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

Lecture – 19
Factorial Designs – V

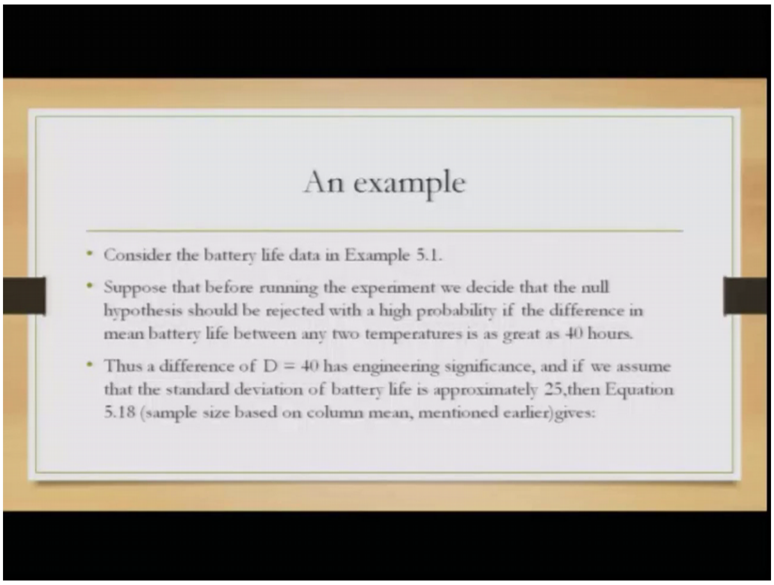
Welcome back my dear friends and students, a very good morning, good afternoon, good evening to all of you. This is the TQM II lecture series under NPTEL MOOC and this is lecture number 19 and we were discussing the factor model design of experiments, how the factors are dependent, how you can build the model what is the total sum of squares, then what is the total sum of errors, what is the total sum of the errors coming from the factors themselves individually, what is the term total sum of squares of the errors coming from the combinations and so on and so forth.

We found out what was and discussed in details, what was the degrees of freedom, based on the total sum of squares and degrees of freedom we found of the mean square errors, based on the means of square errors we found out the ratios, which would give us the f test f values then we perform the f test to either agree or disagree with the hypothesis what I was being framed.

So, let us consider a an example; So, and if you remember that the main two examples which have been discussed in time and again in the last 18th. I would known not say last 18 is basically starting from the 8th to the 18th lecture. Because still the seventh one it was all discussing with the basic statistics probability distribution, then cumulative distribution, then what are the concept of point estimation have this testing alpha beta n or type 1 type 2 error and all these things.

So, and later on we did discuss on the problem the etching and depending on the wattages powers, then we considered very briefly the aluminium smelting then we consider the material versus the temperature for the battery life. So, these were the three main the first and the third was the main focus.

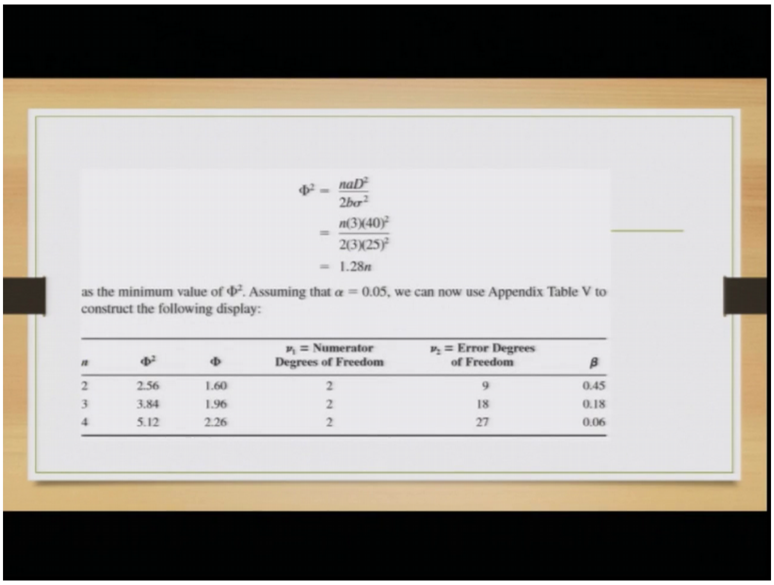
(Refer Slide Time: 02:09)



So, again we consider the battery life example as given example 5.1. Suppose that before running the experiment, we decide that the null hypothesis should be rejected with a high probability, if the difference in the mean battery life between the two temperatures. If you remember the temperature of 15 70 and 125 between any two temperatures is greater is as great as mean life is as great as 40 hours.

So, thus a difference of D which is forty hours is significant and if we assume will be significant, and if you assume that the standard deviation of battery life is approximately about 25. So, then equation which you had considered about the sample size and the concept of the difference coming to the picture, would be given as with the values are as this.

(Refer Slide Time: 02:57)



$$\Phi^2 = \frac{naD^2}{2b\sigma^2} = \frac{n(3)(40)^2}{2(3)(25)^2} = 1.28n$$

as the minimum value of Φ^2. Assuming that $\alpha = 0.05$, we can now use Appendix Table V to construct the following display:

n	Φ^2	Φ	ν_1 = Numerator Degrees of Freedom	ν_2 = Error Degrees of Freedom	β
2	2.56	1.60	2	9	0.45
3	3.84	1.96	2	18	0.18
4	5.12	2.26	2	27	0.06

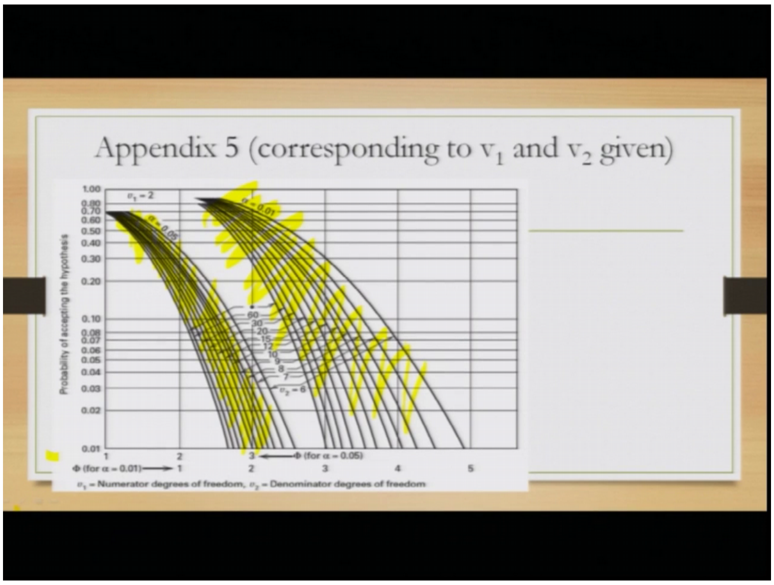
So, capital phi square was if you remember when we are considering the concept of the errors between the effects errors in between the errors coming out from factors b, f fx means factors is the in factor b or coming by the combination. So, if it is coming out from factors a or b.

So, you will find out that in the formula, which ever try highlighting this one. So, the value was n a D square by divided by 2 b sigma square, in the case when you are trying to find out the their the capital phi square for the other factor it will be n d D square divided by 2 a sigma square. Where D was basically the did the differences as we just mentioned.

So, if when we put into the values for this example, it comes out to be 1.28 n. So, as the minimum value of capital phi squared assuming alpha equal to 0.052, we can now use the values on the appendix table 5 to construct the following this space. So, depending up as we change n, the capital phi square would be calculated if the capital phi of. So, square were calculated it will give us the degrees of freedom m m and n depending on where n m and are being counted for the f distribution and we can found all the beta value also.

So, in the first column we have n values, the second column phi square, third column phi then numerator which is the degrees of freedom, the errors of degrees of freedom and the beta values are giving in the subsequent, second, third, fourth fifth and sixth column.

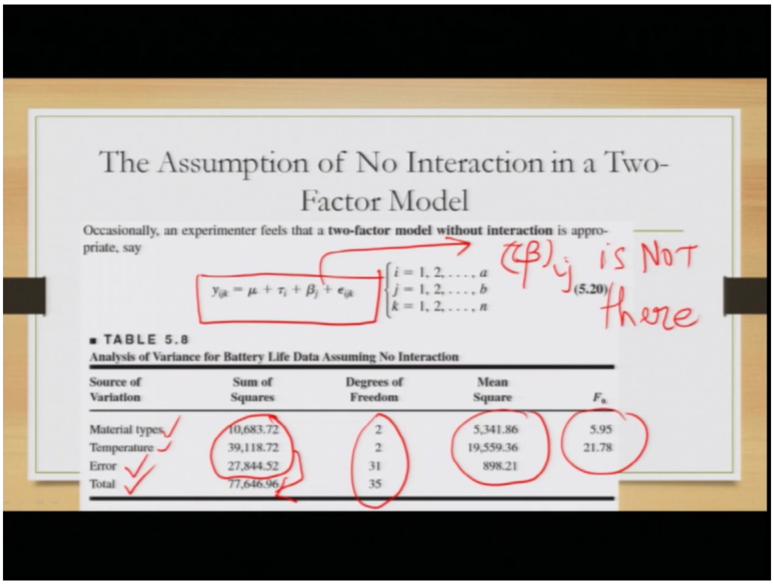
(Refer Slide Time: 04:50)



So, the if we if we find out the corresponding graphs. So, the graphs would be given as this. So, if we have the values calculated. So, we plot basically capital phi for values of alpha and along the y axis you have the probability of accepting the hypothesis. So, it is some it is not exactly an OC curve, but it resembles OC curves in different aspect. So, we have different values of alpha.

So, the first set. So, these and just highlight these are the curves or probability of accepting the hypothesis with respect to the capital phi value for alpha is equal to 0.05 and then I change the color anyway the color did not change. So, this is with the values for alpha is equal to 0.01 and you can find out the values of probability of accepting with respect to capital phi.

(Refer Slide Time: 05:49)

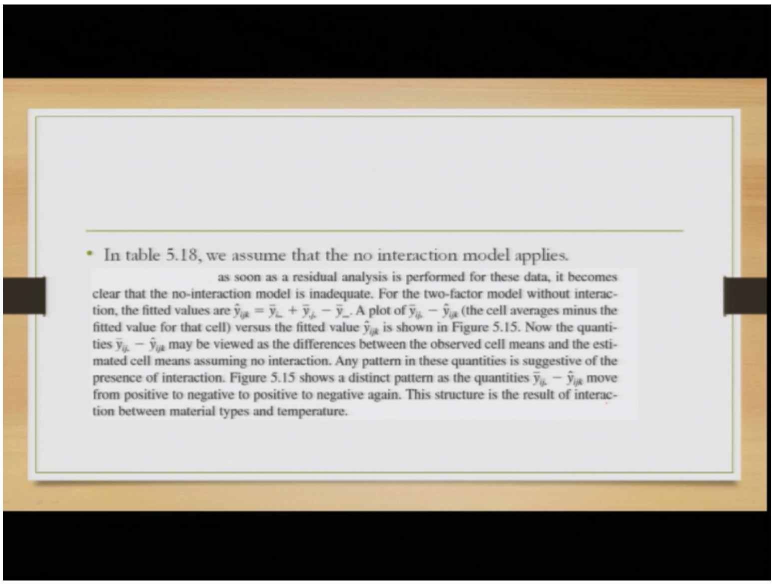


So, assumption of no interaction in a two-actor model; so, we will consider this. So, now, we have considered that if there are no interactions between the factors, it means tau into beta suffix ij would not be there. So, these single factors. So, the effect would be from coming from tau which is for factor a, i is equal to 1 to a and if you are considering the beta value which is beta j, it will be j is equal to 1 to d.

So, occasionally an experimenter feels that the two factor model without interaction is appropriate. So, he will mortal it. So, I will just highlight it; here what is important is tau beta ij is not there. I is equal to 1 to a, j is equal to 1 to b; k is equal to 1 to n. So, analysis of the variance of the battery life assumptions and assuming no interaction; So, there would be material types. So, this is as temperatures effects which is b's.

So, there is no combination of material and temperature. So, the third so, called row is missing, there would be errors there would be total. So, if we add these three values they should add up to the total, degrees of freedom we know the mean square; obviously, will be found out using the summed squares and the degrees of freedom and the f value will be given and you can either reject or accept the null hypothesis.

(Refer Slide Time: 07:23)



So, in the following table we assume that no interactions is there. So, as soon as the residual analysis is performed from this data it becomes clear, that the no interaction model is inadequate because it would not be able to explain all the variations to the at most level of accuracy. For the two factor model without interaction, the fitted values of the curves are. So, you I want to find out the average of the average of the average which is basically mu that, which technically is given by the values of y bar dot dot dot. So, that is one.

Now, I want to basically find out the differences. So, the differences would be technically the error is what? Errors would be the value of the actual y which is y without the hat suffix ijk minus y hat because the estimated value suffix ijk. So, that is the error. So, if I want to basically find out the estimated value. So, this is the estimated value that will be equal to. Now we will take the sum, now the sum would be taken for what?.

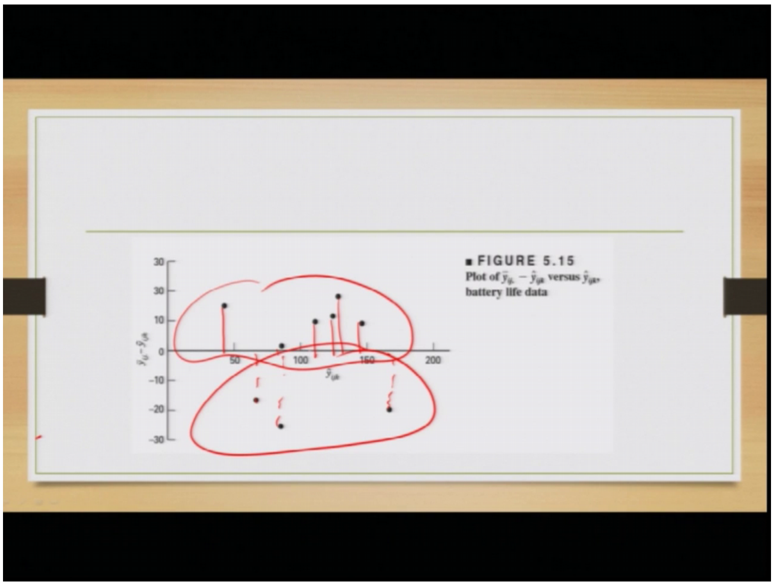
We would take the values of the sum for fixed value of i and all values of j and n, then fixed values of j for all values of i n n, add them up and minus the value of this average of the average of the average, which is given by y bar, which I am just circling I have y bar dot dot dot. So, basically I add up keeping one row and up for all the columns and all ns then keep the column fixed and find out the average and sum it up and find the average for all the rows and all the ns.

So, plots of y bar ij n and n means I am taking for all the values of k which is dot, minus y hat ijk is the cell average minus the fitted value of the versus the fitted value would be shown. Now, the quantities which we found out may be viewed as the difference between the observed cell and the estimated cell measuring may means that assuming no interaction is there.

So, if there is no interaction technically we are assuming the effects of A and B combined to get that is 0. Any pattern in these quantities is suggestive of the presence of interaction. So, if there is no patterns which means they are on a stand on effect affecting the overall reading.

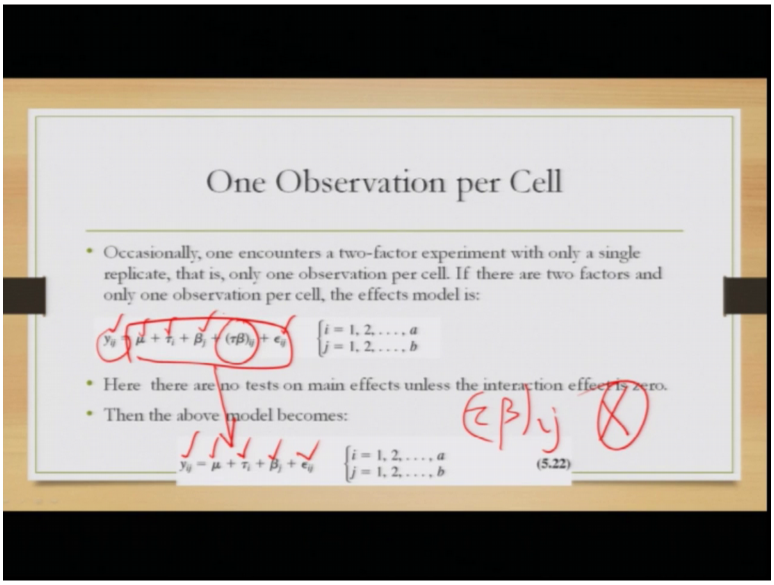
So, figure 5.15 which we will see shows a distinct pattern as the quantities of y bar i j dot minus y hat ijk, which is the estimated value of ijk move from positive to negative and from positive and negative to positive depending on how the assumption and how the variations are. This structure is the result of interaction between the material and the temperature, which or factor a and b which even though are present, we may try to in ignore that, but the results would not basically point to that fact that ignoring would give you the best answer.

(Refer Slide Time: 10:46)



So, if I plot this values of y bar ij dot minus y hat ijk; that means, I am trying to basically take any sell value ij and trying to sum up for all the n values minus the hat this minus this estimated values. So, if this is a change. So, if you see the positive negative. So, this is positive this is positive, this is was positive, positive, positive this is the negative do you see the trend here and see the trend here, it will give you that they as they are moving from positive and negative; obviously, is there some interrelationship happening so; that means, effect of a and b is there.

(Refer Slide Time: 11:30)

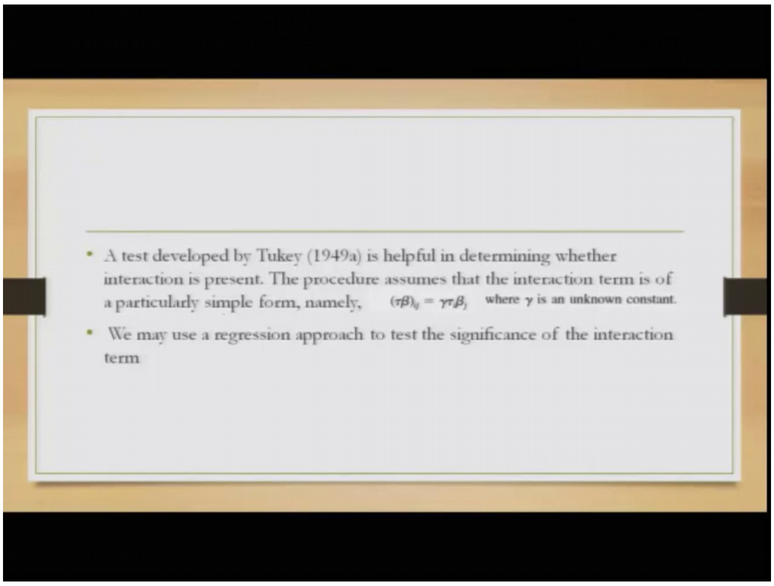


One Observation per Cell; So, occasionally one encounters a two factor model with only a single replicate, that is only one observation per cell is available. If there are factors and only one absorption process cell is available, then the effect model is done like this. So, I will give you the equation. So, here n would not come.

So, this is to be noted; the equation remains the same which is mu which is average of the average of the average minus tau y; obviously, that was the effect happening for only the rows; which is factory a; beta j which is the effect corresponding to the factor b, which is from along the columns and tau beta suffix ij is basically the combinations on the combined factors a and b and; obviously, epsilon ij would be there not epsilon ijk, because those factors are not being considered and k is basically not varying.

Here there are no tests on main effects unless the interaction effect is 0, then the above model when it is put into the factor considering that the interrelationship of the factors is missing which means value of tau ij not there. So, it would not be there so; obviously, if I look this into this equation y ij remains mu remains tau i remains beta j remains error remains what is missing is the interrelationship effect of factor a and b; So, if you can basically extend that for a b c and accordingly, a b c means more than two factors.

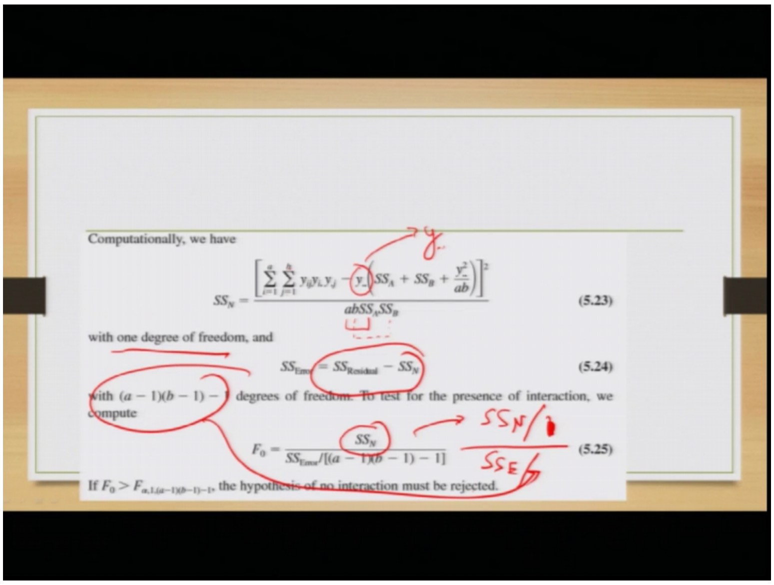
(Refer Slide Time: 13:25)



A test developed by Tukey. So, we remember the Tukey's capital T test is helpful in determining whether interaction is present, the procedure assumes that the interaction tom is of a particular simple form given the factor that it will be equal to. So, what you have interaction is given by tau beta suffix ij.

So, basically I will considers them to be independent of each other in means that, I can multiply tau is beta j is multiplied by gamma; gamma of a very simple constant, where gamma would basically be a unknown rf constant. So, this value will depend on the positive and negative effects of a and b, when they are considered effects are independent in this sense we can just multiply them and to find try to find out the effect. May we may use a regression approach to test the significance on the interaction terms and then do the calculations accordingly.

(Refer Slide Time: 14:24)



Computationally when you find out; So, the sum of the squares corresponding to n; n means capital N is basically you are trying to find out if you remember it is basically equal to a into b into small n. So, that will basically be. So, you will basically have in the do no matter if you consider you have the sum of the squares of a on a standalone basis, some of those case of b on a standalone basis. So, combination is not there so; obviously, they would be multiplied by their corresponding degrees of freedom.

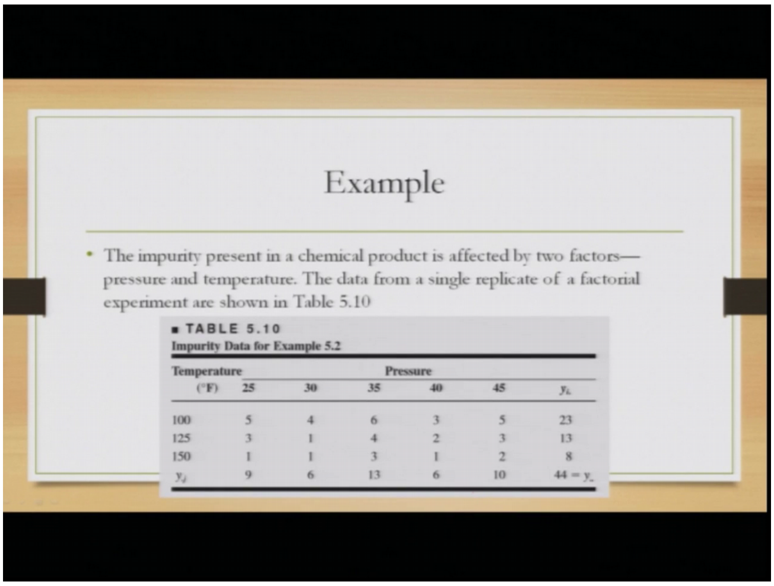
So, SS suffix a would be multiplied by a and this one SS b would be multiplied by its degrees of freedom be small b and in the numerator you will basically have the sum of the squares of a sum of the squares of b plus the y square y square basically the values I am finding out for all the values of is and all the values of j's because here the n value factors or the cells if we consider they were independent so called independent with respect to j I k which is basically from 1 to n.

So, that would be divided by the degrees of freedom a and b, and you will basically multiply it by the sum of the multiplied by the average of the average not average of the average of the averages, 3 times averages is this not being mentioned it is two averages which is y bar bar this dot dot.

So, I am going to find out the sum because this is not the average, I my apologies I should not have used the word average is basically I am trying to find out the sum of the squares multiplying them. And those would be summed up from j is equal 1 to b. So, that is for factor b capital B and i is equal to 1 to a, which is for factor capital A. So, with 1 degree of freedom, we can find out the sum of the square of the errors which is for the residuals minus the n n means that what you find out. With the degrees of freedom for the errors would be a minus b into b minus 1 b a minus 1 into b minus 1 minus 1 because we are considering only 1 cell concept to test for the presence of interaction we find out the f value under the null hypothesis, which is given by the ratio of SS n divided by because SS n would have a degrees of freedom given by 1.

So, technically this is SS and y 1 which is the degrees of freedom as mentioned divided by SS E by its corresponding degrees of freedom; which is coming from here; So, you find out the f value and test whether right or wrong or in the sense that with the accept or reject.

(Refer Slide Time: 17:18)

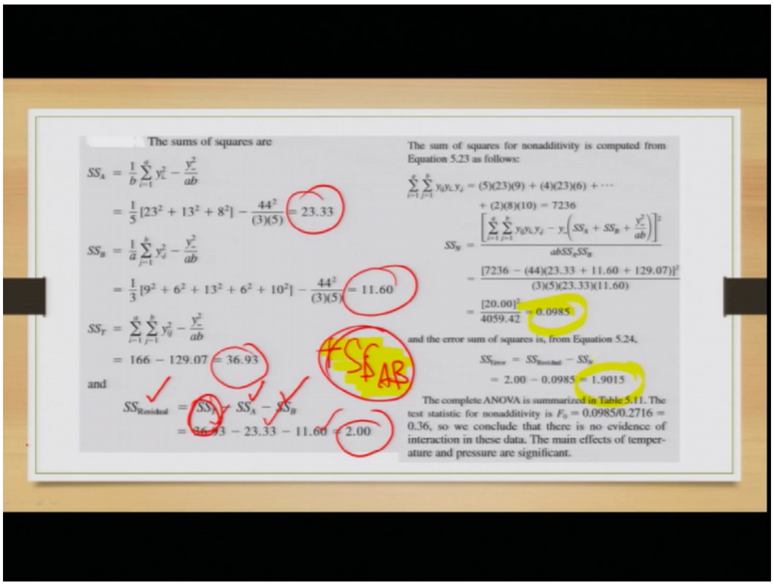


Temperature (°F)	Pressure 25	30	35	40	45	$y_{i.}$
100	5	4	6	3	5	23
125	3	1	4	2	3	13
150	1	1	3	1	2	8
$y_{.j}$	9	6	13	6	10	$44 = y_{..}$

The impurity present in the chemical products is affected by for two factors for the battery example or when you are considering. So, this is pressure and temperature and then the data from a single replicate of the factorial experiment is shown in figure 5.10. So, what you have the temperatures along the first column. So, these are 1 100, 120, 125, 150 and the pressure values are whatever the units are 25, 30, 35, 45 values are given.

The cell values inside the matrix. So, if we consider the second column is 5 3 1 which is basically the pressure corresponding to the pressures of 20 and the temperature values are 100 125, 150. Similarly if I consider the second the third column for pressure 30 the values are 411, similarly for 35 pressure the values are 643 pressure value of 40, the corresponding values are 321 and similarly for the last pressure 45, the impurity data correspondingly for temperatures 100 to 125 and 150 or 532, and then you can find out the y i dot which is keep I fixed sum it up for all the j's and now the one will be y dot j which we keep j a fixed sum it up for all i’s.

(Refer Slide Time: 18:48)



The sum of the squares can be found out they come out to be 23.33.

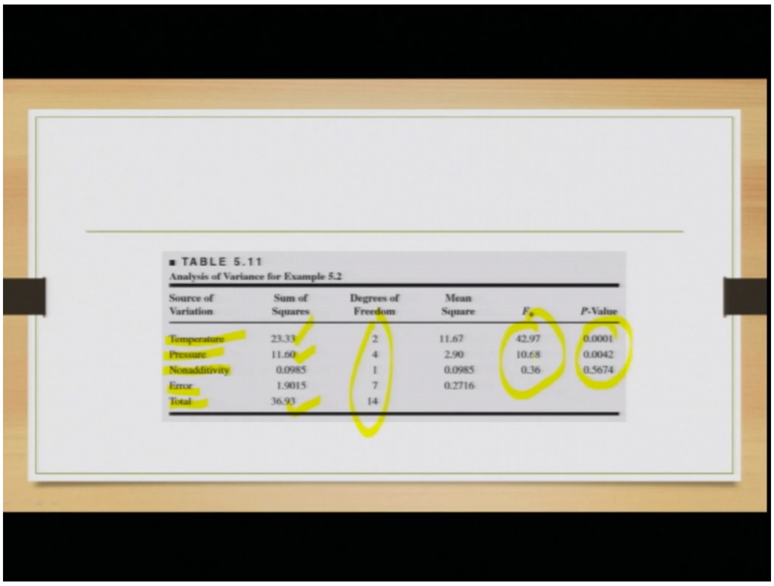
So, this is the value, the sum of the b values who comes out to be 11.60 from here just the simple calculation, which we have done. The sum of the total square basically comes out to be 36.93 and the residual values would be coming out. So, this is sum of the total square 36.93 minus 23.33 and 11.6. So, this is 2.

So, here if you see the residuals technically, the total sum would be equal to SS suffix A, SS suffix B is SS errors or residuals plus; obviously, your term would have been there or if you have considered SS A B, but that is not there if we remember. So, this is not here. So, that is why the calculations are simple the sum of the squares for the non negativity is computed from the equation which you have already considered.

So, that was value comes out to be 0.0985 and the errors comes out to be once we find out some of the squares errors comes out to be 1.9015. So, to complete the ANOVA we perform the test the f naught value comes out to be 0.36.

So, we conclude there is no evidence of interaction in this data the main effect the temperature and pressure are significant.

(Refer Slide Time: 20:20)



■ TABLE 5.11

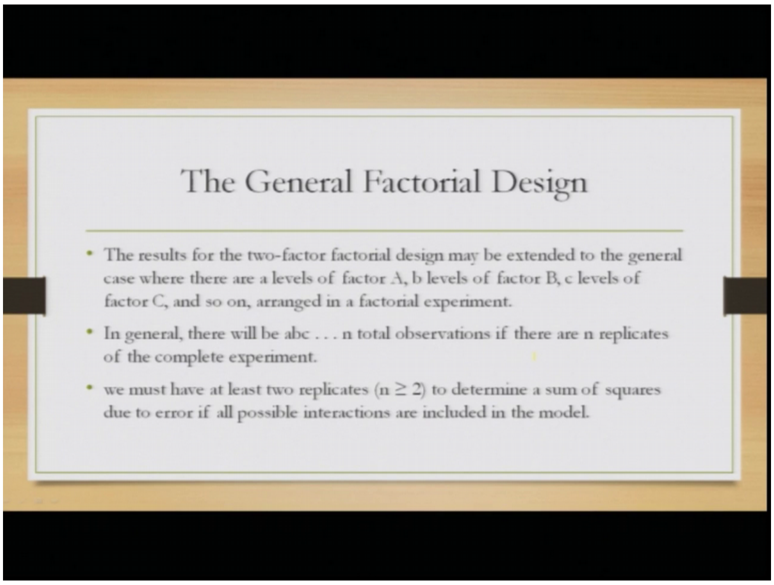
Analysis of Variance for Example 5.2

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Temperature	23.33	2	11.67	42.97	0.0001
Pressure	11.60	4	2.90	10.68	0.0042
Nonadditivity	0.0985	1	0.0985	0.36	0.5674
Error	1.9015	7	0.2716		
Total	36.93	14			

So, this is given summarized in table 5.11, temperature is there pressure is there, they are non negativity concept errors are there total. So, I find out 23.33 has already noted 11.06 and 36.93 degrees of freedom we know. Mean squares can be found out by the sum of the squares divided by the degrees of freedom.

So, you find out the f values which is the ratios of the of the sum of squares and divided by is there and there are my degrees of freedom from there we can find out the p value

(Refer Slide Time: 20:57)



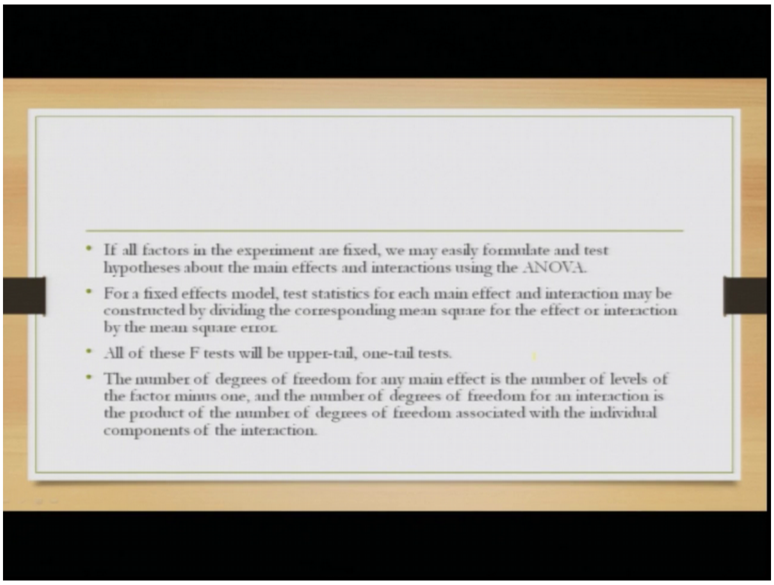
Now, in consider the general factorial design, the decisions for the 2 factor factorial design will be extended to the general case when there are levels of factors as I said A B C so on and so forth.

So, A would be of values from i is equal to 1, i consider i suffix 1 because there would be many such nomenclatures, i suffix 1 are the is the nomenclature corresponding to the values of factor capital A, with an i suffix 1 changes from 1 to small a. Similarly you will have for factor B the suffix given by i suffix 2, the value changes from 1 to small b; then you will have basically have the factor C the nomenclature is i suffix 3 small 3, i suffix 3 and the values of i suffix 3 changes from 1 to small c. So, it can be done accordingly.

In general they would be corresponding to this a then b then c then d efg h. So, and. So, forth the factors and; obviously, there would be a total number of observation we needs with sample size in each factors would be small n; so, when if you want to find out the total number of the observation for all the factors and the samples would be into b into c so and so, forth work till the last value into n small n.

So, n there small ns on the replicates we must at least have two replicates to determine the same the squares due to error if all possible interactions are included in the model.

(Refer Slide Time: 22:39)

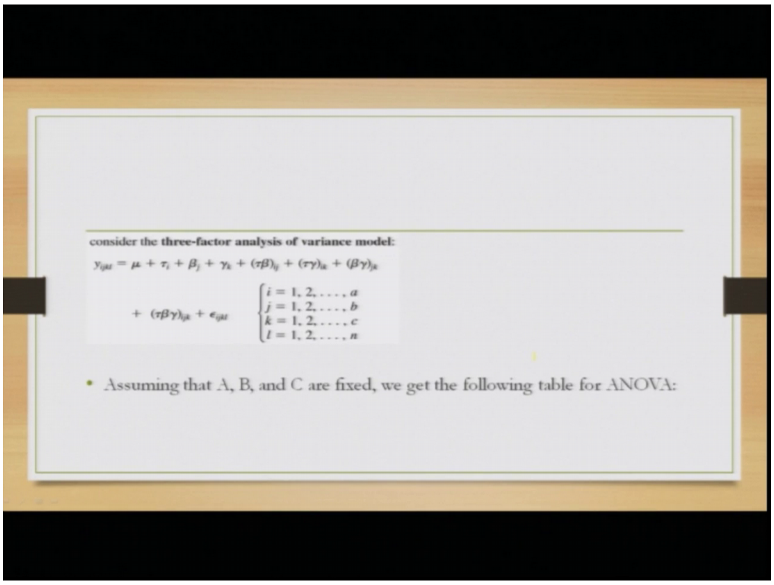


If all factors in the experiments are fixed, we may easily formulate and test the hypothesis about the main effects and direction using the ANOVA model, for a fixed effect model test statistics for each main effect and internationally be constructed by dividing the corresponding mean square for the effect on or interaction by the mean square errors.

So, all that these F test would be basically of the upper tail or one tailed test either greater or smaller depending on whatever, but not in between. So, it will depend on the problem formulation which you have done. The number of degrees of freedom for any main effect is the number of the levels of factor minus 1, because it a minus 1, b minus 1, c minus 1, d minus 1 so on and so, forth and you can find out the degrees of freedom for the errors and for the summit will be for the total, it will be sum of them.

And the number of degrees of freedom for any and interaction is the product of the sum of the degrees of freedom associated with the individual component of the interaction and we can do the calculations accordingly.

(Refer Slide Time: 23:44)



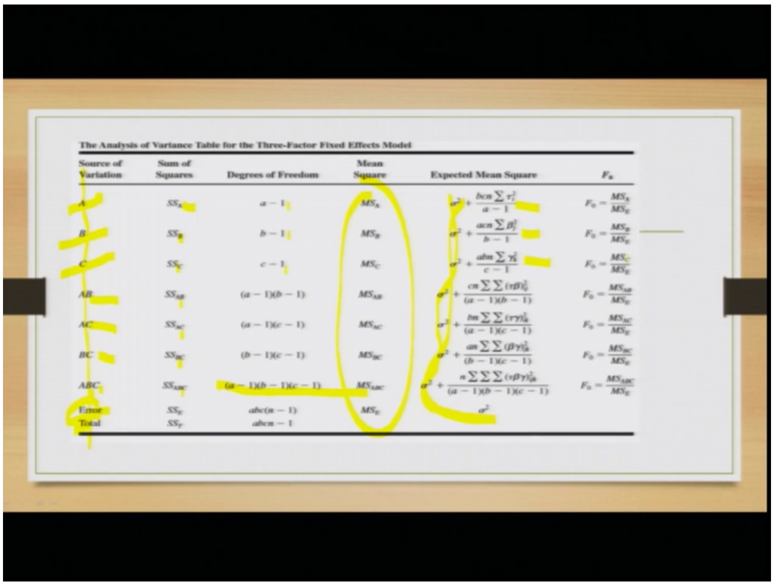
Consider the three factor model, which I have already discussed I have been mentioning them, but I will still explain it on both detail. Now, consider there are the average; So, what you have is basically y suffix ijkl. So, i is basically from 1 to a, j is equal to 1 to b, k is equal to from 1 to c, through I abc are all smalls and l is l is equal to from 1 to small n n.

Now, on the right hand side, it will be mu, mu is basically now listen to me carefully average on the average of the average of the average; that means, I am taking take a average four times, because I am summing it up for i's, for j's, for ks for ls. So, that would technically be y hat which is the predicted value, I actually it should be and you can find it the errors accordingly.

The next terms are individual factors of a which is tau y, individual factor for b which is beta j, in we need factor for c which is basically gamma k and obviously, you would not have any l because that is corresponding to the nomenclature for n.

Now, when I go to the next stage of combinations, combinations will be tau beta ij tau gamma i k, beta gamma jk. So, if we highlight tau beta ij, tau gamma j i k, beta gamma j k and; obviously, if I go to the next level which is 3 will be tau beta gamma suffix i jk. So, this is 1 and obviously, they would be rather epsilon ijk l. So, i values basically are changing from 1 to a, j values are changing from 1 to b, k values are changing from 1 to c and l values are changing from 1 to n. So, assuming that ABC are fixed we get the following tables in the ANOVA

(Refer Slide Time: 25:57)



Now, this looks complicated, but is very simple. In the first column we have all the sources of variations A B C individual, AB AC BC combined two at a time ABC, for error term and the total. So; obviously, sum of all of them would be equal to the total now how do I represent the sum of the squares is very simple, SS suffix.

Now, note down the suffix SS suffix a for a, SS suffix b for b, SS if I am just highlighting the suffixes; SS suffix c for c SS suffix a b which is for the combination a b, SS suffix come a c which is the combination abc SS suffix bc, which is the combination of bc and SS suffix abc which is the combination of abc three of the errors you have the total.

So, degrees of freedom would be a minus a 1, b minus 1, c minus 1 and the and the degrees of freedom when I consider them as a combined state would be for 2 would be a minus 1 into b minus 1 for the first one, b minus 1 into c minus 1 for the second one, b minus 1 into c minus 1 for the third one if I go to the 3 combinations it will be a minus 1 into b minus 1 into c minus 1 as noted.

If I want to go for the errors, errors would be basically ABC, we are already done in the initial case it was A into B back and those was n minus 1 it is ABC into n minus 1 and the total sum of the degrees of freedom would be ABC n minus 1. So, the mean squares again a ms with the suffix would give you the mean square for those factors and the mean squares if I find out for all of them it will be sigma square will be the common part c, where I am moving my highlighter.

And the corresponding addition would basically correspond to the fact which is AB or A or B or C or D C or A C or ABC. So, we can find it out and you can find out the F values. So, with this I will end this 19th lecture I continue more discussion about this right and later on.

Thank you very much. Have a nice day.