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

Lecture - 13
The Analysis of Variance (ANOVA) – VI

Welcome back my dear friends. A very good morning, good afternoon, good evening to all of you I am Raghunandan Sengupta from the IME, Department, IIT, Kanpur and this is the TQM II course under NPTEL, MOOC and this is the 13th lecture. 13th lecture means the third lecture in the third week.

So, already you have taken two assignments and later on just for your information on the assignments with the solutions would be put up on the net, on the NPTEL website. As it has been done for project management and the TQM would be done for the TQM 1 which I have delivered through the NPTEL, MOOC concept.

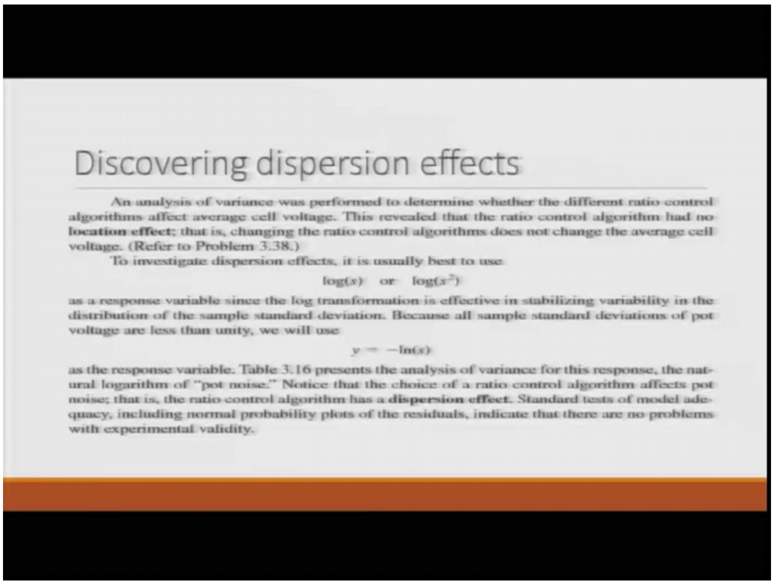
So, if you remember we are discussing about the different tests which can be done therefore, the ANOVA, the Bartlett tests and they later on try to discuss the balanced, non balanced structure models. No, I am not using a word structure models in its true sense, but the models where we want to test the under the null hypothesis that all the mean values are same mean various means for the treatment I am using the word treatment for the example which we did.

And then considering sigma squares are equal and the null hypothesis in the formal case would be mean values are not equal for any two of them. Then variances being not equal to 0 and all and then we also considered tau, tau was the basically the overall dispersion concept of the mean of any one treatment from the averages of the averages or the mean from the mean and we continued discussing those.

And then we later went on to the aluminium smelting problem and discussed maybe not in detail, but give a very good basic idea about how the test considering the capital T tackies statistic could be used to check the validity that the mean values are same for different levels of treatment. If you remember I am again coming back to the initial example 160, 180, 200 and 220 and how whether the power change would have a change in the mean values of the etching and so on and so forth.

So, now we want to basically discover or understand about the dispersion effects dispersion be technically basically the movement in and around the mean value. But obviously, standard deviation could be judged with respect to the median values also, but we generally use standard deviation which is which is respect to the mean value.

(Refer Slide Time: 03:07)



Analysis of variance ANOVA was performed to determine continuing with the last example whether the different ratio control algorithms affect average cell voltage. This revealed that the ratio control algorithm, so based on whether you want to control the overall effects. This is the smelting example not the initial former example we have been discussing long for a long time that is the etching problem on the vapour thin circuits.

So, continuing the reading. So, this revealed that the ratio control algorithm had no location effect that is changing the ratio controlled control algorithm does not change the average cell voltages and we can refer to as I said the montgomery problem which was and in this book it is 3.38.

To investigate dispersion effects it is usually best to use either the log log will scale it down values of the reading which you are going to take a log of SS square. So, as a response variable, one of the transform best transformation is the log transformation in order to find out whether they are as close as possible to the normal distance.

Obviously, one of the main drawbacks which is without any in mathematical background you can say is that because normality considers distribution from the negative infinity to positive infinity, but log values are definitely not in the negative sign. But still the transformation does work and gives us good result when you want to do the transformation and find out the adequacy how close they are to the normality assumptions.

So, if it is usually best to use log of s or log of s square as the response variable since the log transformation is effective in stabilizing the variability because if you see if you have done the log log scale graphs in your colleges in school, so they were basically trying to squeeze in I am using the English word squeeze in the x axis or the y axis proportionally considering the readings which you are going to take.

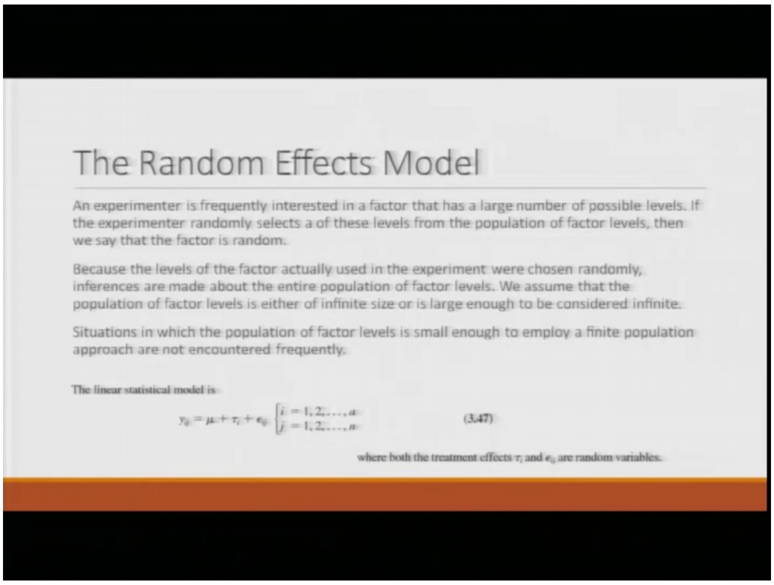
Because in that case if you have say for example, reading of 100 and you a1re using say for example, log base 100 as it is so obviously, you have to measure 100 units on the x axis. But if we use the log scale, so to and onto the base 10, so it would be log 100 to the base 10 and that would basically be 2. So, you can use that value. So, 2 would be the corresponding value on the log scale for 100 with respect this, the value of 100 on the non transformed scale.

Similarly, you can do it for 1000, 2000, 3000 or 10000 whatever it is it will be scaled on accordingly. So, and obviously, it would mean that as you are squeezing in the values making more nearer and nearer so obviously, the dispersed ability or the so called variability would be much less. So, thus stabilizing variability in the distribution on the samples and which is the standard deviation; because all sample standard deviation of the voltage are less than unity and hence we use the transformation of y is equal to minus of log and basically transform it accordingly.

Table 3.16, if you refer to the book represents the analysis of variance for these responses, the natural logarithm, the pot noise or for the smelting process. Notice that the choice of a ratio control algorithm affects pot noise that is the ratio control algorithm has a dispersion effect.

Standard test of models adequacy including normality from normal probability plots of the residuals indicate that there are no problems with the experimental validity. So, they basically agree with the underlying facts on has already discussed.

(Refer Slide Time: 07:06)



Now, we come to the random effect model. So, the random effect model if you consider the regression part and the random effect model there are subtle differences which we will consider as we process, so progress. So, if you consider I will first tell about the regression model which you all know.

So, in regression model we consider y is equal to some constant consider is alpha constant can be there may not be there plus beta 1 x 1. So, beta 1 is basically the so called coefficient which technically gives you the rate of change of y with respect to x 1 provided all the other x 2's to x k's are constant. So, we are considering that there are in the variable variables x 1 to x k which are all independent; that means, there is no covariance between x 1 to x k considering two at a time with both of them are different,.

But obviously, there is a variance for each x x is i is equal to 1 to k and the relationship between y and x can give basically given by the multiple linear regression model. And obviously, as all the x's are normal with a certain fixed mean and certain standard deviation without any covariances and if then y would also be a normal.

And there is an error term also on to the right hand side; that means, you have to add that error term with respect to alpha plus beta 1 hat because that is the expected in the estimated value which you will find for beta 1 then into x 1 plus beta 2 hat which is again the estimated value of beta 2 into x 2 plus dot dot till the last value which is beta k hat which is again the estimated value of the k of the kth beta value x k plus epsilon. So, each reading which you take there would be an error term.

And obviously, the error terms has some underlying assumptions number 1, the average value of the error terms or the expected value of the error terms is 0 it is normally distributed and the variance is constant or on some fixed value or 1. Constant is very important because if the errors terms are not constant means the errors are affecting each other. So, that becomes very difficult to for us to basically handle the problem in the multiple linear regression setup; obviously, there are other models also.

So, with this background let me again come back to the discussion which we are having. So, this is the random effect model the experimenter is frequently interested in a factor that has a large number of possible levels.

If the experimenter randomly selects a of these levels from the population of factor levels then we say that the factor is a random because the levels of the factor actually used in experiment were chosen randomly influences are made about the entire population of the factor levels. So, assume that the population of factor levels is either of infinite size or is large enough to be considered as infinite. So, the actual sample size is huge.

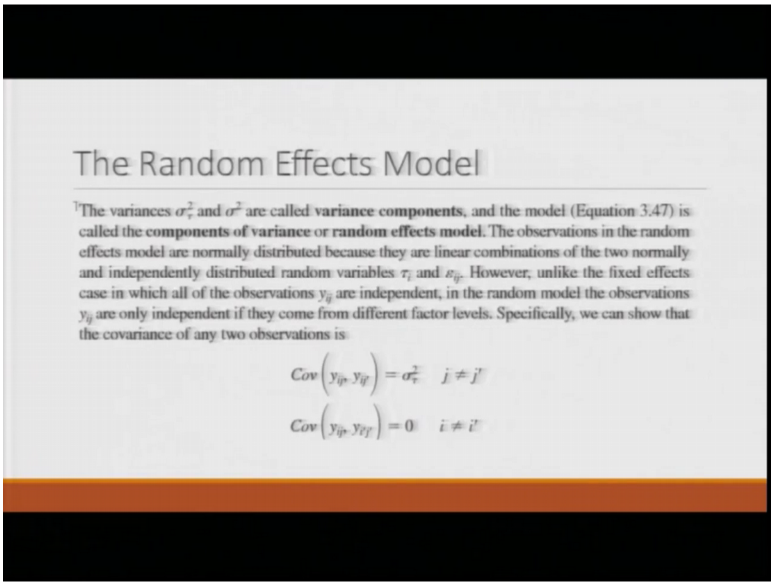
Situations in which the population of factors is small enough to employ a finite population approach are not encountered frequently. So, the linear statistical model which we are considering generally for the random effect models would be like this. So, they would will be y ij, if you remember go back to the example of the etching problem where the voltages were there and there were 4 voltages a was 4 and each sample size in each of the a was basically 1 5.

So, I am repeating this in order to make you like come back to the example which he had been discussing time and again so that will give you much better feel that how things are progressing for this discussion.

So, y ij, where i changes from 1 to a, j changes from 1 to n is equal to mu mu is basically the average of the average or the mean of the mean plus tau i, tau i is basically for each of the m the etching or the treatment they would be some dispersion plus epsilon ij which would basically be dependent on the sample size and or sample number or what is the observation in the sample and also for the treatment or the etching.

So, i is equal to 1 to a, and j is equal 1 to n.

(Refer Slide Time: 11:30)



The variances which is sigma square suffix tau and sigma square are the variance component and the model is called the component of variance or random effect model so obviously, there would be two type of x coming out one for inter and one is for intra.

The observations in the random effects are normally distributed as which is on the very basic assumptions because they are linear models of the two normally and independent distributed variables which are tau y's. And so tau y's is also normally distributed with a certain mean and certain standard deviation and the errors epsilon ij also is normally distributed with a certain mean and a certain standard deviation.

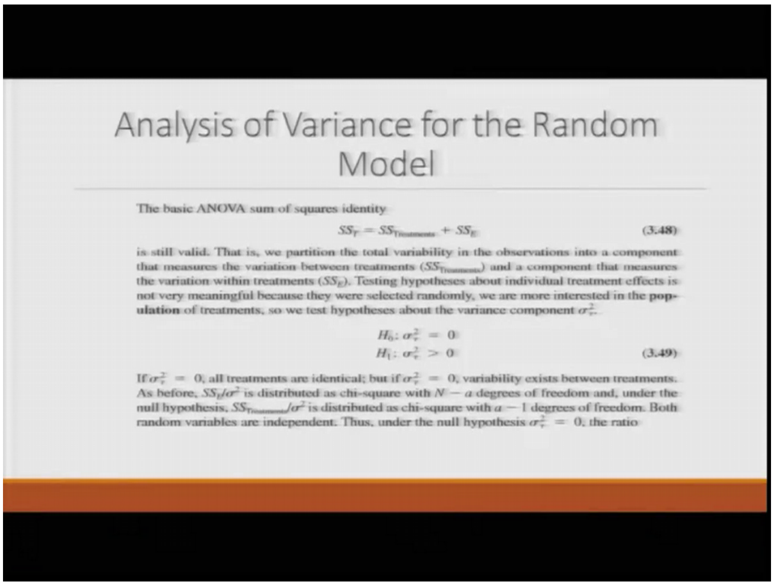
However, unlike the fixed effect in which case all the observations y ij are independent in the random model the observations y ij are only independent if they come from different factor levels, but if they are coming from the same factor levels are they are not independent.

Specifically we can show that the covariances of any two observations now these are what are the what I will discuss if you see the slide the covariances of y ij this ij whatever I am mentioning are all on the suffixes and bit the other values being y ij prime. So, j and j prime are some values in between 1 to n and they are not equal. So, in that case go variances of y ij and y ij prime j and j prime are not being equal is given by sigma square suffix tau. So, this is basically not changing dependent. So, there is no suffixes for tau, so remember that.

And the variances say for example, of the values of the covariances of the values covariance of i prime j prime and i and j, where i and g and i and i prime and j and j prime are not equal to 0 0 which means that if you are considering the way the covariances between two effects or treatments then obviously, that covariances would be 0; that means, one of the effects which you have along the rows is either i is equal to 1 or 2 or 3 till a are not affecting each other.

But obviously, there is some covariances between the observation which we have along the row which is say for example, for i is equal to a there would be covariances and that would basically be fixed which is given by sigma square suffix tau. And obviously, you can say that y naught tau suffix i, so obviously, that would be much more general case.

(Refer Slide Time: 14:15)



The basic ANOVA model would be sum of the squares identity, again same things sum of the total squares would be divided into the sum of the squares for the treatments and the sum of the squares of the error, so add them up.

So, they would be equal to the sum of squares of the total. And obviously, that they coming back to the same ANOVA table which you have considered; that means, the degrees of freedom should also add up and we will consider that also as the underlying the basic principle based on which we will proceed.

So, this is still valid this model that is we partitioned the total variability in the observations into a component that measures the variations between treatments and a component that measures the variations within the treatment. So, that is what I said the inter and intra. So, between the treatments will be given by SS suffix treatments and within the treatments would be given by SS suffix E errors.

So, testing the hypothesis about individual treatment effect is not very meaningful because they were selected randomly. So, we are more con interested in the in the population of the treatment. So, we test the hypothesis that sigma square suffix tau is equal to 0 with respect to the fact that sigma square a suffix tau is greater than 0; that means, we are assuming that the ith, ith suffix for tau are not present here more note that carefully.

So, if sigma square suffix tau is equal to 0 that means, all treatments are identical, but if if tau sigma tau is 0 variability would exist between the treatments as before see in the sum of squares of the errors divided by sigma square would be a chi square distribution if you remember we did discuss that for the normality normal distribution.

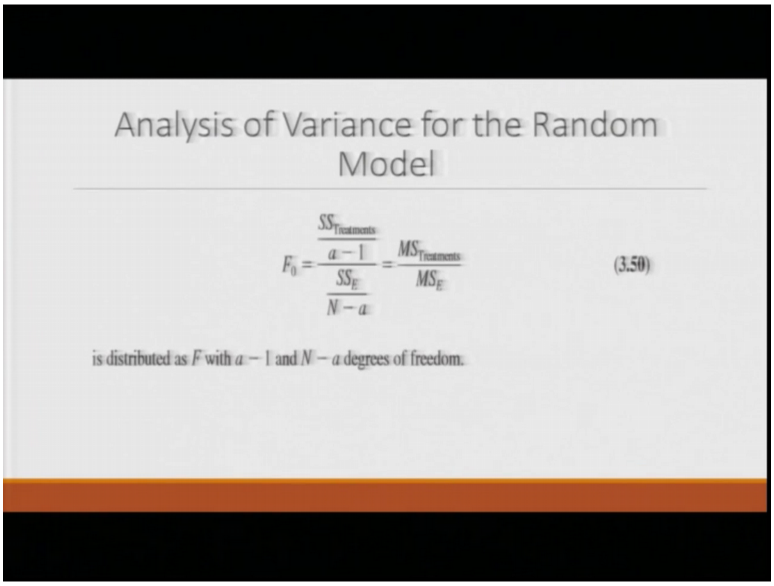
The standard error square which is the standard variance of the of the sample with respect to sigma square which is basically the population variance are basically chi squared distribution with the corresponding degrees of freedom, where the degrees of freedom.

If you remember I am going back to the actual discussion which we did for almost 7 lectures in the beginning related to the distributions. So, those degrees of freedom would be dictated by whether the population mean is known or not known and if you remember we did you go into the concept of using s without the dash s with the dash and all this concepts.

So, again coming back; So, it will have basically a degrees of freedom of capital and minus a we know what is capital N is small n in to a degree with. So, capital N is degrees of freedom and under the null hypothesis the SS of the treatment by sigma square would be chi square. But now with the degrees of freedom of a minus 1 because now we have n number of observations are one degrees of freedom is being lost because you are trying to basically very simply consider that trying to use the average of the average of the mean of the mean of the sample to be the best estimate of that value which you remember mu, where mu was being replaced by y bar dot dot; dot dot means basically the first dot is for summation of all the treatments and second what is summation for all the observations in one treatment.

So, both the random variables are independent as and as I did mention thus under the hand null hypothesis the sigma square suffix tau is 0.

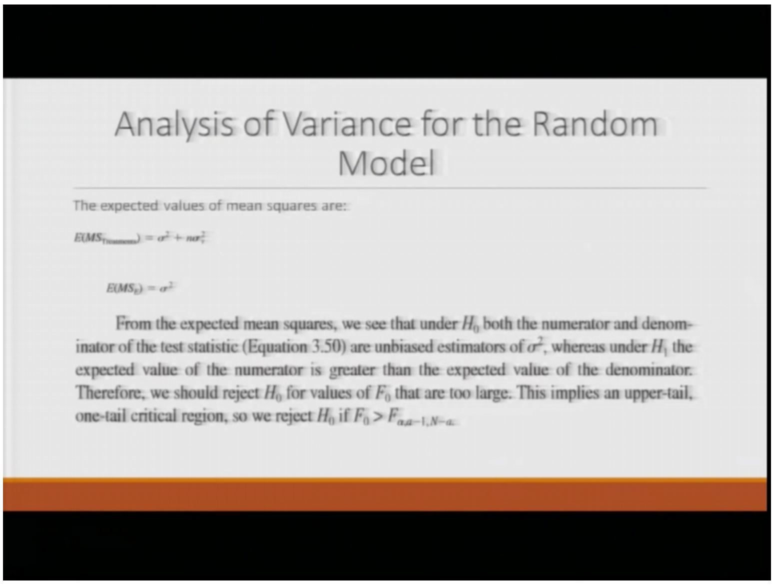
(Refer Slide Time: 17:41)



And in case the ratio is given by in this value which is basically the equation 3.5 we refer to the book, which is the F distribution value if you remember F distribution S the ratios of the so called standard deviations of the populations. So, that is SS treatment divided by its degrees of freedom and in the denominator you have the SS sum of square of their errors divided is degrees of freedom.

So, that comes out to be mean square of the treatment divided by mean square of the errors. So, along and their degrees of freedom would be, obviously, what is the degrees of freedom the numerator is a minus 1. So, that will be the first degree of freedom comma second degree of freedom is capital N minus a.

(Refer Slide Time: 18:27)



So, now considering the further discussion analysis of variance for the random models the expected value of the mean squares would be given. So, if any considering the mean square of the treatment, so there would be two such variances one would be coming from the overall variances with sigma square without any suffix because it implies the errors are independent marked my word because this I have been depending time again plus the sigma square suffix tau into n.

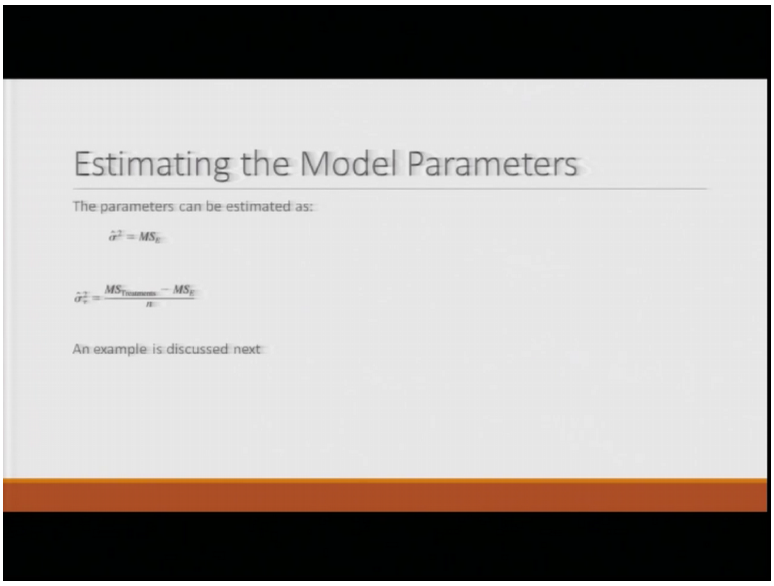
So, that would basically be given and so that that would be considered in such a way that that will give me the total variance and the mean square of the errors would be given by sigma square which is obviously, true because I did mention that the mean square of the errors is sigma square and they are independent of each other and they are constant.

From the expected mean square we see that under H naught which is the null hypothesis both the numerator and the denominator of the test statistics are unbiased estimate of sigma square. Whereas, under H 1 the expected value of the numerator is greater than the expected value of the denominator therefore, we should reject H naught for values of F 0, F z, F z naught F naught means or F o means basically the null hypothesis.

So, that which are the large in values this implies an upper tail one critical region. So, we will reject F under H naught F naught is greater. So, it is basically the greater than sign would be F and mark the suffixes it will be alpha because that is this area still to be covered on to the right because and if you remember still to be covered if you are under the left hand side will be 1 minus alpha depending on if it is a one sided test. And if it is both sided it in that case it will be alpha by 2 and 1 minus alpha by 2 like coming leaving all those discussions once again because I have been repeating that please excuse me in order to make you understand.

So, if you come back to this formula, the suffix is alpha because you are onto the right hand side. And what are the degrees of freedom? If you remember in the numerator the degrees of freedom was a minus 1. So, that would be the first degrees of freedom for the F distribution and in the denominator here at capital N minus a. So, that would be the second degrees of freedom for the F distribution. So, if you note down the formula so that is F which is the F distribution the suffixes are alpha comma a minus 1 comma capital N minus a.

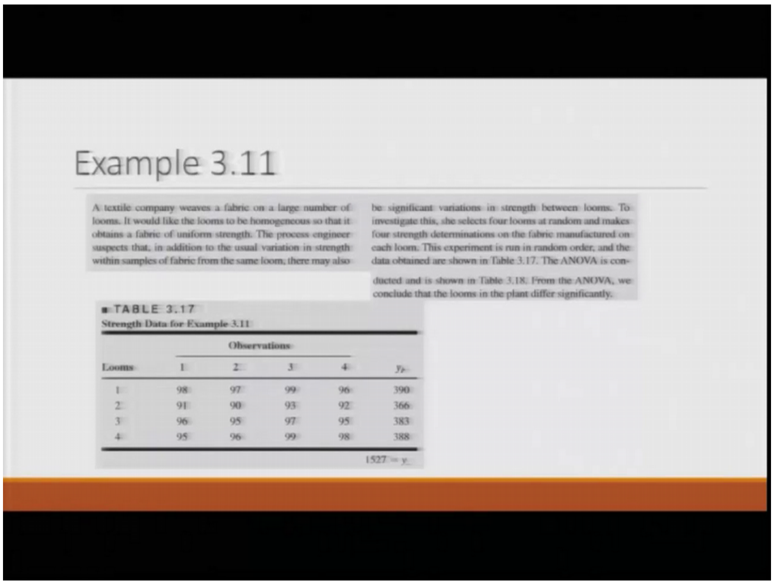
(Refer Slide Time: 20:56)



Estimating the model parameters; So, the model parameters can be estimated using the estimated value which will give me sigma hat square which is the mean square errors and sigma hat square suffix tau would be basically if you remember sigma those were the concept of the of the variances with respect to inter and intra. So, that would come out to be mean square treatment means square error errors divided by n n is the sample size.

So, now, we will just simply consider an example in order to basically go into the depth again. So, repeatedly we will try to discuss the same idea. So, that it becomes very clear to you.

(Refer Slide Time: 21:35)



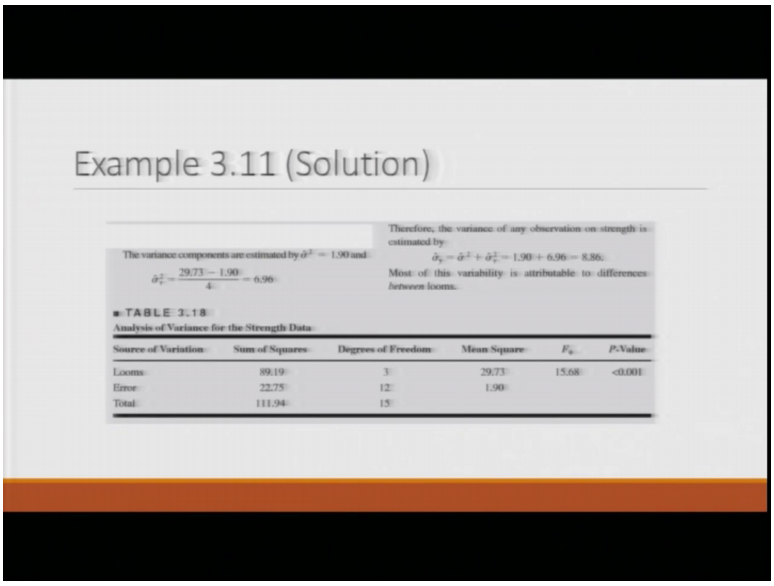
So, the example is as follows a textile company weaves a fabric on a large number of looms. So, there are different looms which of the work is being done, so like the treatments problem. It would like the looms to be homogeneous so that it obtains a fabric on uniform length.

The process engineer suspects that in addition to the usual variations in strength with within samples of fabric from the same loom, there may also be significant variations in strength between the looms. To investigate this the process engineer she basically selects 4 looms at a random which is a 4 and makes 4s strength determination on the fabric manufactured on each loom; that means, N is equal to 4; that means, capital N is equal to 4 into 4 is equal to 16.

This experiment is run in random manner and the data obtained are shown in the table given below in the same slide which is table 3.7 and we want to basically do some tests. The ANOVA is conducted as is as shown and which will discuss, and from ANOVA you can conclude that the plant is in the in the looms in the plant differs significantly. So, the observations are loom 1 2 3 4 the first column the observations are given, where the heading for the columns are 1 2 3 4 which is the second, third, fourth, fifth column in the matrix.

The y i got because you are considering i is equal to 1 then average of that value is going to basically sum we are taken is y not a bar we are taking the sum would be 98 plus 97 plus 99 plus 96 which you will be the further first row. Similarly that value 390, similarly the sum for the second row is 366, for the third row is 383 and the fourth row is 388.

(Refer Slide Time: 23:29)

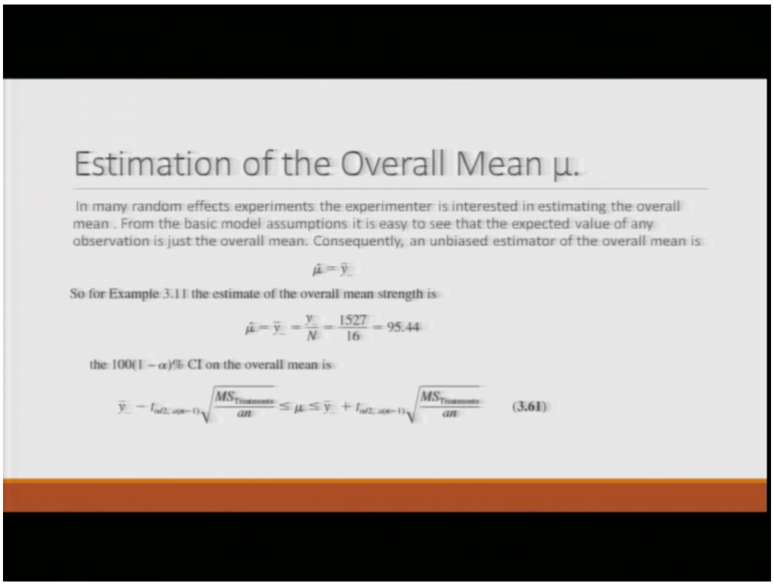


So, various component as estimated would come out to be 1.9 and sigma. So, that is sigma square hat which is the variance of the error and that is constant, sigma square tau value comes out to be 6.96. Therefore, the variance of any observations on strength is estimated to be which is basically sigma square sigma value sigma square suffix y. So, that value would come out to be if you add them up 1.9 plus 6.96 comes out to be 8.86. So, most of the variability is attributed to the difference between the looms.

So, analysis of the variance of this of this data, so in the first column we have the source of variance which would basically be the looms would be the error in each and every group and we basically would be the total 1. So, the sum of the squares for the looms is about I am not reading the decimal parts is 89, for the errors is basically 21 and the total one is 111, so we that the values which you read out are from the column 2. So, degrees of freedom considering that looms would basically be small e minus 1.

So, small e if you remember I repeated it when I was doing the last slide was basically 4, 4 minus 1 is 3 and if you consider the values of say for example, the degrees of freedom for the error. So, what was capital N? Capital N was 4 into 4 was 16, so 16 minus 4 which is capital small a. So, 16 minus 4 is 12. So, that is the value which you have for the error degrees of freedom and the total degrees of freedom for the total part will be 12 plus 3 which is 15. Mean square values are given as 29.73 and 1.9. So, from that the F values comes out to be 15.68 and you can take a decision whether you want to accept or reject.

(Refer Slide Time: 25:28)



So, I have given you the basic concept how you proceed. So, any in any in many random experiment the experimenter is interested in estimating the overall mean from the basic model assumption it is easy to see that the expected value of any observation is just the overall mean. So, that is basically the mean value mu is equal to y bar dot dot that means, average of the average.

Consequently an unbiased estimator for the overall mean as it is given if you see it. So, basically when I am mentioning y bar dot dot that is the best estimate that is replica when you do the concept when you are trying to solve the statistical problem it will give in as mu hat which is the estimated value of mu.

So, for example, for the example the estimator of the overall mean would basically be y bar bar which means you add up all the values divided by capital N which is 16 that value comes out to be 95.44, so the 100. So, the confidence interval considering alpha value the value whatever it can be 5 percent 2.5 percent 10 percent as the problem can be.

Now, the confidence will be I will not go to the value I will discuss why those distributions are kept. So, we want to find out the mean value the confidence level of the mean value which is mu, so on the left hand side would be the lower control right hand side with the upper control.

So, what is the left hand side my left hand side would be mu hat which is y bar dot dot. So, this is value which is given. And this is basically a t distribution because you do not know anything about the concept related to this standard deviation of the population hence you will use the t distribution because you will be utilizing the s concept.

So, now the t values if you remember they are symmetric, so hence if you use one minus alpha by 2 or alpha by 2. Now, mark my words it is alpha by 2 and 1 minus alpha by 2 because they are both sided because this is a confidence level and the t distribution is symmetric for larger values hence we can consider that the plus and minus sign would be replaced in place of 1 minus alpha by 2 and alpha by 2.

So, if you consider it is minus t alpha by 2, now the degrees of freedom would obviously, be what is that capital N minus a. So, that is a into n is capital N and minus a is there which in the in the suffix it is given a in the bracket n minus 1.

And if you remember the values of t distribution technically remember the formula in general you do not have to by heart, it is basically the standard deviation of the sample of standard error divided by the square root of the observations.

So, this is mean square of the treatment divided by a into n square root of that. So, on the right hand right hand side it will be exactly the same only the minus sign which is in front of the t distribution will be replaced by the plus sign because this is on the right hand side because you are taking the both the hand test other two tailed test.

So, with this I will end this 13th lecture and continue discussing more of one ANOVA, the random models and the factor models later on which will definitely give you much more broader perspective about this design of experiments and such that you are able to appreciate that in a much bigger scale.

Thank you very much and have a nice day.