Total Quality Management - II

Prof. Raghunandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

Lecture - 17
Factorial Designs – II

Welcome back my dear friends and students, and this is the 17th lecture for the TQM II under the NPTEL MOOC. And I am Raghunandan Sengupta from the IME department IIT Kanpur a very good morning, good afternoon, good evening to all of you.

Now, if you remember in the 15th fag end of the 15th lecture and starting on the 16th lecture we were discussing about factorial designs in the sense, the relationship between factors to the levels of which their dependence structure is and based on that we will basically have a dependence model, now the word dependence I am using in the very general sense. Dependence model such that we are able to predict or find out the average relationship average.

Because why I am using the average please give me a few minutes I will come back to that, you want to find the average relationship between the independent variables and the dependent variables. Now, if they are related the independent variables are related and if there are two factors; obviously, you will have the effect of x 1, effect of x 2 taken separately and then combined x 1 and x 2 and if there are more than 3; obviously, it will base increase proportionally like first individually all of them then take in two at a time then take in three at a time and so on and so forth.

Now, why use the work concept of average is that if you remember, there is an error term which is very important and we assume the error term has a has a normal distribution with a zero mean and a standard deviation which is fixed, it can be it can be 1 it can be sigma square and the fact that is fixed that it means that it does not depend on time.

So, those things as I as I took your kind permission, I will be repeating it time and again to make it much clearer to you. And then we also consider later on in the 16th lecture that depending on the factors the interrelationship if the efficiencies concept would also be considered.

And how we can find out if there are linear relationship back on a three dimensional scale consider you are in a room and the vertical axis would be basically measuring the y values and the horizontal axis is which are along the breadth and the length of the room would basically be x 1 and x 2 considering there are two different variables of x 1 and x 2; and the linear relationship would be given by the planes which are cutting in what direction they are that would depend on what are the values of so, called beta 1, beta 2, beta 1 2 and all these.

And beta naught; obviously, as I mentioned would be the conceptual values of c which we have when you are trying to basically I would draw the curve y is equal to mx plus c. And if you if you see the plan from the top and considering my values are not changing, then you can find out the relationship between x 1 and x 2 if they are only 2 variables. And if they are none the higher dimensions considering x 1 x 2 x 3 are there you will have hyper planes.

So, with this we will start with the 7th lecture. So, now, if you remember you are trying to basically study the concepts of interrelationship of the example. I am coming back to the example the discussion further on, they were batteries and batteries operated at different temperatures the temperatures were such that you want to find out what is the effect the temperature difference has on the battery operation. And we found out the relationship between, the temperatures and the different levels of in independent structures which were there.

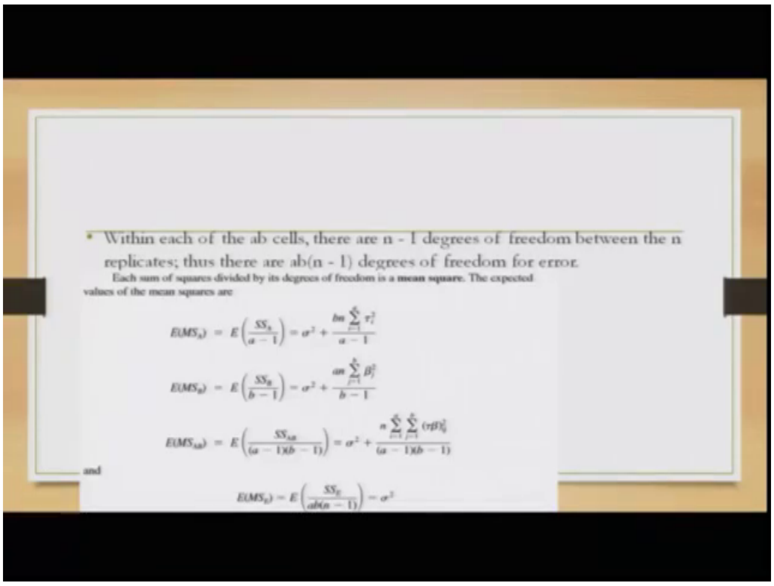
And in order to make life simple for us I did mention, that they would be three variables or nomenclature of trying to basically specify the variables one was i, one was j one was k, i changed from i 1 to small a which was factor for factor a, j tends from 1 to small b which was for factor b and k was basically the number of samples which you had for each group which was basically form 1 to small n.

Hence, the total number of observation which you had was small a into small b into small n. And we also found out that what was basically the averages of the averages of the averages. So, on I am using the word average is three times, which means that I am really trying to find out the sum for all the a s for all is sorry for all the js and all the k s and then; obviously, you will have different nomenclature to denote y bar dot j dot; that means, you took the averages for i s and the k s keeping j fixed.

Then you also had the average of y bar i dot dot; that means, you took the averages for all the j s and all the ks keeping i s fixed and so on and so forth. So, if it was average of the average of the average which was technically the population was value was me mean mu then; obviously, you had y bar dot dot dot such that you basically dot dot dot I am using for the suffixes you took the averages, for all the three the three ranges of i j and k.

So, considering that we found out that how we can find out the total errors divided into three different three different what I should use; the word three different a if concepts of influences, which means there was some total sum of those or the square of the errors, they would be interrelationship inter and intra in order to make in the things understands. One would be along the is when we along the js and one would be the corresponding the relationship between i and the j.

(Refer Slide Time: 06:06)



Now, once you find out the total errors; obviously, the your next task would be to find out the degrees of freedom. And the degrees of freedom, was mentioned very categorically. So, in one case it would be a minus 1 another case would be minus 1 this a and b are small a s and b s in another case you would be a in d to b small a into b in the bracket small n minus 1. So, with this we want to basically go into the concept of finding of the average error. So, this is what we are going to discuss now in the slide.

So, within each of the a into b cells, there are n minus 1 degrees of freedom between the n and replicate. So, basically he had m replicas on n sample size which was there. So, thus there are a b into bracket n minus 1 degrees of freedom for the errors so, which I did mention that few minutes back. So, each sum of the squares divided by its degree of freedom is basically the mean square, the expected value in the mean squares are to be found out because that will give an average error.

So, those values are as given here. So, I will try to highlight it using different highlighting color in order to make the differentiation much more clear and start for you. So, this is the. So, let me mark it. So, this one is the expected value of the mean square of the errors corresponding to factor a because the suffix a is there. So, you had s s suffix a was the sum of the square of the errors corresponding to factor a, and what was the degrees of freedom degrees of freedom was small a minus 1.

Here you find small a minus 1. So, that overall error comes out to be sigma square if you remember was sigma the errors variance which was constant, and if you remember the tau where I am hovering my this pointer is tau i s was corresponding to the deviations with from the actual average for each and every I values or can we can have tau 1, tau 2, tau 3 till tau a. So, that was tau i square summed up for all the values from i is equal to 1 to a divided by; obviously, degrees of freedom was a minus 1 and multiplied by b into n.

Because they would basically have such values for b number of readings for factor b capital B and each being each of the combinations of cells of a and b. Cells means the combination of a and b factors which were there, you would basically have such n small n number of replicates. Similarly when I try to find out the mean square for the case for factor b so, that is given by m s mean square suffix b. So, suffix would basically make things clear to you.

And if I want to find out the average of that, that it would be the some of the squares corresponding to be that is why it is the suffix b is here where I am trying to basically highlight now divided by the degrees of freedom, which is small b minus a. Again this will come out to be sigma square plus now there is a very nice relationship, which is coming out in the other case for the small deviations or so, called dispersion.

Again dispersion being used in a general sense with respect to the mean value mu, for the case when you are trying to find out the effect of the factors capital A, you had tau i s; i is equal to 1 to a.

Now, when you are trying to basically so, that was along the rows. So, as was along the rows when I am considering along the columns which is capital B starting from j is equal to 1 to small b then; obviously, the so, called dispersion concept was brought up by the comb by using the variable of the symbol beta suffix j. So, you will basically find out beta suffix j square, square is because you are trying to I did not mention the squares over the taus of the squares over the betas.

Basically they are and they are considering that your assuming sum of the squares; that means, you are trying to find out the deviations squaring it up and basically semantics up summing it up as you do for the variances part. So, you had a in to n for all the other replicates. So, for the factor a and for the n number of replicas which you have, and some multiplied by summation of beta j squares for j is equal to 1 to small b and divided by this degrees of freedom, which as we all know is small b minus 1.

When I want to find out the interrelationship errors so, it means factors of a and b combine, the actual value comes out to be again very simple logic is coming out again is I will use a different color let me use the use green. So, this would be e m m s which means square suffix a b for the combinations, and here SS AB the sum of the squares suffix a b divided by what is the degrees of freedom that would be along the rows if you consider they would be a minus 1, because you are losing 1 degrees of freedom for that and along the columns would be small b minus 1.

So, the degrees of freedom for both of them would and equal to find out the so, called expected value would be for the case when you divide by the term a minus 1 into b minus 1 as it is pointed out where I am hovering my pen, and the overall m mean squared average values basically becomes sigma square is already there, now do you have to add some terms in this case it would be n. Because n replicas are there multiplied by a double summation of i is equal to 1 to small a and j is going to 1 to small b.

Now, if you remember the overall dispersion of the errors which are coming out due to combination of the factors was given by the term, tau into b suffix ij. So, this is and; obviously, it will be squared because you are trying to find out some of the squares. So, that would be tau b suffix i j square. And obviously, the division would always remain the same. Now in this case it is a minus 1 into b minus 1 and finally, the mean square of the errors would basically be sigma square as we have found out in the in the general case we can prove it very simply.

(Refer Slide Time: 12:23)

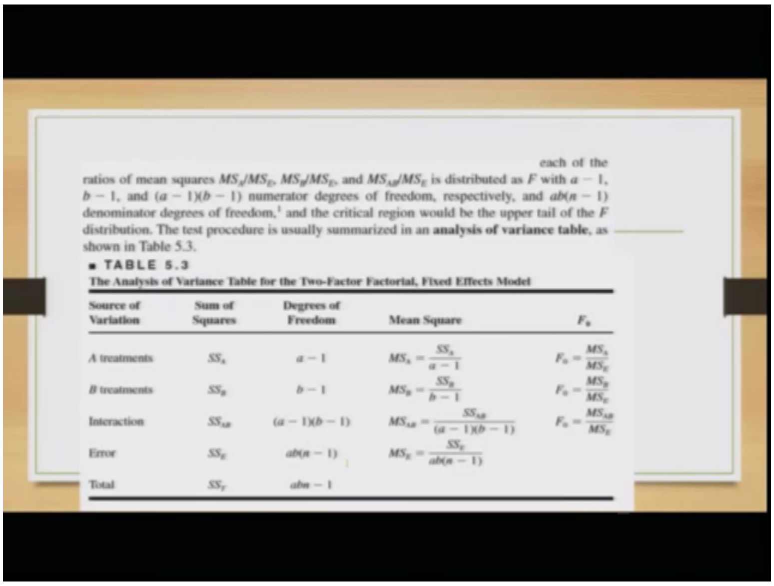

each of the ratios of mean squares MS_A/MS_E, MS_B/MS_E, and MS_{AB}/MS_E is distributed as F with $a-1$, $b-1$, and $(a-1)(b-1)$ numerator degrees of freedom, respectively, and $ab(n-1)$ denominator degrees of freedom,[1] and the critical region would be the upper tail of the F distribution. The test procedure is usually summarized in an **analysis of variance table**, as shown in Table 5.3.

TABLE 5.3
The Analysis of Variance Table for the Two-Factor Factorial, Fixed Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A treatments	SS_A	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$F_0 = \frac{MS_A}{MS_E}$
B treatments	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$F_0 = \frac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$MS_E = \frac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn-1$		

Now, what we want need to find out basically the ratios on the mean squares would give me the values of the distribution. So, what are the ratios of the mean square? Would be mean square for factor a divided by mean square of the errors, another would be basically the mean square of factor b divided by mean square of the errors, errors and another would be the mean the mean square of the combination of the errors of a and b divided by the mean square of the errors.

So, they would be distributed F as F with degrees of freedom corresponding to the fact respectively they are small a minus 1, small b minus 1 and a minus 1 into b minus 1. So, this is what we just saw in the numerators. So, these values would be there in the numerator to give you the concept that what is the degrees of freedom and if I find out the degrees of freedom corresponding to the terms, which is basically the error.

So, that error term would basically be coming out with let us check. Yes, the error term for the degrees of freedom for the error would basically be fair if you remember I did mention small a into small b in the bracket n minus 1. So, the test procedures is usually summarized in analysis of variance chart, as shown in table 5.3. So, what are those I will just explain the sources of variance variations due to the factors which are coming out would be given in the first column.

So, here you would basically have due to factor a or treatment a; first due to then the second 1 would be due to factor b which is the second; sec third would be the interaction that is a into b fourth one was basically the errors and the fifth one would basically with the sum. Now, if I change the scenario I will come to the second third fourth columns respectively later on.

But let me expand the thought process and try to find out if there are three factors. So, one would basically be the overall effect due to treatment a consider a factor a. Next would be I am going along the columns for the hypothetical a thought out example which I am considering; second would be for factor b third would be for factor c, then we will come into the combinations of what are these? Factor a into b factor a into c factor b into c and the last one for those combination would be the factors abc combined together.

And; obviously, the second last one would basically be there for the errors and if you sum up all the errors, it will be the total sum of squares for the whole the problem, which you are discussing. Now, they would basically have the sum of this errors, which will be for this hypothetical thought out example would be denoted by. Now I will main mention SS and with the suffixes it will be s s suffix a, SS suffix b, s s suffix c, SS suffix a b, SS suffix ac s s suffix bc, s s suffix the errors and SS suffix t which is the total.

Now, if you consider the total degrees of freedom, it would be found out accordingly. So, in in the cases if there are abc small a, small b and small c they will be found out accordingly which would be basically a minus 1, b minus 1, c minus 1. Then for the 2 combinations it will be a minus 1 into b minus 1 for a and b a minus 1 into c minus 1 for a into c, and b minus 1 into c minus 1 for b into c.

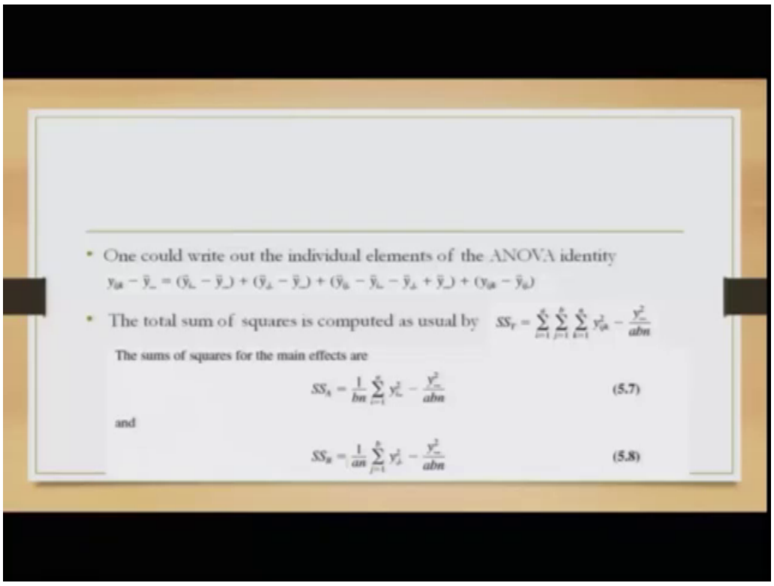
And the last one when you are trying to combine, the all of them would basically be given by the interaction would be a minus 1 into b minus 1 into c minus 1 and if you I consider the total the degrees of freedom would be a into b into c into n minus 1 and obviously, you can find out the errors. According to the degrees of freedom accordingly; Now, if you come to the mean square. So, the values which I just mentioned, was for third order example if there are three factors.

So, now when I am repeating it will make sense if you look at the slide which is they are being shown named when I want to find out the mean square; obviously, it would be the sum of those squares divided by the degree of freedom degrees of freedom as I mentioned for the three factors would be for first three would be a minus 1, b minus 1, c minus 1 and for the combinations you can find out and proceed accordingly.

Similarly, you find out the F values; F values corresponding to factor a would be m s a suffix a divided by m se, then for the second case would ms m suffix b divided by m x suffix e for the third case it will be m s suffix c divided by m s suffix e, and for the other two other three combinations taken two at a time would be m s suffix a b in one case.

In the second case it will be a combination of a c and the third case it will be bc, and division would always remain as the case which I mentioned would be the mean square of the errors and similarly you can find out the f values for the combinations, and the errors and basically proceed accordingly.

(Refer Slide Time: 17:44)



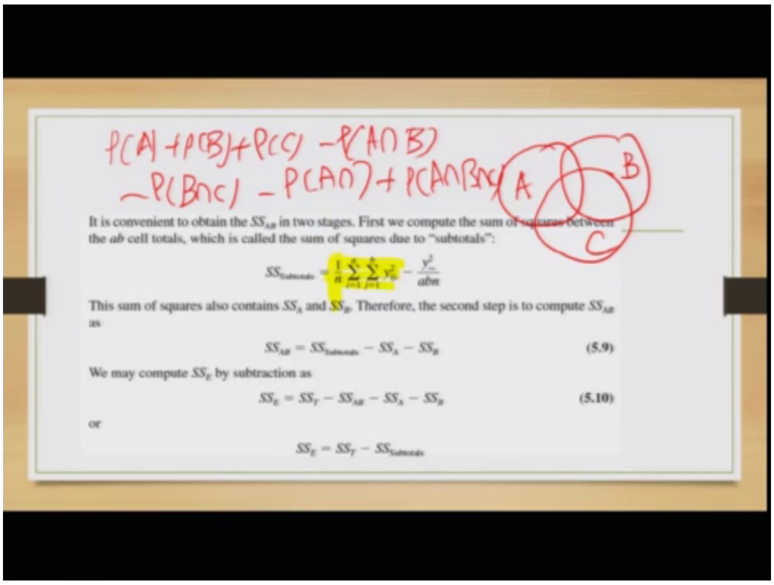
So, one can write if you remember the ANOVA actual identity model as. So, now, there would be differences of each individual cell values with the average of the average of the average for this case. So, it will be y suffix ijk minus y bar dot dot dot because you are summing up for all is all j s and all k s, and that would be the combination of the differences of the errors in this front number 1. I will highlight and mention, the first one is basically when you are summing up for all the j s and ks keeping is as fixed this.

So, that would be the first term minus y bar dot dot dot, third one would be second one would basically be when you are trying to find out the sum off for all the is and all the ks keeping j s fixed; So, which is here. So, I will use a different color to make it interesting and much. So, this one, the next two terms if I consider would basically be I will come to the last term first with a different color. So, it would be again y ij k for each cell minus the sum up for all the n values keeping i and j inter changing from 1 to a and j changing from 1 to b small b.

And the third value this would basically be the differences of the sums taken for ij dot; that means, summing up for all ns and y bar dot dot dot that is for all the summed up for i j and k and the values, which we subtracted would be those combinations which you are already taking on the left hand side that is for the red and the orange; that means, you are finding out the values u d using the plus values for what those are y i dot dot second one is y bar is; obviously, there y bar dot j dot. So, they have to be negated. So, those values would be coming with a negative sign in order to balance the whole equation.

The total sum of areas would be very simply find out found out by finding out the square of the errors. So, that would basically n comma out to be y square dot dot dot divided by a b and n because that is the total number of observations. So, the sum of the squares for the main effect can be found out from a and b and or it can be extended to a b c and you can do the calculations accordingly.

(Refer Slide Time: 20:20)



It is convenient to obtain the sum of the squares are a b in two stage, first we compute the sum squares between the abc is the a b c s in total, which is called call the sum of the squares due to the subtotals and then basically go into finding out the averages of the subtotals talk totally. So, the subtotals would be if you pay attention to the formula, these things would become very clear.

So, what is the formula? Again I will highlight. So, with the blue one or let me change the column because become too dark. So, this is one. So, that will be y square ij dot; that means, you are summing up for all ms; obviously, in the denominator it will be divided by n as it is here which I am just highlighting again with the same highlight a color, but other to make it understand.

And the summations as you know would be basically be fought all the values of keeping i and j so; obviously, it will change for each and i and j. So, it will be summed up form I is equal to 1 to a and j is equal to 1 to b this sum of the squares also contains the SS a and SS b; that means, the SS AB is being basic taking as in a much more broader sense to 8 if you look, deep into the sum of the squares you will have basically the values of a and b.

Therefore the second step it is computing that is to compute SS AB as now what are these? SS AB would basically be the sum of that subtotals minus. So, it is like this I will try to explain. So, when you have to consider this Venn diagram and you want to find out the probability for. So, there are events abc. So, this abc and the abc which is being mentioned for the factorial design problems are different.

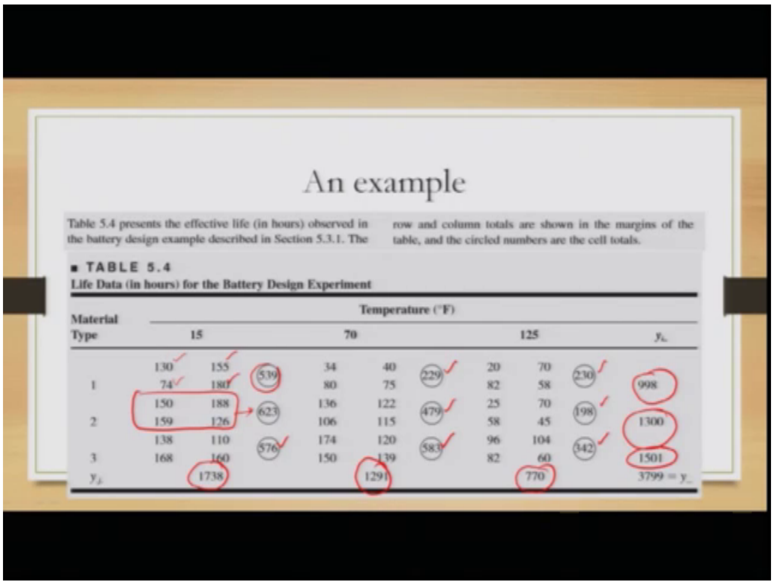
So, they are just events abc which I have drawn. So, when I want to find out the total probability what I do is then? I find out the probability of A plus probability of B plus probability of C minus the intersection of A and B minus the intersection of B and C minus the intersection of A and C then as you have subtracted so; obviously, you will add up A intersection B intersection C. So, this if you consider there in the problem there you will find out.

So, sum of the squares of A B would be subtotals, which is which is already considered the A and B one. So, will basically divide or minus the values of sum of squares of A separately sum of squares of B separately and find out the sum of the squares for A and B taking together. We may compute the sum of the squares for the errors by subtracting this. So, the sum of the square of what is the errors? Errors is basically the difference of the total minus the individual effects of this errors, which is the sum of the squares which is coming up from all the cases.

So, what are these? One would be for a separately no effect of b or no effect of a become mine 1 would be the effect of b for no effect of a and no effect of a b combined and last would not be the effect of a b combined with no effect of a or no effect of b taken individually. So, this would basically the formula. So, you can find out. So, if I basically find out. So, the sum of the errors would be the total errors minus the SS subtotals.

Now, if I go one level higher, abc you will find out in that case the subtotals can be considered in a similar way where they would be effects on abc and then you have to basically subtract the effect of a b a c and b c and go in the same logical sequence as we just discussed.

(Refer Slide Time: 24:23)



Material Type	Temperature (°F) 15			70			125			$y_{i.}$
1	130	155	539	34	40	229	20	70	230	998
	74	180		80	75		82	58		
2	150	188	623	136	122	479	25	70	198	1300
	159	126		106	115		58	45		
3	138	110	576	174	120	583	96	104	342	1501
	168	160		150	139		82	60		
$y_{.j.}$		1738			1291			770		$3799 = y_{...}$

So, I have considered the example. So, this a 5.4 table presents the affected live in hours observed in the battery it is an example, the row and the columns are shown in the end the margins of the tables and the circle numbers are the cell total.

So, basically a material type a as are given and these are the temperatures if you remember in family it was 50 15, 70 and 125. So, for the material type 1 for 2 and 3 the values are given as 130 155. So, this I am reading and just mark it with let me use the red one 130 155 seventy four and 180.

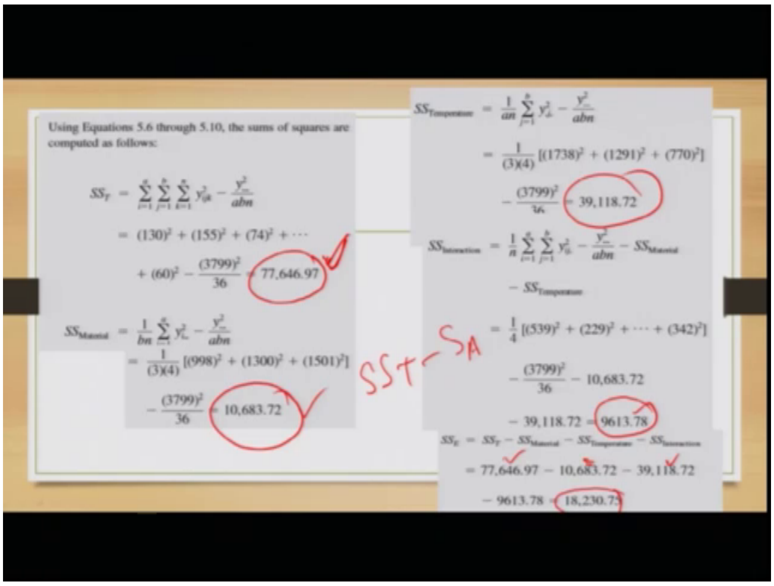
So, if you added up 5 plus 49. So, e is 9, 5 plus 8 13, 20 23 32, 2 1 3 4 and 5. So, this is 539 which you have the total if I add up for the material type 2 the add total comes out to

623 if I have tried to find out for material 3 comes out to be 576. So, and then the total value is 1738 for temperature being fixed which is 15 and for all the 3 different materials.

Similarly, of 228 479 583 for temperature of 70 total value for all the combinations is 1291 and for the last one to dig the last change in the temperature which is 125, the values for from them a total sump sample size for material type 1 2 3 are respective 230 198 342 and the total value is 780.

So, if I find out the sum along the margins which is the last column. So, this 998 means the combination or the sum of addition of all the samples corresponding to a fixed value of material type and for all the three different temperatures, then you have 1300 which is basically corresponding to type 2 material for all the 3 temperatures and 1501 if you consider is basically for all the three temperatures for type three type of material and; obviously, you have the average as well is already written down.

(Refer Slide Time: 26:43)



So, using the equations one can find out the sum square of the totals comes out to be 77646.97 if I find out that of the material type, which is basically the so, called factors which have considered. So, that comes out to be 10000 683.72 for the temperature which is the second factor.

So, if you remember a and b. So, here consider a as the material b as the temperature. So, SS b which is for the temperature comes out to be 39118.72 and if I find out the interaction which between the temperature in the material the average comes out to be 9613.78. So, if you basically separately add them this values of the sum of those and in the sum of the square errors has also to be found out.

So, the sum of the squares of the errors would technically be sum of total, which is this which is here minus S of A which is only for factor a which is given here as 390 ok. So, sorry my mistake this 77.646 is here which is sum total 10683.72 is basically for factor a 39 39118.72 is for factor b.

So, and the value of the combination is 9613.78 which is for a and b combined together which is here and if I find out the sum of the square of the errors comes out to be 18230.75. So, with this you can basically proceed for the next level find out the mean squares find out the F values and do the tests accordingly.

So, with this I will end the seventeenth lecture and continue mode for the discussion for the 18th and so on and so forth. Have a nice day.

Thank you very much for your attention.