Performance Evaluation of Computer Systems

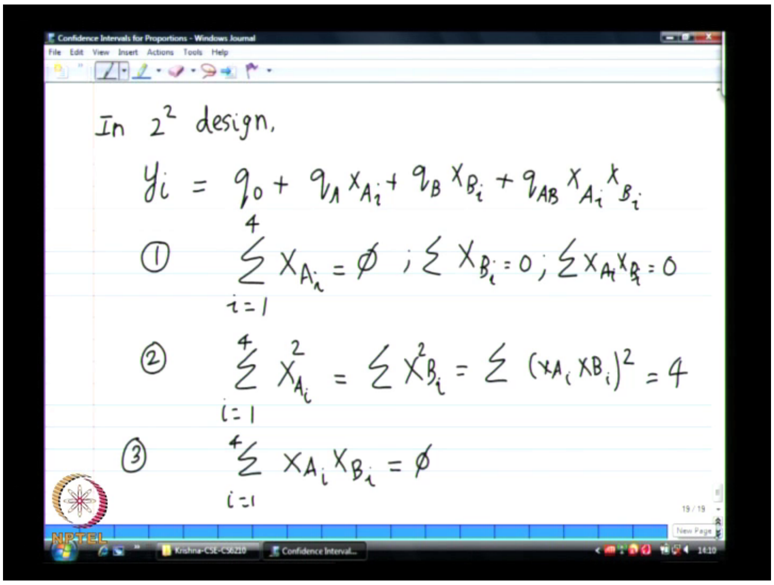
Prof. Krishna Moorthy Sivalingam
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture No. # 35

2k r Factorial design and 2k-p fractional factorial design

Let us come back to 2 square design. So, this is the general expression right y naught y equals q 0 plus qA and so on.

(Refer Slide Time: 00:32)



So, if look at those four values for each of the rows that is y 1,y 2,y 3,y 4 and so on. So I canexpress y i in terms of the corresponding say an X Bi right; so an X Bi, this will be minus 1 plus 1minus 1 plus 1 depending on which row the zero currently looking at. So, thisis generic form(( ))right.So since, you are all asking me that we should prove, why is that some other existing expression was correct, solet us go through the proof on that. And then it is mostly miser than enough to that.

So, some observationsright we say this earlier tooright if we look at X A i across all the four rowsright, what is that summation equal tosorry thisis across across a columnright, soacrossall the fourrows in a given columnrightgo to all the X A i(s) had to 0.And sothus X B i also add upto0right,and X A i B i column also adds upto0,that is aobservation from the table. And if Iadd up the squares,Isquare up each numbers, not that 4.

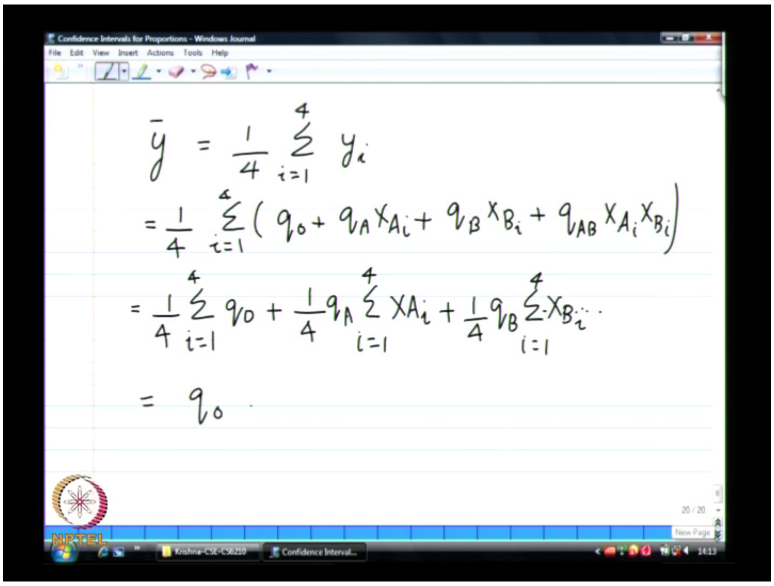
(No audio from 01:53 to 02:04)

This 4.Then how about the dot product of any two columns except that i eth column. We look at column A and B, then what is the dot product?

(No audio from 02:25 to 02:34)

So, likewise A i intothatA i B i column is also 0. So, all theircorresponding column productright go to 0.

(Refer Slide Time: 03:02)



So, if you look at the sample mean,which is simply 1 over 4,all the y i(s)

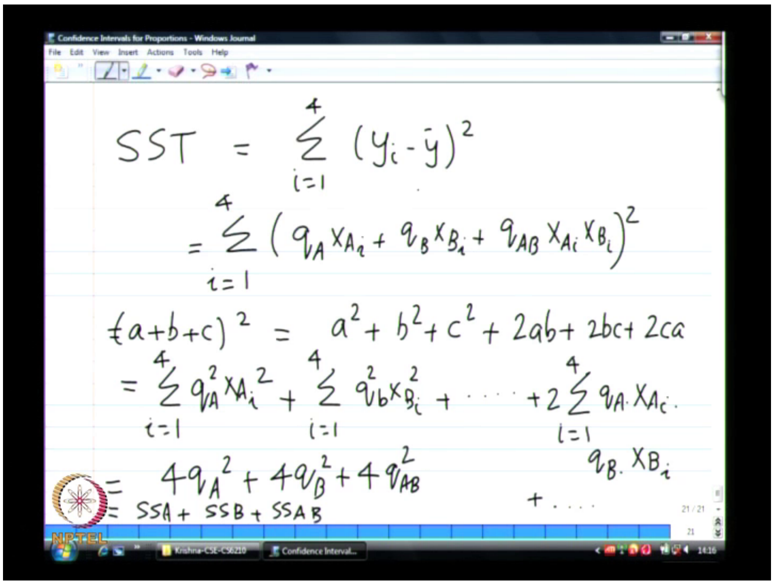
(No audio from 03:05 to 03:18)

So,remember this is defined as followsright. So, y i is q naught plus q Ainto X A i plus q B into X B i and soon. And so the definition of y iright. So, if Iseparate those terms out,so this is 1 by 4 into sigma q naught across all the 4 terms, and again we just (( ))right. So q i gets extracted out, then X A i is what is you know,rememberremaining, and then it is 1 by 4 q B so on.So,what happens in this expression? Sum ofall A i(s) along a column is 0,all B i(s) is 0, A i B i are 0, soeverything has in the first term goes to 0right,that is what we sawin the previous sigma, which is the sigma X B i . So, all thesethreeterms will go to 0and summed up.So therefore, Ihavejust q naught repeated 4times so therefore, mean is q naught that is we derivedthe q naught expression 2, questions.(( )).

(No audio from 05:00 to 05:09)

(( ))

(Refer Slide Time: 05:14)



So, now, we wanted to get that SST expressionright. So, SST is the total variation,so simply y i minus y bar whole square i equals 1 to 4.And since,wejust now do you have y naught equals y bar equals q naught therefore, this is nothing but an expression that looks like this q A in toX A i plus q B into XB i plus q AB X A i XB i whole squared. So, this isbasically, a plus b plus c the whole squared. We have remembereda(( )) this is difficult formula,this is nothing but a squared plus b squared plus c squared plus all the product terms right 2ab2bc.

(No audio from 06:15 to 06:31)

So, applying that you basically have q A squared plus q B squared into X B i squared and soonright. Then plus q A squared B plus X A squared B etcetera;and then we will write oneterm, soplus2 into i equals 4 q A into X A iinto, sobright that is q B into X B iright, that is oneof those terms, other terms Iam not explain. So, basically squares plus thecorresponding product terms.

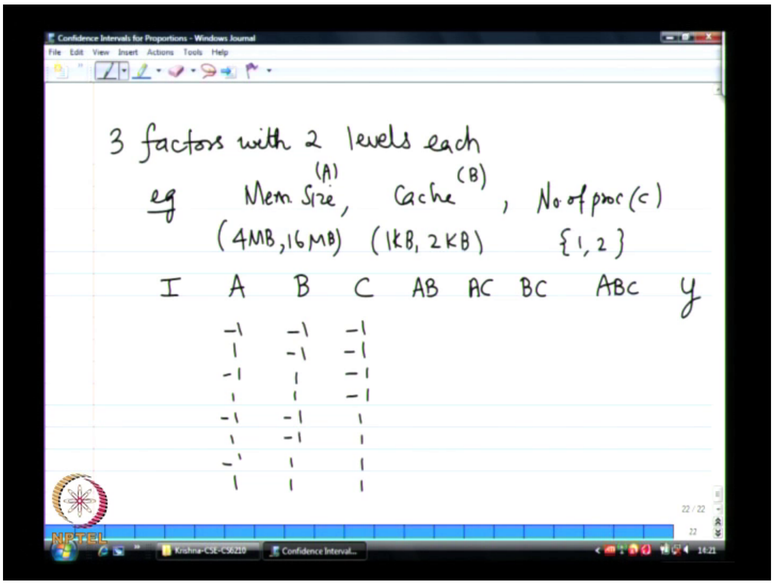
(No audio from 07:27 to 07:37)

So, how does this simplify you know? q A squared, the q A squared comes out,and X A squared is simply 4,we saw thatright. So, 4 q A squared plus 4 q B squaredplus 4 q AB, that is the first threetermsright, from the a squared plus b squared plus c squared, we get thisq A squared 4plus q B squared into 4m squared. Then the remaining dot product terms go to 0 right.XA i and X B i across all the 4rows goes to 0. So, all them get if there is A i in to B i or A i be the everything goes to 0(s), so the last threeproduct terms are simply 0 and that is what we have. So, this is knownas we say SA is how we got that expression.

So, today it is the purpose here is to not fit, the values 4digits find out, which is more important than the other, whichfactor effects your system performance more.Given that each factors has two levels,and you have fixed those twolevelsright;so usually if you have two levels, only you pick two extremesnormallyright; you pick the lowest in the highest, and then see, whether when going from the lowest to the highest part of a declared factor, if it makes more sense than we has we would try to see what happens in betweenright.

Then we will go to next level of experiments, there will had more levels within that same factorsso from 2 GBto 32GB, you see there is a big difference in performance, when Igo to 32 GB compared it to GB, then you would see all at least have 4, 8and16 and see, you know, what is theright spot for stop itand given cost that is the purpose of this particularanalysis. Questions? It just thebiggest questionis, am I going to cover coverchapter 20 or not,because it felt seek chapter20, you got 50(( )). But it is okay, we will enough go through this class. So, this is for atwo factor,Iwill try. So, what if you have threefactors? 3factors with 2 levels, soit is S to q to the power k, sok factor.

(Refer Slide Time: 10:11)

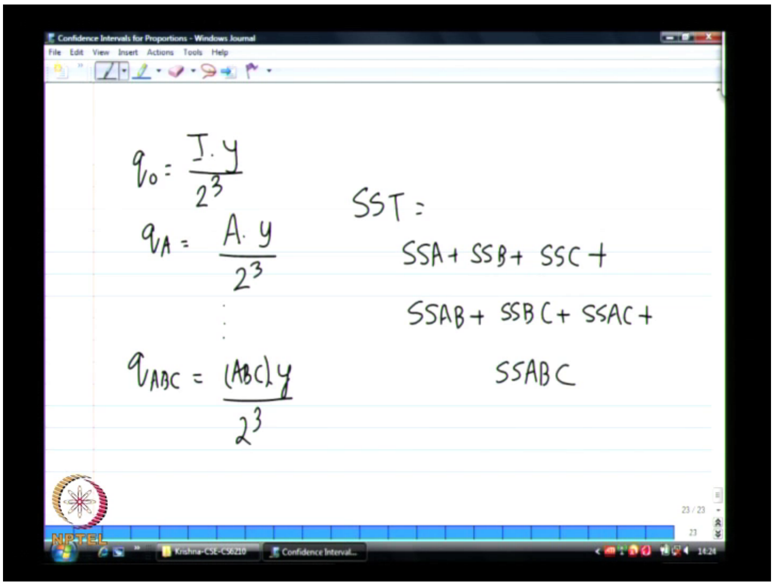


So, if you have 3factors with2 level each. So, for example, Igive you write a system which has memoryis onefactor and there are twolevels, you know say 4 MB and 16 MB, and cache like before, which is 1 KB and 2 KB followed by number of processersright. So, this is say you have factor A, this is factor B, this is factor C, and this is said either 1 CPU or2 CPU, so that is athird factor (( )). So, if Igive you this, and how do we proceed? As before the same principle of generating this table with 2 to the power k rows applies, and then you just haveright. So, you will have, Iam not going to go through the example. So, if Igive you the values, sothis is the column A, column Bright.

So in general, if Ihave k factors, Iwill have k choose 2 effects the whole (( )),then Iwill have k choose 3effectsand soon up to k choose k effects. So there are more effects; in the previous case, there was AB and AB, those are the 3effects, we are interested in.And Igo toright, then we haveA and B have to the interact with each other,then how do A and C interact with each other?Then finally, A, B, C alltogetherhave the interaction witheach otherright. So, this is simply… So, if Ido,I am k choose2, Ihavethree combinations right, so threepossible combinations, let say therefore AB is a column, AC is another column, BC is anothercolumn, ABC is third last column. So, Ihave total of 7 effects.

Now, SST equal SSA plus SSB plus SSC and soon; say in calculated with before us, and then remember with the last column, you know we used to use this y. So, the value would fill up this column is normally, if youlook at my binary value, we have set at 00 1 and so on here it is the. So, will keep the last row constant,or the last column constant,this onewill change every2 rows, and this onechanges every other. So, once Iget this, then it is exactly, then everything else is the same. So, the value for Iright, soagain remember it will be the total column wise dot product. So, Iinto I will give you may q dot; then the column A into Y right,A dot y divided by 4 give me myq A and so on, same same steps as before (( )).

(Refer Slide Time: 13:16)



So, Idot y divided by 4 will be q naught, qA will be column A dot y divided by 4 and soon; q ABC will be that ABC column into 4 divided by into y divided bysorrydot product(( )). Then your SST has more factors SSA plus SSB plus SSC individual factors,then combination factors, nested analysis is detected. So, if you know, how todo (( )) they will do 3in the final, stare missing. So that is, and if you really want Ihave posted samples of exercises we have today, which oneof them includes… there is onlyonequestion in this whole chapter 171 which is basically 3 effects, and we can try that after.

So, this is the end of chapter17. So as per my schedule,I havesupposed to go to single factor, onefactoranalysis, there Ihave set of qualitative variations.So, different CPU types for example, So, Igive you experiments with onlyonefactor would it had several levelsright, but the several levels, whatever 10 levels,15 levels say,10 different CPUs have beeninvestigated, then what is the relative… So, what there you want to find out isrighthow much of the variations is because of the difference in the different factor levels, then how of it is because of errorsright, which you want to make sure that so your SST is simply2 factors; oneis…The

2 components; oneis SSA, which is the variant factorsright, the CPU type it has been varied;and whether the exploit level also that comes to the picture, sothat only SSTequals SSA plus SSCthat is important.Iam not going to go through chapter20(( )).

(( ))

Did Isay 4, sorryeverything in 8(( )). And SSA will be 2 cube2 to the power 3 actually, Iwrote as2 squared last time; SSA will be2 to the power 3, but your q will be squared,2 to the power 3in 2 square plus q to the power cubic squaredright;your q has do not cube only the factorsk factor will come. So, chapter20 Iam skipping, because it is little bitlate to start on a chapter2 days before the final exam. But sowhat,instill what Iam going to dois Istill continue with some other variations, whatever Italk about whether the fig and the final or not at the end ofthe descriptionfor you know, just listen.

(( ))

That what will have do you have to take the general2 k form formulationfrom that we havederived. Yes you mean experimentally, you might have different levels for different factors. If they are all the same, then we have this sort of analysis. If it is…

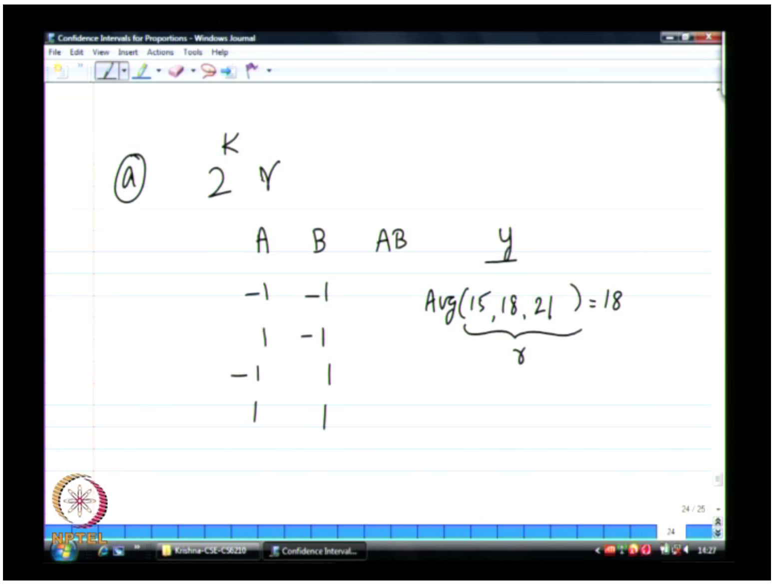
(( ))

(( )) Your question was we have different levels of different levels in eachright. So, in that case, you will still have the full factorial design, but they are going to have the sort of nice analysis when we have different levels.

(( ))

You can design sort of experiments, this kind of analysis, we will have to go back.

(Refer Slide Time: 17:33)



So, somethingis that useful to knowright the first variation is what happens, if Ihave2 to the power k tests or designs experiment types, but they run each other for r applicationsrightthat is it. So, if I 2 to the power k, but each…That what we are going training youright, run multiple replications. So, if they run replication how does my analysis(( ))right. So, with replications, they have only couple of changes to the analysis.You are sowhat we will have isright say AB ABlike beforeright,let us say only2 levels we are looking at, sominus 1 minus 1 minus 1 1right.

(No audio from 18:10 to 18:30)

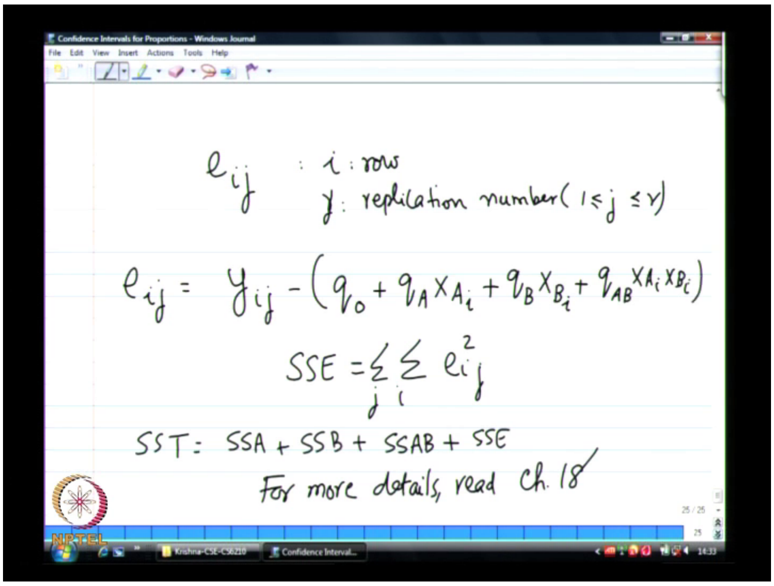
So, it is y columnright. So, what Iwill have is, Iwill have r experiments, results of r set experimentsright, Iwill have r values for every rowright. So, my y will simply be the odd rows of this; then everything is proceeds as before. So, for example, if this is 15, this is 18 and this is 21 made this r equals3rights. So, Isimply take the average of this 3,and that becomes my y, this is 18. So, you will repeat like I did last time, replacing y with corresponding averages, because all the replications. And then of course, later on used to conference intervals and soon. There is a lot mode here only scratching the surface. But now the will depend upon the…

Actual sample taken were actually…

(( )) r replication the standard deviation of those r replication may also (( ))

Which were whether we are using this, theactual configuration of the SST is there.(( ))I will come.

(Refer Slide Time: 19:48)



So, what we will have is for right, you will have a set of eijright for the i ethcombination and thej eth trailright. So, that iswhat your e i j would be. So, there could be e ij(( )). So i is the row, and r and j is the replication number, where j goes from less k. So, we are now y ij becomes.

(No audio from 20:34 to 20:53)

So, each case eij, it is actually defined asyij, which is the actual measured valueminus what do you have been inan…

(No audio from 21:05 to 21:46)

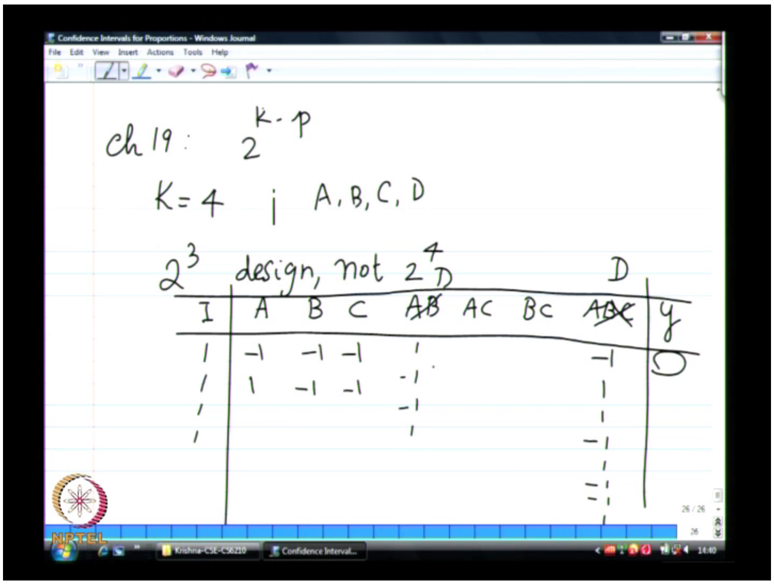
So this actually, will correspond to y i y i bar the average value of y i. So this will be y i. So, this is the error term for every replication, this will be the error term. So, Ihave just,and then todayif look added the, because this is the averageright. So, what will happen is the summation of all errors in a particular will go to 0. So, summation of all errors A j sigma i j, all i j will go to 0then what we are interested in is this, so-called SSE, which is simply Xright, because all rows in a closed(( )). So, if you give the table, we will complete SSE; and then your SST now again, you are not getting in to the details is got this extract component. SSA plus SSB, sothe total variation is actually across all the samplesright;SST is now, existing all the replications that you are taken.

So, the purpose of taking this next to the step isthat to find out whetherSSE is small or not, if SSE negligibly small, soyou do the same analysis like beforeright.If SSE is very large let us say SSE is 70 percent of total variation is because of errors it means that between replications, you have several like there is lot of variations replication itself for a given complication.Therefore, you can say SSE is very large that means, there is the most of the variation is because of experimental errors, it cannot really conclude anything in terms of AB or A;Ireally would want to see A and B having good values or may be AB having larger value, but SSE should go close to something close to 1 percent, 2 percent is accepted. But if SSE is very large, then we will have to simply come and report that whatever Iam seeing here in terms of variations is simply, because of the experimental errors.

And there sometime that is stickers, you know some times you run experiments even for CPU measurement, the same job, same computation and the same system at different points in time; even if there is no other background processing say that background jobs running in the system; sometime you will find out that the actual measurements sent up the quite variation here with thenot correct, some other system CPU process taking rest at the time,which you tryingto do; whatever it is, all of it something like sensors and measuring say you know, values for a sensor, you find that there is somuch variations.

So, the purpose of this is to do if Ido replications, thenIwant to ensure that enough replications are there such that SSE goes to a smaller variation. If SSE is very large, then your experiments are basically inclusive,because you cannot really say that that this particular factor A has a name better impact that some other factor. So, that is just a highlight of the lot more in the chapter, and thenwe just look at oneexample of2 k minus p and after that it will be (( )). So, formore details…

(Refer Slide Time: 25:08)



So, in chapter 19 is (( )) how do I get 2to the power k minus pright, that fractional factorial design, which mean is that I do not wanted to q to the power 2that is tomany, and I wanted to do something less then. So, that I again I just first section. I just briefly mention, what is going on hereright. So, let us say k equals 4right, there are 4, 4 factors – a, b, c, and d. But only one wanted to I wanted 2 to the power 3designright.

(No audio from 25:42 to 25:53)

So, this finally, (( )) 16. So, what is the good combination. So, there are different ways is used; one of them is here first of all find out the which 3or k usual, you decide that 3of those factors have to be a repeated in all the combination, but we have fourth one only in some variations. They do not actually have p equalsright, onehave consistency, but you can have some. So, the way the let us see A, B, C at the selected factors, in d is the (( )). So, we will go were inright A, B, C, AC. So, we will have ourright columns has before, this is my 2 to the power 3designright. So, I have everything as before.

(No audio from 26:40 to 27:04)

So, what you note down. So, I have A, B, C. So, if you want to bring in D. So, again if you look added this is minus 1, minus 1, minus 1right;this is minus 1, and this we said is 1, minus 1 and so onright, say this is 1. And the A, B, C column will looks something like this. I will that 1, minus 1, 1 minus 1, minus 1, 1. So, you can pick, so what happens you have A, B, C;

let us see, you are (( )) you are keen onright. So, you want to know the individual factors of these factors. Would you have 4 other column we choose from…

So, this one, I am tell you interactional A and B are A, B and C, and so on. So, you simply choose any one of them. So, that would be a we can replace that, and then of course, we go on in to.

(( ))

I do not have an answer further right now. I have do not tried this sum. So, when I conduct explants, I want this is the combinations haveright. So, you have been conducting students with A, B, C and D with this 4 different levelsright. So, that is taken, has to why this verses something else. And so, all I want to do is create a set of experiment set with different with variations each of those levels right. So, when the case of A, B and C; I have does in she have done all the combinations that I want, and for those 2 set of combination they just want to. So, what we find here in this case is there; 4 (( )), 4 of those combinations I am testing with a withD at minus 1.

I am not testing with D at 1 for 4 other combinationsright; whereas in the case where the full factorial added the sameadded the same thing what happen repeatedright, I would have (( )) added. So, I am testing for hope the combinations of what I would have normally tested with this minus 1, and the other half of testing with one. And then one about then you run the experiments, what are the values have y that you get with this combinations will have D'sfactors also take in will have right.

(( )) It will be same as minus 1 1.

This is will be there, this is willbe there. See all those saying is that for you see you have to you have to trained create combinations of A, B, Cright. So, you have minus one. So, they should the for the first experiment you will all those 4 factors will be have levels minus 1; whatever that level for (( ))right. Everything it is going to be a same (( )), and your y here will depend up on the D's valueright; if D is 1, you have different result; if D is minus 1, you have different result. That is why y is D is factorial into the value that y that you see in (( )).

(No audio from 30:14 to 30:30)

So, if you replace some other column with D will it make a difference, it will make a difference to the type of experiments you have runright, if it is this one for example, this is 1, 1, 1, minus 1, and so onright. So, there in this way, this have to be D instead of thatthen, less than levels, which you have experimenting in each case is going to be different (( )).