Total Quality Management - II
Prof. Raghunandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

Lecture - 31
Analysis Topics for Factorial and Fractional Factorial Designs – II

Welcome back my dear students, a very good morning, good afternoon, good evening to all of you and, this is the TQM II course on under the NPTEL MOOC series and we are in the 31st class; that means, we have finished 6 weeks of class this is the 7th weeks first class and we will wrap up this course in 40 lectures. And I am Raghunanda Sengupta from IME department IIT Kanpur.

So, if you remember we were discussing about the concepts of the fractional factorials 0.1 point number 2 we did mention time and again I have been mentioning more in the initial classes, that consider the level of importance let me use this word level of importance of any factor is at 2 levels consider is good or bad plus 1 minus 1, 1 and 2, 3 and 4 whatever it is, but here it is given as plus and minus which means plus mean positive effect of a higher level minus mean negative effect, or being of a lower level. .

So, consider there are k number of factors. So, the total combinations based on which you will try to model it would be 2 to the power k degrees of freedom would be calculated accordingly and the design would be of 3rd design 4th design 5th design depending on the accuracy you want so; obviously, it would mean that what is the combinations which you are doing.

Now, and if you remember correctly that for any 2 factor model for the fractional factorial this so, called design of experiments, we considered the factor 1 or a along the horizontal axis, which is x which is my right arm, I am just repeating this and the overall reading of the y axis would basically correspond to the second factor which is b, which is the vertical which is my left arm and all the points which you have inside this so, called first coordinate if you consider the Cartesian coordinate these would give you the information based on which, you can find what is the relationship.

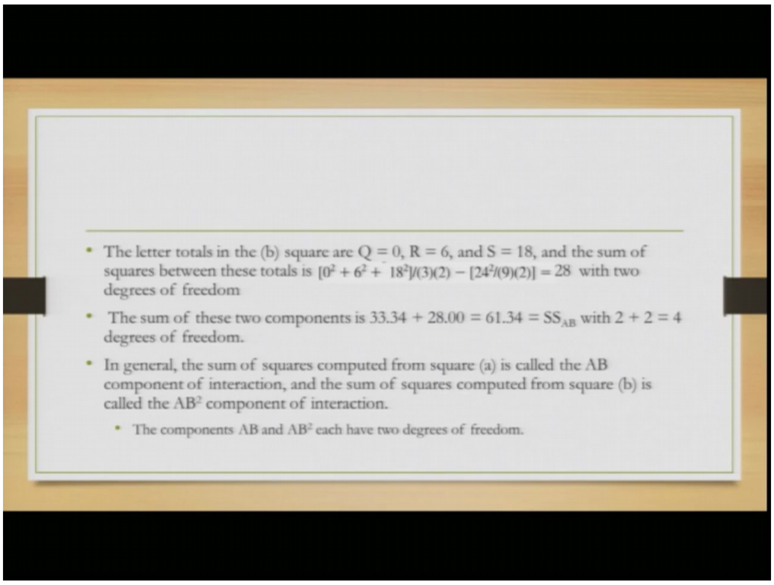
So, technically it would be a square with 4 grid points grid points being of level of 0 0, where the 0 0 if I am considering their lower level for both of them is 0 and 1 1 being for both of them at the highest level. And if it is minus 1 plus 1 the grid points would be minus 1 minus 1 for the lowest level, other 3 points would be correspondingly lead to minus 1 plus 1 plus 1 plus 1 and plus 1 minus 1.

Now, if we expand it as we did discuss in the last class to the factors being 2, but if they are the level of importance basically 3 level of importance means basically high level moderate level low level such that we give them are. So, called points in the quantitative values as plus 1 0 minus 1 we can also give as 1 2 3, but I did mention that why 1 2 3 would not be the right approach in order to basically analyze the problem. And I did give 2 very simple examples from the area of civil engineering and electrical engineering and it can be expanded in any sphere. So, it can be in accounting, it can be in marketing, it can be in psychology, it can be in finance whatever areas they are.

Now, when you consider the 2 factors with 3 level of importance you will basically have again a square of grid, now points would basically be levels each would be a basically at level of minus 1 0 1. So, A would be at minus 1 0 1 B would be at minus 1 0 1. So, you will basically find on the combinations accordingly. Now, if you go to the higher level higher dimension 3 factors. So, consider A along the x axis which is my right arm, B along the y axis which my left arm and the other value C consider would be towards you or would be coming away from you.

So, we will basically have a so called cube or a rubikscube sort of thing cube in the sense if there are only 2 levels, rubric cube would be the sense for if we have more than 2 levels of f x for each like plus 1 minus 1 0 and so on and so, forth. So, with this will consider the discussion and, we did discuss also the fold effect models, where we consider the folding or accumulation on the collation of the effect for the factors such that they give us better predictions, your main your answer is to basically able to utilize the list set of information to get the maximum set of output.

(Refer Slide Time: 04:59)



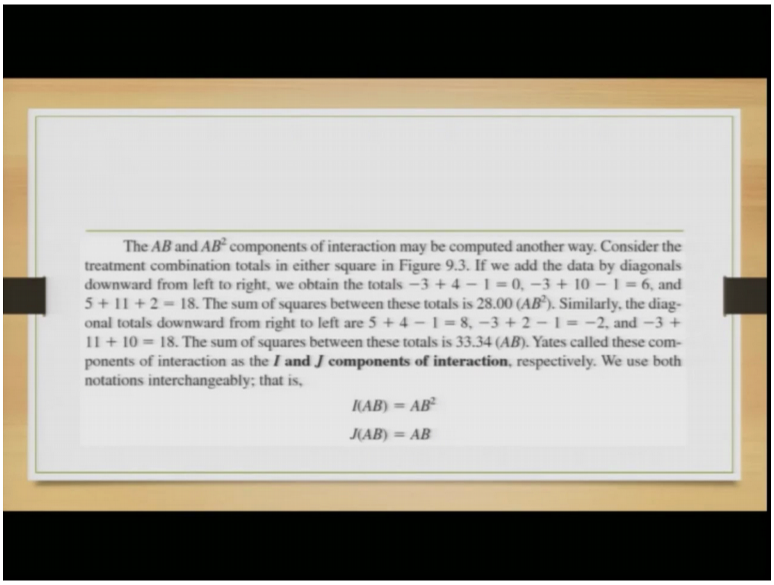
So, the let us obtain in the square. So, we are considering the square of the grids, where the values of Q is equal to 0, R is equal to 6 and S is 18 and the sum of the squares between the totals would; obviously, be you find out the sums as 0 square plus 6 square plus 18 square and divided by the corresponding factors you will get 28 with 2 degrees of freedom, which you have.

The sum of the components would become. So, now, again if you remember that whatever the overall variances, or overall I will not use the word discrepancy, but whatever the difference in the actual reading and the predicted reading was so, they would be two things. The total sum of the squares of the errors or the sum of the squares of the readings the difference in the reading would be given by some component which you can predict and, some component which you cannot predict.

So, it basically some of those squares would be divided into 2 components. So, this is what we will again consider. So, that is the basic notion which we are trying to find out, if you remember I just mentioned that you are trying to utilize the least number of variables or these number of so, called resources to get the maximum prediction value or the maximum output, in whichever words which you use.

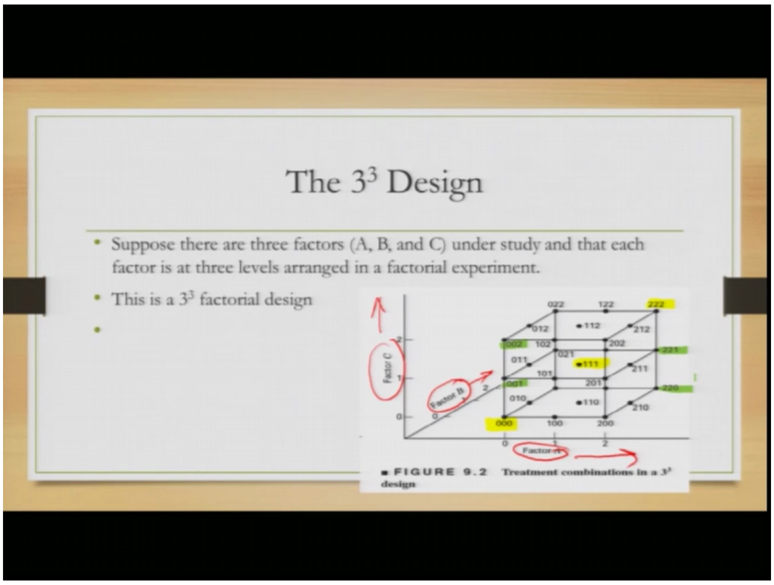
The sum of the 2 components comes out to be 61.34 which is the sum of the squares or A B with degrees of freedom being 2 and 2 depending on because; you have the grids based on the fact that the number of variables are accordingly. So, you basically have the degrees of freedom as 4. In general the sum of the squares computed from the square is called the A B component of the interaction and the sum of the squares component from the square B is called the basically the A B square is basically notion of how you mention it. It is A B square component of interaction the components of A B and A B square each have 2 degrees of freedom and we can basically find it accordingly.

(Refer Slide Time: 07:01)



The A B and A B square component of interaction may be computed another way, consider the treatment combination totals in each square is in figure 9.3. If you add the data by a diagnosis not by the diagonals downwards from left to right we obtain the totals given as 6. So, you are trying to basically find on the total and 18, the sum of the squares between the totals is given as 28.00. Similarly the diagonals towards the downright from left to right are given as minus 2 and 18 and you can find it accordingly, the sum of squares between the totals between which is A B is 33.34 and yet called these components of the interaction as the I and J components of interaction based on which you can use both node notation interchangeably and find out the values as I as A B square and J as A B only.

(Refer Slide Time: 07:57)

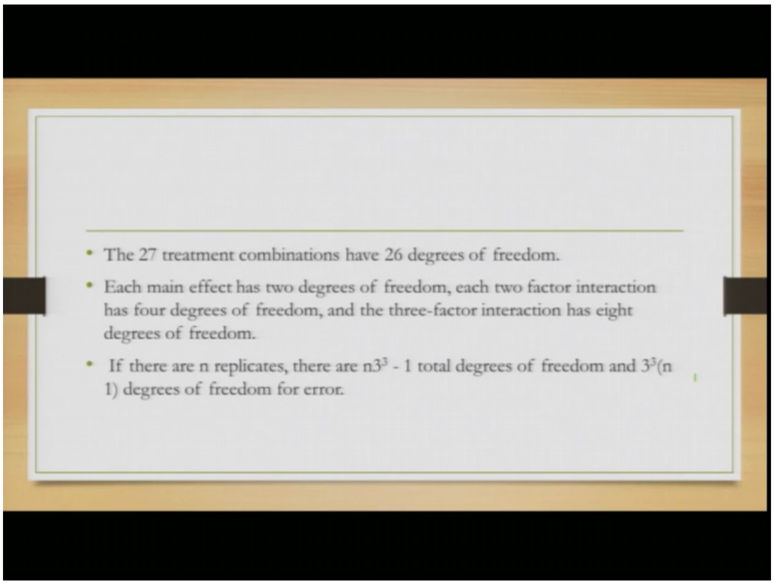


Now, consider the 3 3 3 to the power 3 design which is basically your 3 factors and there are 3 levels of designs. Suppose there are 3 factors A B C under study and that each factor is at 3 levels arranged in your factorial experiment, this is a 3 to the power 3 factorial design, if you see the curve it basically gives you that information. So, basically you have 0 0 0 which means they are at the level of lowest level for each A B C. So, a factor is along and I am basically marking it, let me use a different colour. So, A is along x B is along so, called y and z is as for the C. So, now, use a different colour aural highlight let me use the aural highlight.

So, this is for 0 0 0 which is the all of A B C are at the lowest level, this is 2 2 2 which means that all A B C are they are the highest level. So, 0 0 0 is for the lowest 2 2 2 is for the highest and; obviously, there are intermediate values. So, what is the intermediate values where all of them are the intermediate values would be 1 1 1 and, correspondingly the combinations are you can have 0 0, when I am just hovering I will use another colour to make it a little bit clear let me use a green colour. So, this would be the combination A B at low level and C at a moderate level. This would mean A B at the low level C at the highest level. Similarly if you have 2 2 1 would basically mean A B at the highest level C at the moderate level, again if we have so, I am going in the downward direction now, A B at the highest level c at the lowest level.

So, this need not be 0 1 2 it can be minus 1 0 1 depending on whichever convention you want to use. So, based on this you have the treatment combination of 3 3 design.

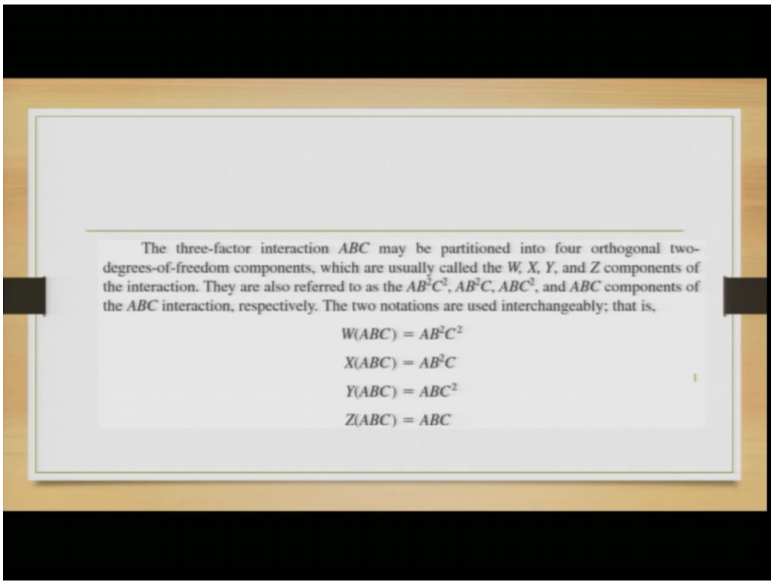
(Refer Slide Time: 10:00)



The 25 treatment combinations has 26 degrees of freedom because, if you count the numbers in this so, called Rubic cube. So, basically in the lower plane you would basically have the combinations would basically start from 0 0 0 and it will go to the extreme end as 2 2 2. So, at the lowest level you will have combinations as 0 0 0 0 0 1 0 0 2, then you have 0 1 0 0 2 0, then you will basically have 1 0 0 2 0 0 and basically proceed accordingly.

So, these are 27 treatment combinations and each has 26 degrees of freedom, each main effect has two degrees of freedom each two factor interaction has four degrees of freedom and these three factor interactions has eight degrees of freedom corresponding to the fact that we have already studied. So, two degrees of freedom for each two there are an n replicates; that means, you have basically trying to replicate it for n number of times the total number of degrees of freedom would n 3 to the power 3, I mean three factors with the 3 levels of interaction, multiplied by n number of replicates minus 1 because, you will always do 1 degrees of freedom. The total degree of freedom for the other case would be the form the replications would be given as 3 3 and bracket n minus 1 depending on the combinations you are going to take.

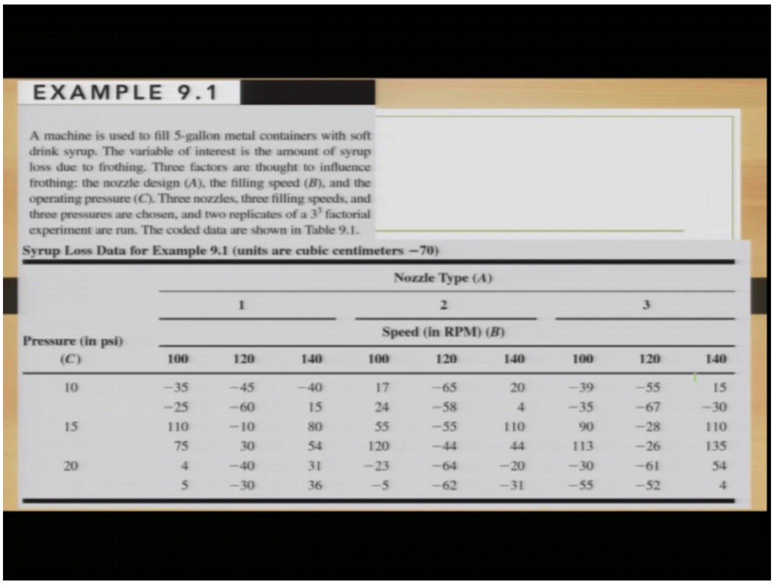
(Refer Slide Time: 11:30)



The three factor interaction A B C may be partially partitioned into four orthogonal two degrees of free components which are usually called W X Y and Z components of interaction, they are referred to as corresponding to how you basically denote as A B square C square A B square C, then A B C and correspondingly you can also consider A B C square A square B C and so on and so, depending on the combinations which you are going to take. .

The two notional notions are used interchangeably interchangeably so, if you use W X Y Z or U V W X Y Z the combinations would be denoted, as technically as consider for w you are denoting as A B square C square. In another case if it is X it could be A B square C, in another case Y it will be A B C square say for example, in another case it is V, V it may be denoted by A square B C. So, depending on different type of conventions you will not the what the conventions depending on the nomenclature will use these combinations accordingly.

(Refer Slide Time: 12:32)


EXAMPLE 9.1

A machine is used to fill 5-gallon metal containers with soft drink syrup. The variable of interest is the amount of syrup loss due to frothing. Three factors are thought to influence frothing: the nozzle design (*A*), the filling speed (*B*), and the operating pressure (*C*). Three nozzles, three filling speeds, and three pressures are chosen, and two replicates of a 3^3 factorial experiment are run. The coded data are shown in Table 9.1.

Syrup Loss Data for Example 9.1 (units are cubic centimeters −70)

	Nozzle Type (*A*)								
	1			2			3		
	Speed (in RPM) (*B*)								
Pressure (in psi) (*C*)	100	120	140	100	120	140	100	120	140
10	−35	−45	−40	17	−65	20	−39	−55	15
	−25	−60	15	24	−58	4	−35	−67	−30
15	110	−10	80	55	−55	110	90	−28	110
	75	30	54	120	−44	44	113	−26	135
20	4	−40	31	−23	−64	−20	−30	−61	54
	5	−30	36	−5	−62	−31	−55	−52	4

A machine is used to fill 5 gallon metal containers with soft drinks syrup. So, you are trying to basically fill up metal containers, the variable of interest is the amount of syrup loss due to frothing so; obviously, when you putting some Fizzy drink like Limca, or Coke, or Cocoa coma or thumbs up, or Pepsi whatever it is, or some aerated drinks. They would be frothing and you want to minimize the frothing end and minimize the loss, their three factors are thought to be influence are the nozzle the design.

So, what is the diameter what is the so, called aperture radius, what is the filling speed high this filling speed, high is the floating know the filling speed low the floating, but; obviously, that will reduce your productivity in the same case if the nozzle diameter is high then; obviously, you can fill faster, but if the nozzle speed aperture is or the radius is slow, then you have to basically fill slower.

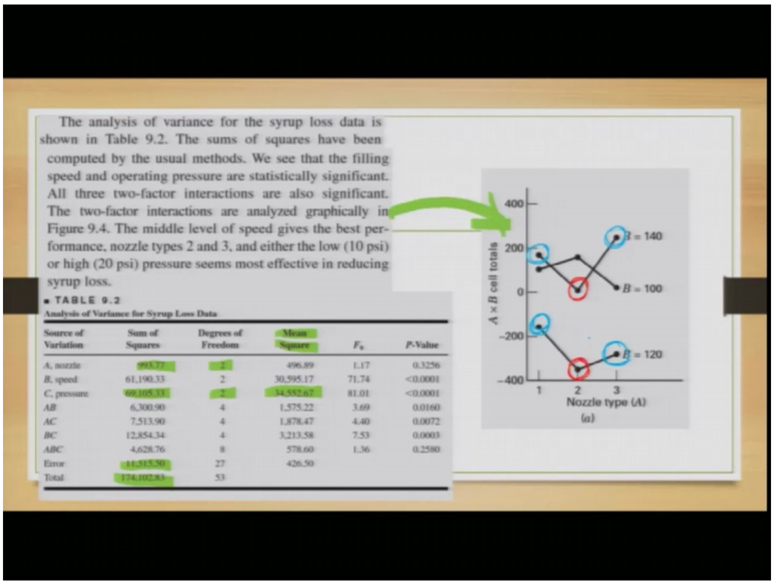
So, the speed would also change and the three being the operating pressure higher the pressure more the frothing, lower the pressure less the frothing so, but; obviously, higher the pressure would mean, that you are able to have a total of quantum of carburetted air, or the carbon dioxide at a much higher level not for lower pressure, but; obviously, it will have a other effect; that means, higher the pressure, more the frothing, more than loss lowered the pressure, less the frothing, less the loss.

So, I am just giving you a quantity qualitative filling 3 nozzles 3 3 filling speed and 3 pressures are chosen. So, now, you see that till now we had considered 2 levels of pressure plus 1 plus minus 1 0 plus 1 minus 1 so, on and so forth. And, now we are choosing 3 levels of pressure. So, they can be denoted as plus 1 0 1 1 2 3 whatever the combination, similarly we can choose the filling speed also at 3 levels and, you can consider the 2 replicates which are there. So, if you consider the three factorial models the pressures in square inch pounds per square inch.

So, is which is the leftmost column. So, these values are 10 15 5 and the combinations which you see, you have basically the nozzle type being a and then you are considering the speed. So, pressure and speed are the variables and you considering the nozzle type A. So, you have basically nozzle type B of 100 121 40. So, consider them as the pressures which are given, speed you can basically find out depending on whatever the combinations which you have. So, if you look at this table. So, it gives you the syrup loss data for example, 9.1 and the units are cubic centimetres and they are there.

So, the values which are given for different combinations are in this table. So, if you consider for nozzle type 1 and the leftmost column 100 so; obviously, you will have the loss given as minus 35 minus 25 110 25 4 and 5 corresponding to the other values you will have the whole gamut in on the data in front of you.

(Refer Slide Time: 15:47)



Then this is the variance for the syrup loss data is shown in table 9.2, the sum of the squares have been computed by a usual methods we see that the filling speed and operating pressures are statistically significant, all these two factors interaction and also significant, the two factor interactions are analyzed graphically in figure 9.4.

So, which is the right hand figure which we have will consider that in a little bit more detail, the middle level of speed gives the best performances as we see because, in that case the frothing way the loss or the loss due to the frothing because, they are spilling what is the minimum is the least. The middle depth of speed gives the best performance of nozzle 2 and 3 and either low, low that is 10 psi or high 20 psi pressure seems more effective in reducing the syrup loss.

So, if you consider the source of the variations. So, source of variations are given in the leftmost column which is a which is A nodal nozzle B the speed C the pressure and, the different combinations are there which you have A B which is speed and speed and the nozzle dimension, then A C would be the nozzle dimension and the pressure, speed and pressure means that what is pressure you are filling that that aerated drink and, the pressure is at what pressure it is coming.

So, it is just a repetition I am saying. If your combination of B C it means basically nozzle is kept fixed dimension is kept fixed and speed and pressure are being varied. If A B C basically are combinations are then; obviously, you will consider all 3 of them and at different combinations and; obviously, there would be a error. The sum of the squares of the error of some of the squares are given, the error value comes out to be 11515 which is this value and the total sum of squares comes out to be 174102 and the degrees of freedom are given accordingly is because, if it is 1 only one factor it is 2, two factors is 2 it will be 2 into 2 it is 4 and when you have basically with all the 3 4 combined factors it will be 3 into 3 into the three factors considered and it will basically have 8 because, it will be 2 into 2 into 2 not the 3 part sorry. .

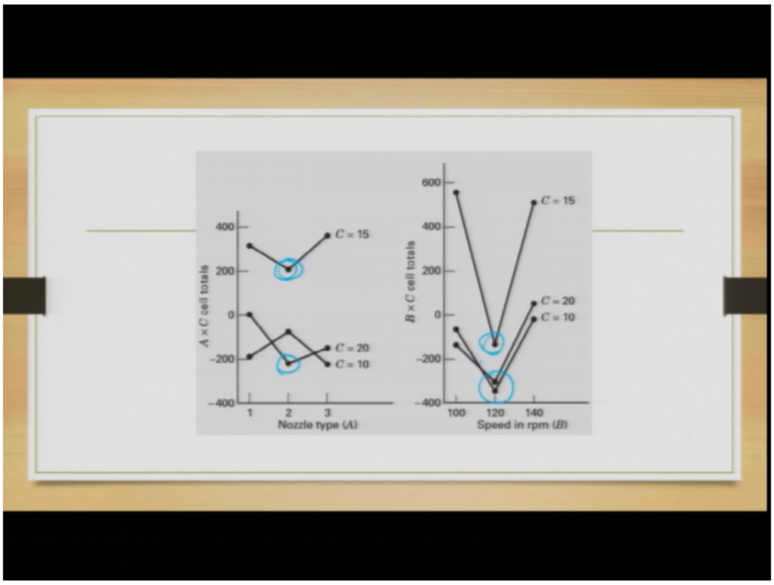
The mean squares would; obviously, calculating you know that you basically do find out the mean squares, which I am highlighting by dividing here this 993 divided by 2 so; obviously, it will come out to 496.89 I am not repeating the decimals. If I consider say for example, the value of 69105 divided by 2 would give you the mean square as this. Similarly once you find out you can find out the F value depending on the left hand side on the right hand side on the degrees of freedom which you have.

So, the P values and the F values are found out and you can basically comment that the significant and non significant. So, F values I will just repeat it which is the second last column starts from 1.17 for nozzle A, only the nozzle effect 71.74 for B which is the speed 81.01 for C which is the pressure and corresponding to A B A C and B C which is corresponding to nozzle and speed combination, then corresponding to nozzle and pressure combination and finally, speed to pressure combinations the F values are 3.69 4.04 0 4 0 and 7.53. So, based on that you can find out the P value.

So, if you look at the graph so, this combinations are the least I am just highlighting and, in this combination it is a high level I am just highlighting few of them, see if you read it I will again repeat repeat read the what is given the analysis of variance of ANOVA for syrup loss data is shown in table which is 9.2 which you have just discussed. And the sum of the squares have been computed by the usual methods and we fit it we see that the filling the speed and operating pressures are statistically significant at all three two factor interactions are also significant.

So, they are the two factors at and interactions being two factors for this affects. The two factor interactions and as graphically in the figure, which I have just marked here in this arrow. And if you see then the middle level the speed gives the best interaction. So, when you (Refer Time: 20:20) table the speed when I am hovering and the red highlighted part, or a circle pattern and the minimum level because, they give them best so, called prediction.

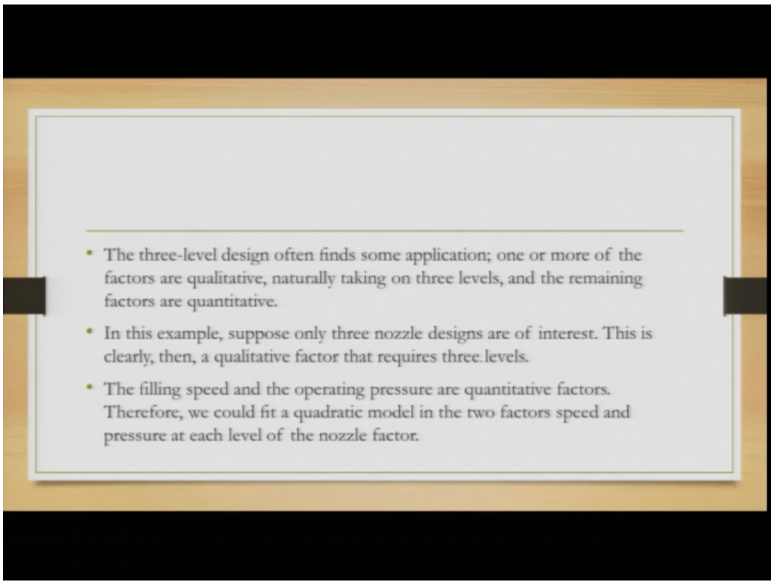
(Refer Slide Time: 20:31)



So, now if I do it for from the nozzle type with the combinations of nozzle type was basically A if you remember. So, if you if I predict the A to see cell totals that combinations are not all speed and pressure being there, then you have again the combinations given for again you find out the second level are the best. So, here again you find out the second level on the best, but these are combinations speed and the combination of B to C combinations or take totals.

So, these would give you the best way how you are trying to find out what is the best effect. The three level design often find some applications one or more of the factors are can be qualitative.

(Refer Slide Time: 21:18)

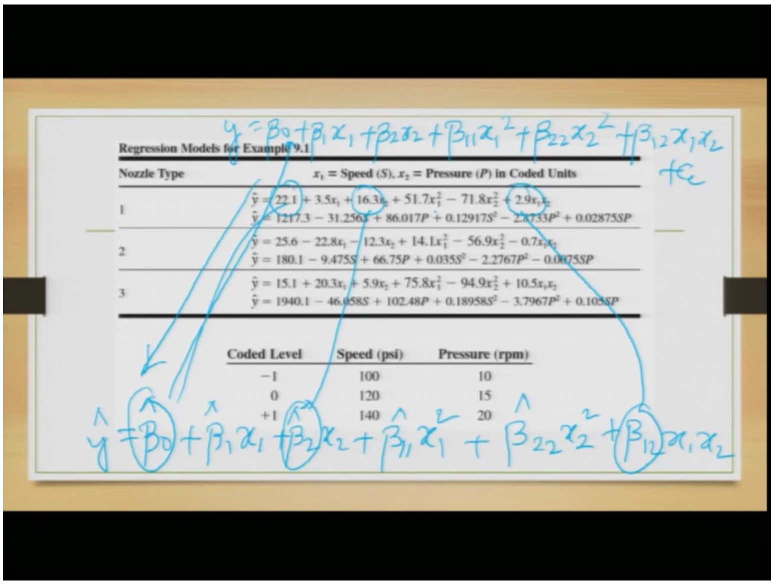


And again; obviously, you may ask the question what happens if all of them are quantitative, but they are at different levels like one is at three levels of importance like minus 1 0 1 plus 1, another can be see for example, at 5 levels of importance may be depending on the things. So, it can minus 2 minus 1 0 1 2 so, how you do that; obviously, the concept with the same, but we will try to basically explicitly, go into the details of simple case and there is a complicated case can be analyzed accordingly.

The three levels so, let me read it the three level design often find some application one or more of the factors are qualitative naturally taking on three levels and the remaining factors are quantitative, in this example suppose only three nozzle designs are of interest this is clearly then a qualitative factor that requires three levels of importance. The filling speed and the operating pressures are quantitative factors therefore, we can fit a quadratic model in the to factor speed and pressure at each level of nozzle factors.

So; obviously, here pressure nozzle dimension and speed are quantitative sorry able to handle it in one wave, with their quantity qualitative we will do it accordingly, but; obviously, if we remember there were some discussions where qualitative factors work one converted into quantitative factors and the work was done according me.

(Refer Slide Time: 22:35)



So, now, we are going to fit a regression model. So, the regression models would basically be with respect to like say for example, you want to find on the nozzle type. And where x 1 in the speed x 2 is the pressure so, we are basically regressing one of the factors on the other depending on what we think because, speed and pressure would basically be more on a control.

So, based on that you want to find out what is the nor nozzle dimension because, nozzle dimension once fix cannot be changed abruptly. So, speed and pressure are the factors which you can change depending on how the experiment, on how the loss of the frothing is going on because, dimension on the nozzle cannot be change instantaneously, once the machine has been designed has been purchases it is fixed.

So, try to understand that. So, the nozzle time are given if you remember of three types, three dimensions, three nozzle diameters and the speeds as I as I am just repeating it x 1 would be the speed x 2 with pressures and in their coded units. So, in all these three cases you will basically have different regression models, but the fact the point is that rather than having a simple multivariate where the multiple linear regression we consider a combination of that.

So, what would the multiple linear regression would be you will basically have the effects coming from x 1 to one degree x 2 into another one degree and the rest would basically be the error part; obviously, you can consider beta naught term depending on the level at which it cuts the y axis. So, if you remember I did mention y is equal to m x plus C equations. So, this value is C, I am basically denoting as beta naught. Now if you want to basically extend that, but that influence. So, you will basically consider the combinations to the second degree.

So, the second degrees would basically be where you were considering x 1 squared x 2 squared and also the combinations of x 1 and x 2 can be considered as a multiplicative factor so; obviously, the model would basically be beta naught I am just repeating it what is the answer is exactly the same once you analyze the problem. So, y would be equal to y, I am again repeating is the nozzle and why the nozzle I did mention that once it is designed it cannot be change.

So, your speed and pressure are the independent variable. So, y which is the normal nozzle dimension is equal to beta naught plus, if it is a regression model where that and all the degree is one it will be beta 1 into x 1 plus beta 2 into x 2 plus epsilon so; obviously, all the assumptions would hold that normality, then the mean value of there are variants of the error and so, on and so forth. And the covariance between the independent variable and the error terms being independent and being 0 and errors being independent from time to time so this factors. If you extend it to the quadratic model it will be y is equal to beta naught plus beta 1 into x 1 plus beta 2 into x 2 plus beta 2 2 into x 2 square plus beta 1 1 into x 1 square plus beta 1 2 into x 1 and x 2 plus epsilon. So, based on that we solve the model so, this is what it would be this is the predicted value so; obviously, the error terms who do not come.

So, this would basically mean the beta naught, I am only highlighting one of them this 3.5 would basically be beta 1 with a factor of x 1 this 16.3 would basically be beta 2 with the multiplied by x 2, this would basically be beta 1 1 multiplied by x 1 square. This would basically so now, there is a negative term. So, this would be basically beta 2 2 with a minus sign with x 2 whole square and, this 1 basically would be beta 1 2 multiplied by x 1 2.

So, I will just write them so technically your equation would be in the simple case y is equal to beta naught, this is the multiple linear equation remember that so; obviously, when you fit when find out the regression. So, it y hat is equal 2 beta naught hat per beta 1 hat so obviously, the error term will not come. Now when I consider the higher model it will be y is equal to beta naught plus beta 1 x 1 plus beta 2 x 2 plus beta 1 1 x 1 square plus beta 2 2 x 2 square plus beta 1 2 x 1 x 2 plus epsilon when you fit it.

So, I am just basically coming here, y naught is beta naught hat which is here plus beta 1 hat x 1 plus beta 2 hat into x 2 plus beta 1 1 hat, this hat means they are the predicted values x 1 1 plus beta 2 2 hat into x 2. So, these values which I have is this consider I am just giving you are the hypothetical values depending on this will be 16.33 simulate the last one 1 2 would basically 2 point. So, we can we will consider this model in more details later on. So, with this I will end the 31st lecture.

Thank you very much for your attention. Have a nice day.