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

Lecture - 18
Factorial Designs – IV

Welcome back my dear friends, a very good morning, good afternoon, good evening to all of you and this is the TMQ II lecture on the course lecture series under NPTEL MOOC; and we are going to start the 18th lecture as shown in the slide here. So, 18 mean basically it is it has gone we have already completed 3 weeks; 3 means 3 into 515 lectures have been completed we are in the fourth week.

So, by the 20th lecture we will finish the fourth which is half of the whole program which is for 40 lectures, which is 20 hours each week as you know we have 5 lectures each of half an hour each and I am Raghunandan Sengupta from the IME department, IIT Kanpur.

Now, in the last two lectures we are the basically deciding and talking about the factors, that affects the singular effect singular it means, their effect there they are effecting on a standalone basis then combinations and we considered the examples later on the last stages, which were which we are discussing was about the battery and its effects the material and at different temperature.

So, there were four different categories of temperatures also. So, we considered also that later part we will considering the totals some of the errors for the totality, and we were considering that the errors would be on a scale for only for a factor, only for b, b factor and for the combinations of a and b and we consider their errors also from total sum for the errors also which are so, called white noise and then combination of all these three errors would basically be the total errors.

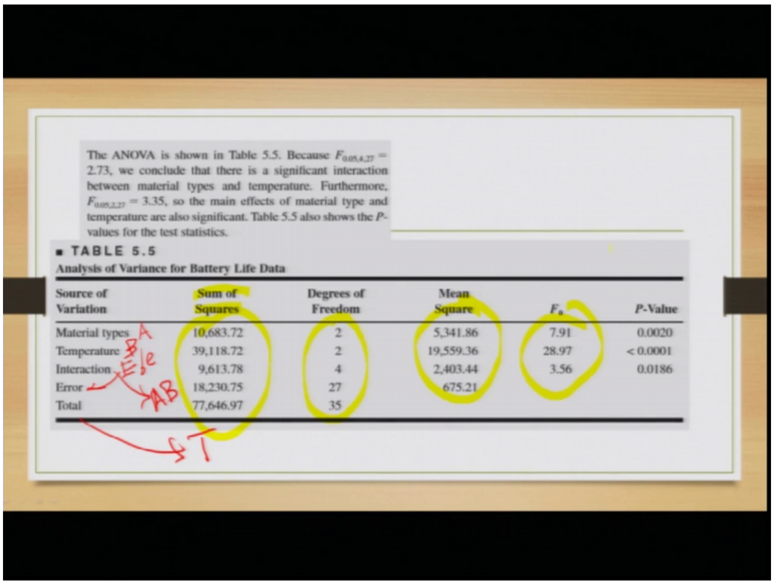
If there is a bc there would be standalone for a, for b, for c combinations of a b, ac, bc and the last one being abc come combined together. And if you consider the overall effects for the so called errors they were basically coming from tau is beta j s for a and b being the factors and for combined it will be tau beta suffix ij.

Now, those I as you know are basically the nomenclature of changing from 1 to a beta from 1 to b and k, which was basically for the sample size the total number of observations was basically k is equal to 1 to small n; and the total observation set of observations was a into b into n and we also saw the degrees of freedom based on the degrees of freedom before that we had fine found out the total sum of the errors and divide, the total sum by the degrees of freedom we found out their mean square errors and once we find out the mean square errors, we can find the expected value and then proceed accordingly.

And another concept which we which I was mentioning time and again was that if you remember the errors coming out from a errors coming out from b or basically some additional over and above the sigma square, which was the errors from the errors that is the standard deviations of the sigma squares from the errors and based on that we can find out what for the example, we consider that how we can find out that given the errors from total sum of the errors from a or and the sum of the errors from b.

And the combination of the errors from of a b and given the total error, we can find out what is the sum of the errors as such. So, we can use simple calculations to find it find it out.

(Refer Slide Time: 03:57)



The ANOVA is shown in Table 5.5. Because $F_{0.05,4,27} = 2.73$, we conclude that there is a significant interaction between material types and temperature. Furthermore, $F_{0.05,2,27} = 3.35$, so the main effects of material type and temperature are also significant. Table 5.5 also shows the *P*-values for the test statistics.

TABLE 5.5
Analysis of Variance for Battery Life Data

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Material types	10,683.72	2	5,341.86	7.91	0.0020
Temperature	39,118.72	2	19,559.36	28.97	< 0.0001
Interaction	9,613.78	4	2,403.44	3.56	0.0186
Error	18,230.75	27	675.21		
Total	77,646.97	35			

So, now, then we consider the ANOVA table to consider it further. So, that ANOVA is shown in table 5.5, we can and also we remember that one should remember that we are using the chi square or the F distribution sorry in order to compare. So, the comparisons would be the mean square of a versus mean square of a mean square of b versus mean square of e e means the errors.

And we can find out what are the under null hypothesis what are the F values and we can compare whether the support the hypothesis do not support the hypothesis and we can formulate the problems accordingly and; obviously, the level of significance which is alpha would be predefined before we start discussing this. So, the ANOVA is shown in let me continue reading its ANOVA is shown is table 5.5.

Because if now these values of 0.05 is basically the alpha value, and 4 was basically the number of observations or m degrees of freedom or m minus 1 depending on how the situation I mean framed and 27 which is the last suffix, which you see under F is basically the degrees of freedom n or n minus 1 that value comes out to be 2.73.

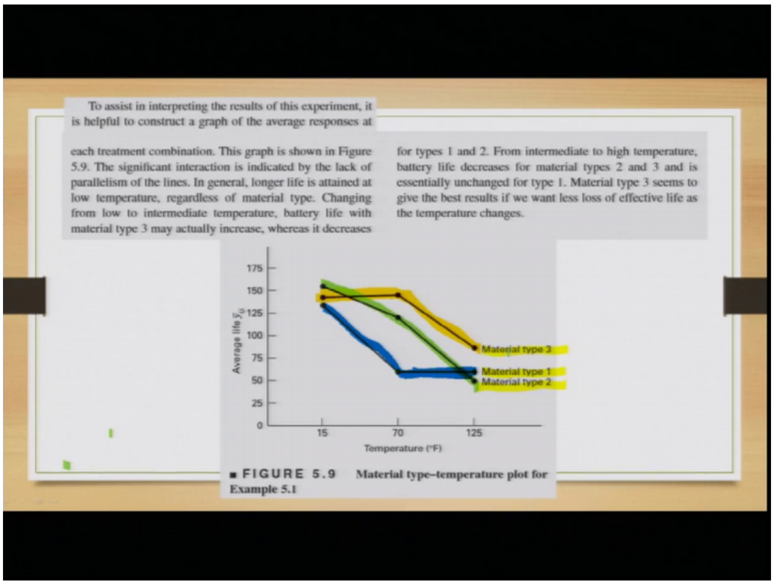
So, we conclude there is significant interaction between the material types, because you remember for the batteries we were trying to find out the effects of the material types on the temperatures combined together on a standalone basis also. So, furthermore the value of f of 0.052 and 27 comes out to be 3.35. So, the main effects of the material type on the temperature are also significant and they are basically given in the table 5.5, where you have in the first column that which is the sources of variance variations.

They are from the material from the temperature from the interaction of the temperature and material and the error. Error is basically we are using the concept of e; so, if I go back and define. So, this is basically A, this is B this is E or small e whichever you denote. So, sorry my mistake my mistake sorry I just jumped 1 step. So, this is basically E would come here and this interaction would be A into B and this is the total one.

So, the sum of the square values are given; So, these are the some of the squares those values are given. So, degrees of freedom are given. So, in if you remember the degrees of freedom was in a minus 1, b minus 1. So, a into 1 a minus 1 into b minus 1; so, based on that we can fine the degrees of freedom.

So, the degrees of freedom have found out the mean square r are found out by dividing the sum of the errors by divided by the degrees of freedom the F values are given and based on that you comment whether they support or do not support the null hypothesis.

(Refer Slide Time: 06:56)



To assist in interpreting the results of the experiment, it is helpful to construct a graph of the average responses at each treatment conditions. This graph is shown in figure 5.9 as stated here and shown; the significant reduction is indicated with the lack of parallelism of this line. So, if they are parallel. So, they would be basically give you a very nice trend in the trend in the relationship.

So, in general longer life is attained at low temperatures, regardless of the material type because the wear and tear decay and whatever concepts we can basically say on the technical part you can basically mention that. So, changing from low temperature to intermediate temperature, battery life with material type 3 may increase. So, what is now important I am going to use the highlighter this is for material 3, this is for material 1 and this is for material 2.

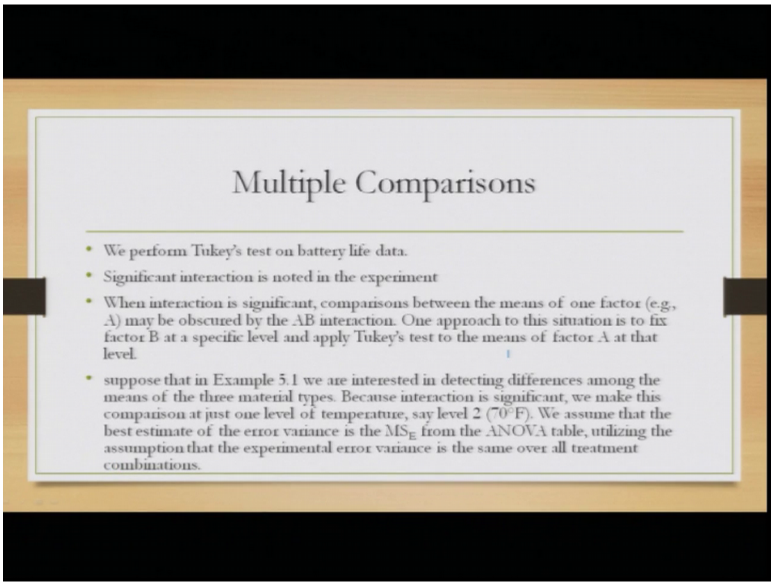
So, if you find out for material 3 I will use 1 color now, not the yellow one. So, I use this. So, for material 3 I am going from the last 1one material 3 it falls like this; that means, from temperature changing from 15 to 70, the average life is almost constant if you see the line which I am highlighting now is almost constant, when it changes from 70 to 125increases the average life changes.

If I consider a material 2 let me use another color green. So, it is not constant, but the fall is almost constant. So, consider this no break is a straight line. So, as temperatures increase from 50 to 70, 70 to 125 the rate of change of the effective average life being affected by temperature is almost constant decreases, and if you consider material 1 let me use another color. So, basically first false and then there is almost constant.

So, from 70 to 125 with this constant; so, let me read it what is written. Changing from low temperature to intermediate temperature, battery life with material three on me actually increase. Increase means this part, I am not going to highlight it, but I am going to going to show it that it almost is in increases whereas, it decreases for type 1 and type 2; that means, this decreases this decreases from intermediate to high temperature; that means, from 7 to 125 battery life decreases 1 material 2 s 2.

For three it decreases s 2 and it is essential unchanged for type 1 which is 2. So, whatever information which I am giving are being validated by the results which have been done in greater details mathematically. So, material type 3 seems to give the best results if you want less loss of effective life as temperature increases it changes. So, we will basically use material of type 3.

(Refer Slide Time: 10:04)

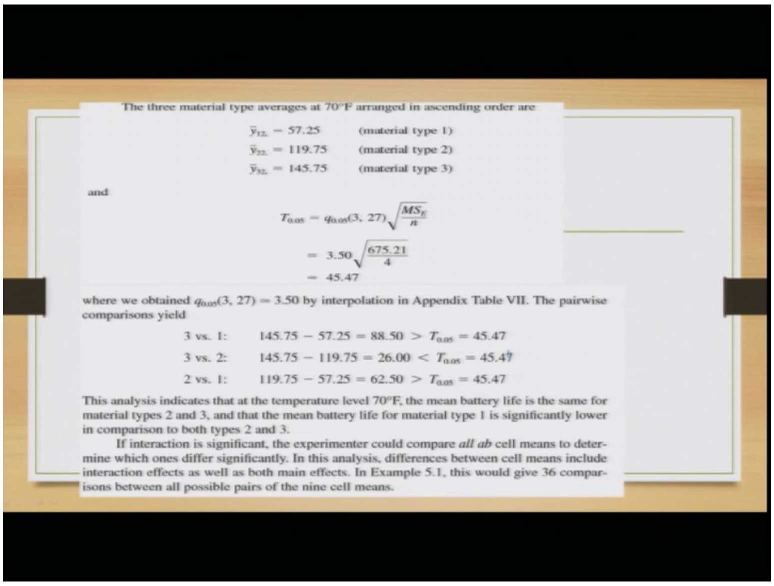


Now, I want to consider multiple comparison, will perform Tukey's test on battery life data significance or significant interactions would be noted down in the experiment. When interaction is significant comparison between the means of one factor example A or B or C may be obscured by the interpretation of A and B combined together. So, it may be possible that B is being obscured by A and B combined together what may be that if you have three ABC C is being obscured by the combinations of A B or B C or AC or maybe whether the overall significance of ABC combine all three together.

So, one approach is to for this situation is to fix factor we had a fig a certain level and apprise the Tukey test to the mean factor of A at that level and try to find out that if A B is fixed or a is fixed what effect does it have on a or vice versa on B. So, keep 1 fixed and try to find out the rate of change on the other. Suppose there is an example which you have discussed we are interested in detecting difference among means of the three material types.

Because in in interaction insignificant we make this comparison at just one level of temperature see in level 2 which is seventy degrees, we assume that the best estimate of the error variance is the MS suffix E from the ANOVA table utilizing the assumption that the experiment, error variance is the same over all treatments combination. So, whatever the combinations are treatments you had it will be the same.

(Refer Slide Time: 11:31)



So, considering that the three material types averages as 70 degrees Fahrenheit arranged in descending order would be, for material type 1 the average is 57.25 for material 2 average is 119.75, for material 3 it is basically 145.75. So, the temperature remains the same and I find out the averages.

So, as the Tukey's capital T value if you remember we have done that q factor multiplied by the Tukey's the ratios, ratios is not the Tukey's ratio the ratio with respect to the mean square divided by the degrees of freedom.

So, they comes out k come out to be the square root of MS E which means square of the errors divided by n, n is the sample size. So, that value is square root of 675.21 divided by 4 and when the Tukey's value at yeah at alpha value of 0.05, which is basically five percent comes out to be 45.47. So, where we obtained now given that we from the appendix if you find out it comes out to be 3.5.

So, now we will do the pair wise comparison. Pair wise comparison are not done three versus 1 then we will do 3 versus 2, then we will do 2 versus 1. So, now, remember we are doing the comparison for the type with respect material type keeping temperature fixed. So, basically a was the factors along the row b was the factors along the column.

So, we keep one fixed and try to basically compare the other factor against each other at different levels. So, once we compare the 3 versus 1 the Tukey's value which comes out to be the difference between the averages which is 114.75 minus 57.25, which is 88.5 is greater than the T value and we say that and then we will basically have a decision whether we support or do not support.

When you accompanying 3 versus 2 the difference is 145.75 minus 119.75 the value counts up to 26, and when you are comparing 2 versus 1 it is 119.75 minus 57.25 the value comes out to be 62.5.

So, this analysis what they mean is, this analysis indicates that are the temperature level of 70, the mean battery life is the same for material types 2 and 3 because for 2 and 3 it is not significant, but for the a case when you are comparing 3 versus 1 when you are comprising 2 versus 1 there is different between that if you are going to clarify, classify 2 and 3 would be and once a category 1 would be at a different category for material comparison.

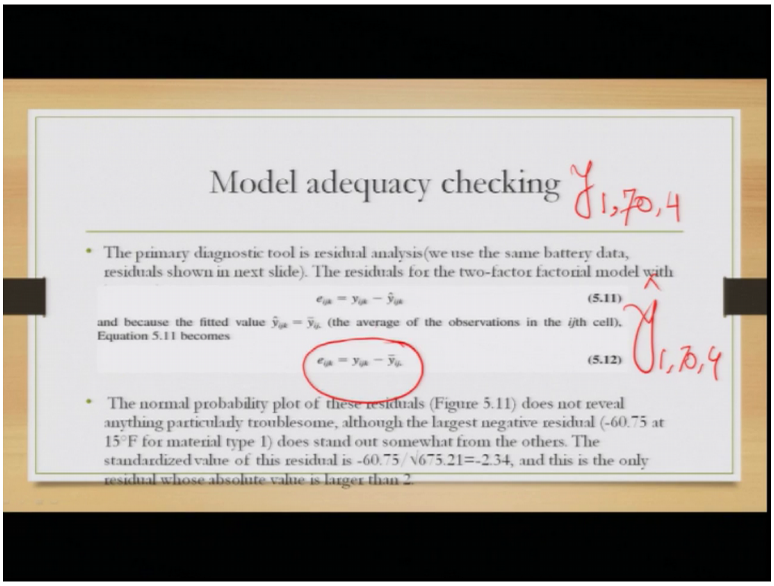
And the mean value so, can let me continue reading it and that the mean way battery life for material type 1 is significantly lower in comparison to both type 2 and 3 which can be compared amongst each other if interaction is significant the experimenter could compare all a b that is a number of observations or a type factor a and k for factor b it was small b.

So, we could compare all the a into b cells. So, each cell had a values, if you remember for n, n which basically is the highest value for k is equal to 1 to n and the e and b cells were basically i is equal to changing from 1 to small a b was the for b was the maximum value for the case when j was changing from 1 to b. So, for this a into b cells, they the cells means to determine which one differs significantly in this analysis different between cell means in include interaction effects as well as both the main effects.

An example which is the material time and the battery the temperature comparison, when you are doing for the battery, this would give 36 comparison between all possible pairs of the 9 cell means which you have. So, 9 cell because they would be comprised comparing the mean values corresponding to the temperatures and also b with respect to the material.

So, my temperature was 15 1 5 15, 70, 7 0 and 125 1 2 5 with respect to the material type 1 2 and 3.

(Refer Slide Time: 15:52)



Now, if you want to do the model adequacy checking and the validation the primary diagnostic tool is residual analysis we use the small same matrix data, residual shown in the which will be shown in the next slide, the residuals for the 2 factor models would be the error; error difference in the values.

So, if you find it the errors would be y ij, which is the actual value for the cell corresponding to i is equal to i, j is equal to j, k is equal to k minus the estimated value for that cell would be y hat suffix ijk. So, say for example, i and j are are a and b and k is equal to n so; obviously, it will be the last cell value such that the actual cell value would be y suffix abn and the estimated value would be y hat suffix a b n.

So, it is basically if I have considering material type 1 temperature 70 and the fourth reading so; obviously, it will be material of type 1 temperature of 70 fourth reading. So, that I will compare with respect to y hat material type 1 temperature seventy fourth reading. So, we can find we can find out the differences and basically fit the value.

So, what is the best fit for y hat would be the average value? So, it means the average of the observations for the ith and the jth cell considering that we are changing and from k from 1 to n; that means, we are taking for that cell corresponding to i and j any fixed value all the observation and trying to find out the averages. So, hence the error term would be the actual value minus the average value for that cell corresponding to any fixed value of i and fixed value ij and all combinations of n which are there.

The normal probability plot of this residuals, does not reveal anything particularly troublesome, although the largest negative residue residual which is minus minus 60.75 at a temperature of 15 degrees for material type 1 thus turned out somewhat from the others that is an outlier.

The standardized value of the residuals would be given and we can find out the value to be say for example, 2.34 and this is the only residual whose absolute value is large and then 2 and we can basically find out some reason for that on the on the technical front.

(Refer Slide Time: 18:24)

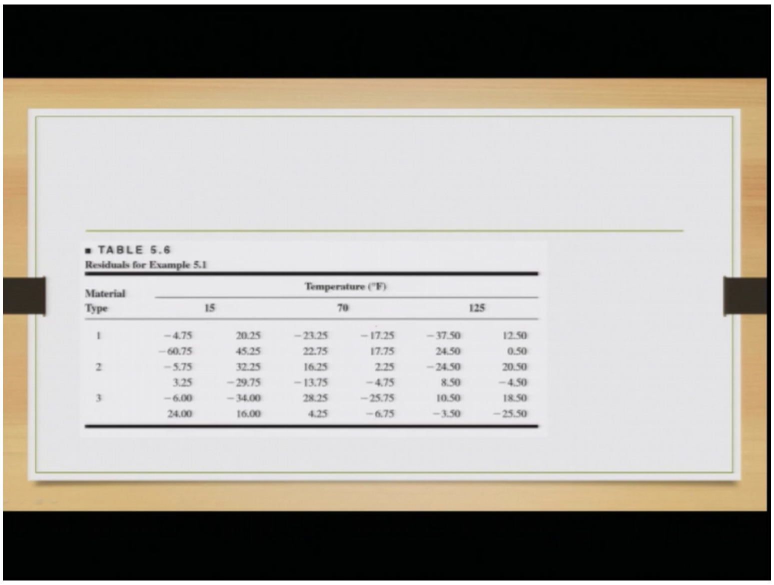
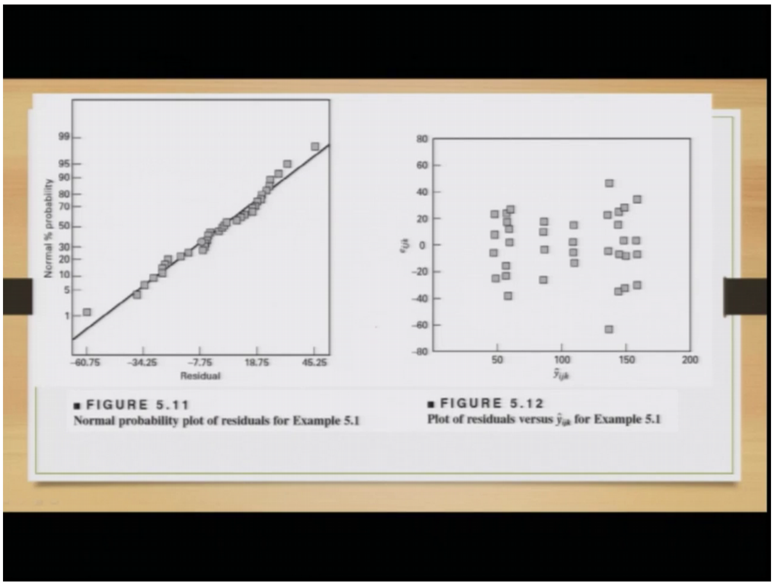


TABLE 5.6
Residuals for Example 5.1

Material Type	Temperature (°F) 15		70		125	
1	−4.75	20.25	−23.25	−17.25	−37.50	12.50
	−60.75	45.25	22.75	17.75	24.50	0.50
2	−5.75	32.25	16.25	2.25	−24.50	20.50
	3.25	−29.75	−13.75	−4.75	8.50	−4.50
3	−6.00	−34.00	28.25	−25.75	10.50	18.50
	24.00	16.00	4.25	−6.75	−3.50	−25.50

The residuals are given. So, the values of the residuals for my we can basically plotted; the material being on the first column 1 2 3 the temperatures being on the first row which is 15, 70 125, the values are given as they are it is minus 4.75, 20.25 I am reading the rows, then minus 23.25 minus 17.25 minus 37.5, 12.5 and we can basically mark these values accordingly.

(Refer Slide Time: 18:55)



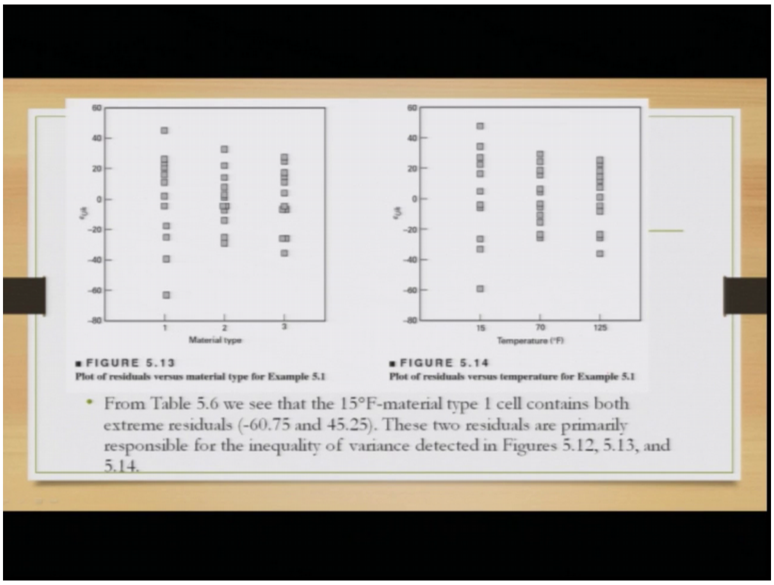
■ FIGURE 5.11
Normal probability plot of residuals for Example 5.1

■ FIGURE 5.12
Plot of residuals versus $\hat{y}_{ijk}$ for Example 5.1

When we do the normal plot of the residuals for this example, the normality plot comes out to be almost the same except one outlier here and the set of outliers, which are here so; obviously, there would be some technical reason for that, but it is just to show that what is the normality plot. Then if you plot the residuals versus that is e error term with respect to the predicted values which is y hat or y bar, which we can basically take the overall distribution correspond distribution over less the worth the variance of.

So, say for example, we find out for 50 y hat value of 50, we can find out for 100 you can find out for 150 and then basically see how the dispersions how the movements or how scattered they are.

(Refer Slide Time: 19:44)

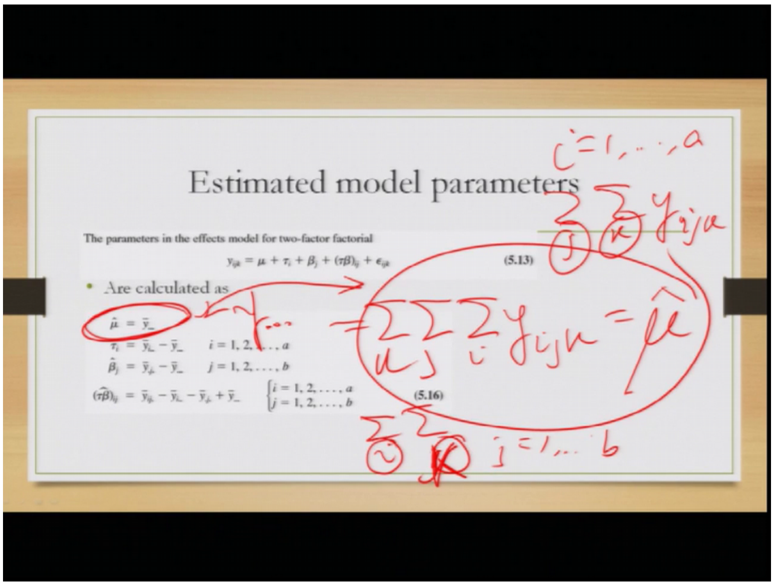


Similarly, we will do the plot of the residuals versus material time; material time being 1, 2, 3, and we will do the residuals plot with respect the temperatures also, which is factor b temperature is 15, 70 and 125 and on the y axis as I mentioned again; I am repeating it is the residual or the errors. From table 5.6 we see or from the information we see that, 15 degree material type 1 cell contents both extreme residuals which is minus 60.75 and 45.25.

So, this is a huge dispersion of the movement; these two residuals are primarily responsible for the inequality the variance detected in the figure 5.13 to 5.13 and 5.14, 5.13 and 14 are in front of you now. So, they would basically give this. So, if I am saying that for 15 and type 1; So, these are the huge amount of variability.

So, they must be I am just hashing it in order to make you understand. So, they would be some ideas you can get about the overall dispersion of the movement. Now, the estimated models of the parameters model would be.

(Refer Slide Time: 21:07)



If you remember we have mentioned the model as y ij k suffix values are ij k, i changing from 1 to a, j changing from 1 to b, k changing from 1 to n would be equal to the average of the average of the via average, which is mu which is technically y double bar or triple bar depending on how many such factors you have is basically the average of the average of the average and you are trying to basically sum it up for three times.

So, when I am trying to find out y i j sum up for all i sum up for all j sum up for all k. So, this technically in the long run should be the mu value. Plus tau i which is basically the so, called movements positive negative movements with respect to the ith factor beta j would basically with the movements positive negative depending on the b th factor and tau b beta in tau into beta suffix ij would basically the movements corresponding the combination of the factor a and factor b and epsilon was the error epsilon ijk.

So, how we can find out? This is very simple if you remember I want to find out the average which I am just noted down. So, technically do with the estimated value, which technically if I put it right it is dot dot dot as shown here. So, this is basically going to make life simple for us to understand. When I find out tau i, tau i would basically be the difference now i is what? I is basically for a factor.

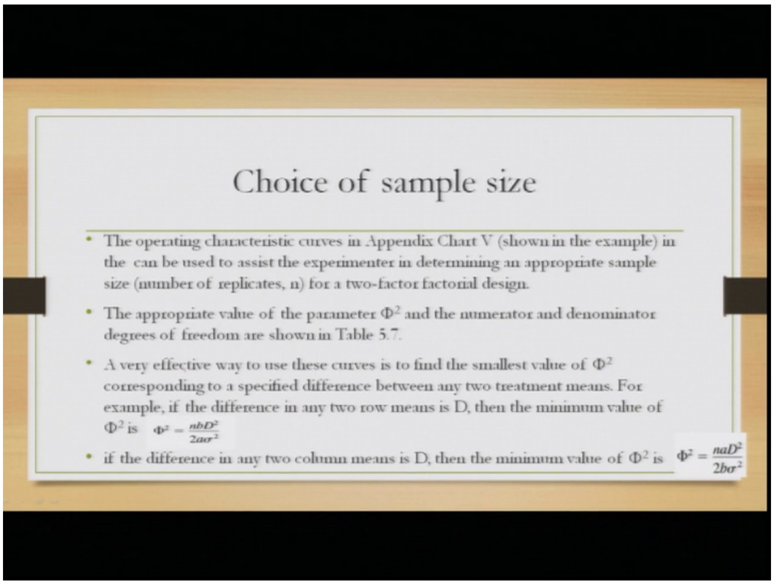
So, for any i I want to basically find the averages for the js and average of the nth combined. So, it will be y bar i suffix dot dot; that means, we are summing up for the second and the third. So, it will be summation of the second and the third which is j and k and i being fixed is equal to anything value i equal to 1 to a.

So, this is y i j k we sum it up, minus the average of the average of the average. So, that will give me tau I similarly when I want to find out beta js or hat value because its the estimate value it would be what? I am keeping j fixed. So, it will be summation for all the i s summation for all the ns.

So, the summation would be for all i s and all j case sorry. So, note down here it is basically summation of i k, note down here it is the summation of jk and here j is equal to 1 to b. So, that value minus the average of the average of the average which is y bar dot dot dot dot dot dot are the suffixes and if I want to find out the estimated value of tau beta suffix ij then; obviously, it will be summation of will go 1 step at a time.

First we will find out the summation for all the ns minus that would be the summation for all the values of j and k keeping i fixed and then trying to find out the all the summations keeping i and k fixed i and k summing up sorry for j fixed, minus the average of the average of the average values, which is this one when i is equal to 1 to a.

(Refer Slide Time: 24:36)



Now, choice of the sample size is important. So, what are the important points to be considered are like this the operating characteristic curves which will be shown for each and every very sample size, m can be used to assist the experimenter to determine an appropriate sample size for a 2 factor model can be increased with the 3 factor models 4 factor model so on and so forth.

The appropriate value of the parameter capital is phi square, which is basically corresponding to the case which you want we which we have considered earlier would be for the case such that, the numerator and the denominator degrees of freedom would be shown and a very effective way to use these curves is to find this operating characteristics curves is to find out the smaller value of the capital phi squares, corresponding to a specified difference between any 2 treatment levels; that means, fif15, 70 or 71, 25 or 15 and 125.

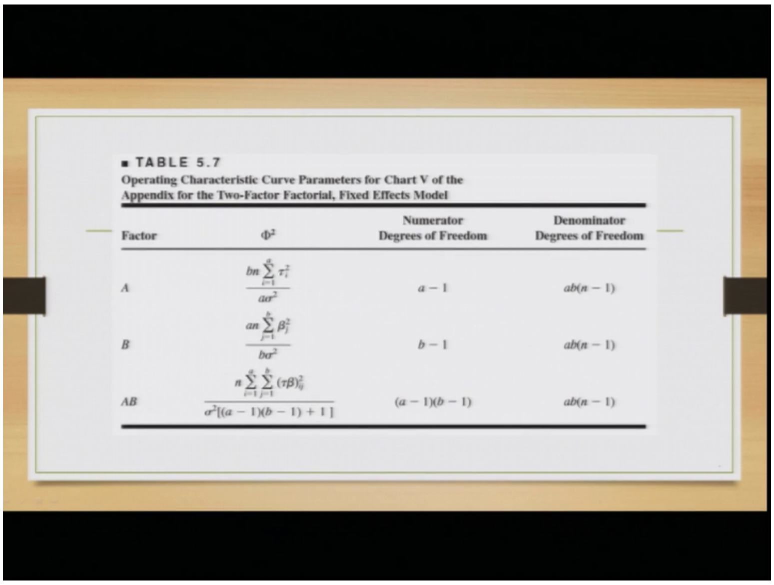
For example if the differences in the in the 2 row elements is the, then the minimum value of phi square would be calculated accordingly which will be n into b because n is basically for the for all this observations b is for all the of all the numbers for factor 2 which is b and d square would be the differences corresponding to to the row means values for the rows means for factor a.

So, that is hence it will come in the numerator which will twice into a into sigma square. Sigma square is there the variances a is the number of observations and y 2 because you are taking the average in the long run. So, it will be the difference if any and in the case when you are trying to find out the differences of the columns, then just replace b by a as you see.

So, this b value which I am circling would be replaced by a, n remains fixed n remains fixed d square, which I am taking as a symbol or the notation for to find out the differences between either the rows of the columns remains as it is d square and in the and in the denominator to remains as it is, sigma square remains as it is and I am and in the first case I find out for the rows.

So, hence it was a, now I find out for the columns hence it is b.

(Refer Slide Time: 27:04)



■ TABLE 5.7

Operating Characteristic Curve Parameters for Chart V of the Appendix for the Two-Factor Factorial, Fixed Effects Model

Factor	Φ^2	Numerator Degrees of Freedom	Denominator Degrees of Freedom
A	$\dfrac{bn\sum_{i=1}^{a}\tau_i^2}{a\sigma^2}$	$a-1$	$ab(n-1)$
B	$\dfrac{an\sum_{j=1}^{b}\beta_j^2}{b\sigma^2}$	$b-1$	$ab(n-1)$
AB	$\dfrac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{\sigma^2[(a-1)(b-1)+1]}$	$(a-1)(b-1)$	$ab(n-1)$

So, operating characteristic curve parameters for the two factor models, for this fixed effect models; so, the factors are given a b a b it can be continued to abc ac bc abc. So, phi capital phi squares values are given in and the degrees of freedoms are given based on the degrees of freedom, you can find out the corresponding average values as required.

So, with this I will end this 18th lecture and continue more discussion about the factor models and expand it, and basically try to consider more examples for a better understand have a nice day.

Thank you very much.