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

Lecture – 11
The Analysis of Variance (ANOVA) –III

Welcome back my dear friends, a very good morning good afternoon and good evening to all of you all the participants for TQM 2 under NPTEL MOOC. And today is the 11th lecture.

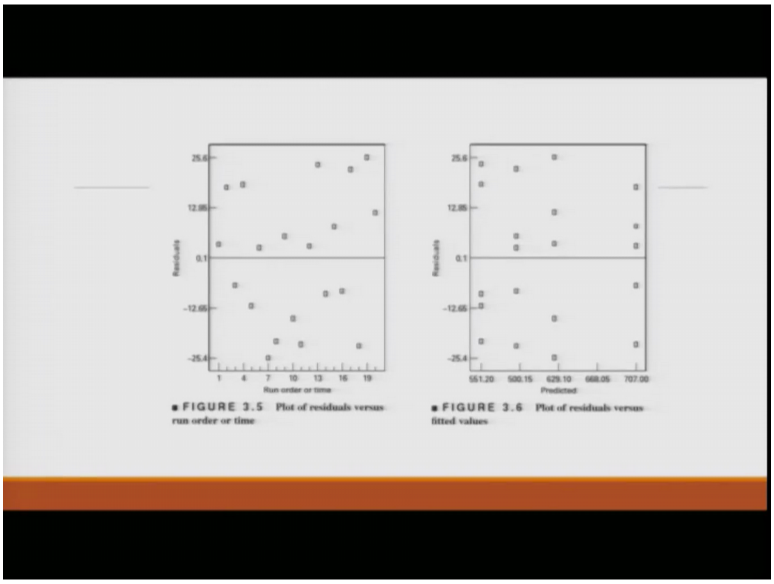
(Refer Slide Time: 00:26)

So, which is the first lecture for the third week and I am sure you have got your this assignments and you are quite comfortably placed considering the concept which are being covered. So, and as you know my I am Raghunanandhan Sengupta from IME department IIT Kanpur.

So, if you remember in the in the last class which is 10th one which was the last class for the second week we discussed the unbalance balanced and how the errors could be there. So, they depend on how? So, I have given you the concept, but; obviously, there would be many ifs and buts depending on how the experiment was conducted outliers being there errors being dependent and all these things, but there are different ways to check whether the experiment is going us as per the norms.

So, some theoretical assumptions should met; obviously, in all the cases all the theoretical assumptions being met in the practical sense may not be possible. So, with this I will continue the 11th lecture. So, if you remember we were discussing on the residual plots the qq plots the normal t and how we can find out some idea about the normal t.

(Refer Slide Time: 01:41)



So, this figure which is 3.5 and 3.6 taken from montgomerys book. So, here we have the run order or time being there on the one for the 4th figure 3.5 along the x axis and along the y axis we have the residuals similarly you have the predicted values on the residuals being plotted in the x axis and the y axis for figure 3.6. So, from based on that; so, it is written that plot of the residual versus the run on time or run order or time is given in figure 3.5 and plot of residuals versus fitted values are even in 3.6.

(Refer Slide Time: 02:16)

See the model is correct and the assumptions are satisfied the residual should be structure less, in particular they should be unrelated to any other variable including the predicted response. A simple check is to plot the residual versus the fitted values this plot should not reveal any obvious patterns and that should be true because if they are random so; obviously, the randomness should be showing up in the in calculations of the values which you are plotting.

A defect that occasionally shows up on this plot is the non-constant variance sometimes the variance of the observation increases the magnitude of the absorption increases so; that means, in case you are taking for 160 180 200 220 as incase the variances increases this would be the case. If the error or background noise in the experiment was a was a constant percent of the size of the observations. So, so this is the second it may become that then if the size of the observations if n increases you are taking more and more observation that would also have an effect on the variances.

If the assumption of homogeneity of variance is violated the f test is only slightly affected in the balanced equal samples of in all treatment fixed models, but; obviously, in the unequal size unbalanced model the effect would be much more pronounced. The approach to dealing with non-constant variance when it when it occurred. So, the above reason is to apply a variance stabilizing transformation and then to run the analysis of variance on the transform data in this approach one should note that the conclusion the and it is a variance applied to a transformed population only not for the unplanned transformed one.

So, basically you change the population do some transformation in order to give some assemblance or normality or bring some of the assumptions in practical sense or in theoretical sense and do the experiment and do the calculations.

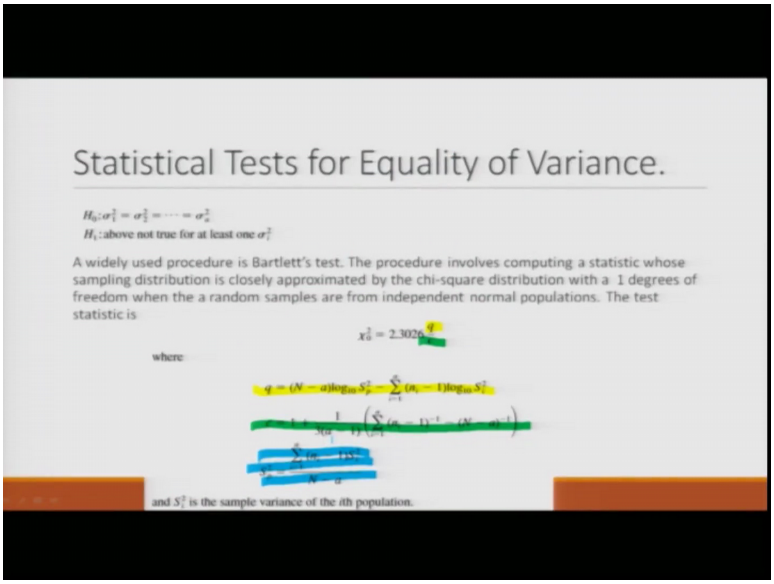
So; obviously, unless assume a variance would be there for the transformed data not for the untransformed data.

(Refer Slide Time: 04:18)

So, few of the transformations for example, if the observations follow the poisson distribution the square root transformation would be used. So, either you can find out why star ij which is the new values which you put in the cell, cell means the matrix 4 by 5 or a by n whatever it is. Square root of y ij or y ij star would basically be 1 plus y ij square root, would be used if the data follows the log normal distribution the logarithmic transformation would used, where yi star would basically be the untrans log value of the untransformed value for binomial expressed as fraction the arcsin transformation would be used which means y star would be I am not mentioning the word value ij would be equal to the arcsin of the square root of y ij.

When there is no obvious transformation the experiment is usually empirically seeks a transformation that equalizes the variance regardless of the value of the mean and then pass it as the experiment accordingly.

(Refer Slide Time: 05:24)



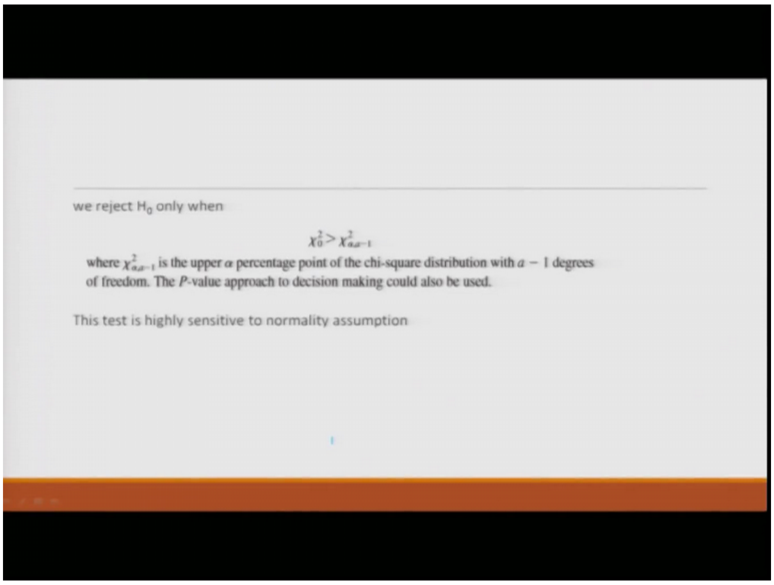
Now, the statistical test for equality of variance which was very important, in the null hypothesis it is sigma 1 squares is equal to sigma 2 squared till the last value zee sigma S square suffix molecule. And an alternative wing being that above is not true for at least one value sigma square I suffix I, where I again I am mentioning varies from 1 to a.

A widely used procedure is the bartletts test the procedure involves computing a statistics whose sampling distribution is closely approximated by the chi square with 1 degrees of freedom when the random samples are from independent normal transformation, the test has a statistics without going into the proof would be given by the chi square under the null hypothesis it is 2.3026 divided by q by c where the q and c values are given. So, it may be difficult for you understand, but when you are solving the problem trying to understand you can safely use these formulas as given in the book and do the corresponding calculations accordingly.

So, this is the q value this is the q value this is the c value this is the c value; and the pooled variance sample variance of the ith population, which n use another is highlighted is this; so, this S square suffix is basically the sam square of the sum standard error for the ith treatment I is equal to 1 2 3 4 till a. And n I minus 1 is basically the sample size minus 1 because you are losing 1 degrees of freedom corresponding to the mean value being not known.

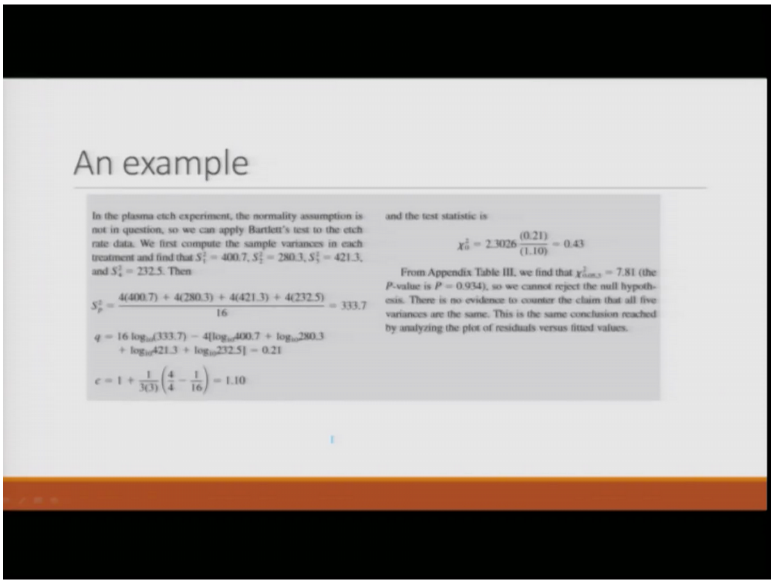
And you will sum it up from in the numerator you will sum it up from I, I is equal to 1 to n and divide the whole thing by the total degrees of freedom for the corresponding pooled sam sample variance would we capital n which is now n 1 plus n 2 plus n 3 dot dot please and n suffix small n I am talking about small n only small n suffix a minus a. So, that would be the total degrees of freedom. So, from that you will get a squares of xp and do that calculations accordingly.

(Refer Slide Time: 07:45)



We reject h naught only when the value so it is on the left-hand side so; obviously, we will reject it or on the right-hand side depending on how you are basically understanding. So, we will be reject a h naught if chi squares non-hypothesis this is greater than chi square alpha comma alpha a minus 1 which is the degrees of freedom. So, where chi square alpha comma r a minus 1 which is the suffix is the upper alpha value of the chi square distribution with a minus 1 degrees of freedom. The P value approaches approach to the decision making could also be used this test is highly sensitivity insensitive to the normality assumptions or else it does not hold.

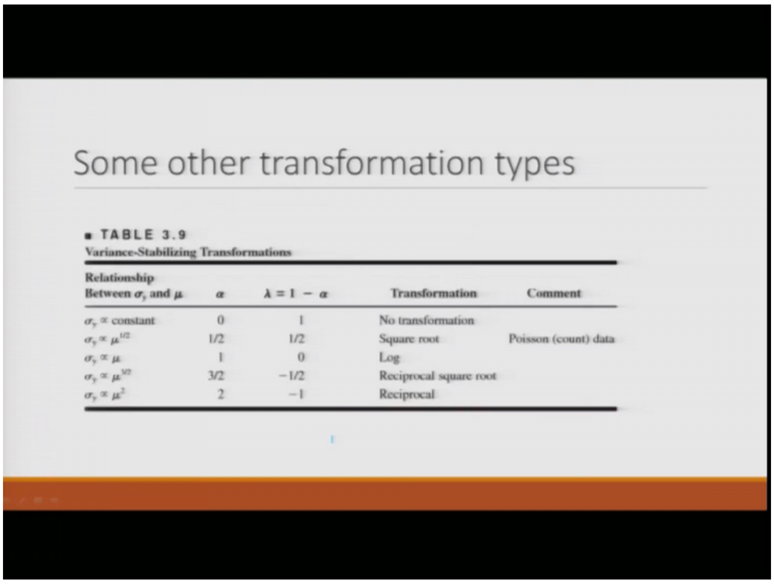
(Refer Slide Time: 08:31)



So, let us consider an example if the plasma etch example which you are consider the normal assumption is not in question. So, we can apply the Bartlett test to the etch rate rate data. We first compute the sample variance for each which is the square root of the standard error. So, Si Si 1 square is r again I am not mentioning the values of the the decimals it is 500, S 2 2 square is equal 280 S S S 3 square is 421 oh it should be basically not S 2 2 squared it is S 2 square, sorry my mistake which was 280, S3 square is 421, S 4 square is 232 then the so called pooled sample variances or standard inner square would come out to be as given as 333.7.

Now, from that we find out the value of q as 0.21 and the value of c is1.10, place that in this values and kind of you find out chi square on the null hypothesis is 0.43. So, from the table you can find out the chi square 4 degrees of freedom being. So, is what a is 4 4 minus 1 is 3 and alpha value is given as 0.05 from that you find out 7.81. So, we cannot reject the null hypothesis. So, there is no evidence to count on the (Refer Time: 09:57) that all the 5 variances are the same. So, this is the same conclusion reached by analyzing the plot of the residual verses fitted values. So, you are basically have seen the diagrams now you are basically trying to corroborate that with your test.

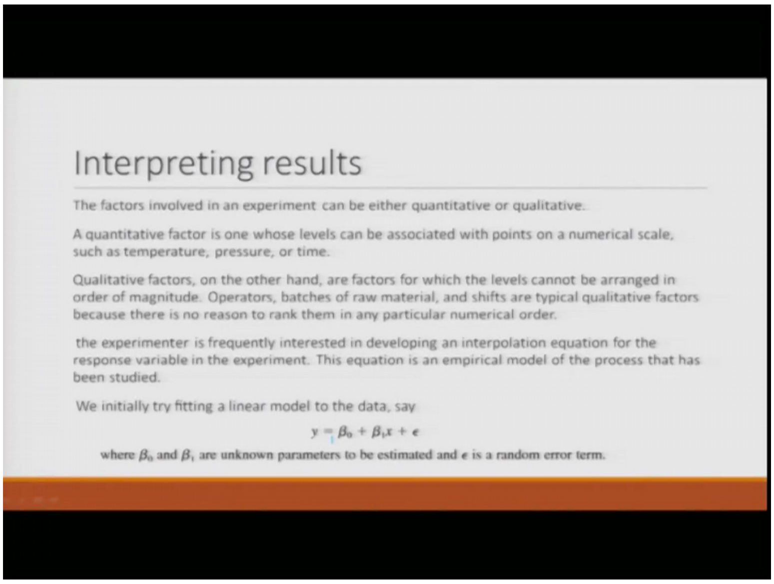
(Refer Slide Time: 10:11)



Other transformations are for variance stability transformations. So, you will basically if the relationship between the variance or standard deviation and mean value if a standard if the standard deviations is basically proportional to constant and the alpha value is 0, and then there is no transformation is utilized if the various standard deviation is proportional to the square root of a mean value, then you use the square root transformation when it is proportional to the mean value then you use the log transformation.

When it is proportional to 3 by 2 power you use the reciprocal transformation. And when it is proportional when it means I am using the word it for the standard deviation, standard deviation is proportional to the square of the mean value then you use the reciprocal transformation. So, these are just for information, but if you get the idea for problem solving the problems you will get everything kick starting.

(Refer Slide Time: 11:17)

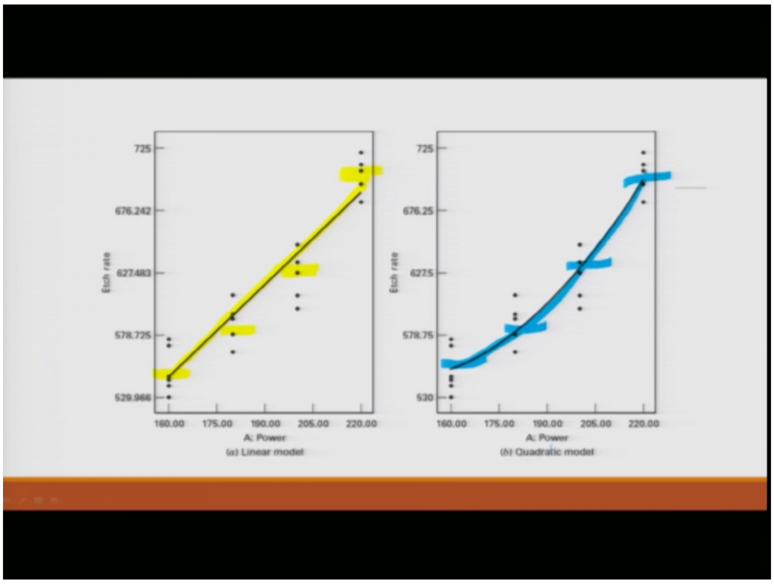


Now you want to interpret the results, the factors involved in an experiment can be the quantitative or qualitative. quantitative factor is one whose levels can be associated with points on a numerical scale such as temperature, pressure time humidity so on. And so, forth quality factors on the other hand are levels which cannot be arranged in order of magnitude. So, it can be operator's batches of raw material the shifts are typical quality factors because there is no reason to rank them in a particular way as we wish. It can be the operators who are coming operator 1 operand 2 operator 3 or it can be say for example, and batches of raw meters which are coming coming from vendor 1 vendor 2 3 so on and so forth.

Ah the experimenter is frequently interested in developing an interpolation equation for the responsive variables in the experiment, the this equation is an empirical model of the process that has been studied and we take basically do the studies accordingly. We initially try fitting a linear model through the data say for example, we say that y is equal to a type of multiple linear regression beta naught plus beta 1 into x 1 if there are one more than one variables plus epsilon if there is only it is a.

If it is a simple linear regression, you will basically a y is equal to beta naught plus beta 1 into x plus epsilon if it is a multiple linear regression it will be y is equal to beta naught plus beta 1 into x 1 plus beta 2 into x 2 dot dot till beta k is equal to xk considering they are k such x values different x values which determine the values of y. So, where beta naught beta one are the known parameters to be estimated and epsilon is basically the nanometer term.

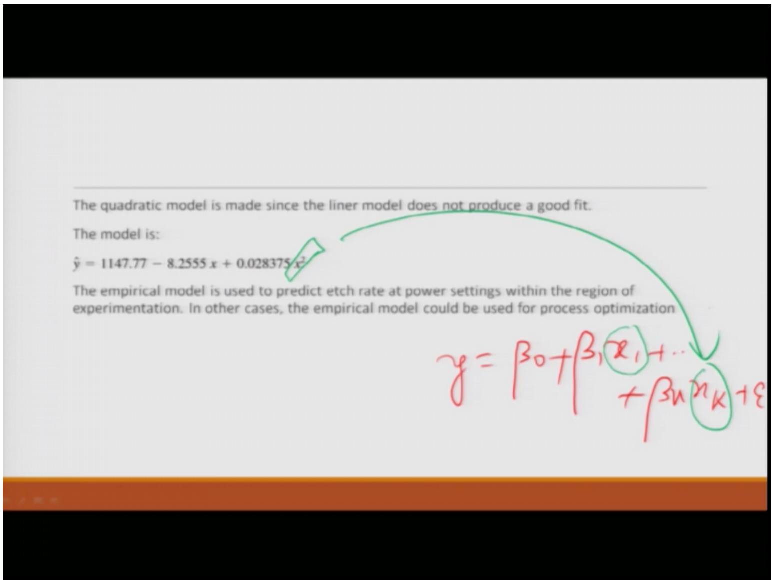
(Refer Slide Time: 13:03)



So, if you basically fit in the data corresponding to the power and corresponding to the linear model and corresponding the power being the quadratic model. So, you are trying to fit the etch rate along the y axis and the power along the x axis, but in the first case you consider powers relationship with the, etch as linear and the second case you consider the power relationship with the etch rate as quadratic. So, you what you are trying to fit you are trying to fit the mean value. So, if you see the diagram here. So, in this case the linear plot is there the mean values are somewhere here, this I am just giving a very rough estimate and if I go for the quadratic model so the mean values are almost falling here.

So, this is just an estimate or estimate which I have done and write to show you in front of you they can be other model fitting also. So, you need not be quadratic it the best it can be cubic it can be square root whatever it is.

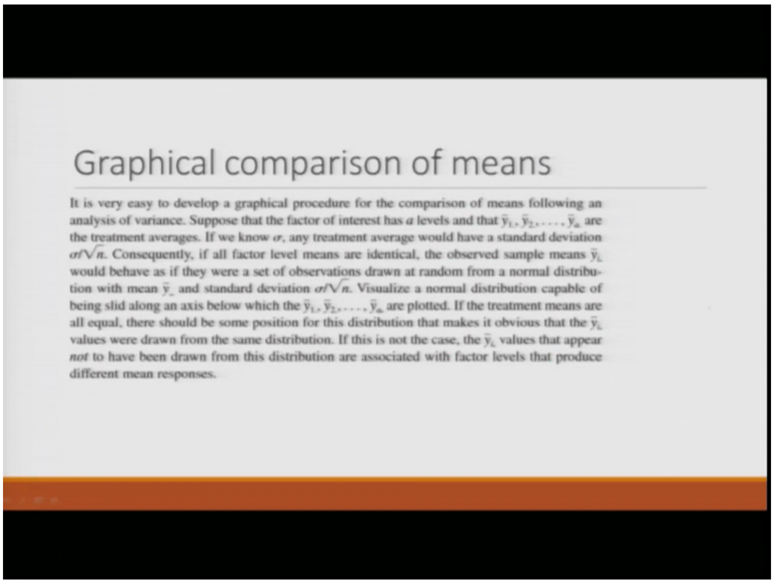
(Refer Slide Time: 14:11)



The quadratic model is made since the lay model does not produce a good fit the model is given. So, this y value which we have taken; that if you remember that we had considered let me put the equation in a right perspective y is equal to beta naught plus beta 1 x 1 plus dot dot dot plus beta k xk plus epsilon. So, these xs which I have and uses some different color . These x s which are using this I am [com considering]considering there are not linear they are quadratic. So, this is where it is coming so something here it can be square.

So, in this example that your quadratic model is the best fit, you find out the values of beta naught as 1147, again I am not mentioning the decimals the value of beta one corresponding to x only comes out to be 8 8 point something and so, called beta 2 corresponding to x square which I consider as x suffix 2 as the second variable. So, then empirical model is used to etch rate at power settings within the regression of the experiment and the model in other cases the empirical model could be used for the process optimization and the work can be done accordingly.

(Refer Slide Time: 15:39)

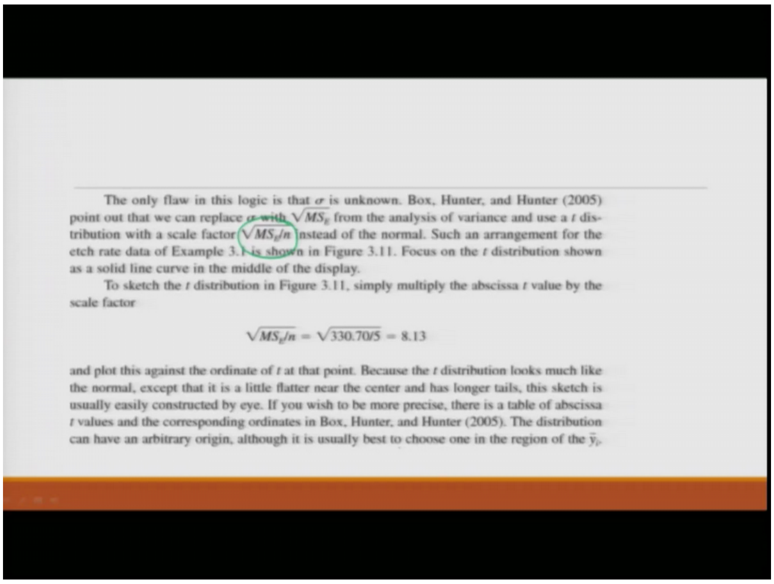


Some graphical comparison of the means it is very easy to develop a graphical procedure for the comparison of means falling an analysis of variance suppose if the factor of interest has and has a levels . So, this a value is if you could remember those were the number of treatments. And that y 1 bar y 2 bar so this 2 1 2 3 4 what I am saying on the suffixes till y a bar are the treatment averages if you remember that is what we have calculated. So, y 1 bar is basically summation of y 1 1 y 1 2 dot dot y 1 n the first one correspondings to the treatment the second number corresponding to the sample observation if it is basically the a 3 pin it will be again I am repeating please excuse me it will be y a 1 plus y a 2 plus dot dot y a n divided by n; so that values are given.

If we know sigma any treatment average would have a standard deviation given by sigma by square root of n, this is basically the actual concept of central limit theorem which I did discuss amongst the first 7 classes if you remember. Consequently, if all factor levels means are identical the observed sample; means y bar suffix I would behave as they were a set of observations drawn at random from a normal distribution with a mean of y bar dot dot, which is the sum of all the that is the average of the average of the grand average.

And with a standard deviation of sigma by square root of n we visualize a normal distribution capable of being slide slipped along the x axis between y 1 bar y 2 bar till y bar y by a bar. I plotted if the treatment means are all equal they should be some position from for this distribution that makes it obvious that y bar I whereas, were drawn from the same distribution if this not is the case, y bar I values that appear not to have been drawn from the distribution associated with factor levels that produce different main responses depending on the outcome of the experiment which is going on.

(Refer Slide Time: 17:44)



The only flaw in this logic is that sigma is known. So, if see if you basically follow the books on the paper of box hunter and hunter, they out that if you replace sigma with the mean square error from the analysis of variance and use a t distribution with the scaling factor or mean square divided by n. So, that is the whole square root because you know what you have sigma square by n square root of that. So, MSE suffix is the mean square you find out the square root and you find in the square root. So, everything is given the square root as given this formula.

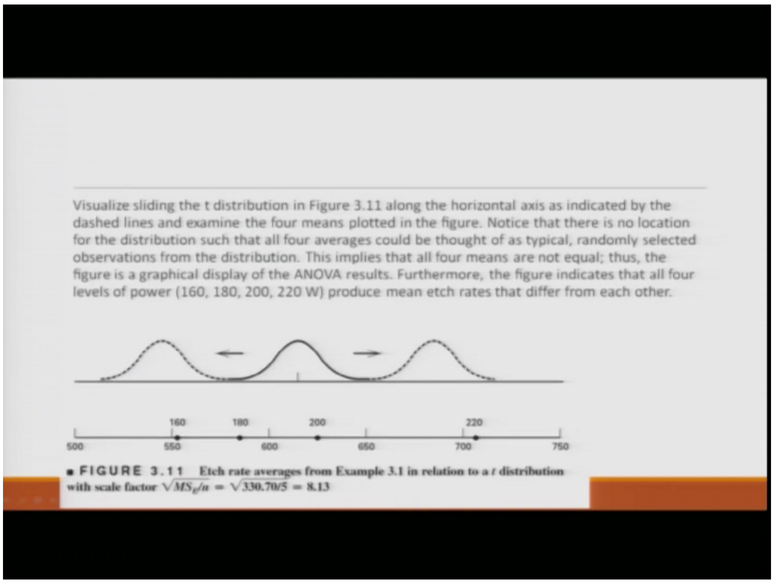
So, instead of the normal case in such an arrangement for the etch rate write data, for a example 3.1 is given excuse me focus on the t distribution shows as the solid line curve in the middle of the display. To etch the sketch the t distribution simply multiply the abscess at t value by the corresponding. So, the value of which I just mentioned is square root of MSE is of divided by n which comes out to 8.13.

And plot this against the ordinate t at that time because the t distribution moves MOOC’s looks more like a non-distribution; again, this is the which had mentioned time and again as the sample size increases that normal in the t distribution means exactly equal to the normal distribution. And another which I did mention when you mention when you are writing the t distribution along the x axis either for the interval estimation out of the hypothesis testing, technically it should be t suffix 1 minus alpha, but in the case as it is symmetric we use minus of t suffix alpha.

This is true for the z distribution also because it is symmetric definite is symmetry, but may not be true for the case of chi square and f destruction; that is why I use on the left hand side chi square, suffix whatever the degrees of freedom is 1 minus alpha and if it is on the right hand side it is chi square, whatever the degrees of freedom comma it is alpha similarly for f f would be m comma n if m and n on the degrees of freedom or m minus 1 n minus 1 whatever the degrees of freedom are comma if it is on the left hand side the value is 1 minus alpha, if it is on the right hand side it will be f with the same with the corresponding degrees of freedom comma alpha and; obviously, if it is both sided for that for the chi square and the f distribution those 1 minus alpha and alpha get replaced by 1 minus alpha by 2 and by alpha by 2. So, this sketch is usually easily constructed by the I.

So, if you have you I want to be more precise there is a table of the abscissa t values and the corresponding ordinates. So, the distribution can have an arbitral origin always usually best to choose one in the region of y bar I and do the calculations accordingly.

(Refer Slide Time: 20:35)

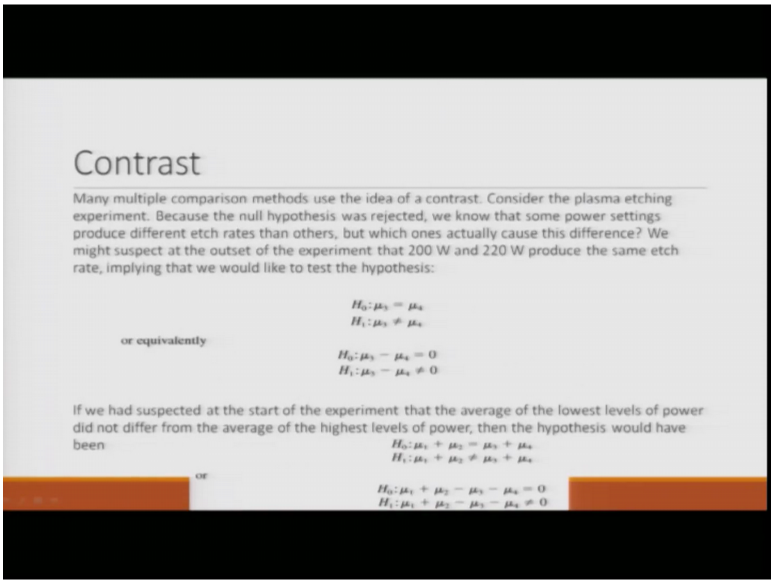


So, here you have basically trying to be seek plotted. Visualize sliding the t distribution as shown in figure 3.11 which is there in front of you in the slide; along the horizontal axis as indicated by the dashed lines and so you move on to the left on the right. And examining 4 means plotting in the figure notice that there is no location for the distribution because either it can increase or decrease; that means, it can get to get more flat or can be more spiked depending on the in the change of the variability.

Notice there is no location for the distribution such that all 4 averages could be thought of typical randomly selected observations from the distribution. This implies that all 4 means are not equal obviously, but it would also mean that if you can do a visual test very simply and being careful the whether the variability is increasing or decreasing or it is constant.

Thus, the figure is a graphical display of the anova results furthermore the figure indicate that all 4 levels of power of 160 180 200 220 produce mean h rates the differ from each other and; obviously, it will give you an idea that the mean values which is basically mu 1 mu 2 mu 3 are not equal. So, at least one of them being not equal; obviously, would negate your null hypothesis. So, from here you can find out the etching and you can do your example accordingly.

(Refer Slide Time: 22:01)



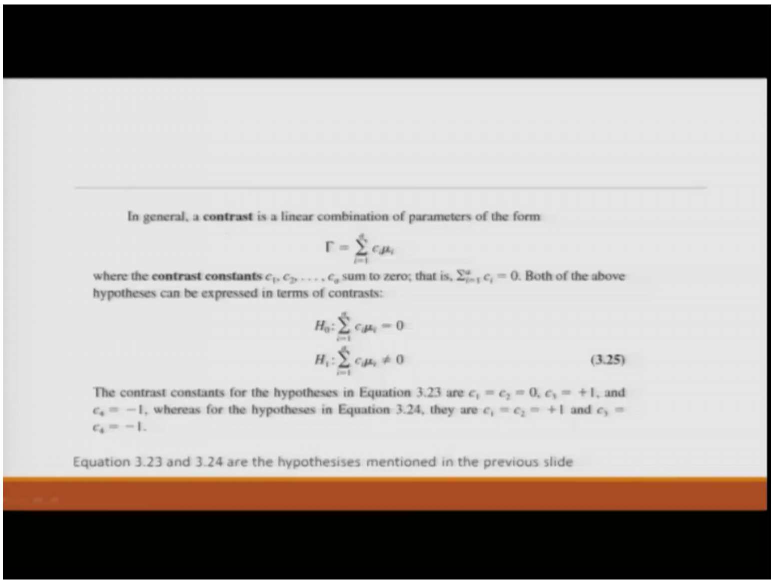
So now, we will consider the contrast and another type of problem. Many multiple comparison methods use the data of contrast. So, when I am basically doing trying to contrast them and compare them and do the analysis consider the plasma etching problem again because the null hypothesis were rejected we know that some power settings produce different etch rates than others and it has been proved in their examples, but which one actually causes the difference is basically the main question we need to answer.

You might suspect that the on set of the experiment that 202 or 220 produces the same etch rate implying, implying that the null hypothesis is can be friend as mu suffix 3 is equal to mu suffix 4, which is the third and the 4th which is the 200 and 220 with respect to the alternate hypothesis which is they are not equal or equivalently the difference between the mu values for the third and the 4th is 0 with respect to the alternate hypothesis then they are not equal to 0.

If we had suspected at the start of the experiment then the average of the lowest value of the power did not differ from the average of the highest power then the hypothesis could have been framed that; 1 plus 2 which is the first and the second is equal to the third and the 4th or and are with an alternative hypothesis could be 1 and average of the first and second is not equal to the third and the 4th or else it can be implied in a different way where you find out the differences between the first and second minus the difference sum of the of the third and 4th; that means, you add up the average of the first and second which is 180 and 1 1 and is 160.

Add that value and from that minus the value average value of 2 to 200 and 220 if it is 0 that is under the null hypothesis if it is not 0 that is the under the alternative hypothesis.

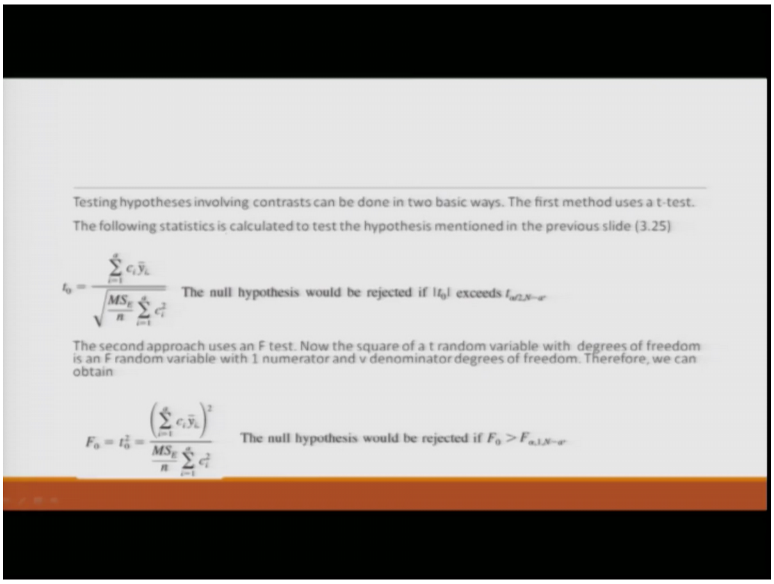
(Refer Slide Time: 23:51)



See generally a contrast is a linear combination of the parameters of the form given that I want to check; that I have basically have a function. So, I want to find out a convex combination of this values of mu 1 and mu 1 mu 2 mu 3 mu 4 or till mu a so in this case if it is from 1 to a. So, it will be multiplied by mu mus would be multiplied by the corresponding c. So, this is c 1 plus mu 1 plus c 2 into mu 2 dot dot till c a into mu a, where the contrast constants are given as c 1 to c a and they sum up to 1.

So, what you are trying to do you are gone to basically be load them and less load them depending on what is the values of c; that is the values of c are cs as 0 both of the can be expressed as null hypothesis the contrast sums are 0 against the alternative hypothesis they are not 0. The contrast constants of the hypothesis in equation ,1 which in 3.2 3 4 that equation done is c 1 is equal to c 2 is equal to 0 and c 3 is equal to plus 1 and c 4 is equal to minus 1 whereas, the hypothesis for equation 3.24 they would come out to be c 1 plus c 2 is equal to 1 with respect to c 3 and c 4 is minus 1 because the sum; obviously, adds up to one in both the cases is 2 0 in both the cases. So, equations are the formulated are basically the corresponding hypothesis in the previous slide that has been done.

(Refer Slide Time: 25:27)



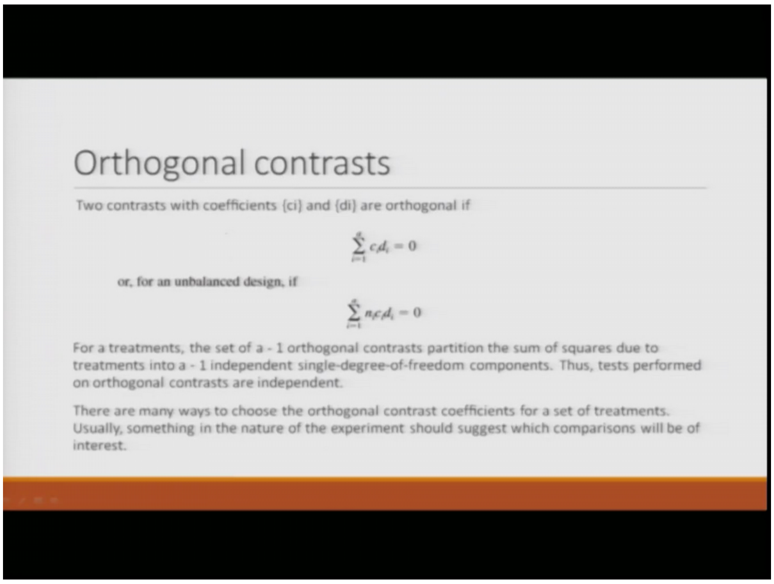
Now, testing the hypothesis involving contrast can be done in 2 basic ways the, first method use your ts test and the following test as test statistics is calculated to test the hypotheses mentioned in the previous line. So, in t naught which is the null hypothesis t values is equal to summation of c I into y bar I because mus have been has to be replaced by the corresponding best estimate, which is mu hat which is equal to y bar and the test statistics; obviously, would have divided by its so called standard deviation divided by the degrees of freedom so that is what is done.

So, a mean squared error divided by n into the sum of the cis would basically be the denominator which you have. The null hypothesis would be rejected if t naught exceeds t. So now, you are taking both the test; obviously, it can should be on the left-hand side and on the right-hand side. So, the degrees of freedom are n minus a values. So, this a values would basically be corresponding to the number of etchings which you have capital n is basically n 1 n 2 n 3 whatever it is you sum them up. And this alpha by 2 value is basically coming to the fact that this is a both sided test not a one-sided test either left on right one remember that .

The second approach uses is basically F test. So now, the square of a t random variable with the degrees of freedom being calculated with 1 numerator and v number of denominators would give you the degrees of freedom. So, therefore, you will have the F would be, if you remember the F distributions can be found not as we have mentioned the chi square the F d distribution and the f distribution in one of those 7 classes first lectures in the (Refer Time: 27:22) end. So, we you can find out the F value and on the null hypothesis, as in the numerator you have the square of the value of cs multiplied by the corresponding estimate of mu hats which is y bar I. So, it will be c 1 into y bar 1 plus c 2 into y bar 2 so on and so forth plus dot dot till c a suffix a; obviously, these suffixes into y bar a divided by.

So, called if you think a simply on this lines it is the mean squared error divided by n multiplied by the square of c c cis which I if you remember the sum of ci should be 0. So, null hypothesis would be rejected if the value of F naught under naught is greater than F now see the values which you have in then the suffixes are F which is alpha. and one is basically the degrees of freedom the other degrees of freedom is n minus a. Now y alpha because you are considering either left or the right-hand side not the (Refer Time: 28:30) we will start the orthogonal contrast.

(Refer Slide Time: 28:33)



So, the 2 are contrast with coefficients cis and dis are orthogonal if if the if you multiply them; obviously, in the values is 0 or for an unbalanced diagram. So, considering that n 1s n 2s n 3s are all different it will be multiplication of n and then then sum it up is equal to 0 so; that means, summation of nici di is equal to 0 ah. So, that is the orthogonal contrast, for a treatment the set of a minus 1 orthogonal contrast partitions the sum the squares due to the treatment into a minus 1. independent single degree of freedom components thus tests performed on orthogonal contrast are independent and you can pass on much better judgment for that.

There are many ways to choose the orthogonal contrast coefficient for a set of treatments usually, sometimes in the nature and the experiment should suggest which comparisons would be of more value with is the orthogonal one or when you are going to basically do the contrast accordingly. So, with this ill end this 11th lecture. And thank you for all of you for attention and we will start on the 12th one and correspondingly discuss more about anova and the design of experience have a nice day.

Thank you.