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

Lecture – 21
Factorial Designs – VII

Welcome back my dear friends, a very good morning, good afternoon, good evening to all of you. And this is the NPTEL MOOC course which is the TQM II course. And as you know this is for a total duration of eight weeks; and we are going to start the 21 lecture that means we are almost halfway through. And we are starting the fifth week, the fifth week first lecture.

So, if you remember we were discussing about the concepts of different factors, and if you and again I will repeat this examples time and again we did one example where power and time and the effect of different type of etching problem was there. Then in the case of battery, the temperature and the material problem was there then was the aluminum smelting and then we gave the formulas that given more than two factors you will basically a factor A, B, C.

Then the combination of two at a time AB, BC, AC then taking three at a time ABC, you can have basically the total sum of squares. The individual factors sum of squares then the error terms and obviously the factors would be in the combinations as I said ABC separately AB, BC, AC taken to take two together and then come combination of three.

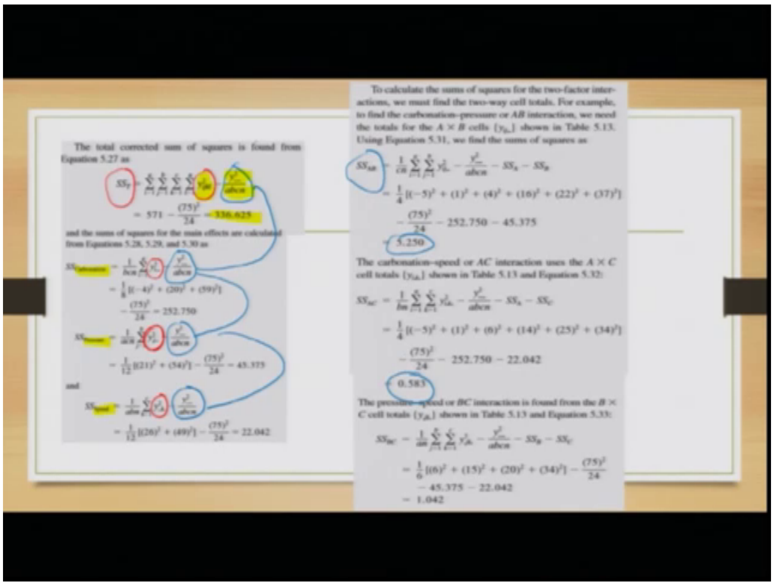
And if you go to four factors obviously, they would be individual ABCD, then combination of two, combination of three and then four combined together. And then we also understood that means, the number of so called observation of factors I would not use the word observation factors for a was small a, for b was small b, for c was small d c, and for d was small d, and the total number of observations was small n then the total number of set of observations for the whole experiment will be a into b into c into d into n.

And combined combining that you can find out the different degrees of freedom for each factors individually combined two together three together and so on and so forth plus the degree of freedom for the error terms.

Then based on the total sum and the degree of freedom we found on the mean squared errors using the mean square errors for the total, for the factors, for the errors, we found out the f statistic and then we combined that said that whether they were significant or non significant depending on the level of significance.

Now, coming continuing with the last problem, so there were basically three factors A, B, C as I said and the total observations was small n. So, if you are trying to find out the sum of the square of the totals which would basically be if you continue the only concentrate on the formulas the values would be coming out automatically.

(Refer Slide Time: 03:20)



So, what you have is so we are some of the squares that means, you have to find out the square. So, here the first term is y square and for each cell. So, how many cells are there if you consider a matrix of higher dimension more than three. So, the first cell would basically the first cell means the factor corresponding to the first cell would be termed as i, the second would be the j for the b, k for the c, and l would be basically for n. So, I changing from 1 to small a; j is changing from 1 to small b; k is changing from 1 to small c; l is changing from 1 to small n.

And if I consider the dots, so technically the dots would be if it is the first dot it means that I am summing up for all the a small a's which is for factor capital A. If the second dot is there which means I am summing up for all the b small bs for factor B capital B. If dot is there for the third place, it means I am summing up for all the small c's which is four cam the factor capital C and so on and so forth. So, it can go for like to d to e to f and so on and so forth. And the last one; obviously, would be corresponding the fact that you are summing it for all set of observations.

So, if you have the sum of the squares, the value of the sum of the squares for the total corresponding to the case, so this was one element y square i j k l. And y square four dots corresponding to i corresponding to j corresponding to k corresponding to l divided by obviously total a set of observations would be as I said abc into n into n. So, based on that I find out sum of squares of total is 366.625. So, then if I go to the factors, so they would be SS suffix capital A, SS suffix capital B. So, if we name them, so here it is carbonation which is factor A, pressure which is basically factor B, and speed which is factor C.

From there, if I want to find out the sum of the square of the errors, so here it will be y square I triple dot because I am summing up for keeping i at fixed and one level I am summing up for all j k l. Similarly, when I go to the second one, it will be dot for the first place, j for the second place, and third and fourth would be dot as mentioned here.

So, if we are I am circling please make a note. If I go for the third factor which is c, which is the pressure it will be dot, dot, k and again a dot, so dot, dot, k again a dot. And if I consider the term which is to be subtracted it will be y square four dots divided by abc n, which is whatever I will try to utilize another highlighting color, let me circle it. So, highlight may make it dark sorry for that.

So, this is what we had abc n, now it will be again y square triple four dots abc n. So, these are matching, this one and this one. So, once you have this, the values which you have sum score total squares is 336.625 for a factor A, it is I am I would not repeat the decimal part it is 253. For the b part B factor is 445; and c part, which is speed is 222. Now, I need to find out the carbonation of ab which will be carbonation and pressure, then bc which you do pressure at the speed, and ac would be carbonation into speed.

And finally, I need to find out all the combinations, combination which is a pressure which is b, and speed which is c. So, let us see. So, here it will be SS AB. So, I will read it. To calculate the sum of squares for the factor in inter interaction as I just mentioned. We must find the two ways l totals for example, to find out the carbonation of the pressure which is or the AB interaction as I mentioned. We need to find the total for all the cells of A into B. So, which is basically y ij dot dot, I am basically keeping i and j fixed at any particular level starting i is equal to 1 to smaller a, j is equal to 1 to small b, but summing up for all the c s, small c s and summing up for all the small n s.

So, as already discussed so using this equation we find out some of the squares of ab which is carbonation and pressure. The value comes out to be up to be about 5.250, I am mentioning the decimal part here. Then if I go for the pressure and speed which is a c then I have the value as point 0.583. And if I go to the o this a c is basically sorry the first and the third which is carbonational speed.

And then I go to the bc introduction which is pressure and speed the value comes out to be 1.042. The calculations the formulas remain the same only you have to put it and try to understand that. And always dividing remembering that when you make the tables, on the leftmost column would be the factors with the factors being mentioned as a slash whatever the factor actually in words it is.

Like in simple example in this example a means carbonation, b means pressure, c means pressure speed. So, you write it down. The second last value in the first column would be the error term. So, double check that the sum of all the in each its itself when you are doing the total sum of squares or the error sum of all the errors would def definitely will be the total errors. You can double check in both ways for bus by first finding out the total sum of square errors then finding out the individual and the combined one. And then utilizing that to find out the error terms or back calculating to find out the error terms as the error terms by matches some of those squares matches in both this case.

So, as we were discussing the sum of the squares would if you have the first column written with all the factors then the second column with the total sum of squares. So, total sum of squares, and what I mean is basically they would individually be for A, B factor, C factor, combinations of AB and C taken two at a time. In the last one would be for the combination of the three taken all at a time, if the considering their fifth three factors.

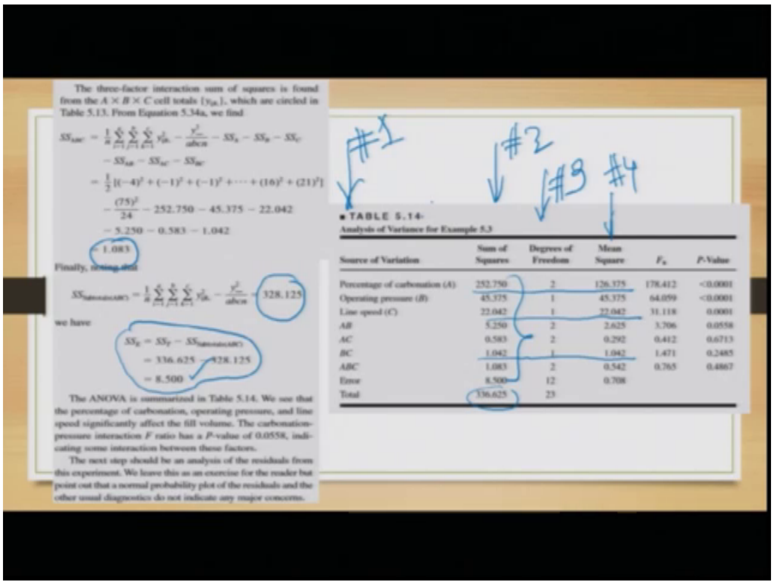
The second last again I am repeating would be the errors, and the last one was the total sum of square of the error. So, obviously, you can double check as I said finding out the sum of all the errors plus the error term would be where some of the squares would be equal to the total sum of squares errors. And then you have basically the third column would be corresponding to all the degrees of freedom, so that you have to basically make a note that how we will find out considering the total number of factors for A capital is small a, for capital B is small b, capital C is small c and so on and so forth; And the total number of observation being small n.

And then you divide. So, we complete for the first column, second column, third column. In the fourth column, you will basically find out the mean square errors. And then in the fifth column considering the ratios in the mean squares you will find out different on the f statistic and then utilize that to solve the problem. So, you basically have a chart and in the chart you plug in all the values to calculate the values.

Now, as I mentioned that you have the SS suffix AB as 5.250, SS suffix AC, I will be only using ABC of the factor name rather than going to detail. So, it will be much easier for us to proceed fast. And also you are in in tune with the word ABC which I am being using time and again and SS BC is 1.042. So, now I go to the next stage. What do you mean my next stage is basically I need to find out the combinations of abc.

See if it is combinations of abc, it will basically mean that rather than keeping the cell value for all the values of initially, say for example, it was a b, what we did is that we found out for all C s and all n s now it is abc. So, it will be for one specific value of a, one specific value of b, one specific value of c which means it will be suffix i that that the sum of the squares would be suffix i, j, k. And the fourth dot would imply that I am finding out for all the set of observations as it is given.

(Refer Slide Time: 12:16)



So, if I do so let me change, so the so if I do the value comes out to be 1.083, and the subtotal basically for A, B, C sky calculating, it can the values comes out to be 328 125 from that you can find out the sum square of the errors. So, this is the errors which I am finding out is 8.50. So, now, basically I had a table before reading, it I will go first to the table and then come back. So, if we note down the first column where I am I pointing out I am writing number one means column, it is all related to say for example, factor A, factor B, factor C, AB, AC, BC, ABC error total. Check the total sum would be exactly equal to the sum of all this.

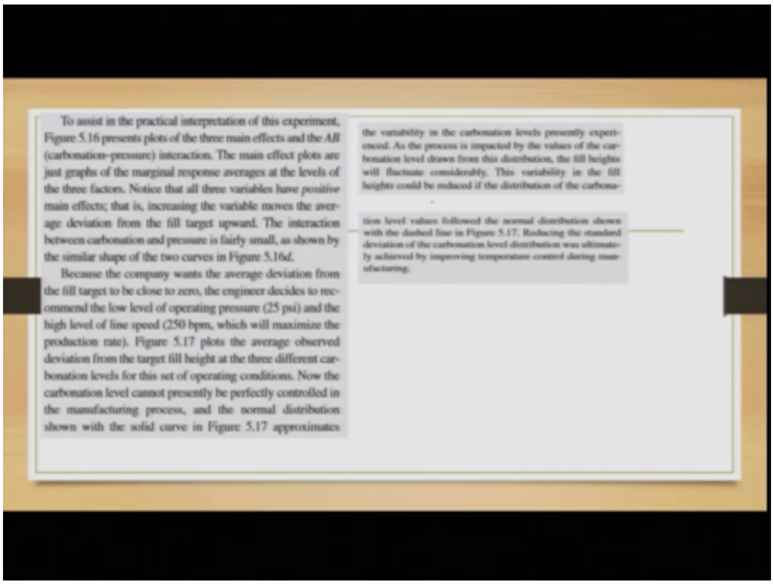
So, this is column number 2 which is the sum of the squares, then you have the degrees of freedom which is column number 3. So, here also you can check the degrees of freedom sum, then you have the mean square that means, you divide column number 2 by column number 3. So, say for example, you are taking this writing so 1.042 divided by 1 comes out to be 1.042.

So, similarly 22.042 divided by 1 becomes 22.042. Then if you find out say for example 252.750 divided by 2 comes out to be 126.375. So, once you have the mean square in column number 4, then if I have the ratios of the mean square giving you the f values, and then depending on the p value on the level of significance you can basically say whether you agree with the hypothesis disagree with the hypothesis.

So, it can be extended for A, B, C, D factors and so on and so forth. So, the ANOVA is summarized in table 5.14 which is here as I am mentioning. We see that the percentage of carbonation which is A operating pressure and the line speed basically which is factor B, and C significant affect the final volume, the carbonation pressure interaction if ratio has a P value.

P value depending on the m the level of significance which I have is basically 0.0558 in it indicating some interaction between these factors. The next step would be an analysis of the residuals from this experiment, we leave this as an exercise and one can solve it very, very, very simply to find out what are the results.

(Refer Slide Time: 14:56)



To assist in the practical interpretation of this experiment, Figure 5.16 presents plots of the three main effects and the *AB* (carbonation–pressure) interaction. The main effect plots are just graphs of the marginal response averages at the levels of the three factors. Notice that all three variables have *positive* main effects; that is, increasing the variable moves the average deviation from the fill target upward. The interaction between carbonation and pressure is fairly small, as shown by the similar shape of the two curves in Figure 5.16*d*.

Because the company wants the average deviation from the fill target to be close to zero, the engineer decides to recommend the low level of operating pressure (25 psi) and the high level of line speed (250 bpm, which will maximize the production rate). Figure 5.17 plots the average observed deviation from the target fill height at the three different carbonation levels for this set of operating conditions. Now the carbonation level cannot presently be perfectly controlled in the manufacturing process, and the normal distribution shown with the solid curve in Figure 5.17 approximates the variability in the carbonation levels presently experienced. As the process is impacted by the values of the carbonation level drawn from this distribution, the fill heights will fluctuate considerably. This variability in the fill heights could be reduced if the distribution of the carbonation level values followed the normal distribution shown with the dashed line in Figure 5.17. Reducing the standard deviation of the carbonation level distribution was ultimately achieved by improving temperature control during manufacturing.

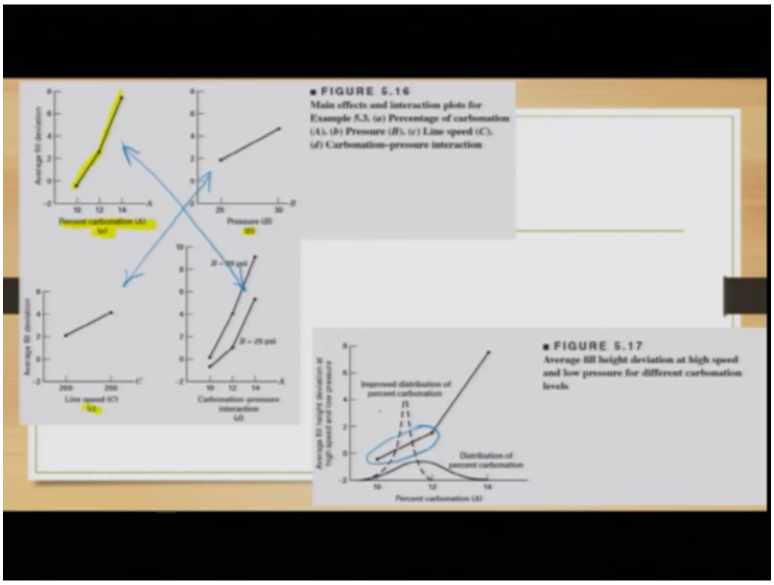
To assist in the practical interpretation of this experiment, we have figure 5.16, which present the plots of the three main effects and the AB which is basically the carbonation and the pressure interaction. The main effect plots are just graphs on the marginal response averages at the level of the three factors. So, at the three factors, I want to find out the marginal effects. Notice that all the three which I will come to that notice that all three variables have will have positive mean effects that is increasing the variable moves the average deviation from the fill target and upwards. So, it basically keeps increasing positive effect.

The interaction between carbonation and pressure is mainly small, so obviously the tan and the line would be almost horizontal basically giving you the impression and obviously, that would be true the tan of that angle of divide d dx or the rate of change of that function value with change of these factors would be almost 0, because that is why the effect is almost 0. Because the company wants the average deviation from the field target to be close to 0, so obviously, the average deviation should be asked as as as as dispersions or the variances would be 0, which means the average deviation from the fill target to be as close as 0, which I am just repeating.

The engineer decides to recommend the low level of operating speed of 25 and the high level of line speed, which is through 250 which will maximize the production rate depending on the answer we are getting. And figure 5.17 plots the average observed deviation from the target fill height at the three different carbonation level for this set of operating conditions. Now, the carbonation level cannot presently be perfectly controlled in the manufacturing process and the normal distribution shown with the solid graphs mean maybe a deviation, but will still follow that which approximates and such that we can find out the variability in the carbonation levels presents experimentally as the process is impacted.

But the value of the carbonation they will drawn from this distribution the fill heights will fluctuate considerably, this variability in the fill height could be reduced with the distribution the carbonation level was followed no normal distribution. So, obviously, it would not be normal, but we would still consider that to be true. So, reducing the standard deviation, the carbonation level distribution was ultimately achieved by improving the temperature control in during the manufacturing processes such that the best fit could be observed accordingly.

(Refer Slide Time: 17:30)



So, if I see the graphs now we are trying coming, so I will go slowly in trying to explaining that. Consider, you have the average field deviation. So, I am trying to find other deviations with the dispersions with respect to factor A, which is in figure as I am pointing out, let me highlight it. This is for a, and factor is a. The graph is named a not the number of observations for pressure B, the graph is B and you have again along the y-axis you have the average deviation. And for diagram c, it is basically line speed which is factor C and here again the average field deviation.

And when I am considering the combinations of carbonation pressure, carbonation pressure is basically be the factor B. But at two different pressure levels, so at two different pressure levels is 25 psi and pound per square inch and another one being 30 psi, the values you will basically get would give you an answer. So, in this case, in the first diagram, which I am highlighting the rate of change is quite high it is almost linear even though what percentage combination changing from 10 to 12, 12 to 14, they are not, because it is more steeper from 12 to 15 then from 10 to 12, but they are almost linear.

Similarly, if I consider the pressure 1, it increases from 25 to 30 while the average field deviation changes from 2 to almost about 4.75 or 5. And if I consider the lines be changing from 200 to 250, the average field deviation change from 2 to all most about the same value as we found out in the pressure. So, in these two diagrams, the change the rate may be different, but the change in the value of the y-axis which is average field deviations is almost the same. Now, when I find out the carbonation pressure the interaction for two different pressures, I find out that in the case when you have 30, so exactly what we have drawn. And the value of 25 psi what the equations would basically mean that the initial rate of increase from 10 to 12 for different psi is slower in read then when you increase it from 12 to 4 as I found out in both these cases.

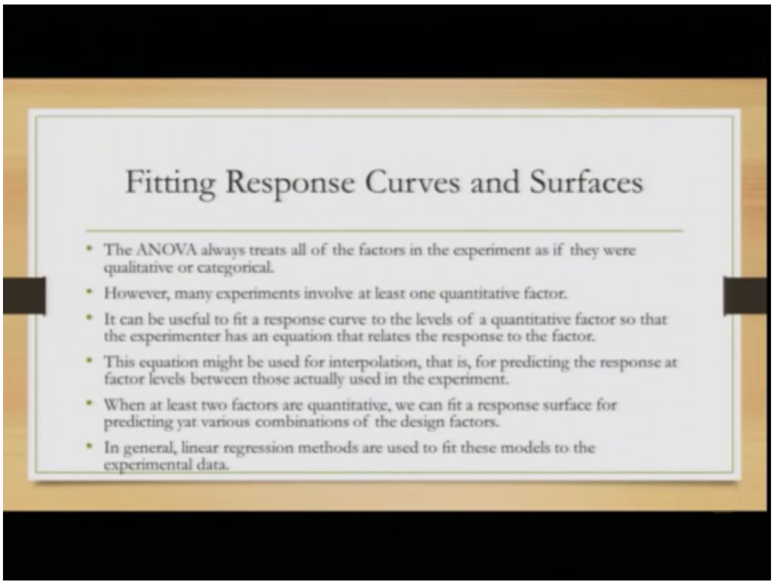
So, it means that the main effect on international plots for example, in diagram a is a percentage of carbonation which is for factor A. The second one is pressure b, which is basically the factor B; and the diagram c is the line speed which is factor C. And then I find out the carbonation in the pressure combinations based on that we have the diagram d as a expect. Now, figure 5.71 as shown basically shows the average fill height deviations, deviations again for high speeds and low pressure for different carbonation levels.

So, as I check you will find out that you take different values. So, the at lower values here, it is a deviation for which it is very small. And we would definitely like the deviation to be as small as possible. So, when we and when we plot the percentage carbonation, and we are trying to find out the improved deviation distribution, the percentage carbonation and the distribution the person carbonation in two different cases. And then we find out that considering normal distribution, we will be able to find out the deviation has really decreased that is what our aim would be.

So, we try to find out the best combination. So, for these three factors which will give the maximum amount of throughput; if throughput is not based throughput means average and if a concern is to reduce the deviation; obviously, we will try to basically concentrate on the standard deviation and the variance

And one thing we should remember that we underlined underlying we had always consider normal distribution to be true which may not be the case in the maximum examples. And another example is that we are always considering the sum of the square of the errors, which intuitively would not is the best way of trying to find out how to minimize the error because square means that we are trying to find out the variances. And trying to minimize the variances would also mean that we are trying to find out the dispersion to be the minimum.

(Refer Slide Time: 21:56)



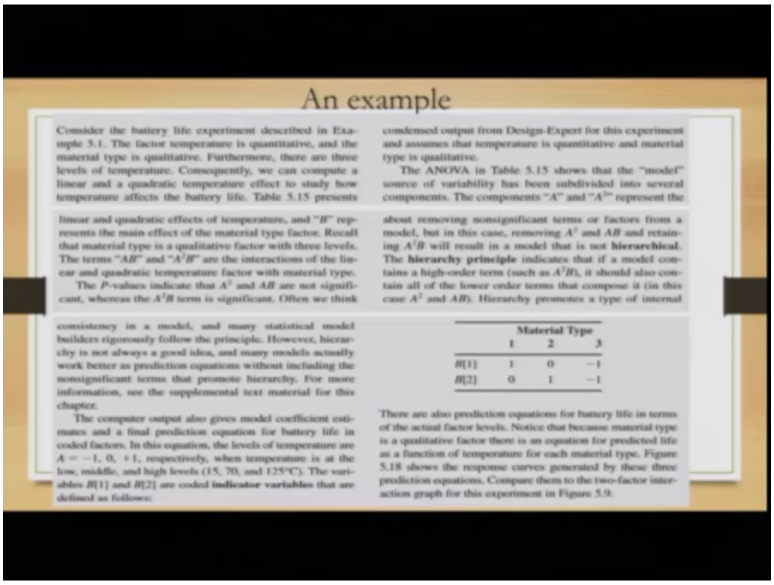
Now, you want to do find out what study something about the fitting response curves and surfaces. So, the ANOVA always treats all the factors in the experiment as if they were qualitatively or categorically designed. However, many experiments involve at least one quant quantitative factors, it can be useful to fit a response curve to the level of quantitative factor, so that the experimenter has an equation that relays the response on the factors and basically builds up the model accordingly.

This equations might be used for interpolation and better prediction that is for predicting the responses at the factor levels such that we are able to understand the effect and the relationship of the factors to the maximum possible extreme using quantitative equation modelling.

So, let me read it for predicting the response at the factor levels between those actually used in the experiment. When at least two factors are quantitative, we can fit a response curve because in that case we will basically utilize the x-axis and the y-axis for those factors and then basically plot the curve accordingly.

So, we can so let me again continue rereading it with your due permission, we can fit a response surface for predicting in the various combinations of the design factors. And in general linear regression models are used to fit these models, because linear regression would be the best way of trying to basically find out the quantitative relationship which we are thinking about.

(Refer Slide Time: 23:23)



So, let us consider an example in detail. Consider the macro level experiment which you are already designed discuss. The factor temperatures temperature is quantitative and the material type is qualitative. So, material type can be same in higher level of say for example, lithium, lower level of lithium considering lithium batteries. But the temperatures are if you if you remember the temperatures are a quantitative in nature and because those are values either in centigrade or Fahrenheit. Furthermore there are three levels of temperature. Consequently, we can compute a linear and quadratic temperature effect to study how temperature affects the battery life.

So, figure five 5.15 the table 5.15 represents the linear and quadratic effect of the temperature and b represents the main effect of the material type factors. So, let us recall that the material type is a qualitative factor with three levels. So, consider the terms as A into B, another would be A square into B depending on the level of significance we think A would have what basically denote the interaction of the linear and the quadratic type. When A and B, it is basically a linear relationship between A and B.

If it A square into B, it is basically quadratic relationship between A and B and again if it can also be put like a into b square also then the quadratic relationship in words remain the same, but the f x would be different because in the first case you are squaring A A value to A square. In the second case, you are trying to basically squaring B to B square.

So, the P values indicate that what is the relation indicates that A square and AB are not significance. Whereas, in the term that we can whereas A square B terms is significant so obviously, you have to basically pass judgments accordingly. Often we think consistency in a model and many statistical models builders rigorously follow this principle. However, hierarchy is not always a good idea.

So, if I want to basically have an hierarchy, then A square B would be the highest effect then AB squared would be the next effect then would be AB and so on and so forth, so that may be true theoretically, but what is the practical essence it is very difficult to find out.

So, however, hierarchy is not always a good idea and many models actually work better at prediction equations without including the non segments terms that promote the hierarchy. For more information, let us see the temperature takes material for this chapter. So, the computer output also gives model efficiency. So, what we have is basically in this equation, the level the temperatures I have been denoted by as a as minus to the lowest level 0 at the middle level and 1 plus the higher level. So, when the temperature levels are consecutively at 15, 70 and 125.

So, basically am I am trying to equate the quantity part to the qualitative part in a way using as plus 1, 0 and minus 1. The variables are coded are coded indicator variables that are defined as follows. So, they would basically be the design experts comments for this experiment. And we would assume that the temperature is qualitative and the material type is also qualitative, that means, we are converting the temperature from the quantitative scale to the qualitative scale.

But obviously, your question would be that when we are converting these temperatures of 50, 70 and 125, so should not it be say for example, because the difference between 15, 50, and 70 is basically about 55. And the temperature difference from 70 to 125 is also 55.

But what if say for example, the different temperature differences from 70, it is now to 225, so how would we do that; obviously, the scaling factor has to be brought into the picture. So, rather than plus 2, we maybe I am tempted and utilize the value of plus 2 to signify that a higher level of quantitative to the qualitative conversion, which we are trying to do. The ANOVA in table 5.15 shows that the model sources the variability have been subdivided into several components. The components A and A square represents about removing a non secure terms or factors from a model.

But in this case removing A square, and AB and retaining A square B will result in the model in different hierarchies which may not be actually possible to understand. Now, hierarchy principle indicates that if a model contains a higher order terms such as A square B, it should be content it should also contain all the lower order terms also. Which means that in case if the step jumps for the quantitative part of our or our equal values, then trying to basically do a one to one mapping from the quantitative part of the qualitative part would be easy.

But if the jumps are different values of different quantums, so say for example, jumping from in the first stage is jumping from 10 to 20. And on the second stage is jumping from 20 to 50, then trying to be single him find out plus 1, plus 2, plus 3 may not be the actual answer for converting the quantity to the qualitative concept. So, hierarchy promotes a type of internal way of trying to model it. So, there are prediction equations from the battery life in terms of the actual factor models.

So, we should notice because material types is the quality factor, hence it is easier to convert the quantitative to the qualitative one rather from the qualitative to the quantitative one.

So, with this, we will end the 21st lecture and continue discussing more about this quantitative and quality part later in the 22nd and 23rd lecture and further on. Have a nice day.

Thank you very much.