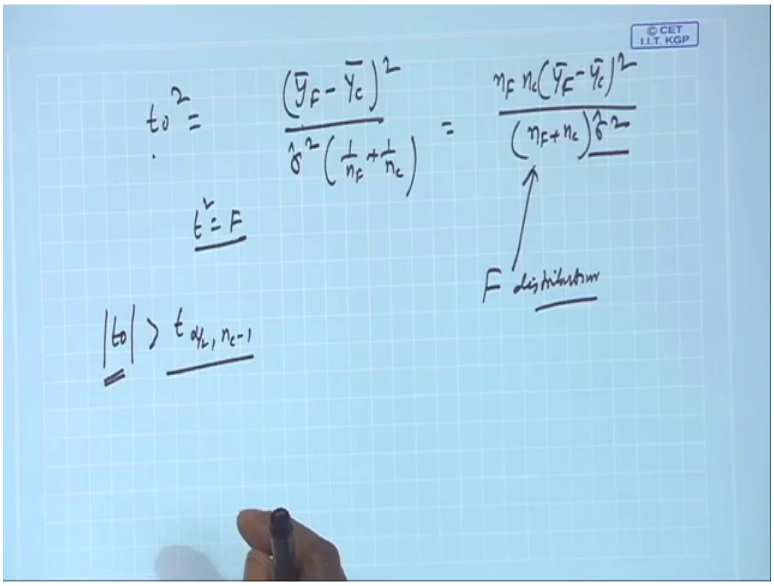
So, obviously, the y bar cap y bar F or y F bar minus y c bar this is random variable. So, having mean is already we have seen that mean value expected bar that minus this will be this and what will be the variance of this y F bar minus y c bar this variance will be we will be sigma square 1 by n F plus 1 by n c.

So, you all know that if y is suppose distributed, normally it is let it be y is distributed with mean mu and variance sigma square when you collect a sample y bar will be distributed with mean mu and variance sigma square by n. That is what is happening sigma square by this because of the constant error variant across all level of x or y the sigma square is the y variability whether it is conducted in factorial point or central point irrespective of this will be error variability.

So, this one we will so, sigma square will get and we will basically use sigma square cap, fine we all know a statistical 0 which is y F bar minus y c bar minus by [FL] this minus expected value of y F bar minus y c bar divided by variance square root of variance of y F bar minus y c bar this is what is t statistics or this is a what I can say this is a quantity which is t distributed and it will be having n c minus 1 degrees of freedom, with followed t distribution with n c minus 1 d o f when under H 0 is true.

So, this quantity is y F bar minus y c bar minus expected value is sum total J equal to 1 to K beta j j divided by variants is sigma square root of 1 by n F plus 1 by n c. Now, when H 0 is true H 0 is sum of beta j j equal to 0. So, this quantity becomes 0 so, this will become y F cap minus y c bar cap by sigma square root of 1 by n F plus 1 by n c.

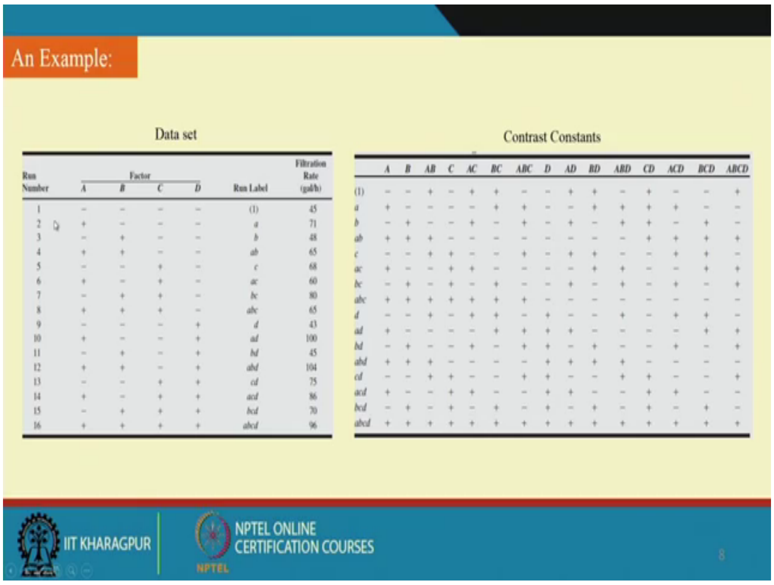
(Refer Slide Time: 23:32)



So, now, what will happen if you square it suppose I make t 0 square so, this will be y F minus y c bar this square divided by sigma square into 1 by n F plus 1 by n c. Now, this is nothing, but n F n c y F bar minus y c bar square divided by n F plus n c into sigma cap square. So, here what will happen, this t square from theory we know t square equal to F so, this quantity follows F distribution ok.

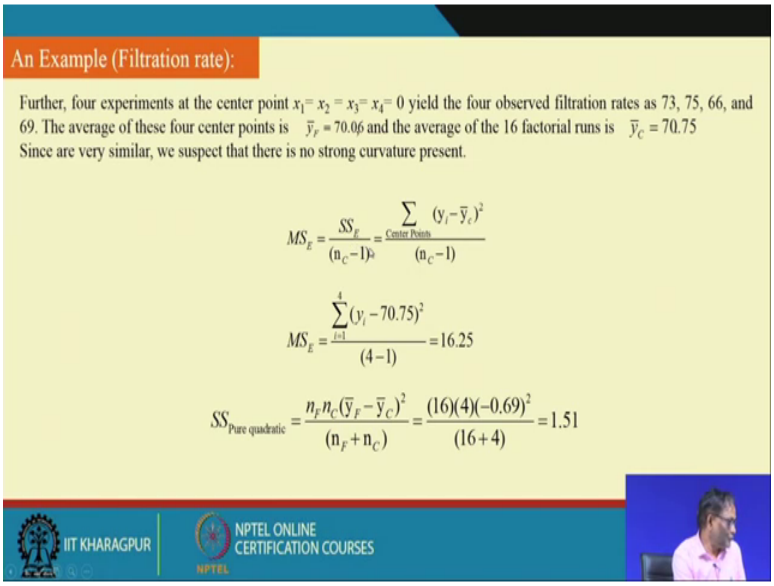
So, that mean either you do you do one thing that absolute value of t 0 greater than t alpha by 2 n c minus 1 and then we say that quadratic effect is actually if this is the case. So, H 0 is quadratic effect is not there if this quantity is less than greater than this then H 0 will be rejected other way it will be accepted. So, I will give you the example here that what we have discussed in last class that filtration rate.

(Refer Slide Time: 25:09)



So, with single replicates so, there are four factorial points and these are the run levels and these are the response values and this is what is our contrast constants for all those effect parameters and this things are known to you.

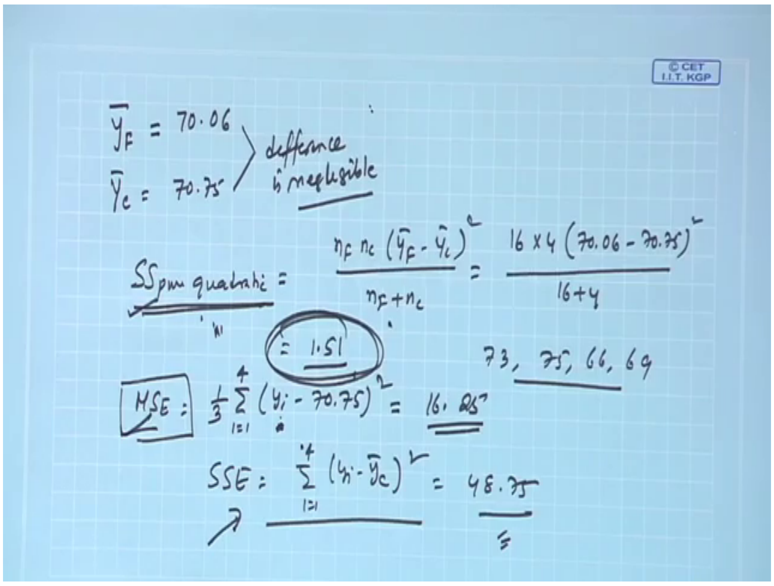
(Refer Slide Time: 25:30).



And now, suppose in addition to this that four experiments were conducted at the central point and the y value filtration rates are 73, 75, 66 and 69.

So, in that case what is the average at the central point, average at the central point means the average of sum of all those filtration rate values divided by 16 and this will become 70.06 and average of this 73, 75, 66 and 69 this is 70.75. So, in that example that y bar F that is the factorial point part it is 70.06 and y bar c central point part is 70.75 and visually also it is the difference is negligible.

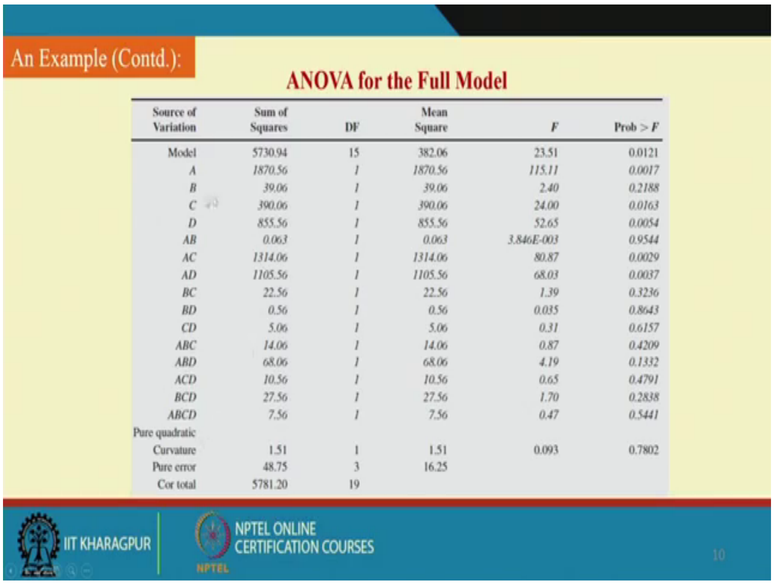
(Refer Slide Time: 26:15)



So, now, what happens you we will use the test what will calculate we will calculate SS pure quadratic so, you know this is n F n c y bar F minus y c bar this square divided by n F plus n c. So, n F is 16 n c 4 and y F bar is 70.06 minus y c bar is 70.75 square divided by 16 plus 4 the resultant value will be 1.51. Now, we want to compute the MSE value so, the at the centre point values can be used to compute MSE this is nothing, but sum of I equal to 1 to 4, 4 centre point values y i minus 70.75 this square divided by 3.

So, this will give you when all centre point values like 73, 75, 66 and 69 is used in this formulation then you will get MSE value is 16.25. Now, this SS and MSE because we because of having centre point and you are able to compute this central point is well and also because of having both factorial and centre points experiment you are able to compute the quadratic part also effect SS. So, now, let us see that after adding the central points when the final what is the final results.

(Refer Slide Time: 28:35)



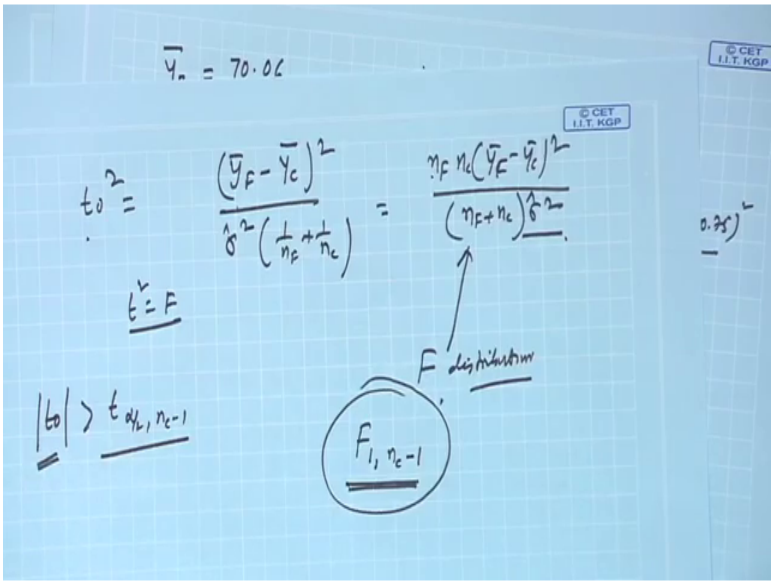
Source of Variation	Sum of Squares	DF	Mean Square	F	Prob > F
Model	5730.94	15	382.06	23.51	0.0121
A	*1870.56*	*1*	*1870.56*	*115.11*	*0.0017*
B	*39.06*	*1*	*39.06*	*2.40*	*0.2188*
C	*390.06*	*1*	*390.06*	*24.00*	*0.0163*
D	*855.56*	*1*	*855.56*	*52.65*	*0.0054*
AB	*0.063*	*1*	*0.063*	*3.846E-003*	*0.9544*
AC	*1314.06*	*1*	*1314.06*	*80.87*	*0.0029*
AD	*1105.56*	*1*	*1105.56*	*68.03*	*0.0037*
BC	*22.56*	*1*	*22.56*	*1.39*	*0.3236*
BD	*0.56*	*1*	*0.56*	*0.035*	*0.8643*
CD	*5.06*	*1*	*5.06*	*0.31*	*0.6157*
ABC	*14.06*	*1*	*14.06*	*0.87*	*0.4209*
ABD	*68.06*	*1*	*68.06*	*4.19*	*0.1332*
ACD	*10.56*	*1*	*10.56*	*0.65*	*0.4791*
BCD	*27.56*	*1*	*27.56*	*1.70*	*0.2838*
ABCD	*7.56*	*1*	*7.56*	*0.47*	*0.5441*
Pure quadratic Curvature	1.51	1	1.51	0.093	0.7802
Pure error	48.75	3	16.25		
Cor total	5781.20	19			

So, you see all the model parameters starting from the main effects to the fourth order interaction effects there sum square the way we have shown you in the last class this is as it is there. So, in addition what happen the pure error that is the curvature and quadratic error that all are pure error is basically this is the error when y is at the central point y values are observed at central point. So, you have seen that the mean MSE value is 16.25. So, that mean 16.25 into 3 48.75 so, that mean y i SSE, SSE at central point is what SSE at central point is sum total of I equal to 1 to 4 y i minus y c bar square.

So, this value is 48.75 so obviously, it has three degrees of freedom so, we have MSE divided by 3 16.25. So, in addition here is the curvature the quadratic curvature pure quadratic curvature 1.51 that I have shown you earlier in calculation. In calculation we have seen that SS pure quadratic curvature is this so, you are now having all information with you.

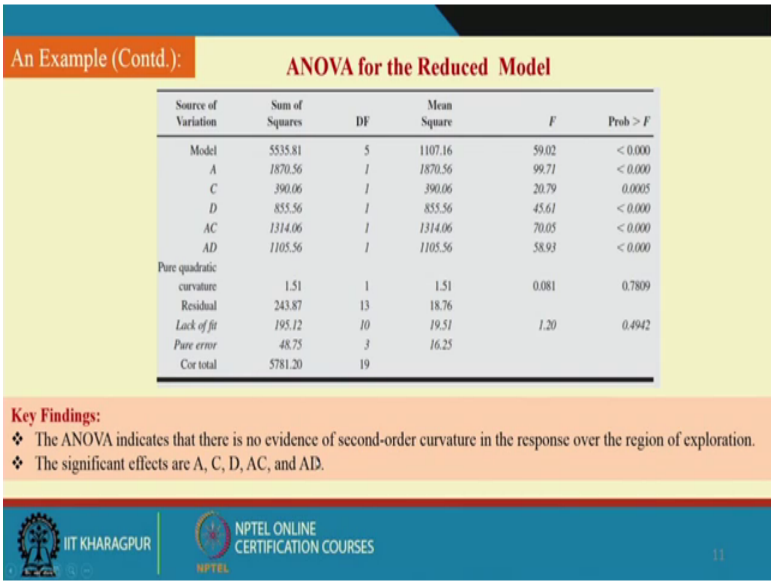
So, we have 16 factorial runs and 4 central point runs in total 20 observations so, that is why the total is this and it's degree of freedom is 19 and all other cases it is clearly computed and finally, you are in a position to compute the F also and then using F statistics you are able to find out what is the what is your significance level. Whether how many are significant or not, now SSE you did quadratic curvature this also with 1 degrees of freedom ok.

(Refer Slide Time: 30:54)



So; that means, when I say that this follows this particular part follows that F distribution and this is F with 1 for the and MSE that whatever will be the error, that degrees that means the 3 n c minus 1. So, 1 n c minus 1 that degrees of freedom this follows F distribution with this degrees of freedom. Here now the traditional way you will do you see which are significant, which are not significant remove the insignificant part you will be getting this is the final table ANOVA table for reduced model.

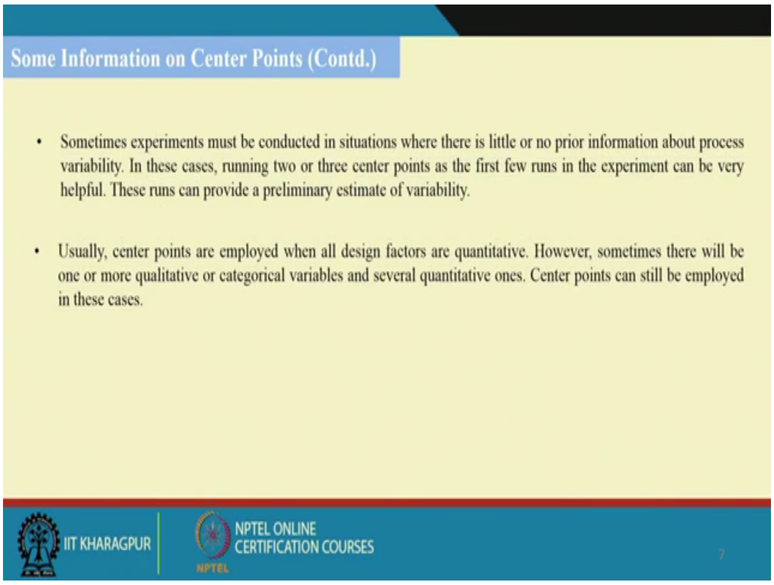
(Refer Slide Time: 31:33)



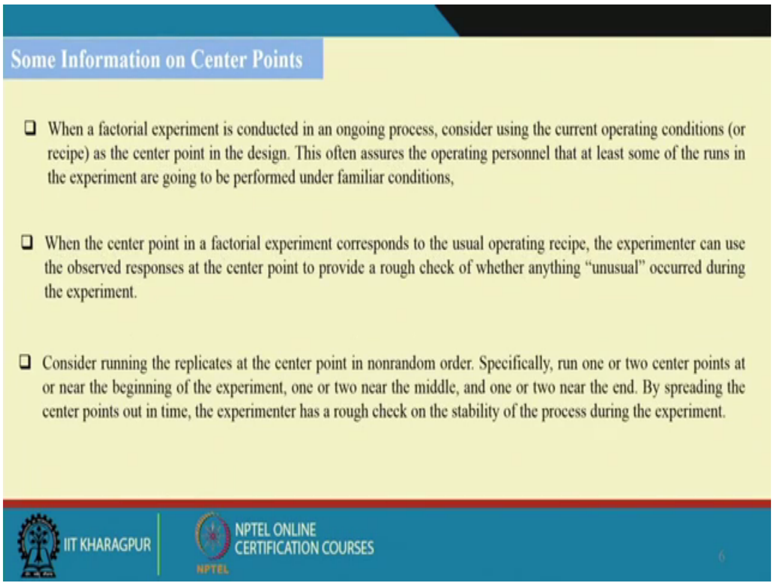
Source of Variation	Sum of Squares	DF	Mean Square	F	Prob > F
Model	5535.81	5	1107.16	59.02	< 0.000
A	1870.56	1	1870.56	99.71	< 0.000
C	390.06	1	390.06	20.79	0.0005
D	855.56	1	855.56	45.61	< 0.000
AC	1314.06	1	1314.06	70.05	< 0.000
AD	1105.56	1	1105.56	58.93	< 0.000
Pure quadratic curvature	1.51	1	1.51	0.081	0.7809
Residual	243.87	13	18.76		
Lack of fit	195.12	10	19.51	1.20	0.4942
Pure error	48.75	3	16.25		
Cor total	5781.20	19			

Here, you see that the ANOVA indicates that there is no evidence of second order curvature that mean the quadratic effect. The significant effect are A, C, D, AC and AD. Almost similar thing we have seen in using the other policy that we have discussed in last class, but here it is better it gives you better.

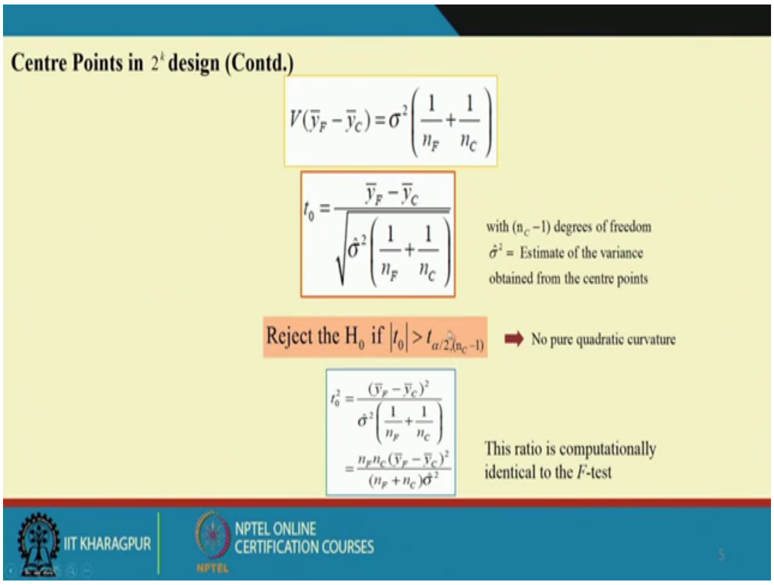
(Refer Slide Time: 32:02)



(Refer Slide Time: 32:03)



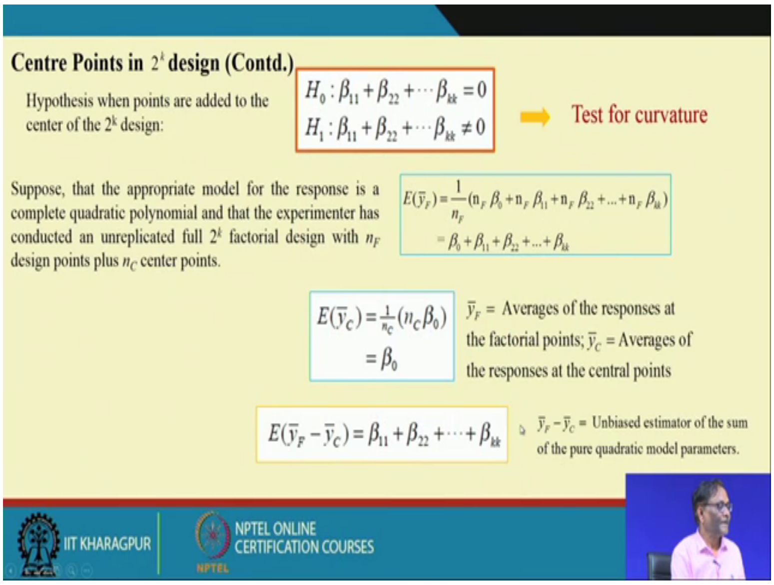
(Refer Slide Time: 32:03)



And another one is that, there is a typographical mistake here reject H 0 if this greater that this is correct.

So, this shows that quadratic if H 0 is rejected then there is quadratic curvature it is not no quadratic curvature. This is not this is quadratic curvature is there if this is less than degree of freedom then there is no quadratic curvature. So, the mistake here is it should be quadratic curvature is there if I follow this so, these are the some trivial typographical mistake you must be able to find out all those things.

(Refer Slide Time: 32:44)



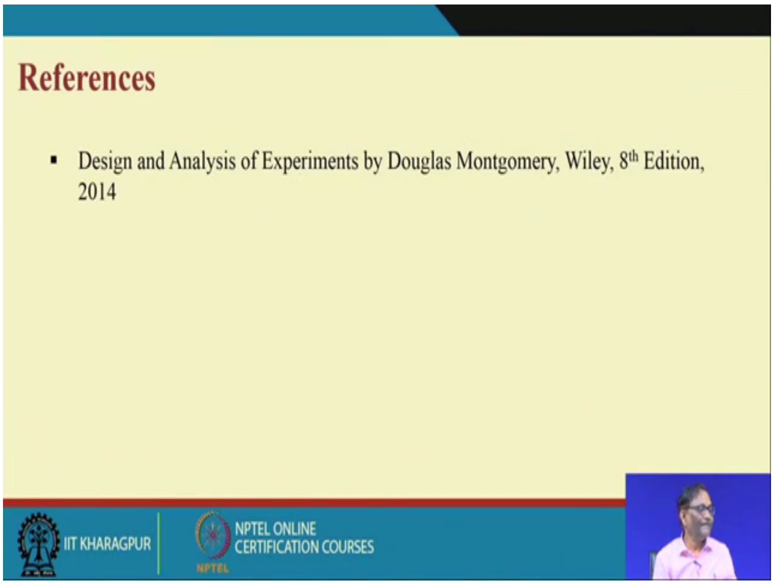
So, another important part here in the central point is that so, when you go for central point experiment. So, some of the hints when a factorial experiments is conducted in ongoing process, consider using the current operating condition as the centre point in the design. This open assures the operating personnel that at least some of the runs of experimental going to be perform under familiar conditions. So, that is in other way I explain in the beginning that the centre points will be familiar condition for the actually you choose the central point for the familiar condition where most of the production runs are made.

When the centre point in the factorial experiment correspond to usual operating recipe, the experimenter can use the observed responses at the centre point to provide a rough check whether anything “unusual” occurring at the experiment or not.

So, now consider running the replicates at the centre points in nonrandom order. Specifically, run one or two centre points at or near the beginning of the experiment, near one or two at the middle and one or two at the end. By spreading the centre points out in time, the experimenter has a rough check on stability of the process during experiment. In the same manner sometimes experiment must be conducted in situation when there is little or no information about process variability. In these cases running two to three centre points as the first few runs in the experiment can be very helpful. These runs can provide preliminary estimate of variability.

Usually, centre points are employed when all design factors are quantitative. However, sometimes there will be one or more qualitative or categorical variables and several quantitative ones. Center points can still be employed in these cases.

(Refer Slide Time: 34:47)



So, with this I show you the reference that we have taken material, lecture prepared lecture material by taking information and available resources in that book.

Thank you very much I hope that you understand the use of central points in 2 to the power k factorial design and you will be able to solve problems when central point data experimental data, along with factorial point experimental data will be given to you.