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

Lecture - 14
The Analysis of Variance (ANOVA) - VII and Introduction to Factorial Design

Welcome back my dear friends, a very good morning, good afternoon and good evening to all the students who are taking this TQM II lecture under the NPTEL MOOC series and I am Raghunandan Sengupta from the IME department IIT Kanpur and this is the 14th lecture which means that you are in the third week. So, with the 14th and the fifteenth you will go into the end of the third week you will take the third assignments and we will continue more discussions more about the TQM II.

So, if you remember we had gone into trying to basically consider the concept of ANOVA how they can be intra inter relationship between the errors. So, we basically divided them into mean square or the total errors for treatments, and total errors as for the errors as such total sum of errors and then we also considered the random effect, where the effects would be coming inside the sample for each so called treatment or etching problem.

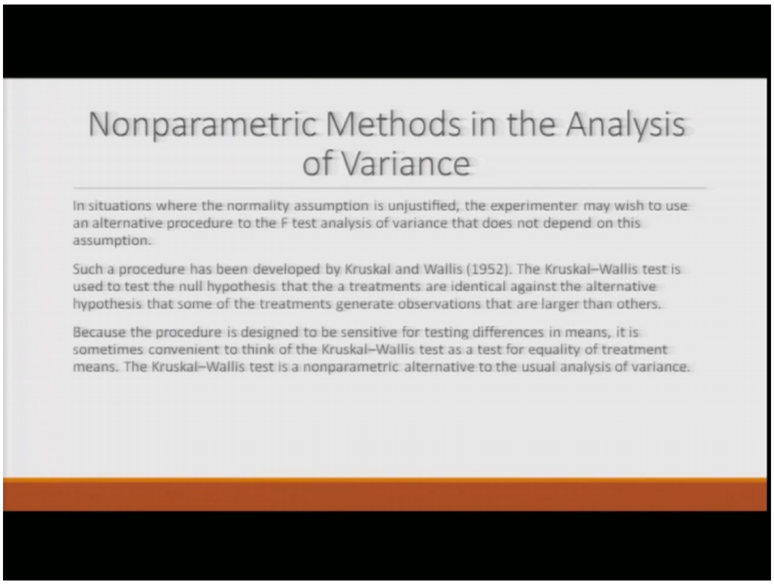
I come back to this example time and again because we have been repeating it trying to discussed it for form of quite a few lectures starting the 9th, 10th, 11th and basically more into the 12th also. And we considered that how we could find out the estimated mean which is the best estimate for mu.

So, we used mu hat which was y bar dot dot. So, those two dots I need not repeat the first dot basically implies with respect to i and second dot implies with respect to j and i changing from 1 to a the total number of treatments and changing from 1 to n total number of samples in each treatments.

We can also consider the concept that what is sigma square, why it should each should be kept constant what is tau y and what does the implication of tau I mean, we consider a different type of test apart from the hypothesis testing the Bartlett test the we considered the test as a test statistic as p value, f value and we also went into chi square concept we consider the capital T Tukey’s test and so on and so forth.

So, let us continue discussion will basically come back to these concepts more. And more and I am sure with many repetitions and the things the basic concept being repeated more than one time and you going through the slides, going through the assignments and as well as basically referring in the book which I told the book you should need not buy because it is a costly book you can refer to the library or get some help from your faculty members teachers, whoever has a copy or has access to this copy you can basically have a look at this MOOC.

(Refer Slide Time: 03:09)



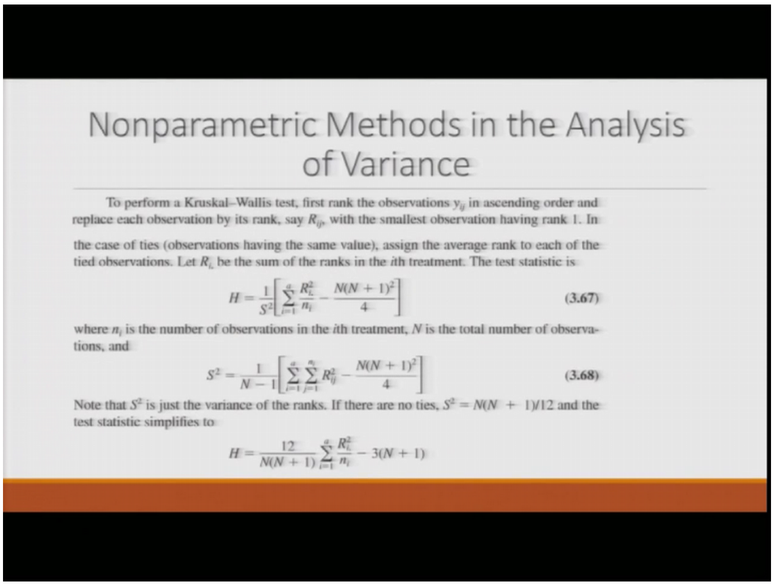
So, now we come to the nonparametric methods in the analysis of variance and ANOVA and I did not mention by the way just small digression no not going into too much in the digression concept. The concept of when you consider the multivariate analysis of variance we consider the concept of manova, we would not be discussing the concepts of manova, but the conceptual framework for manova would be in same names like ANOVA.

So, for the nonparametric methods analysis of variance in situations where the normality assumptions is unjustified the experimenter, may wish to use an alternative procedure to the for the F test analysis that does not depend on the assumptions.

So, if you remember for all the tests whatever we consider the underlying distribution to be normal and based on that we proceeded. Such a procedure has been developed by Kruskal and Wallis. So, the Kruskal-Wallis test is very famous. So, is used to test the null hypothesis that the treatment are identical against alternative hypothesis that some of the treatments generate observations that are larger than the others and they are basically not identical because the procedure is designed to be sensitive from testing differences in in the means, it is sometimes convenient to think of the Kruskal-Wallis test or the KW test.

As a test for equality of treatment basically we are trying to be compare the mean values of the treatment or the etching which was basically four different powers 161 822. The KS test or the Kruskal-Wallis test is a nonparametric alternative through the usual analysis of variance.

(Refer Slide Time: 04:48)



Now, to perform the KW test or the Kruskal Wallis test first rang the observations. So, the observations are w ij and I have already mentioned what the suffix i means what the suffix j means. So, rang the observation y iy ij in ascending order and replace his observation by its rank say r or capital R suffix ij with the smallest absorption taking the rank 1 and so on and so forth going on.

In case of ties with the observations of the same rank on the same value, and sign the average rank to each of the observations; So, let R i dot be the sum of the ranks in the ith treatment. So, for ith treatment they would be number observation. So, again you are summing up for the number of observation which you if you remember, that was basically the suffix j being utilized and for each treatment or each etching we had basically the sample size of n.

So, the test statistic which is the Kruskal Wallis statistic is given by the formulation of 1 by square, I am going to come to the conceptual feel of s square later on. So, 1 by square multiplied by the summation of R i square by n minus n into N plus 1 by 4. So, now, if you consider the Ri, Ri are the basic of the ranks.

So, you are squaring them which means that you are trying to build bring into the picture, the concept that the concept of squared error loss or the concept of equal parallelization for over estimation or under estimation are of the equal quantum and divided by n is basically you have for each i, i is equal to 1 to a the number of etching or the number of treatment you want to basically divide it for each respective values of n is to find out the average values.

So, what you are doing? Doing is that for R i dot you are summing up all the ranks squaring them up and dividing by n i corresponding to that treatment to find out the average square of the so, called errors based on the ranks, and then sum up all the all the values in a way that you will basically find out the difference from the mean value. So, where is the mean values this is what this says; So, minus n into N plus 1 by 4.

So, now if you remember n was basically capital N was basically the multiplicative value of small n into a; that means, the whole set of observations which you have you might multiply those values in such a way that you find out the average distance so, called distance is not a distance in the equate like, in the Cartesian coordinate or if you feel the distance when you are traveling from city a to city b, but which basically.

This is the difference based on the concept of squared losses where an what are the implications of the of the different type of variables, which are used I have repeated that, but I will again read it out for your own convenience. Where n i is the number of observations in the ith treatment, capital N is the total number of observation which you have already observed it to be kept small n into a and S square is basically the square of the standard errors.

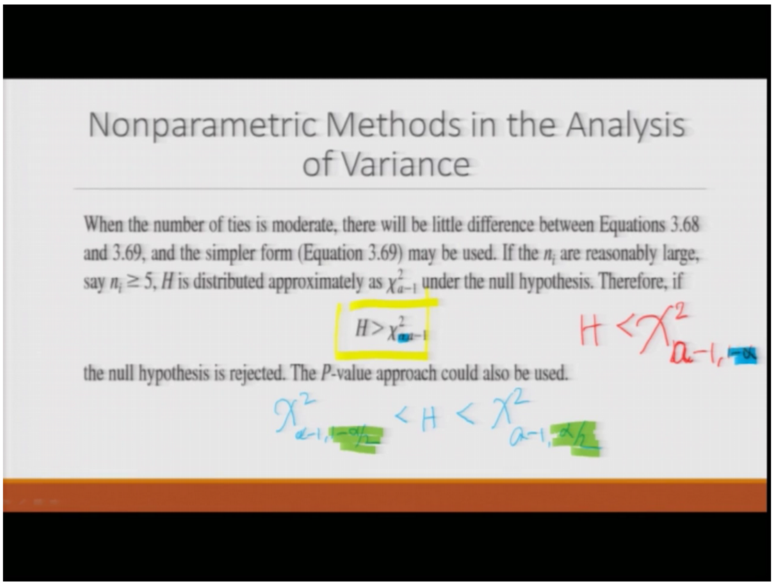
Now, why I am saying that if you look into the formula that will give you and given in information. So, what you doing is the, you first look at the denominator is 1 by capital N minus 1 which is basically you are dividing by the so, called sample size, but you are losing one degrees of freedom. Because one degrees of freedom the reason being that you are finding out the best estimate of mu using y bar dot dot which is the estimate so;

that means, you are losing one degrees of freedom. And the those values which you find out the differences which you find out are with respect to the so, called squared error loss; I am using the word square of loss for the first time, but it will become very clear to you why I do that within another few minutes please have patience I will explain that.

So, once you find out S square. So, here note that S square is the variance of the ranks as I mentioned if there is no ties so; obviously, S square would be N into N plus 1 divided by twelve then the test statistics means basically simplify to the formula which is given here.

So, I am using the word H or the symbol H to denote this test statistics, which is the Kruskal-Wallis test statistics and highlight it using the yellow color. So, this is the test statistics here you know R square suffix i dot is the rank squares sum them up for each a treatment or etching and then square them up and this n i is basically the number of observation which you have for each etching or for each treatment.

(Refer Slide Time: 09:57)



When the number of ties is moderate, then there will be little difference between the equations which you found out where there is some tie and there is not tie.

If n is reasonably large; that means, for each and every etching the number of observations samples you are taking is large, which is greater than say for example, five. So, s would be distributed approximately by chi square with a minus 1 degrees of freedom under the null hypothesis. So, you will basically find out whether that is greater or less than or in between the intervals depending on what you are trying to test, whether it is a right hand test left hand test or not equal to.

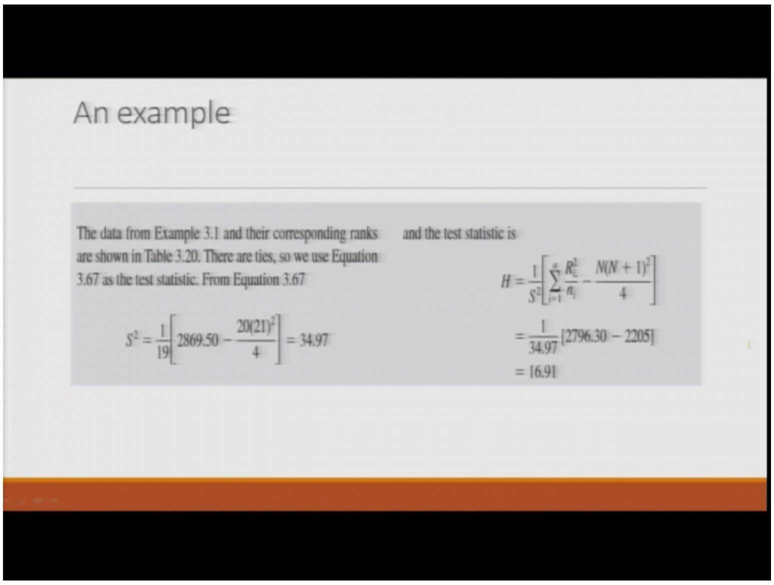
And the test statistics as I mentioned the values is basically chi square; if it is on the left hand side it will be a minus 1 which is the degrees of freedom which is already fixed; and the overall value of chi square will be denoted by 1 minus alpha it is on the left hand side, it will be denoted by alpha if it is on the right hand side as mentioned here. And if it is in between it will be denoted by 1 minus alpha by 2 and the right hand value would be alpha by 2.

So, these things we should be very clear about. So, the null hypothesis would be rejected for the right hand side, if this value which I am going to highlight which are already repeated is H is greater than equal to chi square a minus 1 a, 1 is the degrees of freedom and the suffix value of alpha would basically note at what position level of confidence it is. And we can also use the p approach to find out whether you want to reject or accept the null hypothesis.

Similarly, if you want to may have a concept of less than time so; obviously, the formula and use the color red. So, it will be H I am writing the degrees of freedom first. So, this is the value 1 minus alpha which I denoted I use an another color to highlight it. So, if you denote this and this they are different and if it is both sides. So, we to become I am using the blue color. So, in order to make the differentiation chi square a minus 1 1 minus alpha by 2, H less than chi square a minus 1 alpha by 2.

So, just note down the differences what color should I used I should use the green one let me use the light green one for the time being. So, note this with respect to the blue one where I am highlighting my pointer for the right hand side and this is for the left hand side.

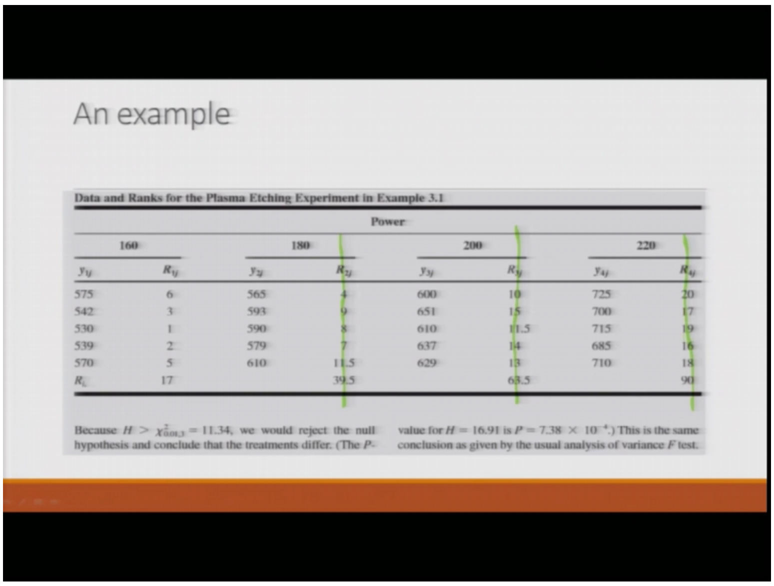
(Refer Slide Time: 13:27)



So, let us consider an example the same one data from example 3.1, and the corresponding ranks are shown in the table you can find it out. There are at ties so, you use the equation as given not to test the statistic, we find out s square to be about 34.97 consider is 35 and the test statistic value once, you put it comes out to be again I will remove the decimal values it is about 17 because it is 16.91.

So, the so, if you are interested in the actual values based on which I just mention the end results.

(Refer Slide Time: 14:07)



So, if the data is collected for the plasma etching problem, again going back to the etching problem for the 160 wattages power, the y y ij that is for the first treatment the jth reading. The jth reading are j is equal to 1 is 575 then for 2 it is 542 5 then next is 530, next day 539, next is 570. The rank values are given R i j you rank them as the first one would now basically get the rank of sixth, then it will be third then first second fifth and the total rank is 17 you add them up, then you find out for the power for 180 what are y 2 js you put them rank them.

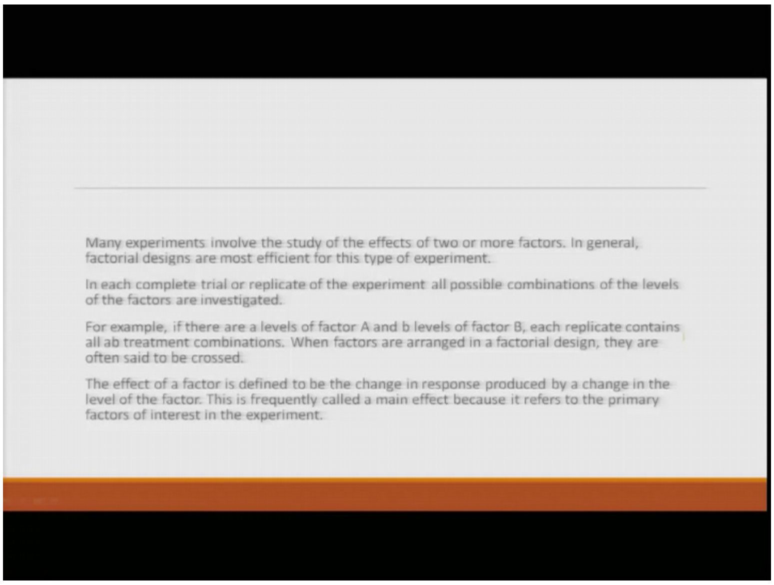
So, those ranks are given when I am just highlighting on the fourth column similarly when I do it for the a is equal to let me use the word not a, i is equal to 3 I rank them I find out R 3 j, when I do it for the last value which is i is equal to 4, which is a value I find on the rank which is the last column which I am just highlighting.

So, because H is greater than H value means the Kruskal Wallis test statistics greater than chi square of 0.01 comma 3 comma 3 because it is a minus 1 and alpha value which you have is 0.01. So, if it is equal to 13.47. So, we will reject a null hypothesis and conclude the treatments do differ and similarly we can basically do the p value test to find it out accordingly.

So, we can do it for the different the example of the smelting process also, which was just discussed not in that detail like the etching plasma etching problem, but it was does it was discussed in the initial stages.

Now, I will go into the brief introduction of introduction to factorial designs, this will be a slow process may not be as fast as ANOVA. Conceptually it may be a little bit tedious for us to understand, but I am sure once you appreciate its actual implication then the utilization will be much easier using the concepts from the book in general.

(Refer Slide Time: 16:22)

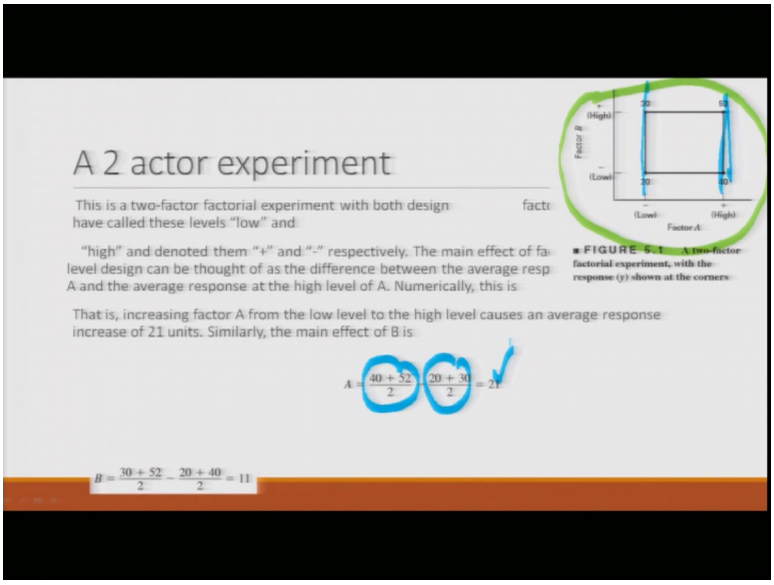


Many experiments involve the study of the effects of two or more factors in general factorial designs are most efficient for these type of experiments, in each complete trial or replicate of the experiment all possible combinations of the levels of the factors are investigated.

For example, if there are levels of factors A and b are the factors b capital B each replicate contents and the A into b values treatment combination. So, you are basically finding on the treatment and the combinations accordingly, when factors are arranged in a factorial design, they are often said to be crossed depending on what the combinations are.

The effect of a factor is defined to be the change in the response produced by a change in the level of factors this is frequently called the main effect model because it refers to the primary factors of interest for the experimenter in the experiment what he or she is trying to do.

(Refer Slide Time: 17:18)



So, A 2 factor model and basically the general implication on the background. So, this is called a factor 2 mod m factorial experiment with both design having being called the levels of low and high so; obviously, there would be low high, low high it can be low medium high low medium high, but we are considering only 2 factors. So, the combination is technically 2 by 2. So, the high is denoted with the plus and the low is denoted by the minus or vice versa depending on however, we quantify.

The main effect of a factor level design can you brought off and or thought off as an as the difference between the average respectively of A and the average responses to the high level off of A. So, basically we have some stimulus some response is coming out from the stimulus, they can be of high and low and you want to basically test the factorial effect design in order to basically make some meaningful conclusions.

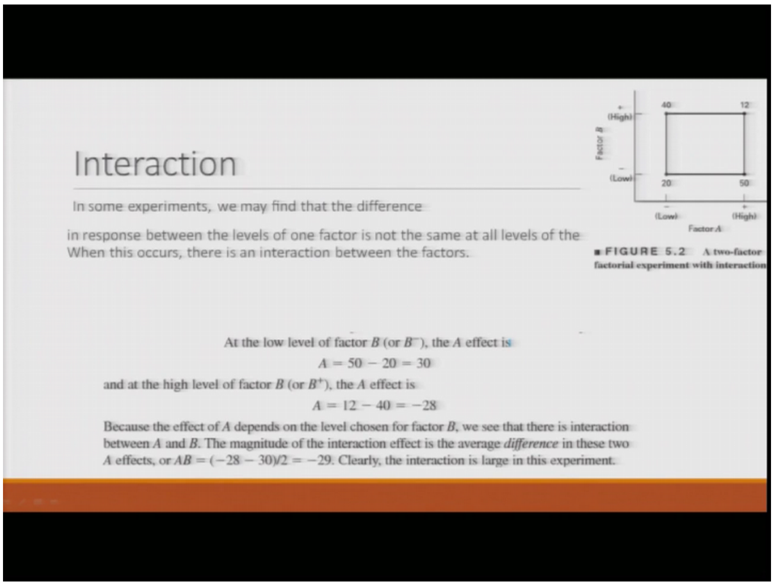
Numerically this can be thought of like in the diagram which I have. So, they are basically 2 factors, factors A and factors B we are basically just highlight it with the green highlighter.

So, this is the diagram it can be combination of more than 2 factors also. So, the low factors for factor A, is there for the low has a factor value of low which is given a minus value. So, technically we can denote A with a minus, but we are not going to make it things complicated, and they would be a high value for corresponding to A which can be denoted by A plus. So, the values of A minus and A plus or minus and plus for A values are 20 and 40.

Similarly, when you go to the factor B the combinations are again B minus B plus or low value and high value for B and those values correspondingly are 20 and 30 and we when we combine them. So, they can need not be symmetric always. So, the combination gives us a extreme right point of our value of 52. So, when we kind find out the values of A.

So, the A values average would basically be you want to find out the what is the average value A, A would be basically 20 40 plus 52, which is on the right hand side. So, I will use a different color. So, these are the values. So, this is what is the average here minus, the value which we have is 20 plus 30 you divide by 2. So, the total value comes out to be 20 21 for A similarly you can find it for B also.

(Refer Slide Time: 20:19)



In some experiments we may find out that the differences in response between the level of one factor is not the same at all levels correspondingly.

So, say for example, the level is 20, for B consider that is 30 for B when the responses which we will have for a may not be of the same quantum differences or vice versa; that means, for A we have 2 different levels of responses and the corresponding combination of B or the response of B are at and the lower level of v and at the higher level of frame may be different. So, this is what is what we mean.

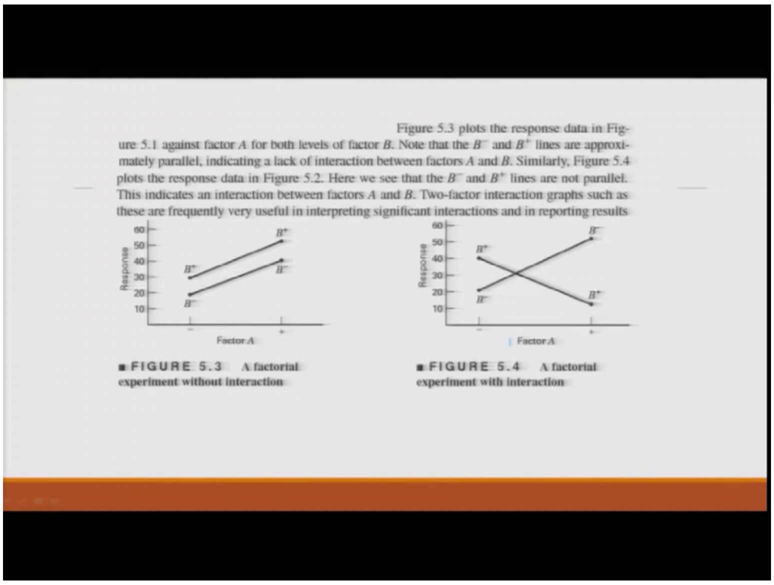
So, in some experiments may find that the difference in response between the levels of one factor is not the same at all levels. So, that is what I mentioned. When this occurs there is an interaction between the factors. So, the factors would be and the low level of actor be, we would basically have the effect of a would be 50 minus 20 is basically the difference of the high and the low which is 30 and the high level and a higher level of B which is B plus I did mention that the lower level of B, would be B minus or minus for B.

So, the effect of A would now be basically negative, which is twelve minus forty which is minus 20 80 because. So, it means that as being is or as A increases there may be a negative effect also that should also be considered. Because the effect of A depends on the level chosen for factor B, we see as see that there is a interaction between A and B, the magnitude of the interaction effect is basically the average difference in these two values of A which is basically when we try to find out the of the.

So, called combination of A, and B it will be combination of the values of the difference of the maximum value which is now was basically for B plus was minus 28 and minus of that value would basically be the lower value of B what is effect average value F effect does it have on a which is minus 30.

So, the average value comes out to be minus 29. Clearly the interaction is quite large because 29 is quite a high value; it would be quite large in the case of this example. Why I am saying large because we see the values of A and B they are definitely in the range of maximum 50 60 not more than that so; obviously, a value of 29 leaving aside the sign of that is quite appreciable.

(Refer Slide Time: 22:48)



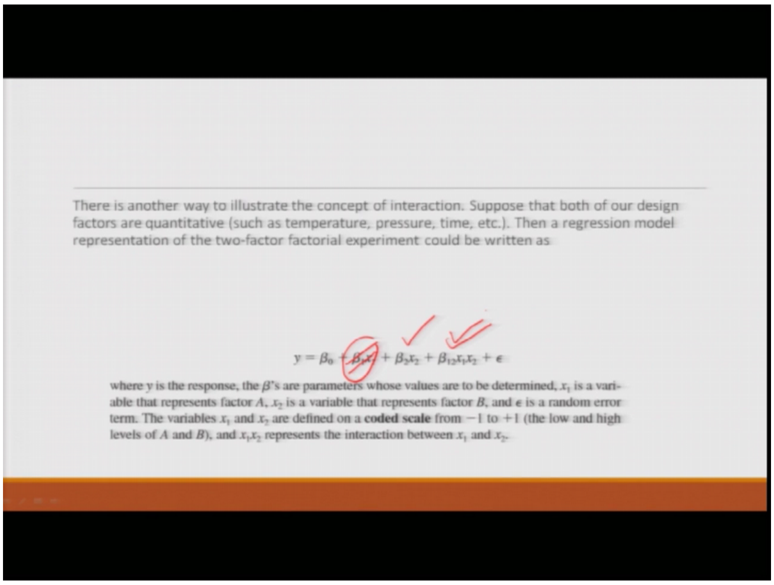
Now, in figure 5.3 which is on the left hand side, where I am I am holding my left hand palm. So, that plots the response data for figure 5.1 against the factor A for both levels of factors of B. So, which means that for B minus initially A had a minus value of minus 20, which is this blue. The A plus value would basically B now given my say for example, 40. So, this means this line corresponds to the case and use the (Refer Time: 23:22) pen with the color that will be red color.

So, this red line which I am drawing, I am trying to basically follow the black one corresponds to B minus only, the lower values is basically for A minus, factor A minus and upper value is basically which is 40, which is for A plus. Similarly when you change the levels of factors of B from B minus to B plus. So, they considering there are 2 steps B minus B plus A minus A plus, the value of B plus changes from 30 to 50 and the corresponding values of A would basically change accordingly and you can find out the values change for A with respect to B also in the other way around.

Now, if we consider the similarly if you consider the figure 5.4, it plots the responses where there a inter relationship; that means, the movements are not in the same proportional manner, they would be negative in inversely proportion concept should also be there. Similarly which means let me read similarly figure 5.4 plots the response data of figure 5.2; here we see that B minus and B plus lines are not parallel they intersect, this indicates an interaction between the factors A and B, the 2 factor interaction graphs such as those that are frequently used are very helpful in interpreting the significant interaction of reporting results, which would basically otherwise be missed.

So, you want to basically draw some meaningful quantitative and mathematical analysis from them.

(Refer Slide Time: 25:04)



There is another way to illustrate the concept of interaction suppose that both of our design factors are quantitative such as temperature pressure, time, humidities, wind speed whatever it is. Then a regression model represents representation or the 2 factor model can be represented as like this.

So, there are 2 factors which is x 1 and x 2 so; obviously, we will have one will be independent for x 1, one would be independent for x 2 and 1 would basically the combination. So, this will be y equal to beta naught or alpha corresponding to basically the no effect coming from either x 1 x 2 or combinations of x 1 and x 2. The next 3 terms will be bit the leave aside the coefficients; coefficients can be made beta 1, beta 2, beta 3 whatever it we like.

So, the effects are coming from x 1, which I am just highlighting with the red color this one, it is not a cross is just highlighting. The next one is basically will I put the tick mark is for x 2 and where I will put a double tick mark is basically for effect coming from x 1 and x 2 combination and obviously, there would be an error term. Now, you will be thinking that what combinations do I have in case in a very general format when there are 3 factors. Consider obviously, there would be on a univariate level; univariate level means each has basically one in singular effect x 1 separately, x 2 separately, x 3 separately.

Now, if I go to the 2 factors they would be combination of x 1 into x x 2, another would be x 1 into x 3 and the third one would be basically be x 2 unto x 3. Now I go to a higher level where the 3 factors are also there; obviously, it will be combination of x 1 x 2 x 3 combined. So, if as I go up the combinations will be considered according me and you can make our model theoretically very nice, but the main problem would be how we are able to solve the problems accordingly.

So, here it states that y is the response beta B are the parameters who values are to be determined, and x 1 is the variable a x 2 represent the variable that is a difference in factor B. Epsilon is the error term and the variables x 1 and x 2 are defined and on the coded scale from minus 1 to plus 1 so; obviously, you can scale them and if while x 1 into x 2 represents the interaction between x 1 and x 2 taken together.

So, with this I will close this lecture which is the 14th lecture and continue more about the factorial design in the 15, then (Refer Time: 27:41) minus 4. And for any queries as I said that you are most welcome to send it on the forum myself and my TA's will we try a level best to answer to all the queries and please take care and do the assignments properly, because that they will give you a lot of good feel that what we are going to going to consider later on. And, what we have considered in the in the so called lectures, would actually be coming up in the examples in the assignment as well as end term. With this I will close this class and have a nice day.

Thank you very much for your attention.