Statistics for Experimentalists

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

Lecture - 36
Factorial Design of Experiments: Example Set (Part C)

So now we will construct the analysis of variance table or ANOVA.

(Refer Slide Time: 00:21)

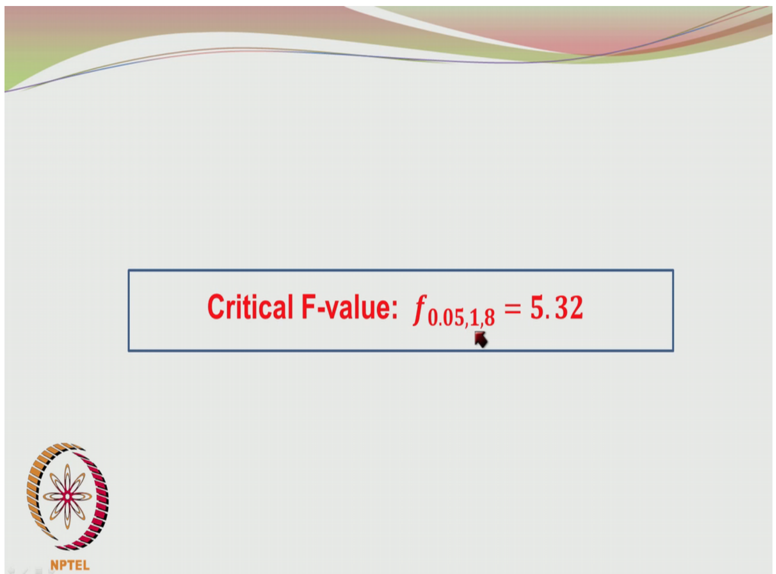
Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	f_o
A	1	110.25	110.25	$\frac{110.25}{0.3675} = 300$
B	1	172.1344	172.1344	$\frac{172.134}{0.3675} = 468.4$
C	1	0.25	0.25	$\frac{0.25}{0.3675} = 0.680$
D	1	116.64	116.64	$\frac{116.64}{0.3675} = 317.39$
Error	abc(n-1)=2.2.2.(2-1) =8	2.9397	0.3675	
Total	abcn-1 = 15	399.2734		

In the analysis of variance table we have 5 columns, the first column is the source of variation from different factors and their interaction, the degrees of freedom associated with each of the factors and their interactions, the sum of squares linked to each factor and its interaction, then mean square is obtained by dividing the sum of squares by the degrees of freedom. So mean square is equal to sum of squares divided by degrees of freedom.

And you also have the contribution from error as the source of variation last but not the least, and the degrees of freedom for error would be abc*n-1 that would be 8 for the present case, as we have only 2 repeats per round. And you can see that the sum of squares and the degrees of freedom are used to find the mean square for each main effect or the interaction, you have only 1 degree of freedom for each of them for error it is 8.

So here you divide 110.25/1 you get 110.25, and then you also divide the sum of squares due to error which is 2.9397 divide it with 8 and you will get 0.3675, so you have 110.25/0.3675 which is approximately 300. And similarly, the next sum of squares for B would be 172.1344 and that you divide by 0.3675, the mean square error contribution you get 468.4. For C it will be 0.25/1 which is 0.25 that divide by 2.9397/8 which is 0.3675, so you have 0.25/0.3675 as 0.68. Similarly, you find it for D, and you can see that except for C all other f values are pretty high.

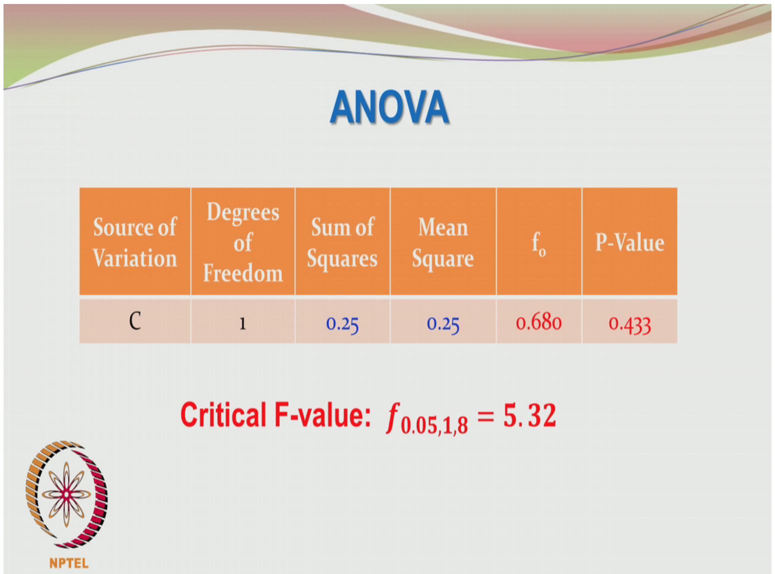
(Refer Slide Time: 02:35)



The critical f value we are using the level of significance as 0.05, and you have one numerator degree of freedom, and 8 denominator degrees of freedom, so the critical f value would be 5.32. You compare the actual f value with the critical f value, and see whether the f value is exceeding the critical f value, critical f value is 5.32, so 300 is obviously >5.32, 468.4 is even more so, 0.68 corresponding to factor C is however < 5.32.

And for factor D 317.39 is higher than the f value corresponding to the critical one. So you can see that factor C is not significant here, so analysis of variance has been used to find that C is insignificant.

(Refer Slide Time: 03:50)



So the C insignificance is highlighted in this slide, what we mean by C being insignificant is that factor C is not influencing the response, and hence it is agreeing with the null hypothesis which said that factor C is not important. On the other hand, the null hypothesis which said factors A is not important, factor B is not important, those hypothesis are rejected because of the f test. You also look at the P-value, you can find for factor C the P-value is pretty high at 0.433 and so.

The null hypothesis would stand to be accepted for this case, the critical probability value was 0.05 and any probability or P-value lower than 0.05 would lead to rejection of the null hypothesis. However, for factor C the P-value came to be 0.433 which was >0.05, and hence we accept the null hypothesis. What is this P-value? The P-value is the type 1 error, then what is type 1 error? The type 1 error is probability of wrongly rejecting the null hypothesis.

If the probability of wrongly rejecting the null hypothesis is very small then we reject the null hypothesis, if the probability of wrongly rejecting the null hypothesis is pretty high then we accept the null hypothesis. In this case the probability of wrongly rejecting the null hypothesis is as high as 0.433, since this probability is quite high we cannot reject the null hypothesis which says that factor C is insignificant.

(Refer Slide Time: 06:05)

ANOVA

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	f_o
AB	1	18.438	18.438	$\frac{18.438}{0.3675} = 50.17$
AC	1	43.745	43.745	$\frac{172.13}{0.3675} = 119.03$
AD	1	118.81	118.81	$\frac{118.81}{0.3675} = 323.30$
Error	abc(n-1)=2.2.2.(2-1) =8	2.9397	0.3675	
Total	abcn-1 = 15	180.993		

And we can also calculate the P-values for factor A, factor B, factor C and so on, but let us first complete the ANOVA table for the interactions. We have the binary interactions, we do not have all the binary interactions for 4C2 would be 6 binary interactions, but we are showing only 3 binary interactions, because the other 3 binary interactions are aliased with these 3 binary interactions.

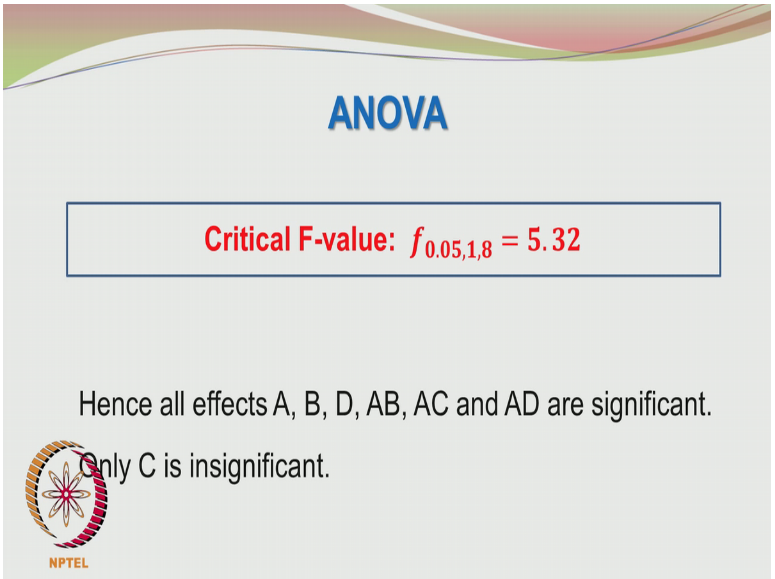
So when you look at AB again the same procedure is followed, sum of squares divided by degrees of freedom you get the mean square, we are talking about the single combinations so we have a single degree of freedom. And here 18.438/the mean square error which is again 0.3675 it is constant for all the effects and their interactions and that comes to 50.17, this is obviously higher than 5.32, and hence we reject the null hypothesis.

So what we are doing here is we are finding the f value, we are comparing it with the critical value, the critical value is 5.32, and then we either accept or reject the null hypothesis. And that is one way, we know the critical value and if the f value is much much different from the critical value, then we are confident that we made the decision correctly. But if the critical value is 5.32 and the f value is 5.33 or 5.30, then it is a original case.

So it is very difficult to confidently either reject or accept the null hypothesis, but usually we do not get values very close to the critical value, we are able to reject the null hypothesis or accept it by a comfortable margin, but anyway we will be finding the P-values and then we will see, very rarely we will find the P-value to be 0.051 or 0.049 and so on. So when you look at all these, it appears that all the 3 interactions.

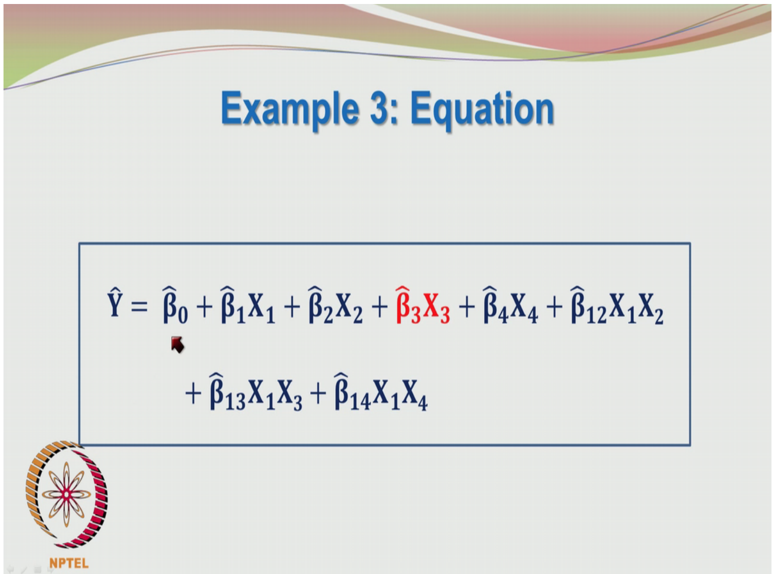
All the binary interactions AB which is aliased with CD, AC aliased with BD, AD aliased with BC, all these binary interactions are significant by looking at the f values, they are lying in the rejection region and hence we can reject the null hypothesis which says that the binary that these binary interactions are insignificant. We are rejecting the null hypothesis which says that the binary interactions are unimportant.

(Refer Slide Time: 09:07)



So among the different effects we have considered A, B, D, AB, AC and AD, only C was considered to be insignificant, and all other effects are significant. We are not looking at ternary interactions ABC here, because that ternary interaction ABC is aliased with the factor D, ternary interactions BCD is aliased with factor A, B is aliased with ACD. So we are having ternary interactions, and these ternary interactions are also aliased with the main factors. So we are not able to obtain the ternary interactions separately.

(Refer Slide Time: 09:57)



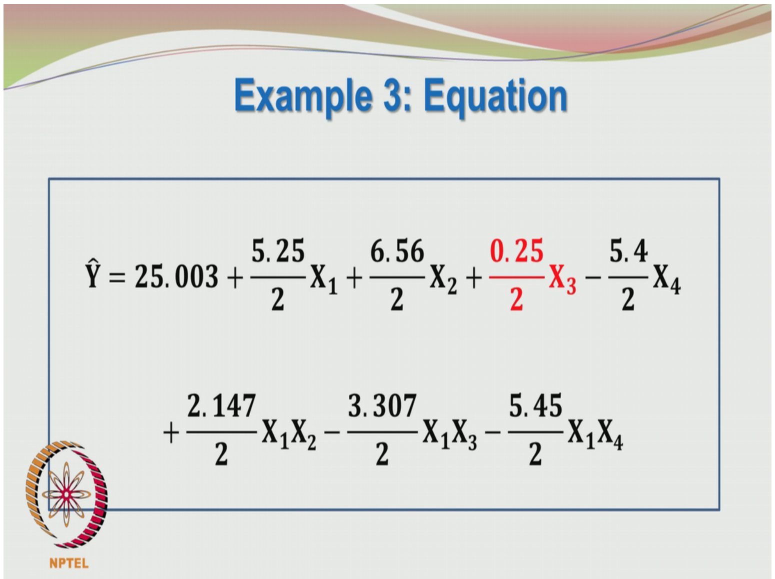
So we can write down the model equation, and the model equation is given by this form, here you have the intercept and this is due to factor 1 or factor A, due to factor B, factor C or factor 3, factor D, interaction between A and B, interaction between A and C, and then interaction between A and D. So based on the ANOVA table, we have seen the beta 3 X3 is a candidate to be removed from the model, because factor C is not significant.

Well I have not given the P-value in this table, yeah it does not seem to be there, but you can calculate the actual P-value corresponding to the f value by using the Excel or another spreadsheet application or you can go to the probability tables for the f distribution, and try to find out the P-value that would be difficult. Because finding the actual probability value using the f distribution is somewhat difficult, only standard values are given for f tables like 0.050, 0.025, 0.1, 0.01 etc.

Only selected limited values are given for the probabilities, but if you want to find a probability corresponding to a certain f value that would be difficult, for example it may be 0.31, whereas you are having only 0.05, 0.1. So but you want to find out probability of 0.31, you will not be able to get, so in this sense this spreadsheet becomes a very valuable tool, so what you do is simply put in excel, f dist 300, numerator degrees of freedom and denominator degrees of freedom.

So for example I may put f dist 300 1, 8, and I will get the probability okay. And that probability value will be very small because the f value is so high, and it is lying well inside the rejection region. So that is how you find out the value of P, I request you to try out a few cases using your spreadsheet or any other statistical software that you may have access to, and find the P-values.

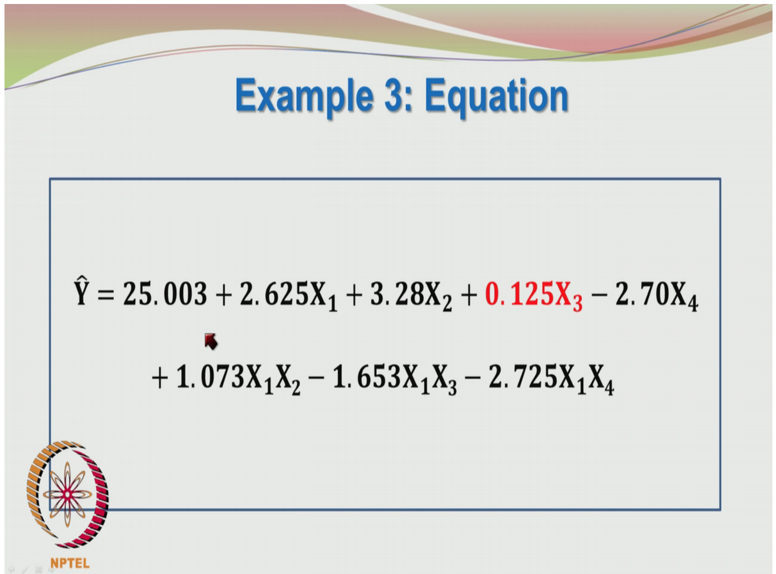
(Refer Slide Time: 13:26)



$$\hat{Y} = 25.003 + \frac{5.25}{2}X_1 + \frac{6.56}{2}X_2 + \frac{0.25}{2}X_3 - \frac{5.4}{2}X_4 + \frac{2.147}{2}X_1X_2 - \frac{3.307}{2}X_1X_3 - \frac{5.45}{2}X_1X_4$$

So plugging the values which we found, we have for the fractional factorial design. The predicted Y is sum of the 25.003 +5.25/2 X1+6.56/2 X2+0.25/2 X3-5.4/2 X4+2.147/2 X1 X2-3.307/2 X1 X3-5.45/2 X1 X4. So except for factor 3 or factor C, all others are present in the model, and also you dividing the effect by 2 account for the jump from -1 to +1. So the effects are divided by 2 here, and they are coefficients of the model equation.

(Refer Slide Time: 14:25)



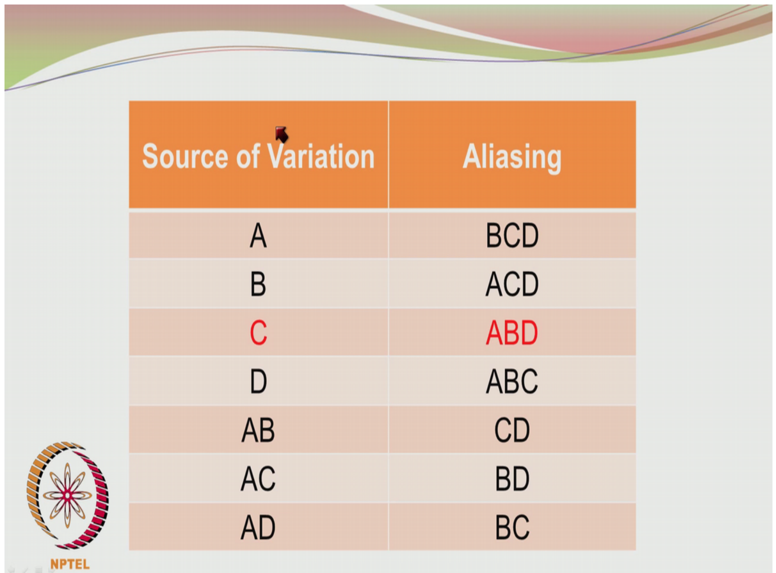
So these are the values, and you can also see the coefficient for factor C is much smaller than the coefficient for other terms. And since these factors are coded as -1 and +1 their actual values do not matter, and so the coefficients are also on the same basis, that is another advantage of coding your experimental data. But without doing a proper analysis of variance, we should not conclude certain factors are important and certain factors are not important, even after coding.

Because it may be 0.05case or 0.04 case, and we might have arbitrarily accepted null hypothesis and rejected the term from the model, whereas it may be a marginal case. So to be very sure and quantify your answers without any subjectivity, it is important that we do the ANOVA table, and for doing the ANOVA table you need to have some estimate of the error variability. Sometimes if you do not do repeats, you think that I can look at the model equation or inspect the model equation.

And then find out which terms are important, which terms are not important that is not recommended, do repeats so that you can get an idea about the error variability, and then you can carry out the analysis of variance where the means square of a particular effect is divided by the square of the error. And then you can make the correct conclusion and also quantify why you rejected some factors, and why you included or did not reject the other factors.

So this is very important for you to do repeats, and hence find this error variability, do the analysis of variance exercise, and get the P-values or see whether the f value statistic is lying in the critical region or, whether it is lying in the acceptance region or in the rejection region.

(Refer Slide Time: 16:51)



Source of Variation	Aliasing
A	BCD
B	ACD
C	ABD
D	ABC
AB	CD
AC	BD
AD	BC

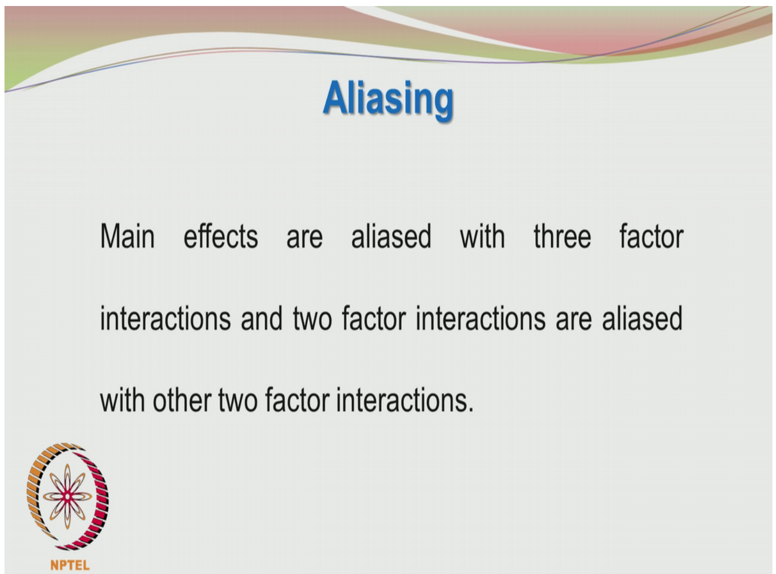
So we have the aliasing table here, where the list the source of variation which we could deduct in the ANOVA table. Even though we are talking about factors A, B, C, C was found to be insignificant, and D, AB, AC and AD. We are also talking about the aliases as well, and that is what the statistical software will also report, so you have BCD, ACD, ABD, ABC, CD, BD and BC. So here if factor C is insignificant, we may automatically think ABD also would be insignificant.

Montgomery has an interesting discussion on this, what he says is the combination C and ABD is insignificant, but this combination may involve a very high value of C and a very low value of ABD. So that there even though they are individually powerful, when they combined they become weak, because C is highly active and ABD is highly negative. And when you take the combination of these 2, the net effect may be insignificant may appear to be insignificant.

And C on its own, and ABD on its own perhaps where exercising very strong effects, and but in the opposite sense and hence you cannot afford to ignore them. There is one argument somebody may put forth, but that is very rarely the case. Montgomery also says that usually the simple explanation that both C and ABD being insignificant is usually the correct explanation for this, even though you can find out some special or extreme cases, where this may not be applicable.

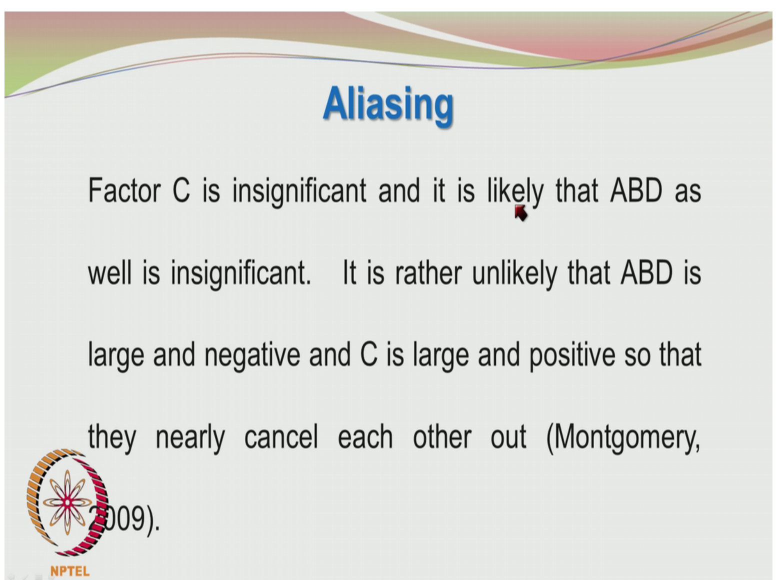
So net effect is if factor C is found to be insignificant from the partial factorial design, the factor aliased with it here this ABD may also be deemed insignificant. But we never know whether that is the correct assumption or not until we carry out the second fraction also.

(Refer Slide Time: 19:29)



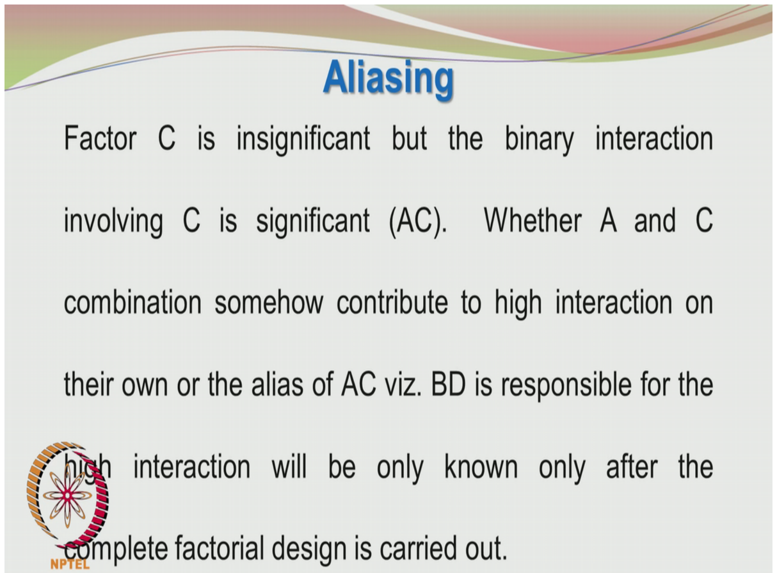
So to repeat the same thing, main effects are aliased with 3 factor interactions, and 2 factor interactions are aliased with other 2 factor interactions.

(Refer Slide Time: 19:40)



To emphasize factor C was considered to be insignificant, and it is likely that ABD is also insignificant. It is rather unlikely that ABD is large and negative, and C is large and positive, so that they nearly cancel each other out, as said by Montgomery 2009.

(Refer Slide Time: 20:00)

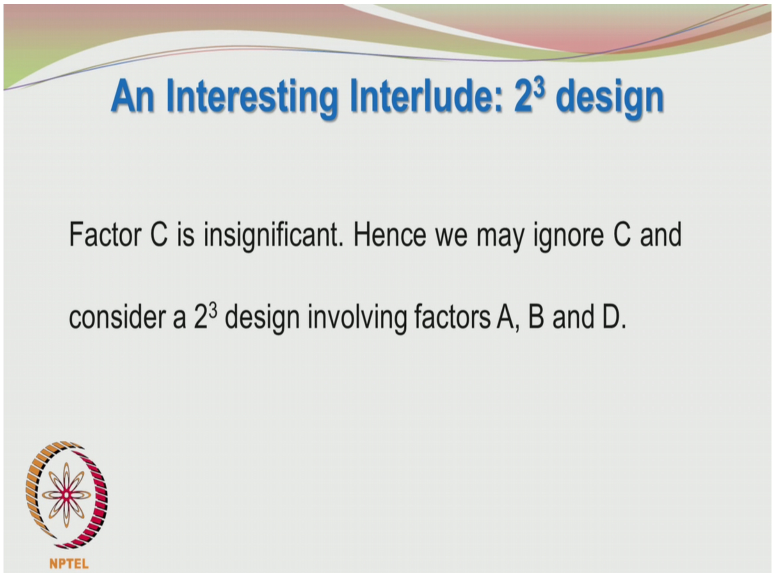


You also look at it, factor C is insignificant but the binary interactions involving C namely AC was significant, that let us verify. We saw that factor C was insignificant very high P-value, so the null hypothesis that C is not having effect was accepted, so you have this C case, and the f0 value was very low 0.68. On the other hand, AC is having a f0 value of 119.03. So somebody may ask naturally, if C is insignificant how come AC is exerting such a large effect, if I expect if C is having a large effect then I can expect AC to be having a strong effect.

Because we know A is also having a strong effect, but C is not having a strong effect or it is insignificant, then how come it is showing up so strongly, something is not correct? Actually if you recollect our discussion AC is aliased with BD, so rather than the effect of AC showing up as significant, it is the contribution of BD which is actually contributing to the high value of the f okay, so the high value of f is 119.03, AC factor is also aliased with BD.

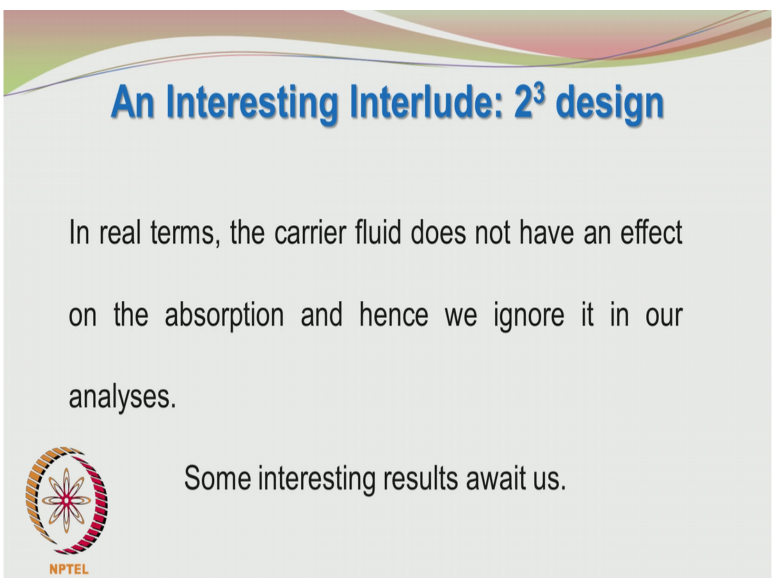
So it may be the BD interaction which is actually showing up as sufficient sum of squares, so that the f value is quite high. So to repeat the contents of the slide, factor C is insignificant but binary interactions involving C is insignificant, whether A and C combination somehow contribute to high interaction on their own or the aliased of AC, namely BD is responsible for the high interaction will only be known after the complete factorial design is carried out.

(Refer Slide Time: 22:23)



So factor C is insignificant, we may ignore C and consider the design as a 2 power 3 design, involving only factors A, B and D. So what we are doing is we are using the first fraction itself involving 8 runs has 2 power 3 design involving A, B and D. Let us see what results we get.

(Refer Slide Time: 22:52)



The carrier fluid does not have an effect on the absorption, on the absorption rather and hence we ignore it in our analyses.

(Refer Slide Time: 23:02)

2^3 Design Results after Ignoring C

Effects	Value	Coefficient	Sum of Squares
A	5.25	2.625	110.25
B	6.56	3.28	172.1344
D	-5.40	-2.7	116.64
AB	2.147	1.0735	18.438
AD	-5.45	-2.725	118.81
BD	-3.307	-1.6535	43.745
ABD	0.25	0.125	0.25
Total			580.2674

When you ignore C, you get the following effects, you find that C is not there, and you have A, B, D, AB, AD, BD, and ABD. And you get the values and the coefficients and the sum of squares.

(Refer Slide Time: 23:23)

Comparison of ANOVAS

Partial 2^4 Design

Effects	Value	Sum of Squares
A	5.25	110.25
B	6.56	172.1344
D	-5.40	116.64
AB	2.147	18.438
AC	-3.307	43.745
AD	-5.45	118.81
C	0.25	0.25
Total		580.2674

Truncated 2^3 Design

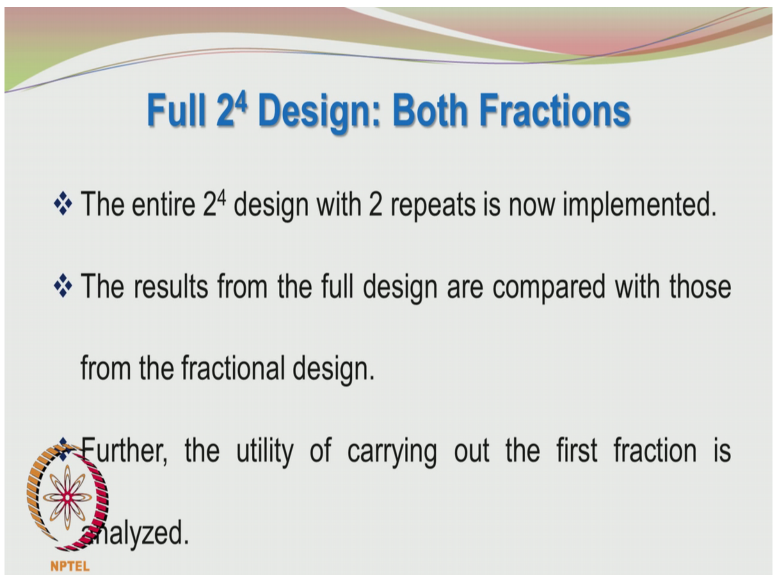
Effects	Value	Sum of Squares
A	5.25	110.25
B	6.56	172.1344
D	-5.40	116.64
AB	2.147	18.438
BD	-3.307	43.745
AD	-5.45	118.81
ABD	0.25	0.25
Total		580.2674

Now let us compare the partial 2 power 4 design with the truncated 2 power 3 design, where C has been deleted. And the results are not as surprising as they seem to be okay, the effects of A are identical, B is identical, D is identical, you have AB that is also identical. And AC, instead of AC you are having BD, and their values are also identical. AD you have same -5.45, and instead of C you have ABD which is 0.25, that is because it is the same response.

And instead of considering C you are deleting it compare completely and then you are putting factor D, and since C is no longer in the picture you are having BD, and you have ABD. And these where exactly the terms which were aliased with the AC and the C, BD was aliased with AC, and the ABD was aliased with the C. And even the sum of squares is identical as you may notice from this slide, the total sum of squares is also identical.

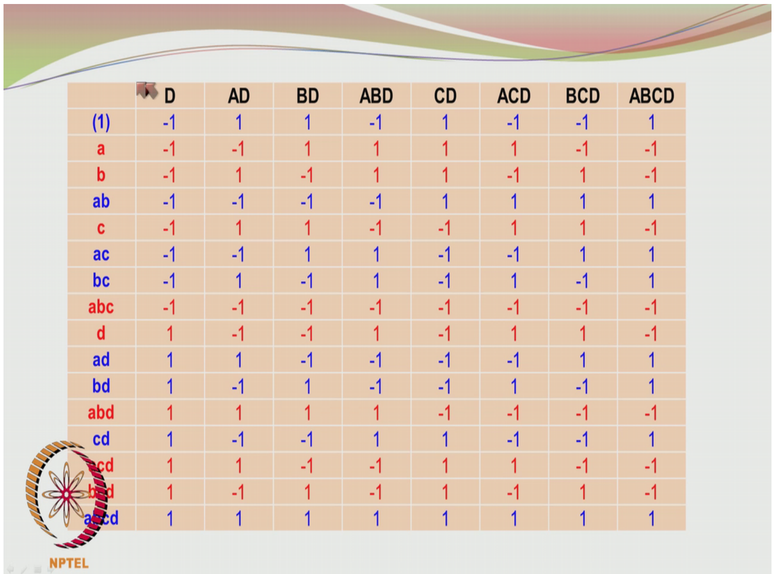
So essentially the truncated 2 power 3 design is telling us the same thing as the partial 2 power 4 design, but it is giving prominence to the aliased factors for involving B and D, and AB and D, because C is no longer in the picture.

(Refer Slide Time: 25:01)



Okay, that was a brief interlude which you can do without doing any additional runs using the statistical software, I use MINITAB here. And now the next and final step is to do the complete 2 power 4 design, we already have seen once you have done the first fraction based on I=ABCD or D=ABC generator, we can construct the next fraction based on I=-ABCD. In other words, we have to look at ABCD column in the design matrix, and then take all the -1's in the ABCD. This we have already seen previously, I am just bringing it to your attention.

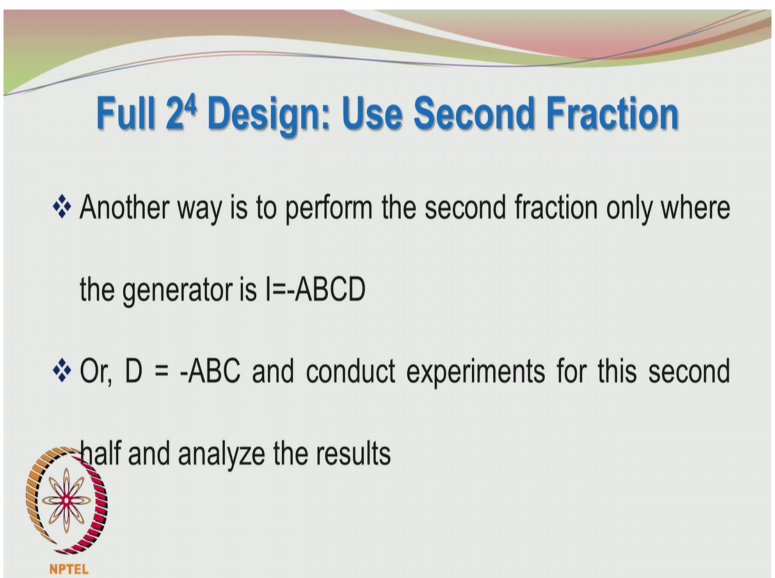
(Refer Slide Time: 25:58)



	D	AD	BD	ABD	CD	ACD	BCD	ABCD
(1)	-1	1	1	-1	1	-1	-1	1
a	-1	-1	1	1	1	1	-1	-1
b	-1	1	-1	1	1	-1	1	-1
ab	-1	-1	-1	-1	1	1	1	1
c	-1	1	1	-1	-1	1	1	-1
ac	-1	-1	1	1	-1	-1	1	1
bc	-1	1	-1	1	-1	1	-1	1
abc	-1	-1	-1	-1	-1	-1	-1	-1
d	1	-1	-1	1	-1	1	1	-1
ad	1	1	-1	-1	-1	-1	1	1
bd	1	-1	1	-1	-1	1	-1	1
abd	1	1	1	1	-1	-1	-1	-1
cd	1	-1	-1	1	1	-1	-1	1
acd	1	1	-1	-1	1	1	-1	-1
bcd	1	-1	1	-1	1	-1	1	-1
abcd	1	1	1	1	1	1	1	1

So whatever +1's we had in ABCD constituted the first fraction, and those settings are given in blue color, the remaining would be all -1's and those settings are also given in the red color. And those settings would be now experimented in the second half, all these are -1's. So we will perform experiments corresponding to the settings where -1 is present in ABCD.

(Refer Slide Time: 26:26)



So we can use I=-ABCD or D=-ABC and conduct experiments for the second half using the settings listed in the table I slowed a couple of slides back, and then we can analyze the results.

(Refer Slide Time: 26:41)

Setting	Organic	Packing	Aqueous	Solvent	% Extraction
(1)	-1	-1	-1	-1	19.22
ad	1	-1	-1	1	19.53
bd	-1	1	-1	1	22.58
ab	1	1	-1	-1	38.28
cd	-1	-1	1	1	21.78
ac	1	-1	1	-1	27.48
bc	-1	1	1	-1	26.43
abcd	1	1	1	1	24.72
(1)	-1	-1	-1	-1	17.52
ad	1	-1	-1	1	19.23
bd	-1	1	-1	1	23.08
ab	1	1	-1	-1	39.59
cd	-1	-1	1	1	22.17
ac	1	-1	1	-1	26.86
bc	-1	1	1	-1	26.25
abcd	1	1	1	1	25.34

First Fraction

So this was the first fraction even though they are coded in red and blue, please do not confuse them with the first and second fractions. I am just showing the first repeat in red color, and the second repeat in blue color for the first fraction, maybe I should have used different colors. So this is the first repeat, second repeat for the first fraction.

(Refer Slide Time: 27:03)

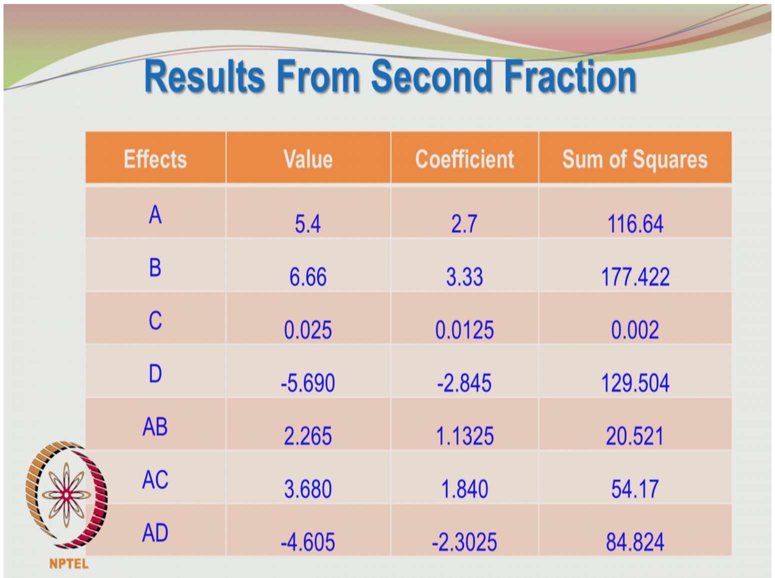
Setting	Organic	Packing	Aqueous	Solvent	% Extraction
d	-1	-1	-1	1	21.26
a	1	-1	-1	-1	26.86
b	-1	1	-1	-1	26.67
abd	1	1	-1	1	25.11
c	-1	-1	1	-1	18.98
acd	1	-1	1	1	19.94
bcd	-1	1	1	1	22.05
abc	1	1	1	-1	39.05
d	-1	-1	-1	1	21.50
a	1	-1	-1	-1	26.19
b	-1	1	-1	-1	27.05
abd	1	1	-1	1	25.20
c	-1	-1	1	-1	18.64
acd	1	-1	1	1	19.93
bcd	-1	1	1	1	22.19
abc	1	1	1	-1	39.26

Second Fraction

Then you look at the second fraction, second fraction are corresponding to the table entries of -1 in ABCD column. And so these settings are same but they are again repeats, and you can see the responses given here D=-ABC, let us check it out. This is D that is –ABC, so this 3 will be -1 but you are having +1 here that is because D=- of the product of these 3 entries, the product of these 3 entries would be +1, so -of -1 would be +1.

And if you look here D=-ABC, ABC is here +1, but it is -1 here because D=-ABC. So that is fine, it is consistent everything is correct, so you are having the second fraction given here.

(Refer Slide Time: 28:07)

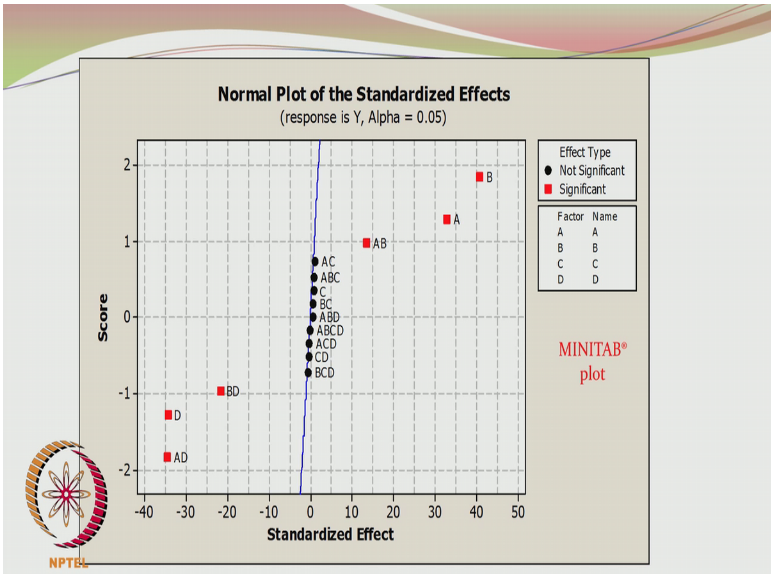


Effects	Value	Coefficient	Sum of Squares
A	5.4	2.7	116.64
B	6.66	3.33	177.422
C	0.025	0.0125	0.002
D	-5.690	-2.845	129.504
AB	2.265	1.1325	20.521
AC	3.680	1.840	54.17
AD	-4.605	-2.3025	84.824

And results from the second fraction we can analyze them separately, and you find these values reported here. By now I think you should be confident to carry out these calculations on your own, and the results are anyway here, so you can carry out the calculations and verify. What you do is you find the values of the effects and then you divide by the 2 to get the coefficients, you also find the sum of squares.

And the second fraction you have A, B, C, D, AB, AC, AD, you do not have the 3 factor interactions, because they are aliased with the main factor interactions, and the 2 factor interaction are aliased with other 2 factor interactions. And so you have only these entries here.

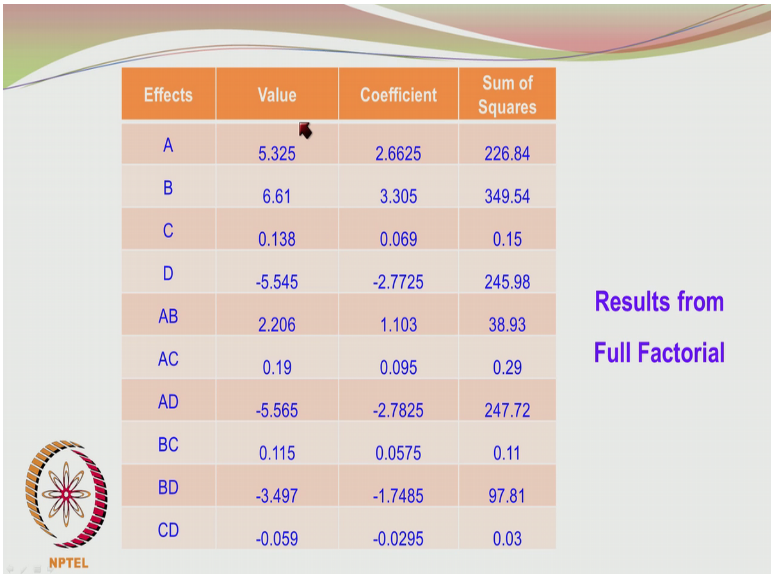
(Refer Slide Time: 28:53)



Now what I have done is combined the 2 fractions, first fraction and the second fraction. The moment I combined the first fraction and the second fraction, I get the complete 2 power 4 design, and hence I can now analyze all the effects. So here you have A, B, C, D, you also have AB and CD, you can see AB and CD. You are also having the aliased effects that means even the aliased effects in the fractional factorial design are showing up in this plot means this is the full design.

There is no aliasing here because this is a full design combination of 2 fractions, and you have BCD which means it was earlier in the fractional factorial design BCD was aliased with A, so now you are having A separately BCD also reported, so this is a full design. And you can see in the full design, so many of these are insignificant, only a few 3+3=6 of the total number are actually significant, this itself is giving us a clue that performing the fractional factorial design was a wise idea.

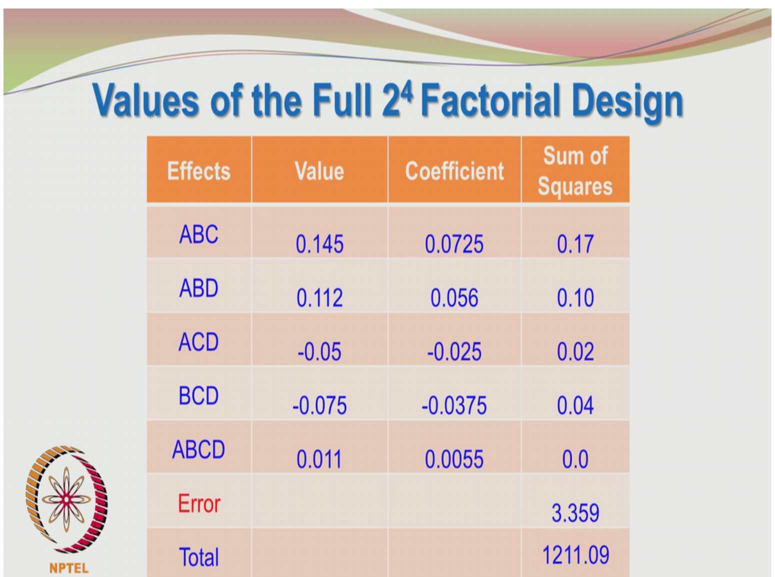
(Refer Slide Time: 30:11)



Effects	Value	Coefficient	Sum of Squares
A	5.325	2.6625	226.84
B	6.61	3.305	349.54
C	0.138	0.069	0.15
D	-5.545	-2.7725	245.98
AB	2.206	1.103	38.93
AC	0.19	0.095	0.29
AD	-5.565	-2.7825	247.72
BC	0.115	0.0575	0.11
BD	-3.497	-1.7485	97.81
CD	-0.059	-0.0295	0.03

So the results from the full factorial are represented here, the value is given here, effects value is given here, and the coefficient is divided is obtained rather by dividing the value by 2, and the sum of squares are also included. The sum of squares values are now higher, because we are combining the 2 fractions to get the full design.

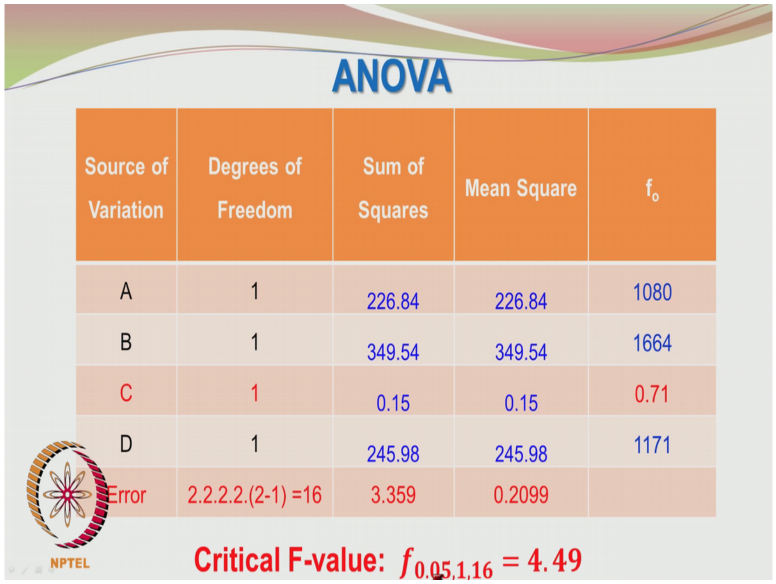
(Refer Slide Time: 30:34)



Effects	Value	Coefficient	Sum of Squares
ABC	0.145	0.0725	0.17
ABD	0.112	0.056	0.10
ACD	-0.05	-0.025	0.02
BCD	-0.075	-0.0375	0.04
ABCD	0.011	0.0055	0.0
Error			3.359
Total			1211.09

And so I gave you the values for the single factor and the binary interactions, now you are have the ternary interactions and the full ABCD interactions. You can see that the values are pretty small, the coefficients are divided by 2 and hence even smaller, and more noticeably the sum of squares are pretty much negligible, when compared to the 100's and 30's you had previously, you can see that the sum of square values are very, very small. You also have the error sum of squares which is 3.359.

(Refer Slide Time: 31:16)



Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	f_o
A	1	226.84	226.84	1080
B	1	349.54	349.54	1664
C	1	0.15	0.15	0.71
D	1	245.98	245.98	1171
Error	2.2.2.2.(2-1) =16	3.359	0.2099	

And when you do the ANOVA, we get the source of variation, degrees of freedom, sum of squares, means squares and f value. And as before we find the critical value corresponding to 0.05 level of significance, 1 degree of freedom in numerator 16 degree of freedom in the denominator corresponding to the error degrees of freedom, we get 4.49 we compare these f values with 4.49.

And see which are lying in the acceptance region and which are lying in the injection region. So factor C is lying in the rejection region, this is old news we already saw it from the first fraction.

(Refer Slide Time: 31:58)

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	f_o
AB	1	38.93	38.93	185.5
AC	1	0.29	0.29	1.38
AD	1	247.72	247.72	1180
BC	1	0.11	0.11	0.52
BD	1	97.81	97.81	466
CD	1	0.03	0.03	0.14
Error	2.2.2.2.(2-1) =16	3.359	0.2099	

And now we see that BC is also lying in the rejection region and so is CD, so any binary interaction associated with C is also lying in the acceptance region okay. So coming back you look at C, C value of f is pretty small at 0.71. So this is lying in the acceptance region, all the other terms listed here all the other f values listed here are lying in the rejection region. Now when you look at it again you find the binary interactions associated with C namely BC and CD are having pretty low f values.

So they are all lying in the acceptance region, what is the critical f value? Critical f value is 4.49, so even AC is having a low f value it is lying in the acceptance region, that means you can accept the null hypothesis that the binary interaction between A and C is negligible. Again the AC interaction is involving the factor C, so any factor binary interaction or even ternary interaction with C becomes negligible from this analysis, BC is negligible, CD is negligible but other binary interactions are significant okay.

(Refer Slide Time: 33:31)

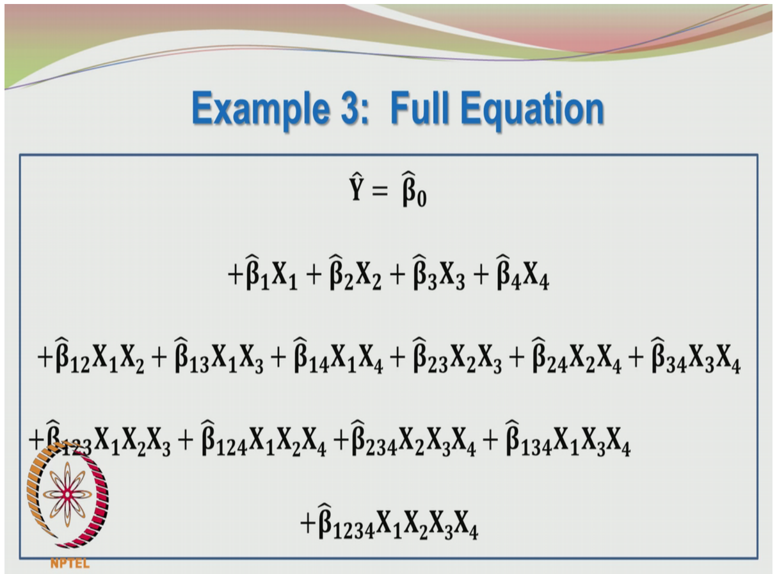
Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	f_o
ABC	1	0.17	0.17	0.81
ABD	1	0.10	0.10	0.48
ACD	1	0.02	0.02	0.1
BCD	1	0.04	0.04	0.19
ABCD	1	0.0	0.0	0.0
Error	2.2.2.2.(2-1) =16	3.359	0.2099	

Critical F-value: $f_{0.05,1,16} = 4.49$

So ternary interactions are also there and you can find that all the f values are pretty small, and the critical f value is 4.49. So all these things the ternary interactions and quaternary interactions will vanish. As a matter of curiosity, what was the error sum of squares previously? And what is the current error sum of squares. So previously the error sum of squares was 2.9397, and now the error sum of square is 3.359 that is increased because we have considered the second fraction also.

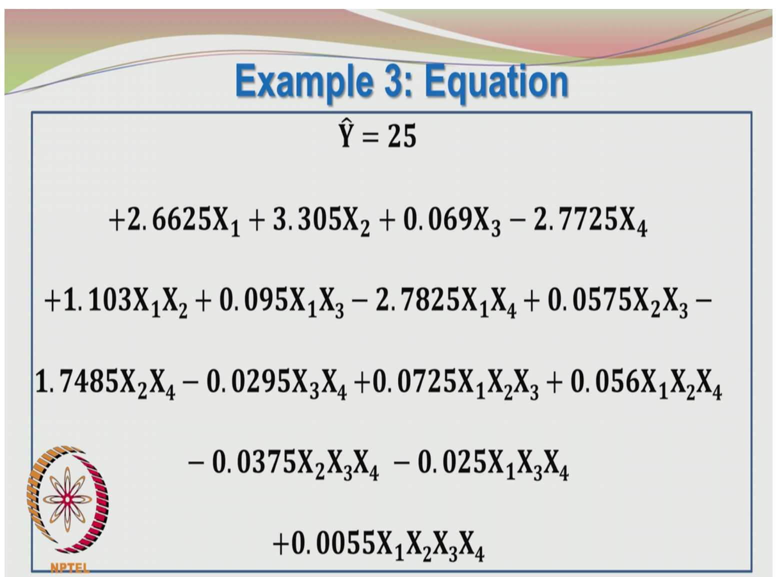
But it is not linear, it will not be exactly half the error in the first half, and exactly half the error will be coming from the second half it is not like that, because error is random so you cannot predict it how it is going to behave. So the error contribution from the second half seems to be lower than the contribution from the first half, but this is expected.

(Refer Slide Time: 34:36)



And so you have the full equation here which is pretty huge, and you can identify all the binary ternary quaternary and even single factor contributions.

(Refer Slide Time: 34:46)



So I am just plugging the value of the coefficients in this model equation.

(Refer Slide Time: 34:52)

Effects	Full2^4	Fract 2^4	Trun 2^4
A	5.325	5.25	5.25
B	6.61	6.56	6.56
C	0.138	0.25	-
D	-5.545	-5.40	-5.40
AB	2.206	2.147	2.147
AC	0.19	-3.307	-
AD	-5.565	-5.45	-5.45
BC	0.115	-	-
BD	-3.497	-	-3.307
CD	-0.059	-	-

And so this table compares the coefficients not the coefficients the effects of the different designs, so these are the basic values for the full 2 power 4 design, and these are the effect values for the fractional 2 power 4 design, the first fraction and the truncated 2 power 4 design. So you have 5.325, 5.25 and 5.25 these values are matching. And this value is different slightly from the fractional factorial design, C is 0.138 and this is 0.25 but anyway C is insignificant.

So we do not really care about what number it is taking. D is -5.545 for the full factorial design and -5.40, AB is 2.206 and this is 2.147. AC the values are quite different the fractional factorial design wrongly attributed so much of effects negative effects to AC, but the full factorial design sets the things in order by saying that AC is not significant it is only contributing to 0.19. But it is BD which is the factor aliased with the AC which is contributing to so much of effect.

But anyway we could have even guessed it with the fractional factorial design alone, because in the fractional factorial design C was shown to be insignificant, even C is shown to be insignificant AC is also likely to be insignificant, and hence this was this could have been attributed even in the fractional factorial design to the binary interaction BD, because AC is aliased with the BD, so that is what you get.

And then you have AD interactions which are comparable, BC anyway this was not considered in the fractional factorial design because BC was aliased with the AD. CD was aliased with AB okay, but anyway even though we did not find these interactions, they were eventually proven to be insignificant in the full factorial design. For example, BC is insignificant, CD is insignificant, they were not detected in the fractional factorial design.

The truncated design also found BD, because we ignored factor C in that analysis, and that interaction effect value is comparable to the full factorial design for BD.

(Refer Slide Time: 37:46)

Effects	Full2^4	Fract 2^4	Trun 2^4
ABC	0.145	-	
ABD	0.112	-	0.25
ACD	-0.05	-	
BCD	-0.075	-	
ABCD	0.011	-	

So I think now we are coming to the end of the story, all the ternary interactions and the quaternary interactions are insignificant as shown here okay.

(Refer Slide Time: 37:55)

Factor(s)	Full2^4	Fract 2^4	Trun 2^4
A	226.84	110.25	110.25
B	349.54	172.1344	172.1344
C	0.15	0.25	-
D	245.98	116.64	116.64
AB	38.93	18.438	18.438
AC	0.29	43.745	-
AD	247.72	118.81	118.81
BC	0.11	-	-
BD	97.81	-	43.745
CD	0.03	-	-

Sum of Squares

So we also now can compare the sum of squares, and these are listed in the table here. The full factorial design has a larger sum of squares, because it is using both the 2 half fractions.

(Refer Slide Time: 38:15)

Sum of Squares

Effects	Full2^4	Fract 2^4	Trun 2^4
ABC	0.17	-	-
ABD	0.10	-	0.25
ACD	0.02	-	-
BCD	0.04	-	-
ABCD	0.0	-	-

And you can see that the full factorial design gives very less sum of squares to the ternary interactions and the quaternary interactions.

(Refer Slide Time: 38:25)

Fractional Factorial Design Model

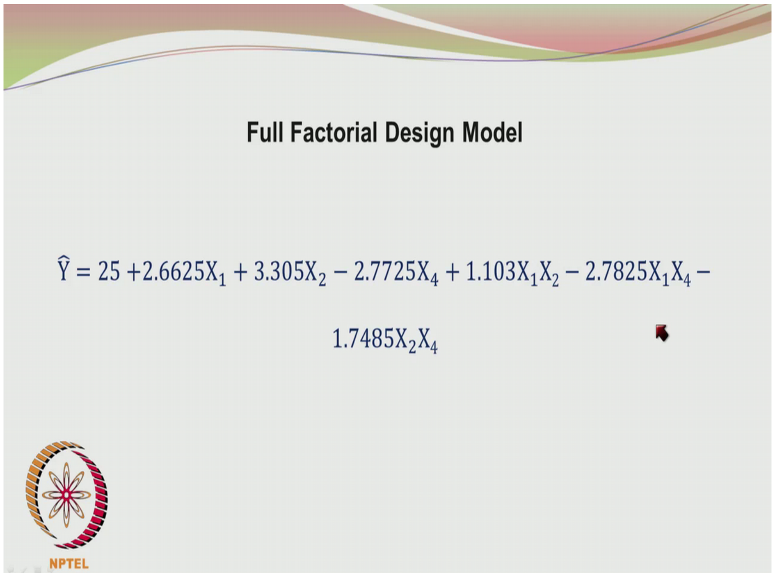
$$\hat{Y} = 25.003 + 2.625X_1 + 3.28X_2 - 2.70X_4 + 1.073X_1X_2 - 1.653X_1X_3 - 2.725X_1X_4$$

Truncated Design Model

$$\hat{Y} = 25.003 + 2.625X_1 + 3.28X_2 - 2.70X_4 + 1.073X_1X_2 - 1.653X_2X_4 - 2.725X_1X_4$$

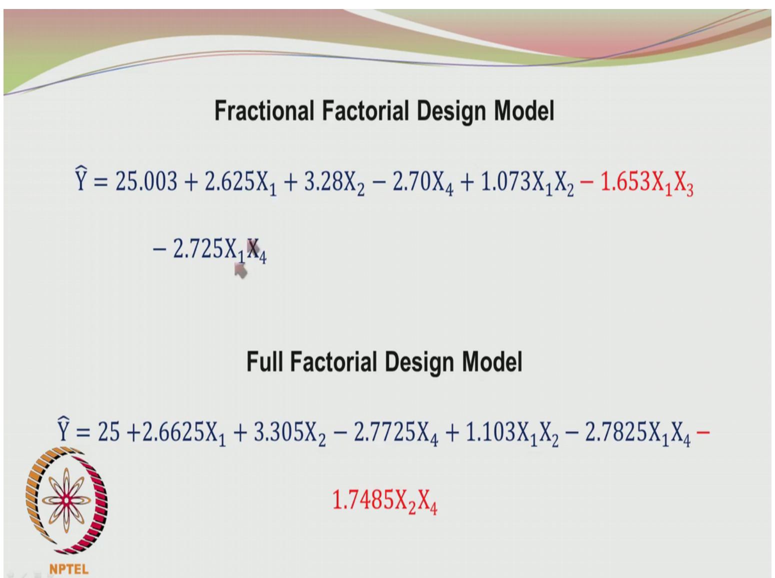
So now we can list the fractional factorial design model and the truncated design model. Everything is identical because the same data was used, but only thing is this factor C was ignored instead of AC interaction you are showing it has BD interaction, but the coefficient is one and the same. The same data is used but C is thrown out, and we are considering only A, B and D, so instead of AC the aliased effect BD is represented here.

(Refer Slide Time: 39:03)



And the full factorial design model is given here, after checking out the insignificant terms, we get this model.

(Refer Slide Time: 39:12)

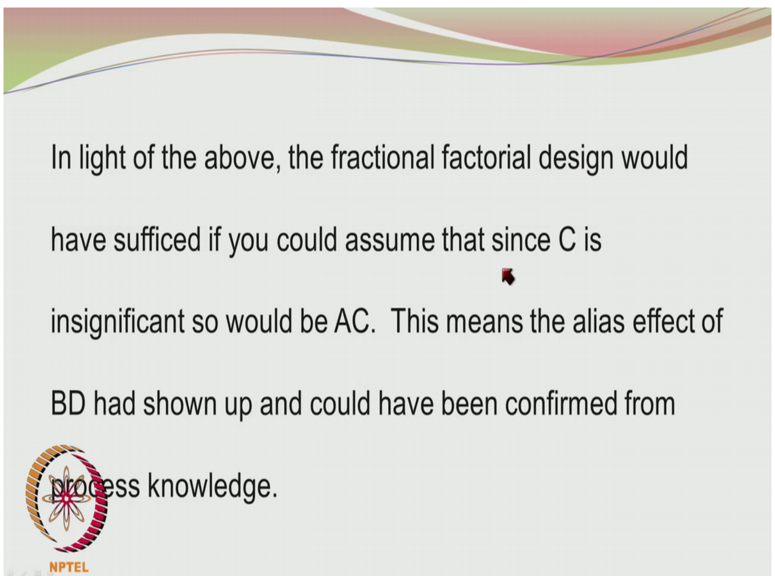


So now we can compare the fractional factorial design model with the full factorial design model. These 2 are comparable, main effect A is pretty much there, main effect B is also pretty much comparable, because the main effects are aliased with 3 factor interactions we do not have to worry about them. So they are pretty accurately represented even in the fractional factorial design model. And then you see factor D, it is also accurately represented.

Factor C was anyway thrown out, because it was not significant, either in the fractional factorial design or in the full factorial design it was thrown out. And then you see the AB interaction, they are comparable. Only problematic case is AC interaction and BD interaction, they are not accurately represented here, but that is AC interaction is aliased with BD interaction. But when you know that C is insignificant, you can say that BD interaction is actually shown up as AC interaction in the fractional factorial design model.

Otherwise, the fractional factorial design did a very good job.

(Refer Slide Time: 40:31)



I request you to go through this, even the fractional factorial design would have sufficed, if you could assume that C is insignificant so would be AC. You could have concluded this from your process knowledge, and hence instead of taking AC even in the fractional factorial design, you could have taken the aliased effect BD as the main contributing binary interaction in the AC, BD combination, AC binary interaction, BD binary interaction.

Even though you are finding out AC, it is actually the aliased factor BD which is showing up, this you could have probably guessed from your process knowledge. So we will conclude now, and we have covered a quite a bit of ground, I request you to go through the problems try to solve them on your phone to the extent possible, or even check some of the calculations partially to the results shown here. This way you can develop confidence in your problem solving and analysis technique.

See what else you can think about the fractional factorial design, and see how it may be modified to suit your purposes. What would have happened if you have taken 1 quarter fraction of a 2 power 4 design? What effects, you would have identified? What effect proudly could have been omitted even in the quarter fraction? So please do these calculations, and see what results you are getting okay, thank you for your attention.