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

Lecture - 16
Factorial Designs – II

Welcome back my dear friends, a very good morning, good afternoon and good evening to all of you. And this is the TQM II lecture series under NPTEL MOOC and we have already completed three weeks of lectures and this is the 16th lecture which is the first one in the fourth week. And by the time when use you are seeing this lecture, you would have taken three assignments and I am sure they have been very useful to you considering the cost concept we are covered.

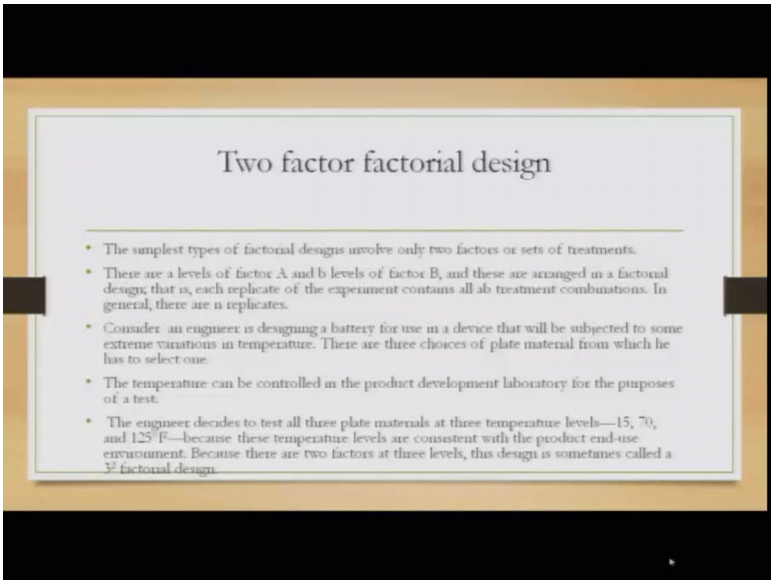
And we will slowly go into more depth; obviously, no theoretical concepts, but more of a conceptual concept how you can utilize this factor models and a later part also different concepts for design of experiments and the TQM 2. And as you know I am Raghunandan Sengupta from the IME department IIT Kanpur. So, with this we will start the 16th lecture.

So, if you remember that, we were trying to find out the efficiencies in the last in the just last few minutes of the class on the fifteenth one, where you considered our two diagrams one set of diagrams was basically the combination for the single for the linear model, how y would change with respect to x 1 and x 2 and considering you had a plane in the two dimension and if we look from the top with y being fixed, you can find out the combinations of x 1 and x 2.

And later on we consider that if you find out the combinations of the factors with respect to the efficiency for the single and the combined model, the efficiency would increase and if you we saw for two factors the efficiency one point five and increased linearly accordingly.

So, this is only true for the linear model; obviously, for non-linear models or more than three factors; obviously, the ideas would be much different. So, continuing our discussion for two factor factorial design models.

(Refer Slide Time: 02:04)



The simplest type of factorial design involves only two factors or set of treatments. Treatment again the word is being used for the, the levels of differences of the effects which you have. So, consider those effects again going back to the example of the etching problem so, those who are basically the levels of wattages of the power.

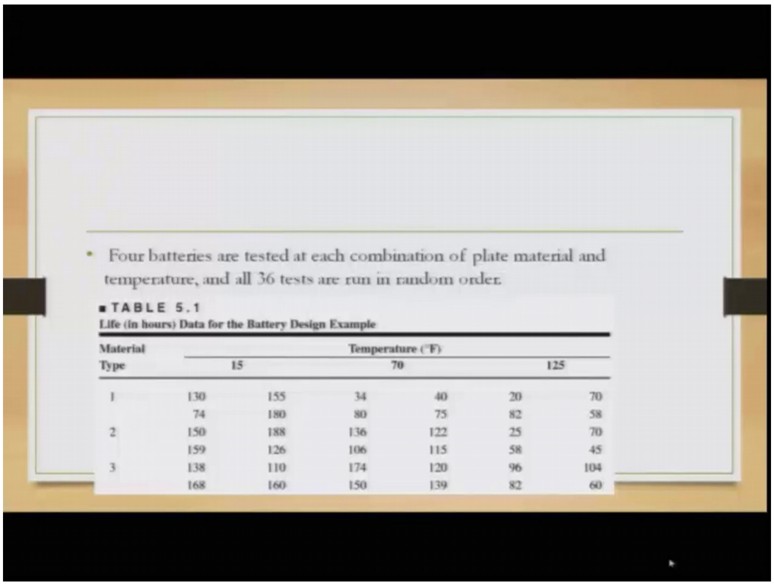
So, there are levels of factors which was A and another factor was B and they are arranged in factorial design, that is each replicate of the experiment contains all the combinations of the treatments in general there are n such replicates.

So, it can depending on the numbers. Considering engineer is designing a battery for using a device that will be subjected to some extreme variations in the temperature. So, there are three choices of plate material from which we have to select one so, that temperatures can be controlled in the product development for the purpose of the test.

And as I mentioned there are three levels; the engineer decides to test all the three plate materials at three temperatures, which are 15, 17 and 125 So, in the other example the etching problem we had four different etching, 160, 180, 200, 220 here the temperatures are 15, 70, 70 and 125 centigrades because these temperature levels which you just mentioned are consistent with the product end use environment because there are two factors at the three levels.

Hence, the design now becomes a problem of a combination of three to the power 2 because there are three vectors and the combinations you are going to take dissolves in 3 square which is 9.

(Refer Slide Time: 03:51)



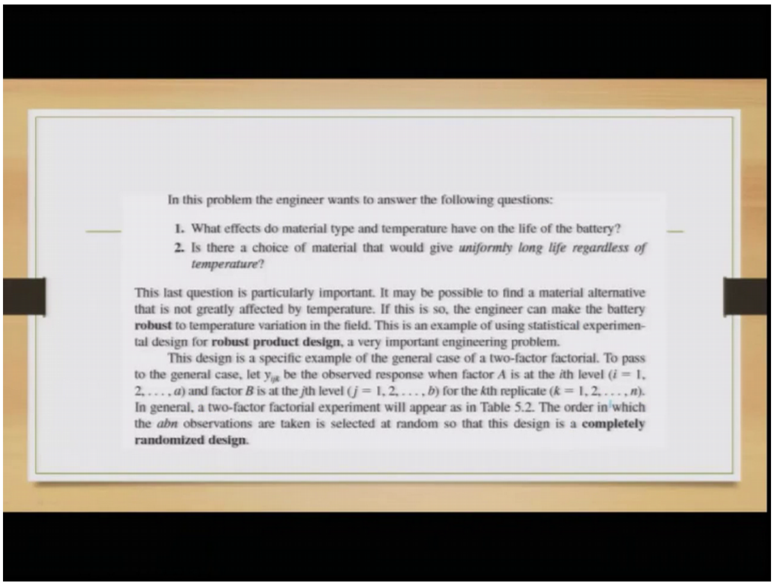
Material Type	Temperature (°F) 15		70		125	
1	130	155	34	40	20	70
	74	180	80	75	82	58
2	150	188	136	122	25	70
	159	126	106	115	58	45
3	138	110	174	120	96	104
	168	160	150	139	82	60

The four batteries are tested at each combinations of the plate material. So, now, there are four batteries; if you remember in the other example for the etching bond, the for each such etchings or treatment there were five different examples or sample size. So, the table 5.1 shows the life in hours data from the battery experiment. So, the temperatures are given in the topmost row, which is 15 70 and 125 degree centigrade or Fahrenheit and the material types are basically given as 1 2 3.

So, the combinations at temperature 15 for me type one is 130 155; For 2 and then the next value 74 180. So, if you read the values in the first column corresponding to the 15 degrees of temperatures, those values are R 1 I have already read, the second values would be 74 if you see the first column. So, I will just try to highlight it in order to make it a little bit more relevant. So, these are the venues. So, this was 130 then 174, 170 I am not highlighting and I am just pointing it out, 150 159 138 168.

Similarly, for temperatures for 70, those values would be I am using a different color. So, it would be 155 180 so on and so forth then the next value is 188 in the last value is 160, for temperature of 125 I will just read the last column in order to make you understand how the overall matrix is, you would basically I have let me use a different color. So, that would be 70 is 58 the next value of being again 70, 45 104 and 60.

(Refer Slide Time: 05:37)



In this problem the engineer wants to answer the following questions. So, they can be more than 2 questions 3 4 and more and so on and so forth, but the 4 2 important questions are as follows, what effects do material type and temperature have on the life of the battery and is there a choice of material, that would give uniformly long life regardless of the temperature. So, you want to basically check whether the temperature is emitted or not.

This last question is particularly important because temperature fluctuations should be there. So, if temperature structure is not affecting the life; obviously, the battery can be used in different environments. It may be possible to as I am reading it, it may be possible to find a material alternative that is not directly affected by the temperature. If this is so, the engineer can make the battery robust to temperature variance in the field and be rest assured that it can be utilized either in the arctic, it can be used it in the tropical climate can be used in the desert and so on and so forth.

So, I am given an example for to make it much more practically oriented. So, this is an example of using statistical experiment design for robust product design, and it is very important engineering problem, which needs to be solved an engineering design in any chemical distillation plant or you want to design it for as in this case for a battery, you want to design for a bridge whatever the factors may be different, but the concept remains the same.

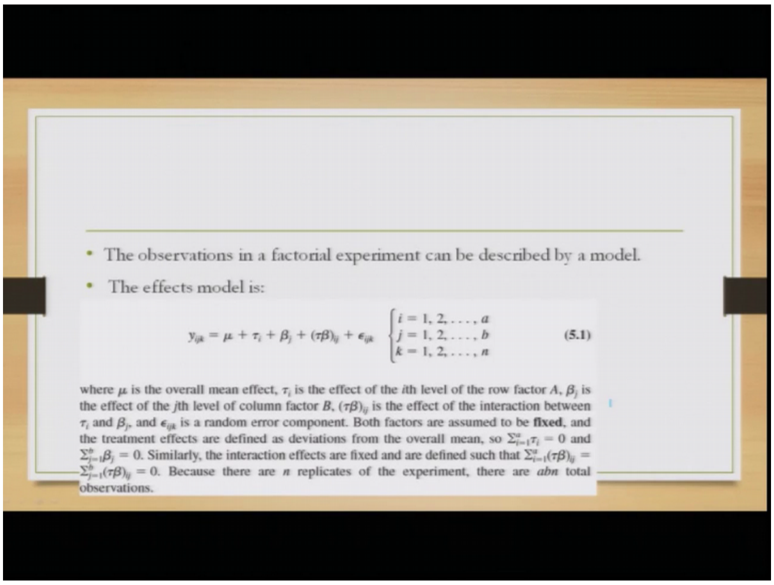
So, this design is a specific example of the general case of two factor models to pass to the general case. So, let us consider these variables as defined. So, you will basically have y which is the effect we want to find out, now the it will have basically three suffixes suffix i, suffix j and suffix k where i would basically be response when factor a is in the ith level. So, these are the factor levels and i changes from 1 to small a, then there would be a factor b in the j th level they would basically change from j is equal to 1 to small b, for of the replications which are being done.

So, for each a and b combinations there would be a replications. So, those replications are given by the suffix k changing for 1 to n. So, say for example, for the replication the kth one or k one th one k 1 is a particular value of k you will basically have a combination of a and b coming out for the which are the two factors.

So, in general a two factor of a factorial experiment will appear as shown in in in table 5.2 in the book, in order in which the combinations would basically be they would be small a number of a capital A factors, small b numbers of capital B factors and each basically would have a replications would it will be given by the value of small n.

So, total combinations or observations are a into b into n and they are selected in at random. So, that this design is completely randomized design of experiments, which we are trying to do.

(Refer Slide Time: 08:41)



So, the observed is basically a factorial experiment can be designed by the following model. So, what is the model? You will have y suffix ijk combinations of all the 2 both the factors and the levels at which they are, which will be given by the mean of the mean of the mean I am using the mean for the three times in the sense that I am trying to find the averages of the averages for both the factors and for all the combinations of the effects which you have; plus tau one tau I would basically be now i you remember basically changes from 1 to small a; that means, you are trying to find out the combinations corresponding to the factor a.

They would be beta j would basically will consider the values of the effects considering i is fixed and j is combined is changing. So, if I consider i along the row and j along the column. So, in the i case what it means that I am keeping j fixed what is the factor be fixed I am trying to basically find all the so called average for any fixed value of capital B, and when I am come combining j only considering j only it means that I am keeping capital A factor fixed and trying to find out the average of the effects along the columns.

Then you will basically have the effects which is coming by the combinations of tau and beta which is. So, it will be tau beta suffix i and j all the combinations and; obviously, there would be an error corresponding to the case when you are considering factor a factor b and the different levels of combinations. Combination is no those effects which you are doing for the experiment. So, again I am repeating i is equal to 1 to small a j is equal to 1 to small b and k is equal to 1 to small n where.

Now, I will explain what I have said that I will again repeat it, where mu value is the overall mean effect for which I use the averages of the averages of the averages. So, technically it may not be the right word, but is just to make you understand that I am taking the averages for all that 2 factors along with the fact that there are different levels of combinations which are n in number. So, tau suffix i is the effect of the ith level on the factor value a which I just mentioned beta suffix j.

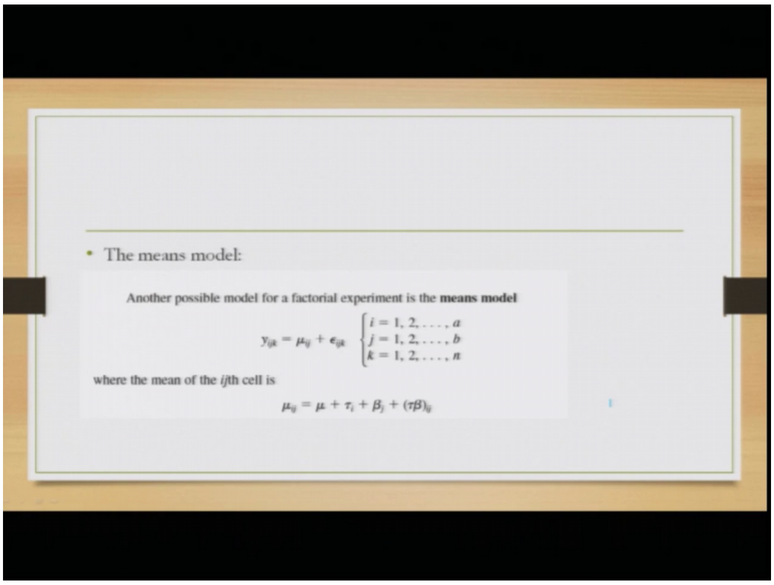
So, tau suffix i was there, beta suffix j is the effect of the j th level of the column factor again as I mentioned row from and the columns for the b s a b are the capital A and B. So, beta j is the effect of the j th level of the column factor, while tau into beta suffix ij is the effect of the interaction between tau y and beta j. So, the combinations I am taking for both the factors combined and epsilon suffix i j k is the random error component corresponding to each combinations of i j and k.

Both factors are assumed to be fixed, and the treatment effects are defined as deviations from the overall mean. So, what we will consider that, when you are considering the effects of as only keeping bs fixed and combining the f effect of bs only keeping as fit then we have some assumptions and what are these assumptions? They are the sum of the errors corresponding to the fact that bs are fixed bs means capital bs are fixed hence the sum of all the tau i s is 0.

Similarly, if you would look at the other picture keeping a fixed and trying to basically find out the combinations of the of the so, called effects along the columns it will be beta suffix j the average value is 0. Similarly, interaction in order to make our life simple we will consider the error terms in such a way such that, the combinations of tau and beta of suffix i and j sum them up for both the values of i and j comes out to be 0, which means that if I keep tau fixed and find out the averages for all the betas for j s or keep betas fixed and find out all the averages for the values of tau is they come out to be 0 this is the assumption.

Because there are n replicas of the experiment, there are total number of small a into small b into small n number of total observations.

(Refer Slide Time: 13:07)

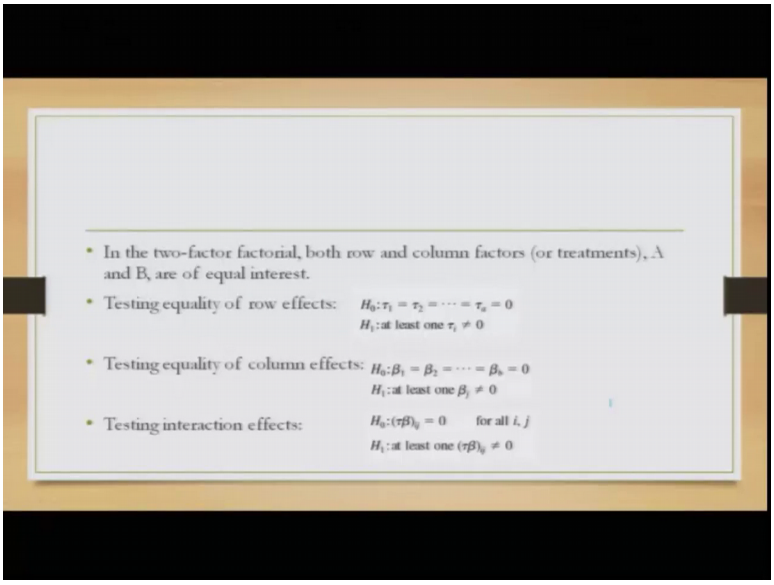


So, the if I find out the mean. So, if you whenever I find on the fiend beam, on the left hand side you behave at basically y suffix ijk, when I try to find out the averages then on the right hand side we knew we know the averages of the taus was 0, average of the betas are 0 and average of the tau betas considering i is fixed j is changing or j is fixed i is changing it is 0.

Hence what you have is basically the total effect considering that, we have been able to assume these three important assumptions of the average values being 0 would be mu is equal to mu ij; that means, the average for any i and j corresponding to the fact that we are doing the experiment time and again would be on the right hand side mu without any suffix, which is the average of the averages of the average.

Again I am using in a very layman sense, but I am sure you will get it plus tau y plus beta j plus the effects of tau beta suffix ij because we are intrinsically assuming another assumption the error term has an average of value 0.

(Refer Slide Time: 14:17)

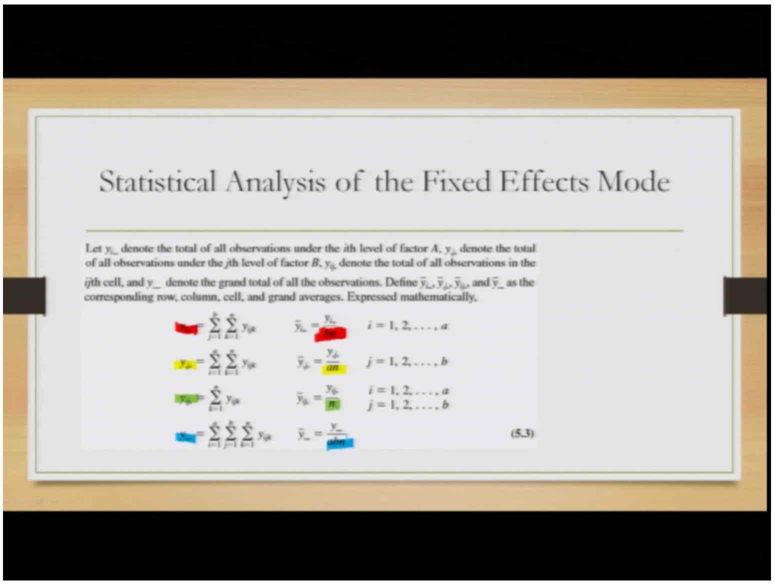


So, in the two factor factorial model both rows and column factors that is corresponding the treatment A and B are of equal interest. So, there if there are equal interest we consider the first test as testing, the value of the row effects to be 0.

In the sense we will consider the null hypothesis as tau 1 is equal to tau 2 till the last value which is tau is the of effects is 0 and the null hypothesis is corresponding to the all term the alternative hypothesis my I am I am sorry for that the alternate hypothesis is that any two of them would not be 0. So, at least one of those taus is would not be 0, when I am taking the column effect I will basically have the null hypothesis as beta 1 is equal to beta 2 is equal to dot dot till the last value which is beta suffix b.

Because there are total amount of combinations small b for the effect of factor capital B is 0, and the alternative hypothesis is that for this for the column effect would be that at least one of them, that is beta j s is not equal to 0. And if I am trying to basically test the interrelationship effect between the combinations of rows and columns, it will be tau beta suffix ij is equal to 0 for all combinations of ij with respect to the alternative hypothesis that one of them is not equal to 0, but at least one.

(Refer Slide Time: 15:42)



Now, the statistic analysis of the fixed every model; So, let y i now this is what I will repeat time and again please make it very clear to all of you. So, y i suffix dot dot. So, which means that I am trying to basically find out the effects of the average of the b s, and average of the ns; basically y i dot dot basically means the total of all the averages under the ith level factor.

Now, if I have y dot j dot it means the total average effects for all the combinations of i s and all the combinations of k for any fixed value of j, which is for the factor B. And if I write it as y and which is one mentioned y dot j dot dot gives the total combinations of all the observations under the j th level of factors and if I give the suffixes as y suffix i j dot it basically means, you know the total of all those averages in the i th and the j th cell considering the effects of all the so, called the experiments, which I am doing considering the different combinations of a and b and the level of I and at the level of j.

So, and finally, y dot dot dot three dots basically means the combinations of all the averages, which I did mentioned as average of the average of the average. So, this basically means the grand total of grand total of all the observations. So, we will define now correspondingly as now listen to me carefully, they would be very simple to understand, but the nomenclature may not be confusing to you that is why I am making it clear would be y bar that is the averages and what are the suffix.

Suffix would be i dot dot that is the first case, second one would be y bar again average it is dot j dot. The third case is y bar and what is the suffixes is, dot dot k and the last one would be dot dot I would not use the word dot dot k my apologies it will be basically y bar ij dot, which means you are taking the averages of anyone combination of a and any one combination in b finding out for all the effects and the final one is y bar dot dot dot which is the average of the average of the average.

So, as the model comes and how the calculations are done, I will explain it very briefly it is very simple to understand just you have to add up and find out the averages accordingly. So, what are those? The first one which I just mentioned; So, I will use different highlighters in colors to of make you understand. So, let me use the red one first.

So, y i dot dot basically means you are taking the averages for j and k. So, this is j being average from 1 to b, k being average from 1 to small n and when and if you find out the averages what would be the denominator? Denominator would be basically be the sum total of all the numbers which you have, when you are taking the second dot and the third dot. So, that will be b into n as mentioned here.

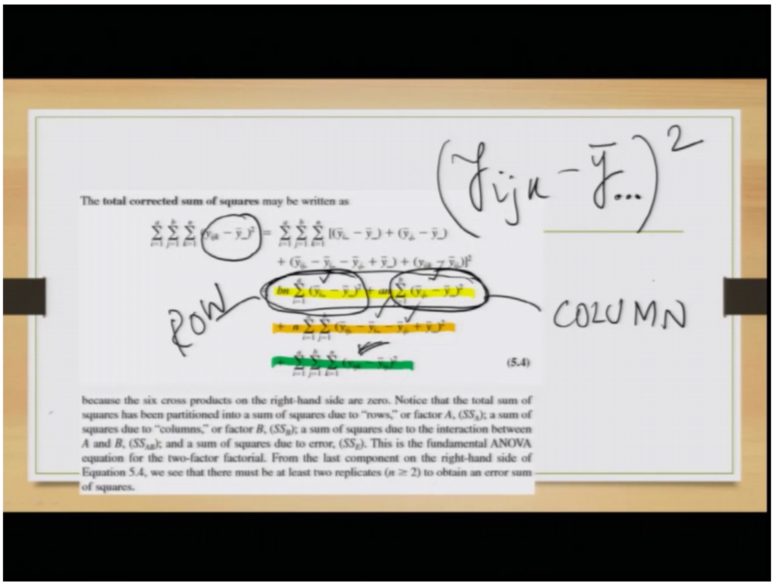
So, i is changing from 1 to small a. So, this is clear; then I go to the second combination of color in order to make you understand let me use the yellow color. So, now, it is y dot j dot; that means, you are trying to find out the averages for all the i s and all ns. So, this is what is given i is changing from 1 to a, the k is changing from 1 to n. So, if I am using the words ijk basically means, that the I is changing from 1 to small a, j is changing from 1 to small b and k is changing from 1 to small n.

So, here the averages are taken for the combinations of all the i s and all the ns or all the k sorry. So, the average would be divided by a into n, if I am taking the average as y i j dot; that means, I am taking the average only once because there is only one dot. So, that would make things clear to you. So, I am trying to find out the average corresponding to k is equal to 1 to n hence it will be divided by n because that is the total number of the observations which I am taking.

And if I combine them for all the three combinations ijk, I have basically find out the combinations of the averages for all these three. So, which is what is given sorry I used in the third case let me just change it in order to make you understand. I will again go back to the third formula. So, this is what should be the case.

So, this is what is there y i j k for a k is equal to 1 to n since it is divided by n, and the last one let me use the color blue. So, if it is dot dot dot basically I try to find out the averages for all the three a b and n. So, this is a b n and I am using the last value. So, i is equal to from 1 to a is changing from 1 to b and i k is changing from 1 to n. So obviously, there are three summations the average would be found out for the sum of all the combinations which is a into b into n.

(Refer Slide Time: 21:18)



Now, I need to find out the total sum of the errors. So, I divide them into technically there would be three such combinations one for a, one for b and one for the samples observations. So, the total corrected sum of the squares would be found out by first the formula you will make you understand, it will be and you use the pen on the highlighter, and use the color black in order to make you understand.

So, this is the average which I am finding out which is y i j for any cell. So, consider ijk is basically a three dimensional matrix any cell would basically have a location or address given by ijk minus the average of all of them combined together.

So, that is why it is y bar dot dot dot, I want to find out the square of that sum them up and find out the total sum of the squares for all the three combinations; combinations being a b and the number of observations we are taken. So, y squared because I am trying to basically minimize the sum of the squares. So, if you do the calculations accordingly and if you have the formula, they would be divided into the three sets of errors. So, this I will highlight without going into the detail combinations of the calculation is very easy, but still I will skip that.

And now you use the highlighter let me use the color yellow. So, this is easily discernible. The first would basically with the color combination here, next would be the color combination of the so, called orange and the third would be the color combination which I am using that red green one.

So, now, I will again switch to the color black. So, it will be now this part the first part which is yellow; if you see it basically be the combinations of the errors corresponding on 2 parts number 1, I am keeping i as fixed and summing any up for all j s and k s. So, this is the first part and finding out the difference from the average of the average of the average which I mentioned.

Second part would be I am keeping j fixed that is a column fixed and summing it for all the is and all the ns and finding out the difference and the squares correspondingly with the average of the average of the average which is here. The second so, this takes care of the yellow highlighted one for the so, called orange highlighted one, there would be one combination. So, what is the combination it will be?

Average of all the ns keeping i and j at the fixed level with respect to the so, called average of the average but when you are taking the average of the average they would be technically be 2 averages; one is when a keeping y fixed and i fixed and another case you have to keeping j fixed.

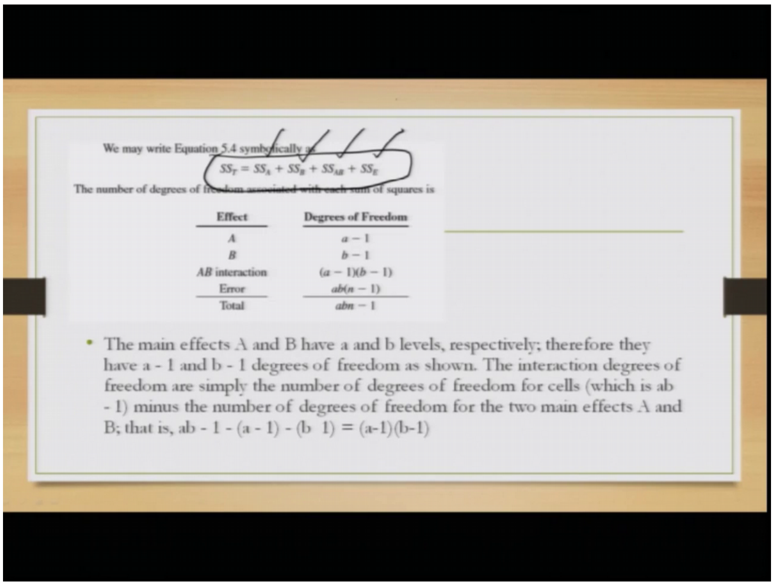
So, i fixed here and j fixed here. So, this gives you the second errors and the third error would be the corresponding error, which you are trying to find out for each cell with respect to the average of the average of the average which is here which I already found, but it would basically be the average of the average corresponding the fact that I am keeping I and a fixed and trying to find out for all the ns.

So, because the 6 cross products on the right hand side are 0. So, notice that the total sum of the squares has been partitioned into some of the squares due to rows some the squares due to columns. So, some of the rows to due to the rows is here the first part, for the columns is this part. So, this is for the row, this is for the column means I am trying to find out for all the bs and capital B and rows is for all the a s.

So, the rows would be basically be given by the fact that the suffix s s suffix a sum of the squares suffix a, for the columns would be s s suffix b and sum of the squares third products is sum of the squares due to interaction between them which is ss s sum of the squares of the of the combination suffix a and b.

And the final one would basically be the error part which is coming which is ss suffix e. So, e I am using all general to find out the error terms, this is the fundamental ANOVA equation for the 2 factor factorial model, for the last component on the right hand side of the equation 5.4 which here does discussed, we see that there must be at least 2 replications that is n is equal to 2 to obtain an error sum of the squares in order to solve the problem.

(Refer Slide Time: 26:18)



So, we can write the total equation as this, sum of the total squares would be for the rows, for the columns, for the combinations and for the error part. So, now, I have given is divided into two parts now pause here.

If I do it I am not going to go into the details, but try to explain it more qualitatively if there are three factors what it would be? It will be coming out for a which is the first factor, b which is the second factor c is the second the third factor, then I will basically partition into combinations like this.

Errors due to combination of a and b, errors due to combination of a and c, errors due to combination in b and c and error due to combinations of abc and finally, they would be error for the errors also sum of the square of the errors also.

So, again the degrees of freedom for a would be the total combinations of column of the rows, rows that is minus 1 is a small a minus 1 the degrees of freedom for the column would be small b minus 1 degrees of combinations for a and b combined would be a minus 1 into b minus 1, and the total sum of the errors would basically be an error term would be total number of combination is what a into b into n minus a and b because you have a number of factors in the row and b minus factors in the column.

Now, if we would go to basically the combination of three, the degrees of freedom would be listen to me carefully it will become clear it will be a minus 1, b minus 1, c minus 1 interactions taken of 2 at a time would be a minus 1 into b minus 1, c minus 1 into c minus 1 and b minus 1 into c minus 1 and the total combination of three taken under time would be a minus 1 into b minus 1 into c minus 1 and finally, when you go to the last part the total combinations would be errors would be basically abc and in the bracket it would basically be combinations as given.

So, their main effects sorry I will just. So, the main effects would be because the 6 cross products sorry. So, the main effects are a and b have a and b levels respectively therefore, they have a minus 1 and b minus 1 and the degrees of freedom the interaction degrees of freedom are simply the numbers of the degrees of freedom for the cells, which would be a minus a in to b minus 1 and the total combination of the errors or degrees of freedom for this example would be a b in the bracket you have n minus 1.

And as I have said for the three factors and four factors it can be done accordingly. So, with this I will end this 16th lecture and thank you for your attention and have a nice day. And, I will continue considering the concept for the two factors and explained it more details with a small problem. Have a nice day.

Thank you.