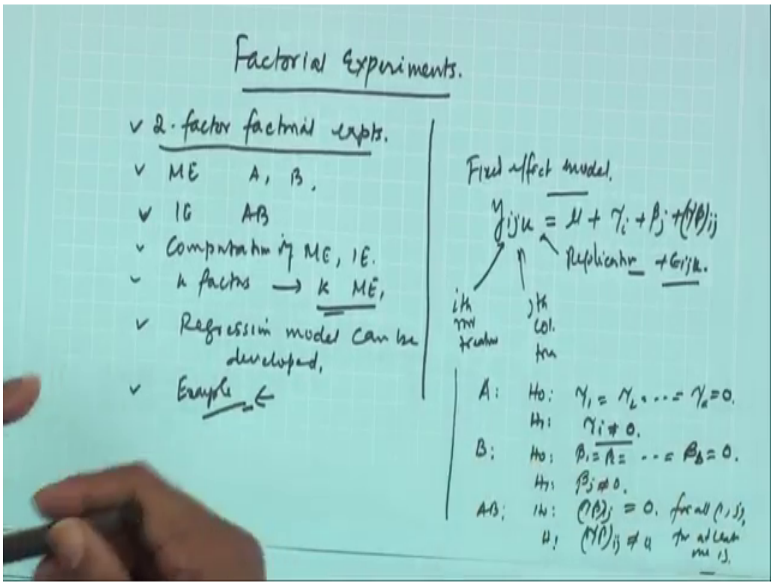
Design and Analysis of Experiments
Prof. Jhareswar Maiti
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture – 31
Statistical Analysis of Factorial Experiments

Welcome we will continue with Factorial Experiments.

(Refer Slide Time: 00:20)

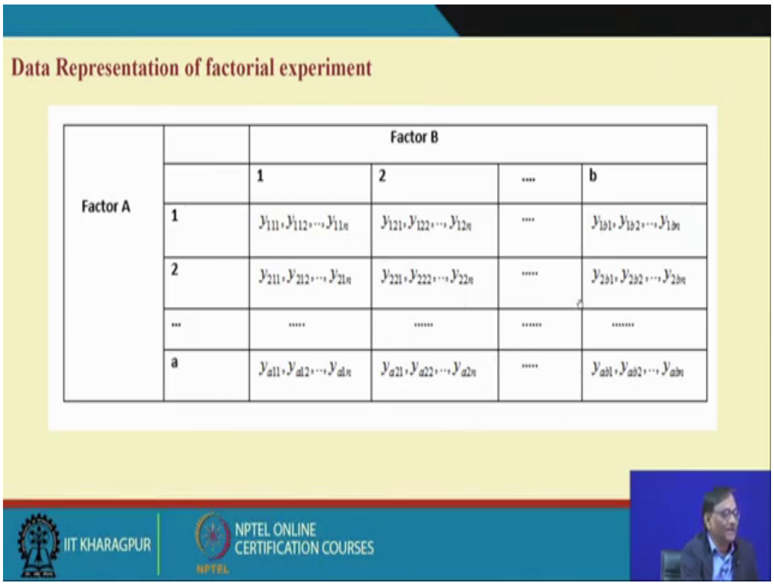


In last class we have discussed 2 factor factorial experiment and we discussed the main effects A and B interaction effects A B then how to compute the their computation of main and inter interaction effects M E and I E.

We have discussed also that there may be K factors and then K number of main effects and K C 2 number of 2 way effects case 3 C 3 number of 3 way effects. So, like this in addition for A 2 factor factorial experiment I have shown you that how the regressions model can be developed model can be developed. And then I have given you 1 example of facts factorial experiment with 2 factors, A data experiment hypothetical data was shown to you and in general what will be the layout for the data actually data table or the data table also shown to you in this lecture.

What I will do I will give you some statistical analysis of the experimental data and show that, we will see whether the model is able to capture the variability in the response variable and accordingly able to determine also, whether the factors main and interaction effects are really contributing in explaining the variability of the observed y or the response variable.

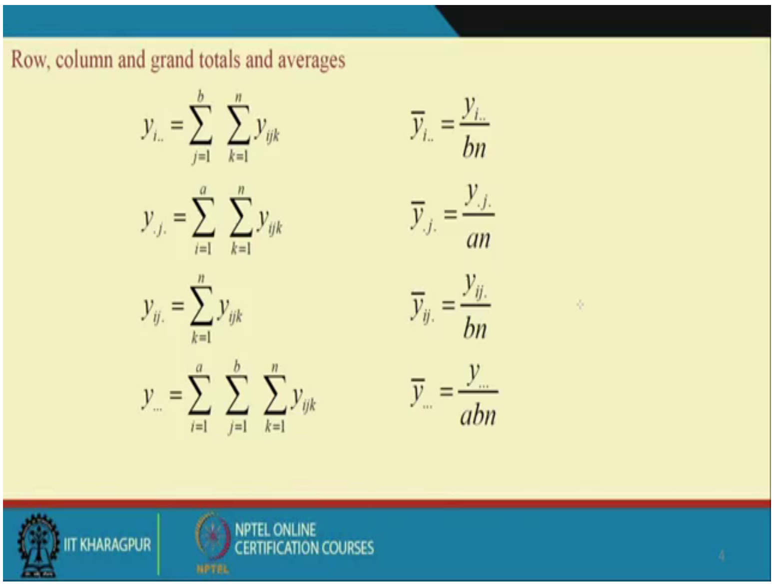
(Refer Slide Time: 02:29)



Factor A		Factor B			
		1	2	....	b
	1	$y_{111}, y_{112}, \ldots, y_{11n}$	$y_{121}, y_{122}, \ldots, y_{12n}$	....	$y_{1b1}, y_{1b2}, \ldots, y_{1bn}$
	2	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$	.....	$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$
	...	.....	......	......	....
	a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$	.....	$y_{ab1}, y_{ab2}, \ldots, y_{abn}$

So, in general there will be your general observation y i j k i for ith ith row treatment j for jth column treatment and k for replication.

(Refer Slide Time: 02:54)



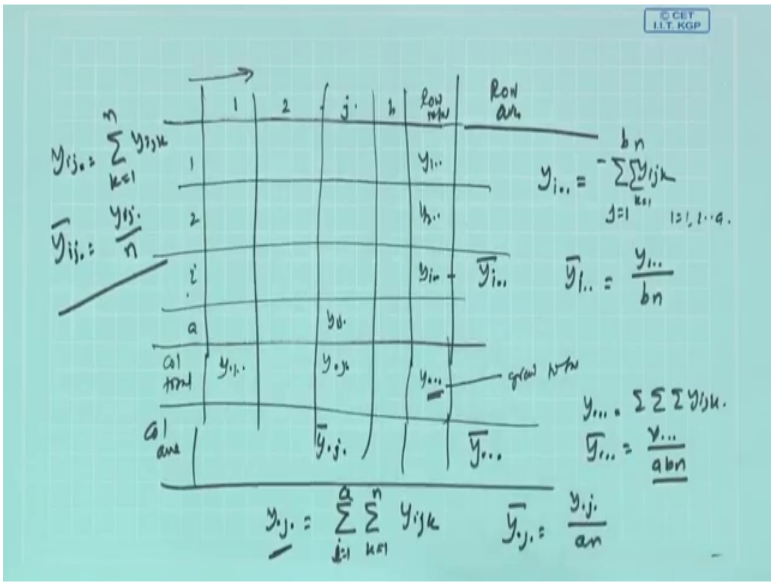
Then also we ahve seen that this y i j k equal to mu that is grand mean plus tau i that ith row effect treatment effect, beta j jth column treatment that is the column treatment effect, plus tau beta i j interaction effect plus epsilon i j k that is what is the error term and this is what is our may fixed effect model this is what is our fixed effect model.

And we have set hypothesis with reference to A we say H 0 that tau 1 equal to tau 2 dot dot dot equal to tau A and H 1 we say tau i not equal to 0 tau i tau 1 tau i this equal to 0 and tau i not equal to 0 for at least 1 tau i. Similarly for B we put H 0 beta 1 equal to beta 2. So, like this equal to beta B equal to 0 and H 1 beta j not equal to for at least 1 j.

And another 1 for A B we put H 0 that tau beta i j this equal to 0 for all i j combination and H 1 we say that tau beta we say tau beta i j not equal to 0 for at least 1 for at least 1 i j. That is what we have seen in the last class. So, now, giving the table we want to compute giving the table means giving the table of data or the observations assumed, we want to compute the different in different parameters or statistics related to the experimental data.

So, y i dot dot this is basically the row total ith row total y dot j dot this is the column total jth column total, y i j dot this is basically the that row column combination that the cell total and y triple dot this is the grand total.

(Refer Slide Time: 05:30)



So, with reference to with reference to our data what is this we have 1 2 like ath rows 1 2 like B columns we have here ith row we have here jth column. So, this is y i dot that is basically row total this will give you y dot j column total and then y dot dot dot. So, y dot

dot dot this is your this, i dot dot i dot dot j i dot j dot this is your grand total this is related to row total. So, the first 1 will be y 1 dot dot y 2 dot dot something like this similarly here y dot 1 dot and this way you can total.

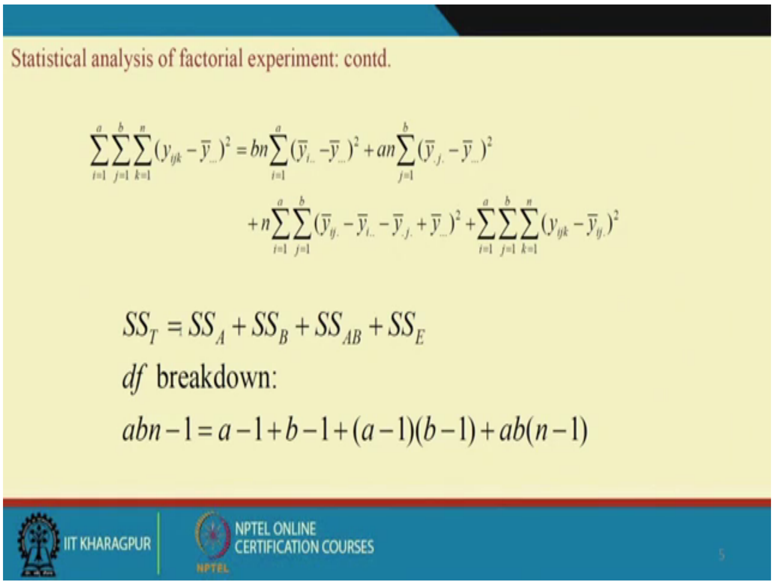
And then you can calculate row average row average which will be the general 1 will be y i dot dot average you can calculate column average which is general 1 dot j dot this average and you can calculate grand average grand average y dot dot dot this is your grand average.

So, then what will be the formula for i for your, i y i dot dot this total these basically going from column wise from 1 2 b. So, sum of j equal to 1 2 b y i j and k. So, definitely k will also be there k equal to 1 to n and i equal to 1 2 a similarly for column case y dot j dot this will be your k equal to 1 2 n j equal to i equal to 1 2 a y i a y i j k. So, this is the total now what will the average. So, y 1 dot dot y i dot dot average will be y i dot dot this total by b into n total by b into n what will be the this average column average this will be y i j dot divided by a n. So, what will be the grand average grand total grand total will be all 3 sum y i j k and then grand average will be y grand total divided by a b into n.

So, this is the computation of different averages row average column average row total column total grand total grand averages and this is what is given here in the slide also, y i double dot y dot j dot y i j dot like this 1 more thing I am not given here that what will be the y i j dot y i j dot this is nothing, but the y i j dot this is nothing, but you have n observation k equal to 1 2 n y i j k that mean the cell total.

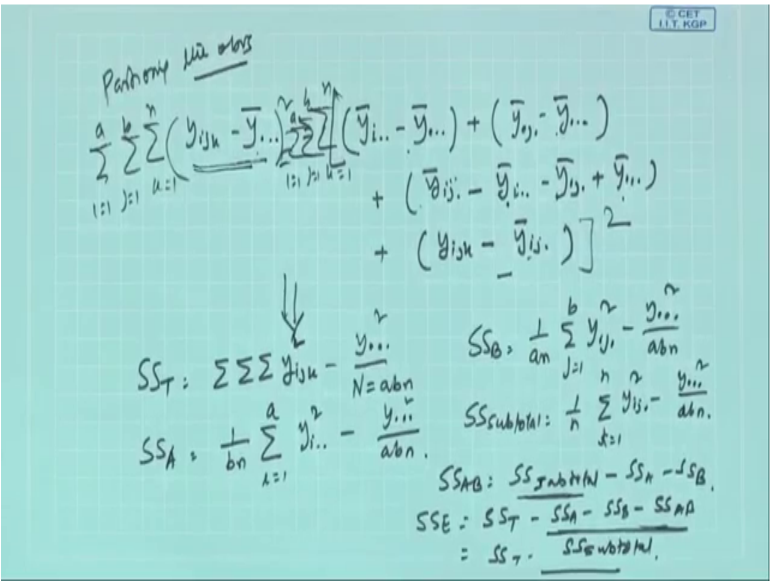
And then the cell average i j dot this will be a cell total divided by how many observation you have counted there n observation counted there. So, these will give you this will give you the required statistic values statistic values for computation of S S T sum square total sum square effect main effects interaction effects and errors all those things.

(Refer Slide Time: 10:07).



So, now I will show you the partitioning the variability.

(Refer Slide Time: 10:18)

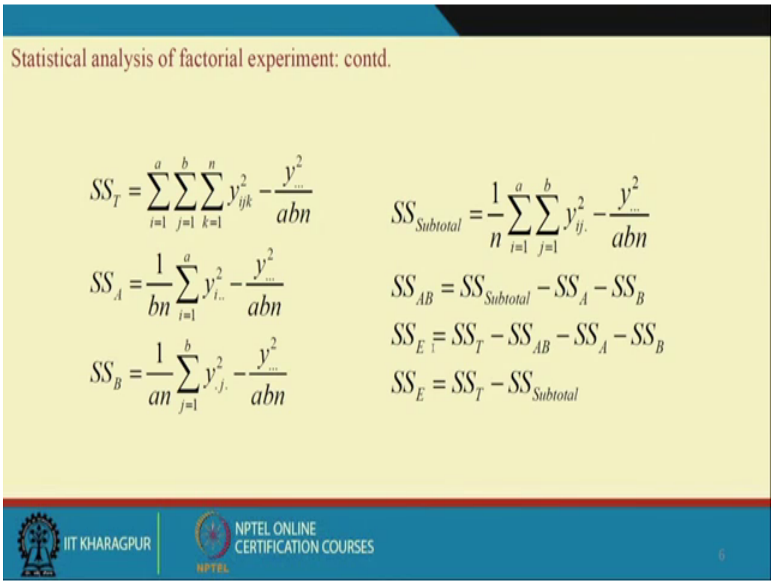


Observation i j k it can be subtracted by it is mean component, this can be written like these y i dot dot bar minus y triple dot bar plus we can write these as y dot j dot bar minus y i y triple dot bar can be written, plus we can write that y i j y i j dot minus this bar minus y i dot dot bar minus y dot j dot bar plus, y triple dot bar plus we can write y i j k minus what is there y i j dot bar. Then this left hand side and right hand side will become equal. So, these and this will cancel out these and this will cancel out these this will cancel out, these and this will cancel out, these and this will cancel out, then this side then this side will be same.

So, this is known as partitioning the observations. So, you can square it take sum over i j k take sum over i j k k equal to 1 to n j equal to 1 to b i equal to 1 to a k equal to 1 to n j equal to 1 to b i equal to 1 to a like this. So, the resultant after algebraic manipulation and taking the sum of that sum 2 way coming cross values will be deleted and then ultimately you will get this equation.

So, left hand side is the S S T; that means, i j k this then right hand side the first 1 b n into these this is S S A that is the sum square for the factor a, second 1 will be sum square for factor b, third 1 will be interactions and fourth 1 will be error. And then the degree of freedom will be accordingly broken down that for the total number of observation is a b n 1 minus 1 will give you degree of freedom for S S T then a minus 1 plus b minus 1 a minus 1 for a b minus 1 for S S B A minus 1 into B minus 1 for S S A B and a b into n minus 1 for S S E. So, this is what is the sum squares, for different effects?

(Refer Slide Time: 13:42)



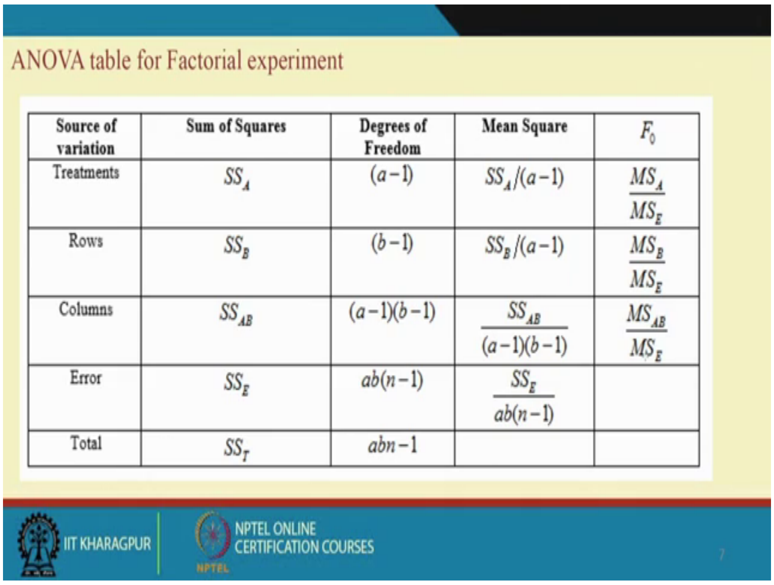
Now, we will use some other combine that formula for each of computation, because you will see that the first these formula here the computation will be complicated laborious here the formula the earlier the manual computation point of view as well as general computation point of view; S S T can be that sums square sum of square of all observations minus that correction factor y triple dot square y a b n. And S S A will be the square of the rows sum square of rows divided by the number of that values observations minus this correction factor i am saying correction factor this is basically correction factor per mean.

Then similarly S S B similarly this 3 you have got similarity earlier, but then another 1 is S S subtotal which is the cell total. So, that cell total will be considered and they are squaring some square will be considered and will be divided by the number of such sum square in a every cell 1 by n, and then it will be subtracted by correction factor it will give you the S S subtotal this S S subtotal will be used to compute S S interactions, which is basically S S subtotal minus S S A and S S B minus S S B then S S E will be calculated using subtraction method and that is that is what is to be done.

So, if I say how to calculate S S T you will calculate S S T by considering y all the observation square and take their sum minus the grand total square by a were n equal to a b into small n. How you calculate S S A you basically consider b n observations then sum i equal to 1 to A y i dot dot square minus y dot dot dot square by a b n. The similarly how do you compute S S B; S S B it will be 1 by A n sum of j equal to 1 to b y dot j dot square minus y triple dot square by a b n. How you calculate S S subtotal. So, this will be 1 by n sum of k equal to 1 by n y i j dot square minus y triple dot square by a b n. Then how you will calculate S S A B; S S will be S S subtotal minus S S A minus S S B.

Then how to calculate S S E S S E will be S S T minus S S A minus S S B minus S S A B. Now S S A B is S S S S subtotal is this 1. So, that means S S T minus S S subtotal. So, these are formula you will be using for computation.

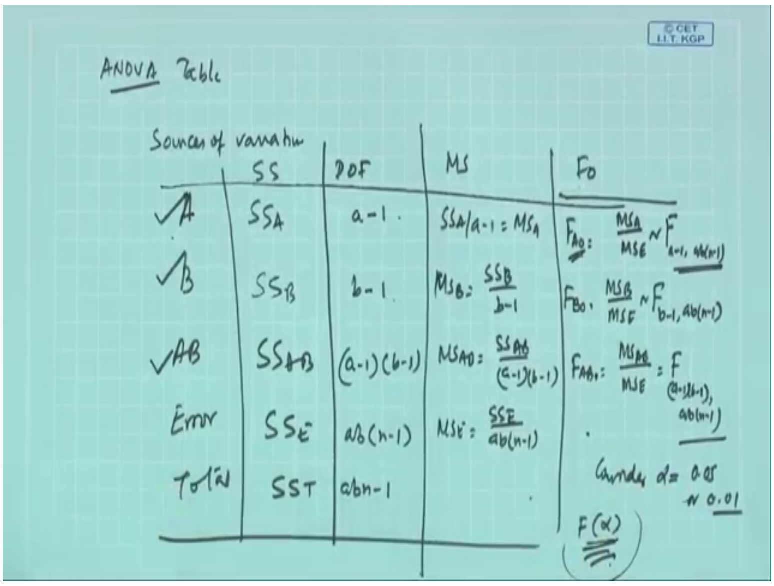
(Refer Slide Time: 17:34)



ANOVA table for Factorial experiment

Source of variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	SS_A	$(a-1)$	$SS_A/(a-1)$	$\frac{MS_A}{MS_E}$
Rows	SS_B	$(b-1)$	$SS_B/(a-1)$	$\frac{MS_B}{MS_E}$
Columns	SS_{AB}	$(a-1)(b-1)$	$\frac{SS_{AB}}{(a-1)(b-1)}$	$\frac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$\frac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn-1$		

(Refer Slide Time: 17:44)



Now, once you compute the sum square you will be you will be going for ANOVA table what will be the Anova table here Anova table will be first sources of variation sources of variation sources of variation 1 is basically with respect to no this is basically that sources of variation is 1 is A treatment, B treatment, A B treatment error and then total. So, here is so it should be A treatment sources of bin factor A factor B, interaction A B error and total.

So, here is the we have written treatment it is basically treatment A this is treatment B interaction A B error and total this is this is the issue. So, we have written rows and column, but it is not the treatment A treatment B treatment interaction A B. So, anyway I am showing you here writing around here again. So, then what will be S S this will be S S A this will be S S B this will be S S A B this will be S S error S S total.

Now what is the degree of freedom degree of freedom this will be a b a b n minus 1, there are a minus 1 degree of freedom for S S A B minus 1 degree from S S B and A minus 1 into b minus 1 will be for in interaction and a b into n minus 1 will be for error.

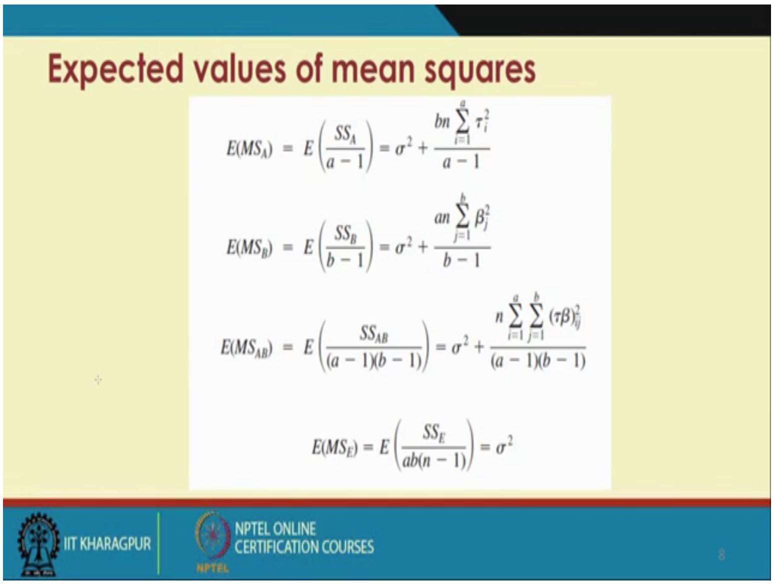
So, these these these added up subtract from these will be will be getting this, then you calculate M S that is S S A by A minus 1 is equal to M S A. So, M S B will be S S B by b minus 1 M S A B will be S S A B by a minus 1 b minus 1 and M S E will be S S E by a b into n minus 1.

Then you find out F 0 here you please understand we have 3 hypothesis is here null hypothesis related to A related to B related A B. So, F 0 for A so F A 0 if arrived this will be your M S A by M S E. This follow f A minus 1 and M S E is a b minus a b into n minus 1 that is the f distribution.

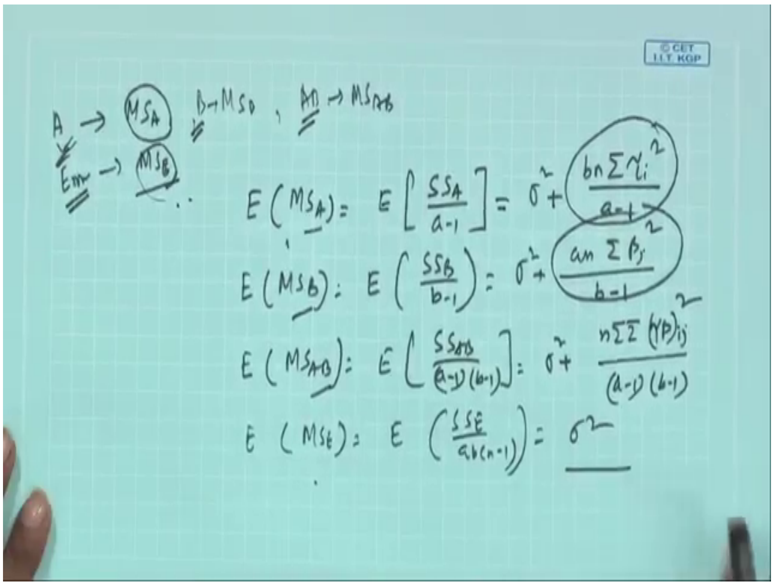
So, similarly M B 0 will be M S B by M S E this will follow if b minus 1 a b into n minus 1 you know (Refer Time: 20:54) f distribution. Then F A B 0 will be M S A B by M S E this will follow if a minus 1 into b minus 1 and a b into n minus 1 this degree of F distribution with this numerated degrees F (Refer Time: 21:12).

Then what we will do basically after that you will calculate the correspond you consider alpha let consider alpha equal to may be 0.0 5 or 0.0 1 easily this 2 will consider. Then find out the corresponding F alpha with required degrees of freedom and compare with the computed F 0 and if the computed F 0 is more than the threshold F then you reject null hypothesis otherwise accept the null hypothesis.

(Refer Slide Time: 21:53)



(Refer Slide Time: 21:59)



So, this is your anova table here interestingly that we have computed M S factor M S A for example, for factor A you have computed M S A and also for error you have computed M S E. So, if error A factor is computed contributing then it is expected that M S A value will be more than expected value of M S E.

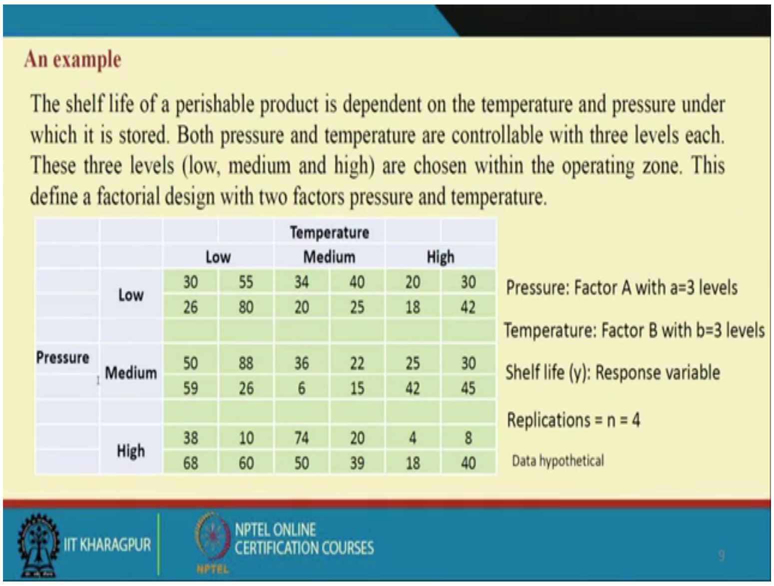
Similarly, your M S B also so if this factors effects are significantly contributing, then the, their contribution should be more than the error M S E is giving you the random

random variation value. So, it should not be just random like error random variable it should be it should be something more.

So, as A result expected value of M S A if you compute it is nothing, but expected value S S A by a minus 1, this quantity will become sigma square plus b n tau i square divided by a minus 1. Similarly expected value of M S B it will be expected value of S S B by B minus 1. So, this will become sigma square a n sum of beta j square divided by b minus 1. Similarly expected value of M S A B it will be expected value of S S A B by a minus 1 and b minus 1 and this quantity will become sigma square plus n double sum tau beta i j square, divided by your a minus a minus or 1 into b minus 1. And interestingly M S E this is expected value of S S E by a b into n minus 1 that value is sigma square.

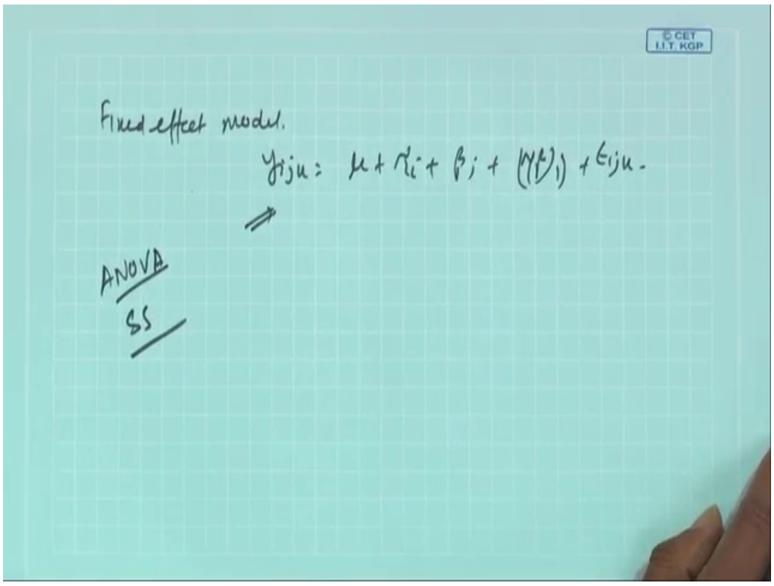
So, if each of the factors are contributing effects are contributing then this should not be 0 for A factor this should not be 0 for B and this should not be 0 for A B they should be significantly more otherwise that they will all be equal to M S E.

(Refer Slide Time: 24:47)



So, we will see one the example what we have set in last class here. The y is the shelf life of a perishable product it is stored under different temperature and pressure medium and then we have seen that what is the shelf life? And this data are obtained based on the experiments this is the, these are all hypothetical data and it is basically experiment we assume that this experiment we have conducted and as data will be random in the sense that the experiments are conducted using proper randomisation. And if this is the data now whether this model whether this model is this model mean y i j equal to the model we have discussed earlier.

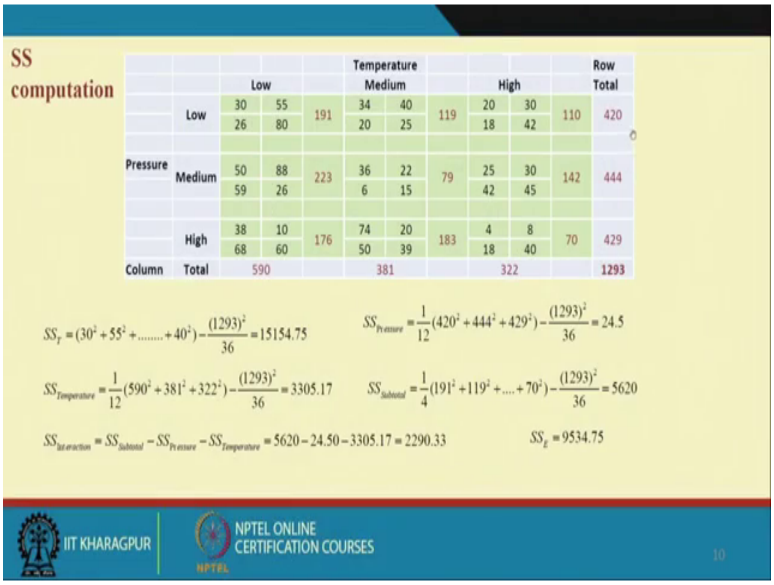
(Refer Slide Time: 25:42)



Now, whether the fixed effect model fixed effect model that is y i j k equal to mue plus tau y plus beta j plus tau beta i j plus epsilon i j k this is able to explain these data or not.

Primarily we are interested in tau i with this this this parameters whether they are really contributing in explaining the variability of y i or not that is the issue. So, we are using anova. So, using anova we are an anova mean that partitioning some squares we are seeing this. So, let us compute this.

(Refer Slide Time: 26:26)

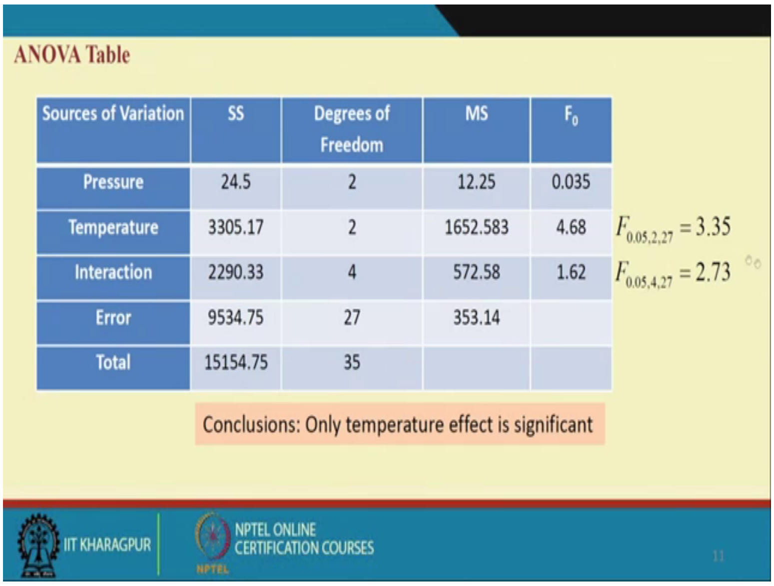


So, what are the things we require for computation we first require row total. So, with reference to low the row total is 420 with reference to pressure medium row total is 4 40 with reference to pressure high row total is 429. So, there are 3 rows because pressure has 3 levels similarly there will be 3 columns because temperature is 3 levels. So, column 1 total is 590 column 2 total is 381 column 3 total is 322. So, this is y i dot dot this 1 is y dot j dot, another total will be cell total y i j dot is first cell 191 second cell total 222, third cell total here like this 176.

Similarly this total this total all totals are computed. And once total are known with means you are in A position to compute the S S because you have seen the formula. Now what is S S T S S T is that sums square of all the observations here A equal to 3 B equal to 3 and n equal to 4. So, 3 cross 3 cross 4 is 36. So, all 36 as in experimental data is values are squared and then we have taken the sum minus grand total is 1 2 9 6 this is the grand total 1 2 9 6 grand total square divided by a b n that is 36 this is giving you the S S T value.

So, in the A for S S T in S S pressure we have 3 row totals 420 square 444 square for 20 9 square and divided by b n. So, n is 4 b equal to 3. So, 12 and minus the correction factor it is giving you 24.5 in the same manner temperature square sum square is 3 3 0 5.1 7 S S subtotal considering 1 2 3 4 5 6 9 subtotals are there this subtotals are there. So, we have consider 1 by n and all those subtotals minus this is giving you this value. And then S S interaction is the subtotal minus temperature and pressure it is giving you this value and S S E is basically S S T minus your S S S S T E is these minus S S S S subtotal is S S subtotal is 5 6 2 0. So, this is giving you this value.

(Refer Slide Time: 29:16)



Sources of Variation	SS	Degrees of Freedom	MS	F_0
Pressure	24.5	2	12.25	0.035
Temperature	3305.17	2	1652.583	4.68
Interaction	2290.33	4	572.58	1.62
Error	9534.75	27	353.14	
Total	15154.75	35		

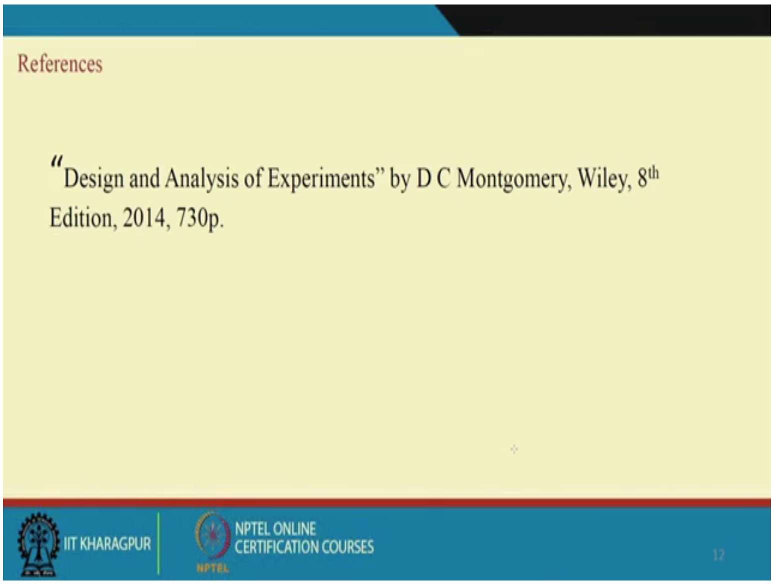
Conclusions: Only temperature effect is significant

So, all those values when I put here we got and with the appropriate degrees of freedom we got that M S for pressure is 12.2 5 temperature is these interaction is these error is these and F 0 values for pressure it is very low for temperature it is 4.6 8 for interaction it is 1.6 2.

Now, the pressure and for pressure and temperature the threshold threshold F will be F if we consider alpha equal to 0.0 5 F 0.0 5 2 7 20 7 is 3.3 5 and for interaction it will be 427. So, it is 2.7 3. So, as A result what is happening here you see that temperature effect is significant other effect are in significant. So, what is our conclusion; conclusion is only temperature effect is significant pressure and this effect is not significant.

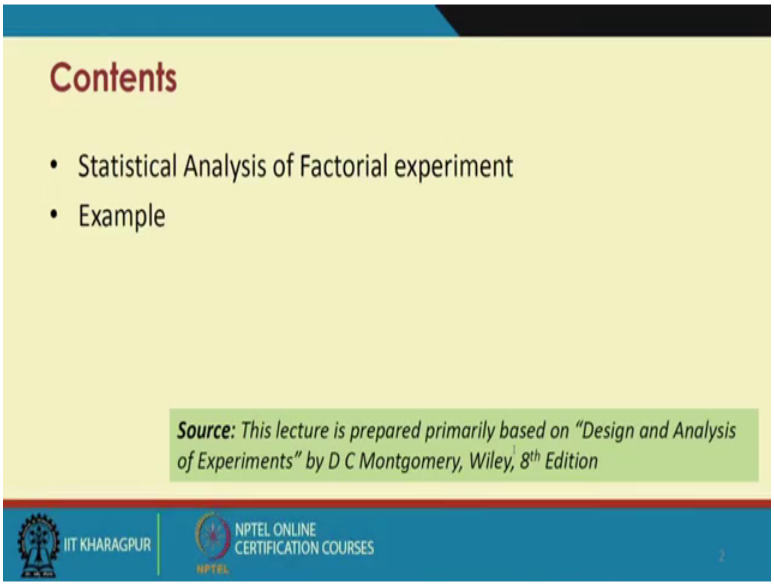
So, this temperature factor to be controlled and then the life can be life preparation will good, can be controlled also.

(Refer Slide Time: 30:24)



So, these things we have the theory portion and also the all the conceptual things we have taken from this book and as I already told you in the beginning that we have we have developed the PPT based on the Montgomery book.

(Refer Slide Time: 30:38)



I am sure that you will be A you are in A position now to given A factorial experimental data set you are in A position to identify do conduct the anova study and conclude, whether what are the factors that are contributing, whether the main effects are contributing or interaction effects are contributing or main interaction both effects contributing and what are the effects contributing what are not contributing.

So, in next class what will see we will see the estimation of model parameters model adequacy test; test of assumption and the sample size regular mean?

Thank you very much.