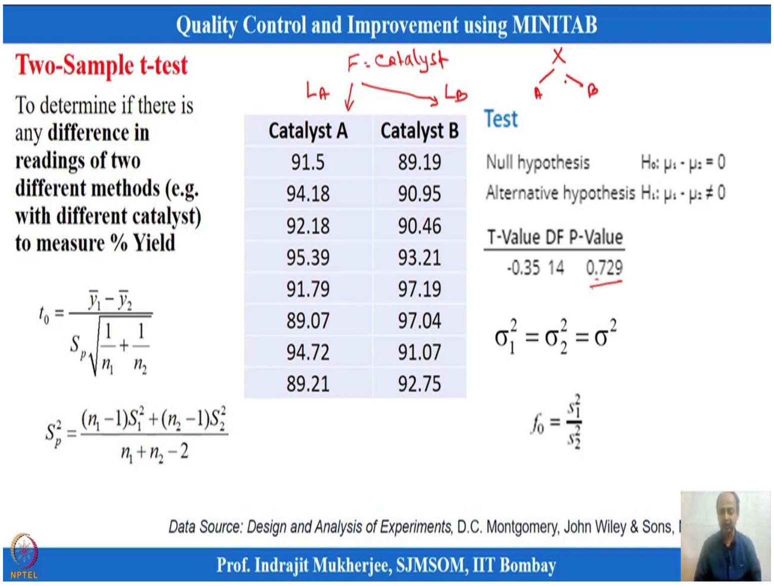
Quality Control and Improvement with MINITAB
Prof. Indrajit Mukherjee
Shailesh J. Mehta School of Management
Indian Institute of Technology, Bombay

Lecture - 20
Paired t Test and ANOVA

Hello and welcome to session 20 on our course on Quality Control and Improvement with MINITAB. I am Professor Indrajit Mukherjee from Shailesh J. Mehta School of Management, IIT Bombay. So, last session what we were talking about two -sample t-test and in what scenarios to use them. Today, we will extend that concept of two-sample t-test and in what scenario it is conducted, we will try to see.

(Refer Slide Time: 00:44)



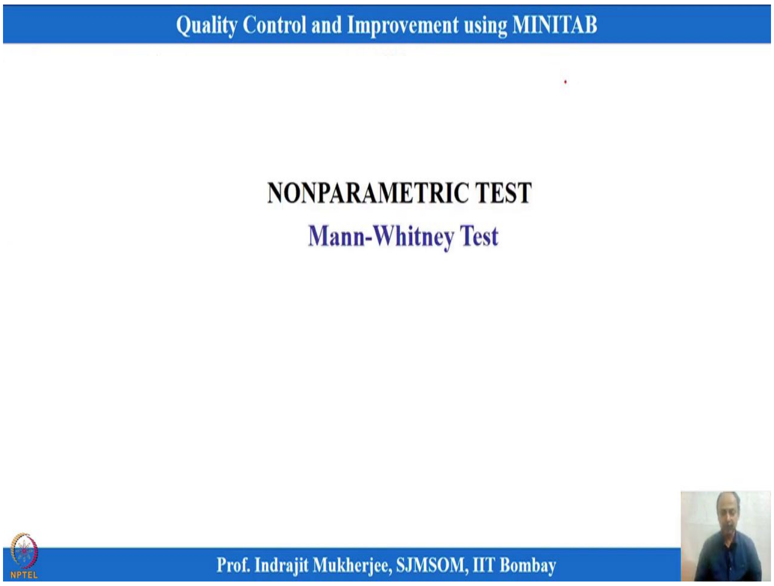
So, this was a Two-sample t-test example that we had dealt with; there are two catalyst over here and we want to improve the yields over here and we want to see that which catalyst is effective which is not effective. We saw that both catalysts are effective.

So, in this case what happens, because of the p-values $(p > 0.05)$ that we are getting over here both the catalyst are equally giving the same means or we can say that . So, I cannot reject the null hypothesis basically over here ok.

So, there is no way improvement has happened with a different catalyst. So, this is basically a one factor at two level experimentation. So, what is one factor two level experimentation? This is factor will be equals to catalyst and there are two levels level A and level B like that.

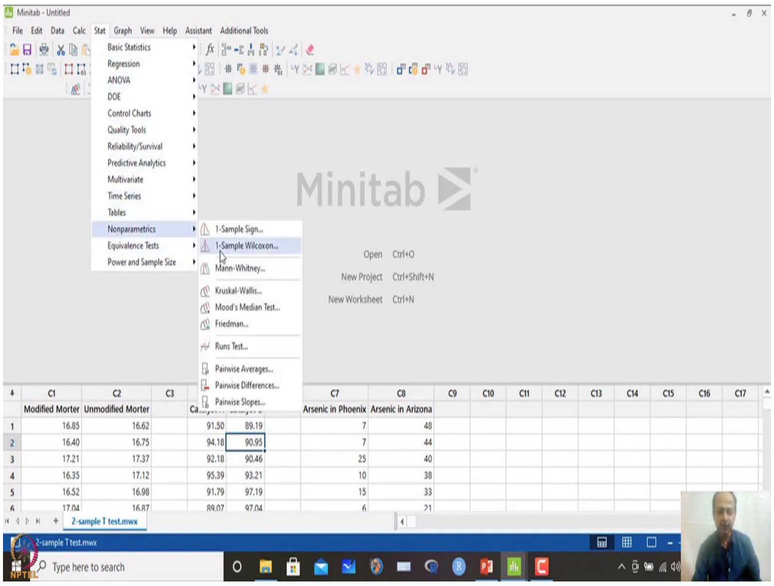
So, when we have one factor and we have over here, catalyst A and catalyst B. So, this experimentation was conducted like that and we wanted to see the effectiveness of catalyst over here. So, two-sample t-test is the starting point of experimentation what we do ok.

(Refer Slide Time: 02:11)



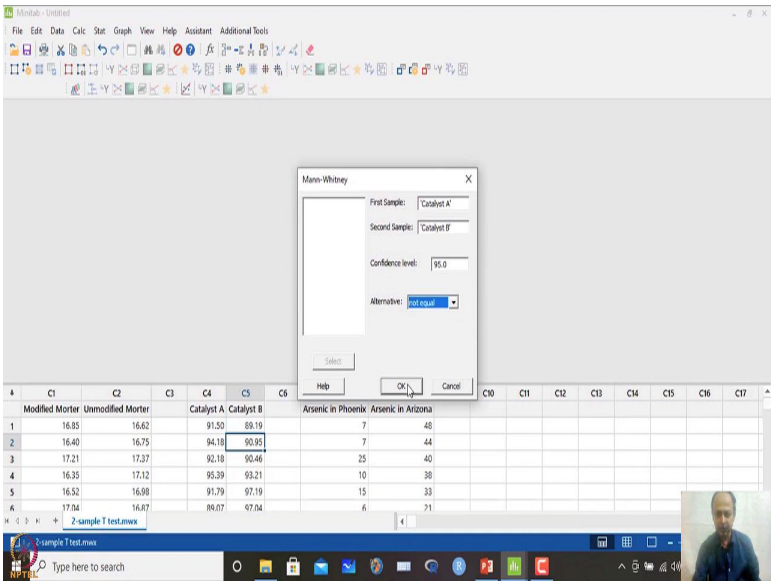
So, in case the distribution assumption fails over here; like catalyst A and catalyst B what we have to do that? We can just check ok. So, I will show you the non-parametric options to this. So, two-sample t-test, we have proved that everything is working over here, but assuming that distribution assumption fails over here.

(Refer Slide Time: 02:30)



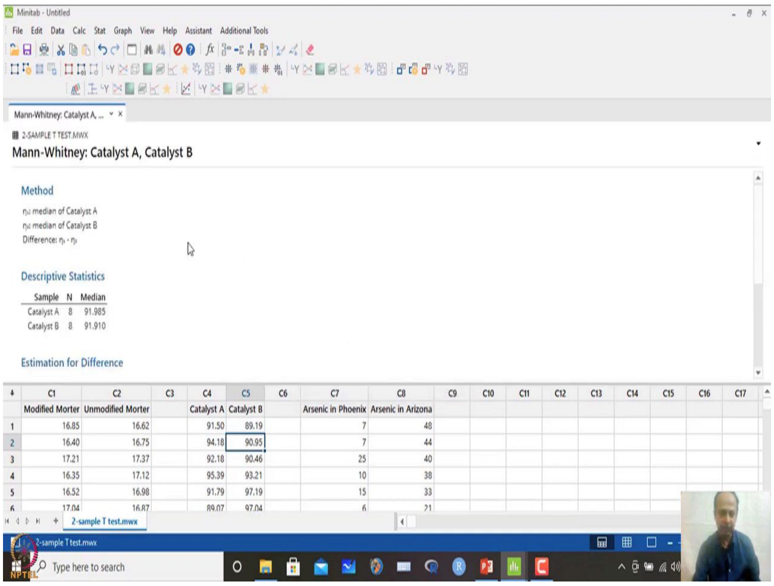
And in that case, what is to be done we want to see. So, two-sample t-test file we will open and. In that case, this is the catalyst A and catalyst B, we can write over here. Catalyst A and this is catalyst B and we want to see the difference between these two. So, you have a non-parametric option over here to check this one. And this case, Mann-Whitney Test is the option that we have where median will be compared

(Refer Slide Time: 03:02)

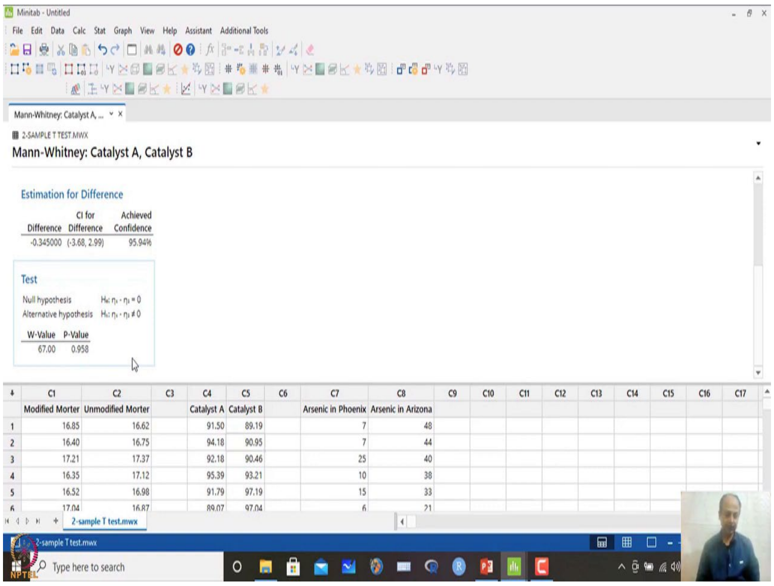


So, one will be catalyst A and the second will be catalyst B, 95 percent confidence level and not equals to condition we want to check. So, I will click ok over here.

(Refer Slide Time: 03:11)

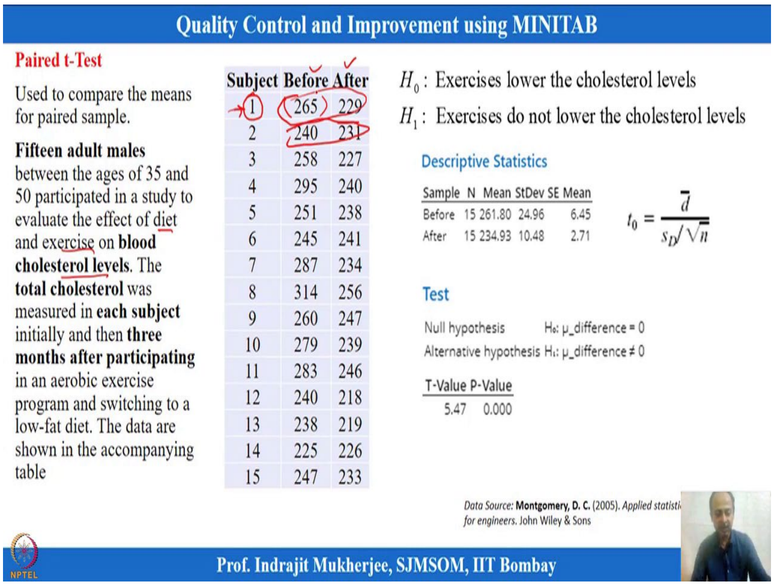


(Refer Slide Time: 03:12)



And these W statistics and p-Values will indicate whether the significant difference exists between level one and level two basically (catalyst A and catalyst B) using median values over here. So, in this case what we are seeing is that we do not see any difference between that. So, non-parametric test also confirms that there is no difference between catalyst A and B ok and the then, we can move ahead with this. So, this is non-parametric for options.

(Refer Slide Time: 03:49)

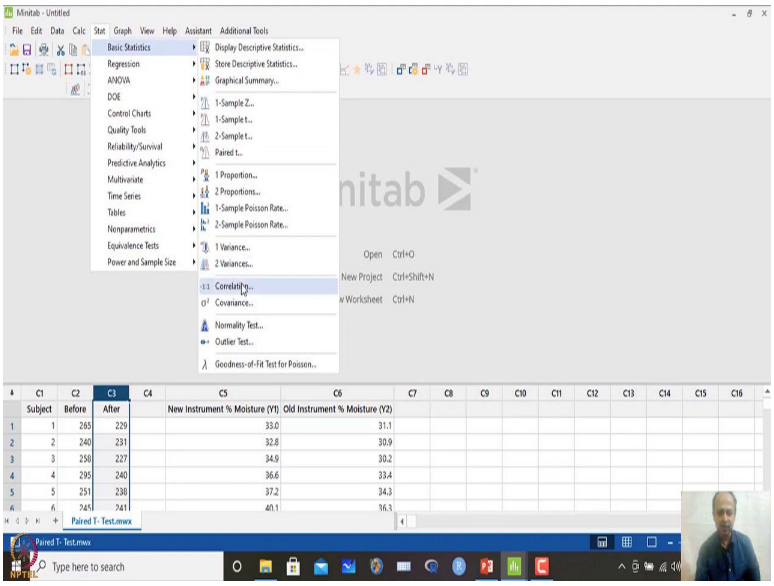


Now, one important thing that we also need to consider when we are doing experimentation and doing two-sample t-test is the independency of the samples that is required. But scenarios can exist that some of the sample observations that we have taken over here are correlated.

As an example, effect of diet and exercise is evaluated on 50 participants. So, before diet after exercise data was recorded. So, we want to see that whether exercise has a positive impact on this lowering the cholesterol levels.

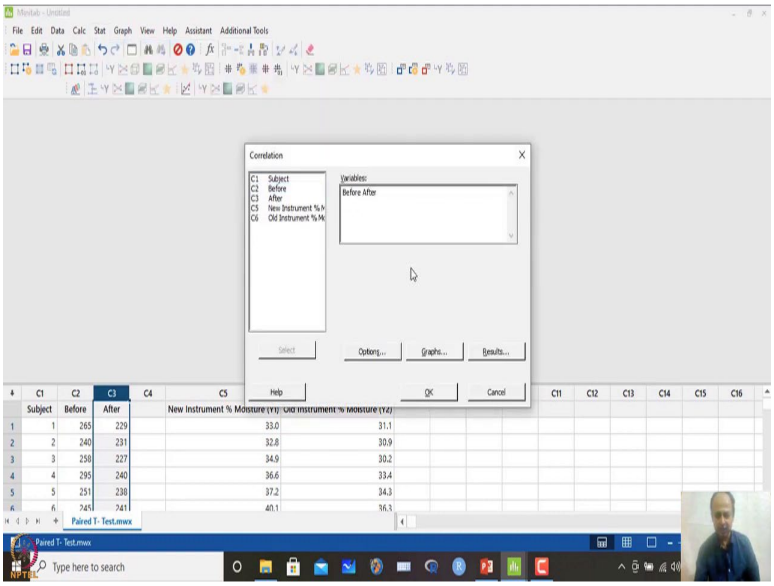
So, total cholesterol was measured in each subjects and then three months after participation, this was again recorded. So, in this case this data over here, before and after experimentation this before and after the that control and all these things exercise is expected to be correlated like that. So, whenever data is generally from the same subject, in that case what is expected that this is a paired data. Basically, this is a paired data from a single sample. So, what is expected is that the data will be correlated and for that two-sample t-test is not the appropriate test. So, here paired t-test is generally recommended where the data is highly correlated or significantly correlated.

(Refer Slide Time: 05:27)



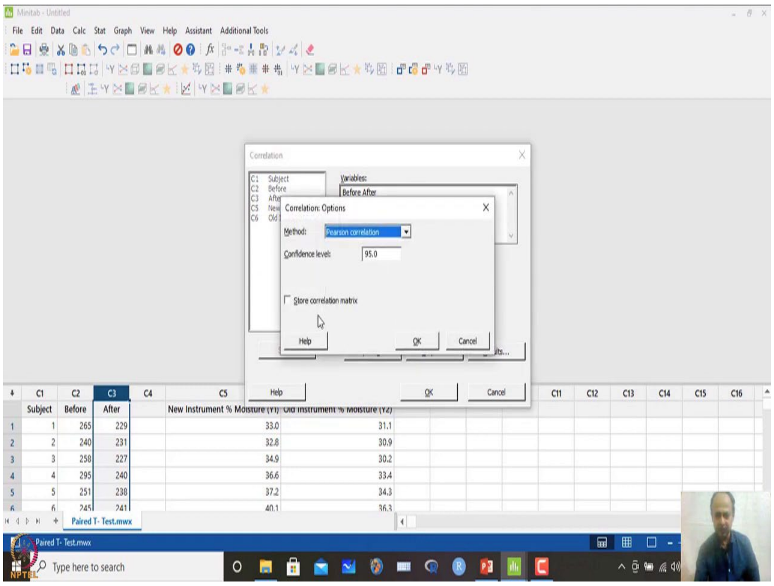
So, how do we see that one? So, we will go to the data set. This is the data set that we are having before and after over here, C2 and C3 column that you see over here, these are the two data set what we wanted to see. And let us see the correlation between this data. Earlier we have seen independency here. Also, we will see basic stat and correlation between the data set.

(Refer Slide Time: 05:46)



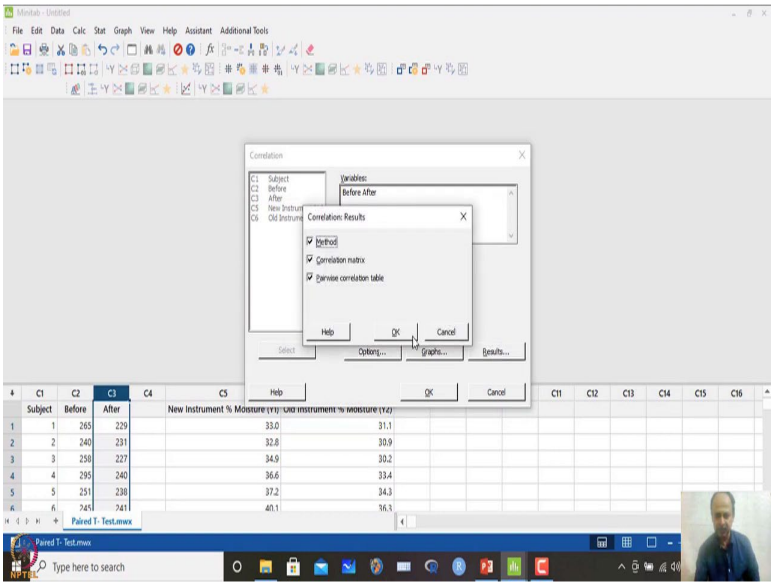
And we expect that there will be high amount of correlation that exists between the data.

(Refer Slide Time: 05:51)



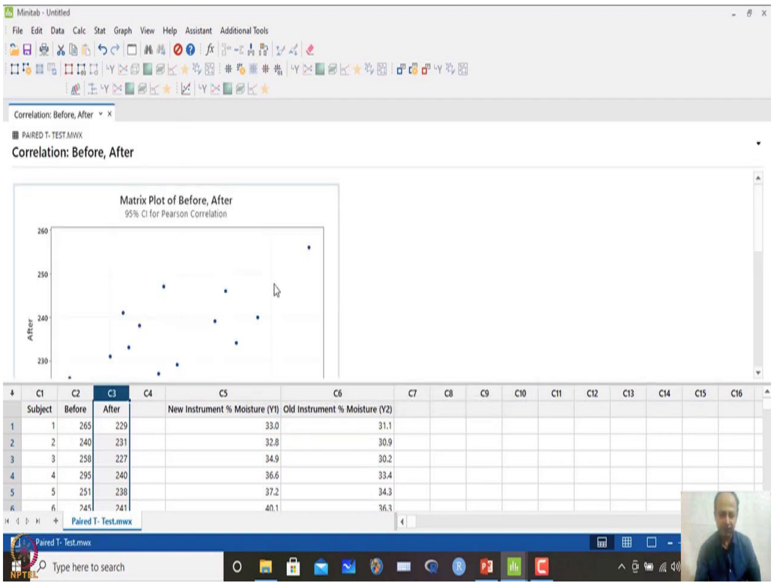
And in options, what you do is that?

(Refer Slide Time: 05:54)

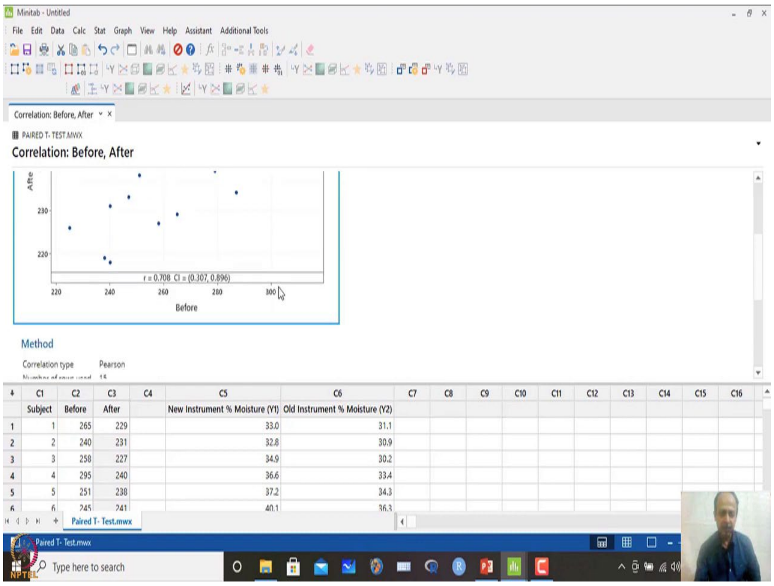


We are keeping this in results we have pair-wise correlation table we want.

(Refer Slide Time: 05:58)

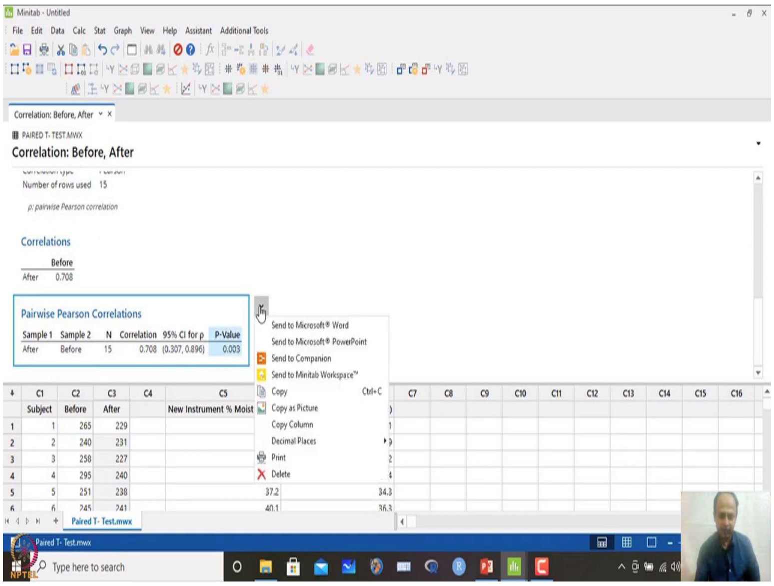


(Refer Slide Time: 05:59)



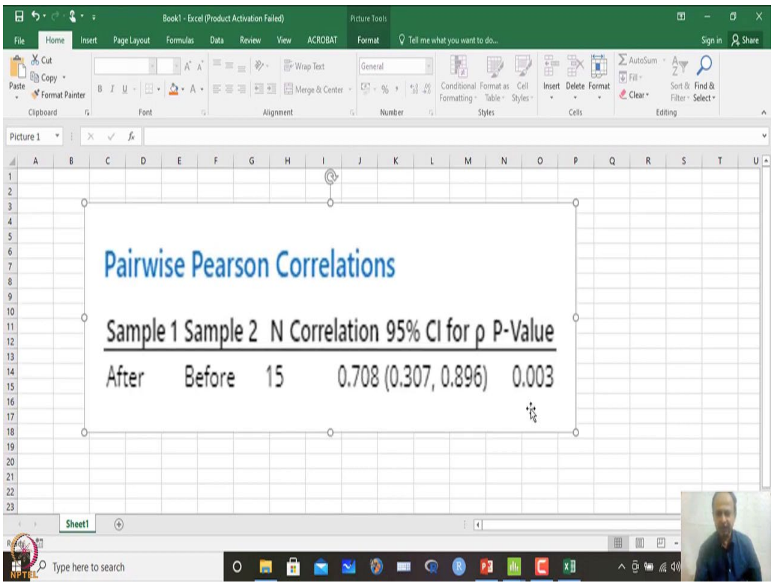
Because, p-Value is required so, when you do this the r value is around 0.7.

(Refer Slide Time: 06:03)



And what you see is that p-Value that is reported over here. So, if I can enlarge this one; I am copying copy as image over here and then I produce then and then I just place it in excel sheets ok, sorry.

(Refer Slide Time: 06:26)



So, this will be placed in excel. Let us say and we want to see over here; whether the whether a fair amount of correlation exist or not. So, I will just enhance this one and I am just copy pasting this image over here. And for that I will just enhance. So, that you can see and the p-Value corresponding is you can see that this is 0.003.

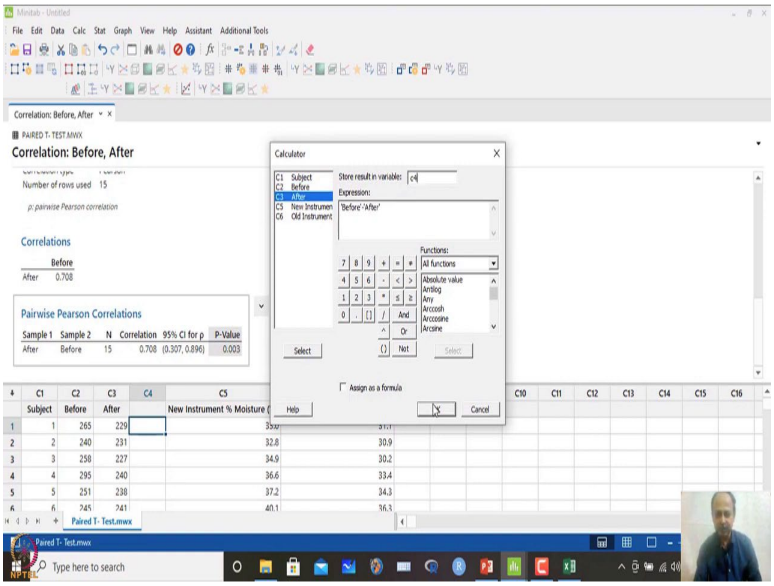
So, this it says that sample 1 and sample 2, which is before and after and before, they are highly correlated over here. So, in this case statisticians suggest is that we cannot go for two-sample t-test over here. Most appropriate test over here is known as paired t-test that is the most appropriate test to check whether there is any difference between these two levels that is before and after like that ok.

So, what they have suggested is that we have to calculate the difference and based on the difference and standard deviation of the difference what we can do is that there is a statistics which is used over here and what you see the formulation of this. So, this is the formulation what you see over here.

So, this is the formulation that is used and d is the difference between this before and after over here. So, the difference average will be taken and then, standard deviation of the difference and square root of number of observations that I have that is 15 observation of paired observation that I am having over here.

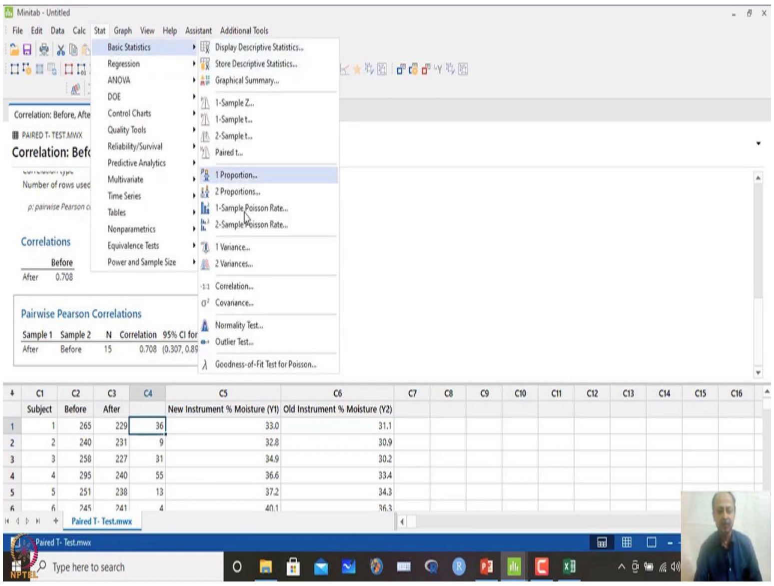
So, based on this I will calculate the t statistics calculated value and corresponding p-value will be calculated like that ok. So, we have ensured that there is high amount of correlation and in that case, what is to be done is that I will go for paired t-test over here.

(Refer Slide Time: 08:01)



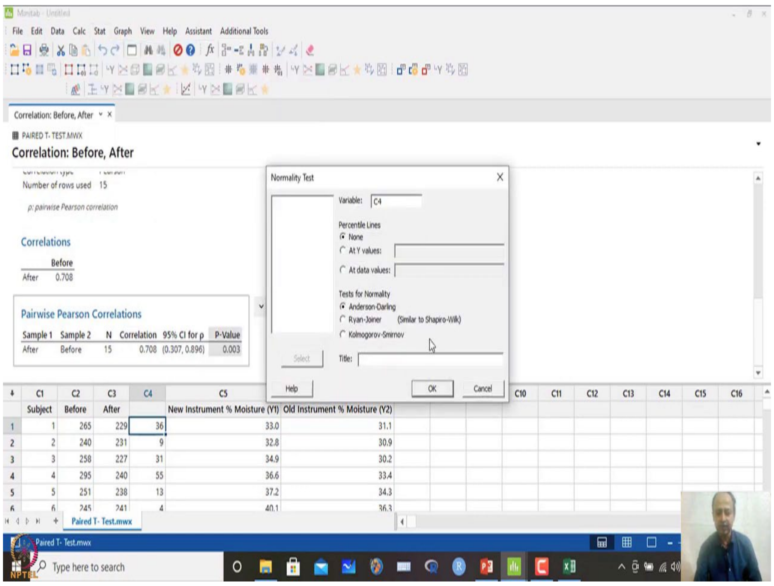
So, this difference can also be calculated. So, if you go to calculator and use calculator over here to calculate difference and save it in C4 column number. So, this is where I will save and I click ok I will find out the difference that is there over here.

(Refer Slide Time: 08:14)

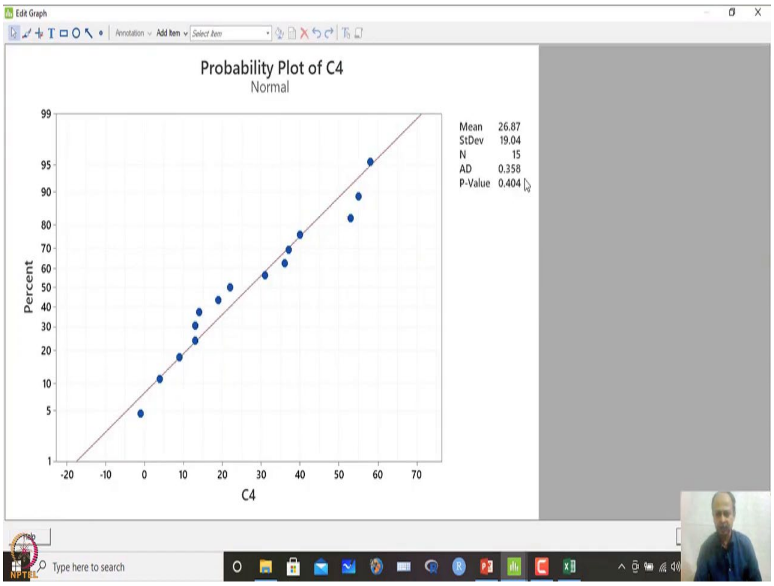


And the assumptions of this paired t-test is that difference should follow normal.

(Refer Slide Time: 08:28)

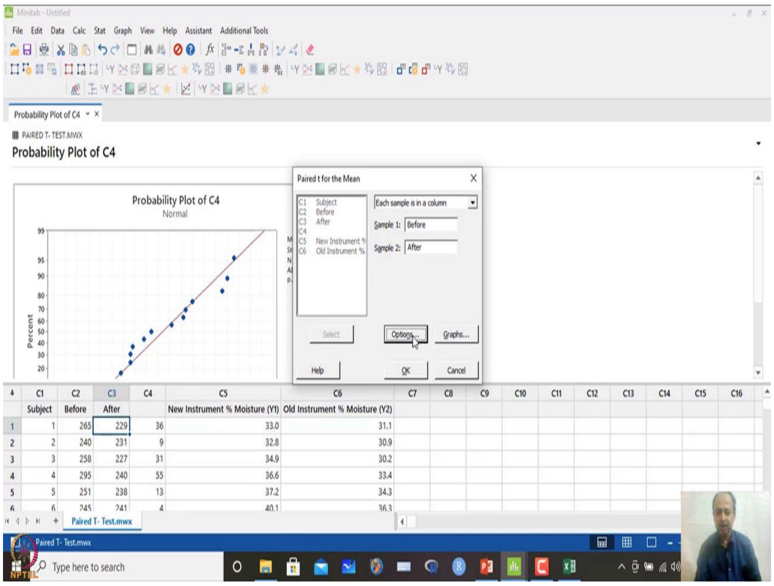


(Refer Slide Time: 08:32)



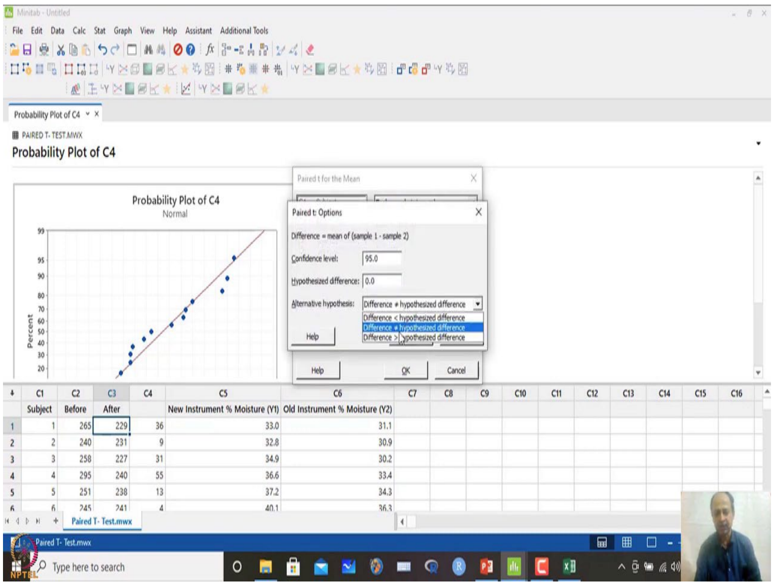
So, that can also be verified over here basic stat normality test like that. So, C4 can be identified and we click ok and we will get a p-values for this and what we can see is that p-value is 0.404 and that satisfies our condition that this difference all of also follow normal distribution.

(Refer Slide Time: 08:50)



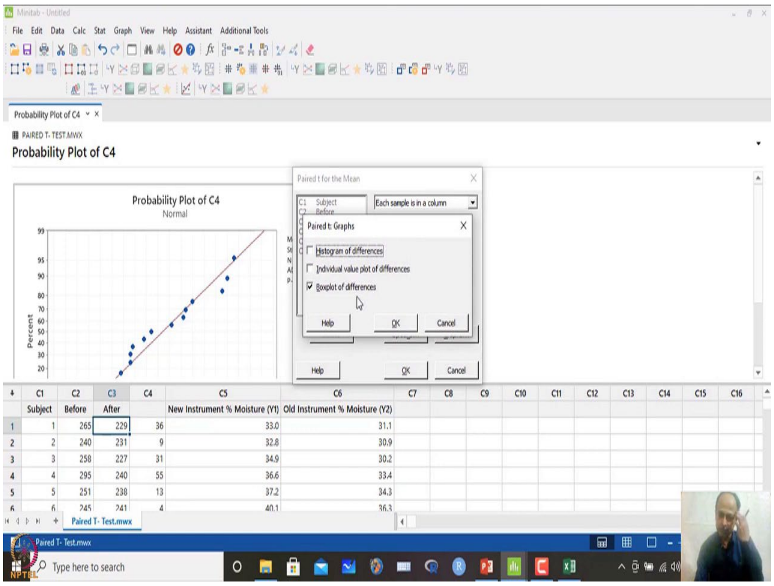
So, I can go for a paired t-test. So, immediately what I will do is I will go to basic stat and I will go for this option paired t-test over here. And in this case, what I will do is that sample 1; this option will be blank whenever you click this one each sample is in a column. So, this is the or you can give difference over here summaries data in difference is given. So, here what I will do is that I have already before and after observation.

(Refer Slide Time: 09:09)



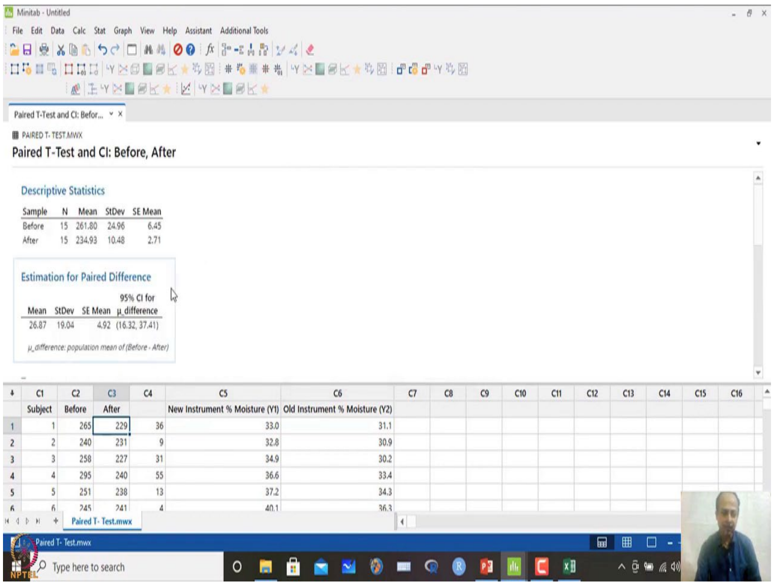
So, then what I will do is that in options, I will not change anything I want whether the difference is equals to 0 not equals to conditions; both sided test I am doing over here.

(Refer Slide Time: 09:18)



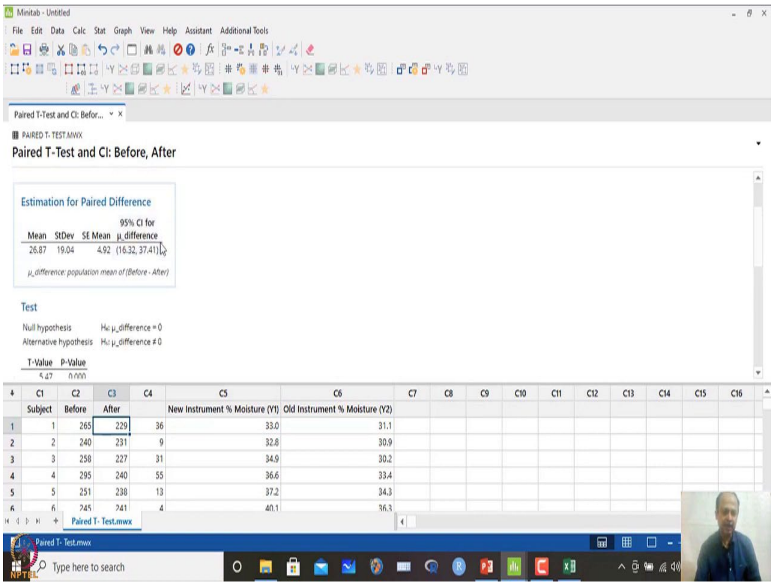
I click ok and in graph also, if you want to see the box of difference like that box plot of difference like that you click ok.

(Refer Slide Time: 09:24)



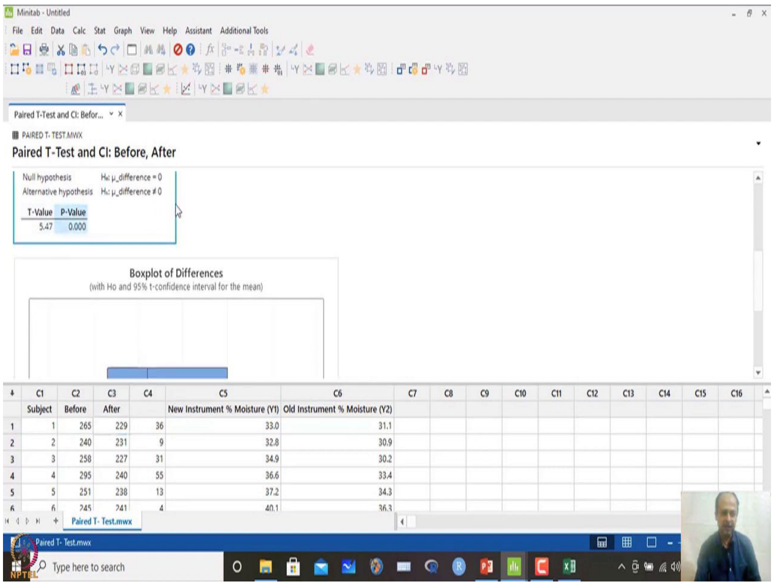
And you click ok like that and you will get all the results.

(Refer Slide Time: 09:30)

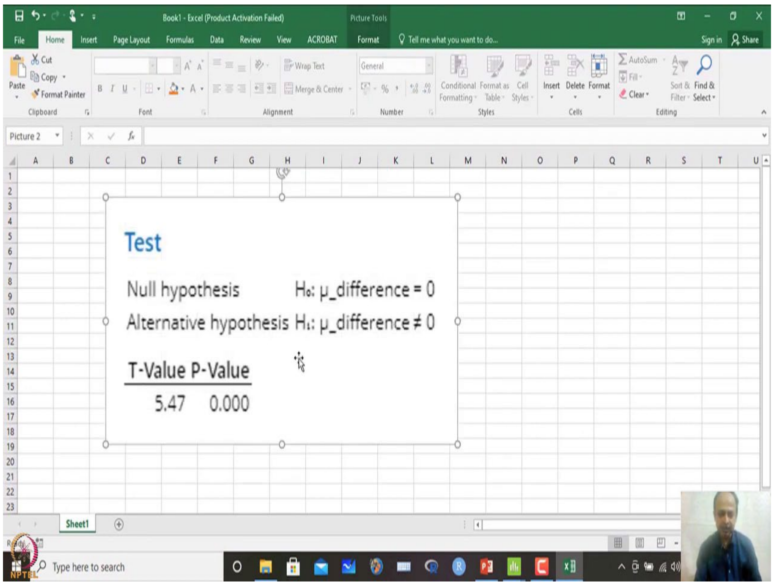


So, before the standard deviation mean is given over here and then, confidence interval is given for the estimation of the means pair difference estimation over here and that is the confidence interval that is given over here.

(Refer Slide Time: 09:41)



(Refer Slide Time: 09:51)



And then we have a key statistics, which is given over here which is the test result which I can copy from here and I can paste it for viewing like that. So, I will paste it and I can just enhance this image over here. So, t-test says that p-Value is around 0. So, approximately and that is less than up to three decimal place it is 0.

So, in this case, what we can assume is that there is significant difference between these two data that before and after data set that we are having. So, in this case the difference is significant over here. So, difference reduction over here is significant that we can observe like that. So, mean is also given over here. So, before 261 after 234.

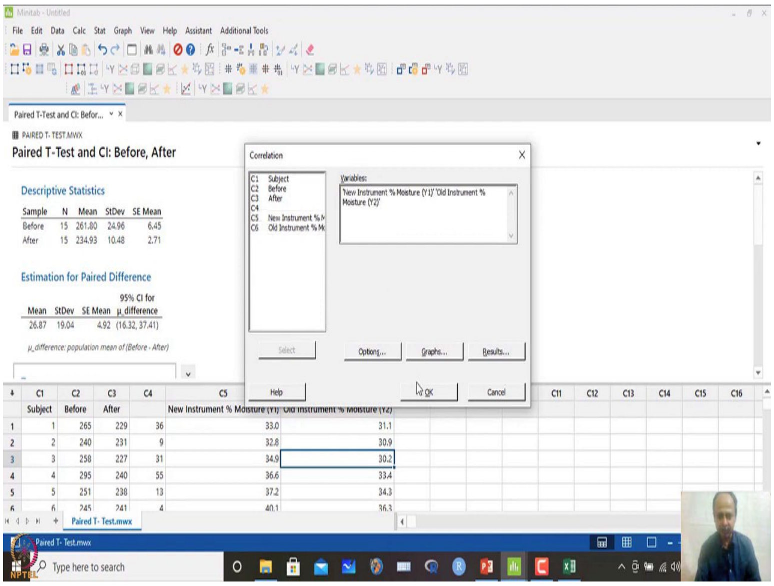
So, basically, it has reduced cholesterol if we are doing exercise and dieting. So, in this case this is accurate. So, because we have taken paired t-test over here. We can also do two-sample t-test, but accuracy level somewhat decreases. In case correlation is high, we will lose some information we will lose some.

So, over here what is suggested is that; we will go for paired t-test rather than two-sample t-test, which is recommended statistically which has been shown that paired t-test is more effective, but if there is not much correlation over here, I will not go for paired t-test I will go for two-sample t-test like that.

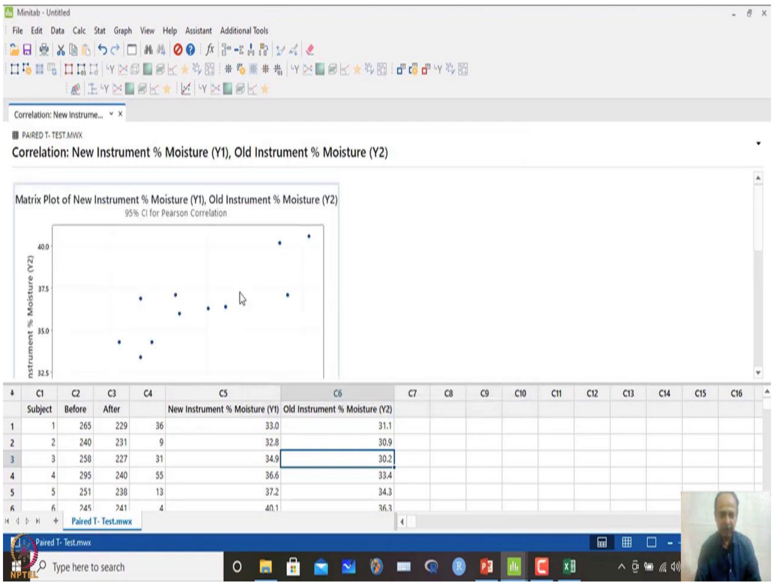
So, there is another example over here what we can see is that this is moisture meter reading with some instrument over here and with an old instrument, I want to see whether there is difference or no difference between old one, because I want to implement a method new instruments. Old methods may be taking long times like that 5-6 hours. Over here in minutes we are getting the reading, but whether the readings are accurate or not.

So, old way was accurate that was the condition over here and this is a new reading for a given sample. So, old reading and new readings are provided over here, I want to see whether there is significant difference between these two readings or not. So, in this case what I will do is that? First, I will go for basic statistics, correlation coefficient whether it exists between these two variables. So, correlation I will go.

(Refer Slide Time: 11:50)

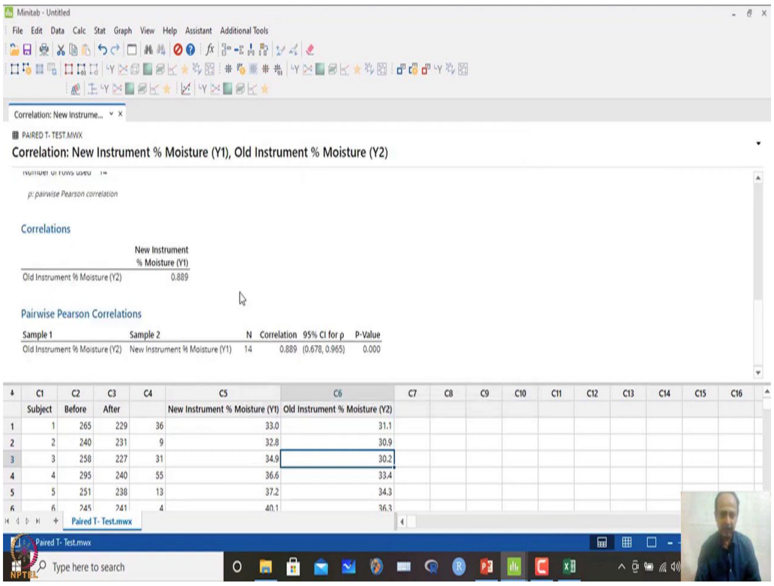


(Refer Slide Time: 11:54)

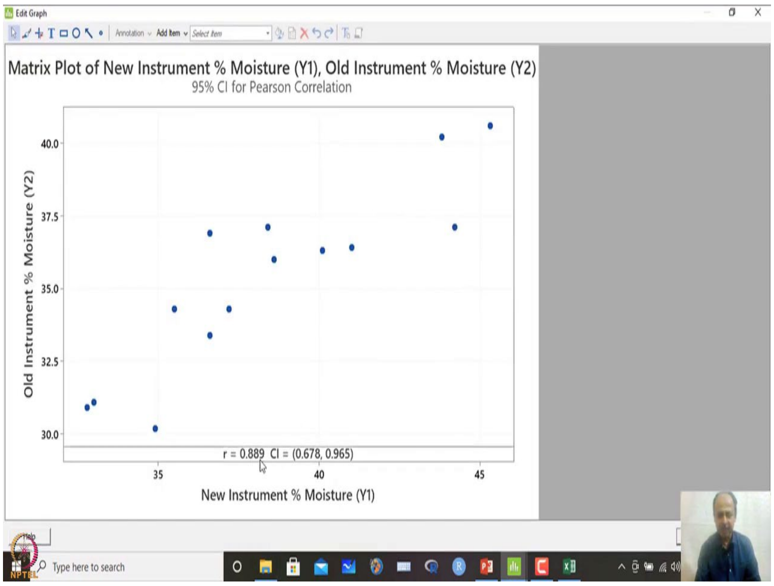


And I would like to check a new instrument and old instrument correlation coefficient.

(Refer Slide Time: 11:55)

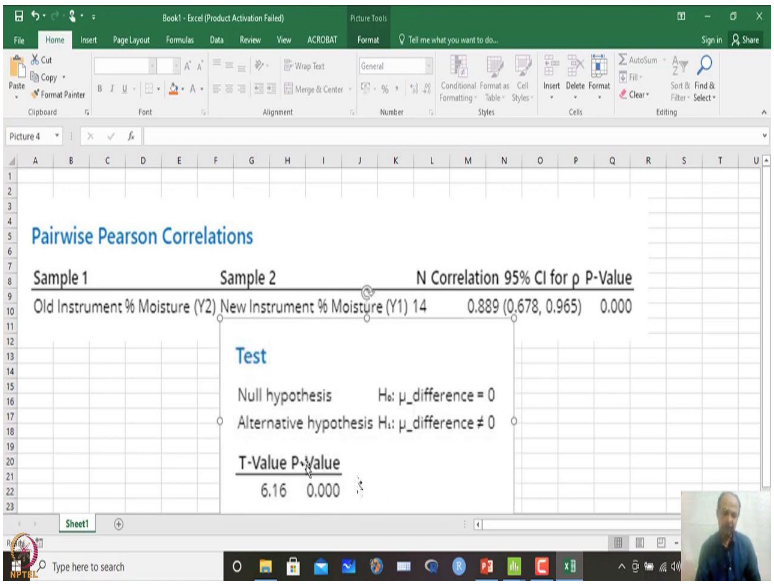


(Refer Slide Time: 11:56)



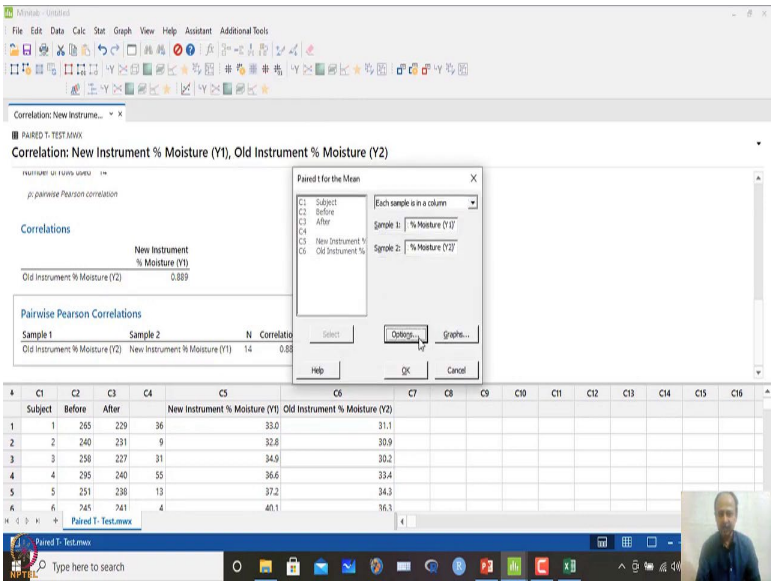
And what I see is that r values is 0.889 approximately. So, if we enhance this image over here, what you will find is there is high amount of correlation that exists over 0.889 that is showing in this relationship diagram. And if you go to the p-Value of this what you will get this p-Value is around 0. So, this value is around 0. So, I can copy this one and paste it over here for your convenience to see.

(Refer Slide Time: 12:20)

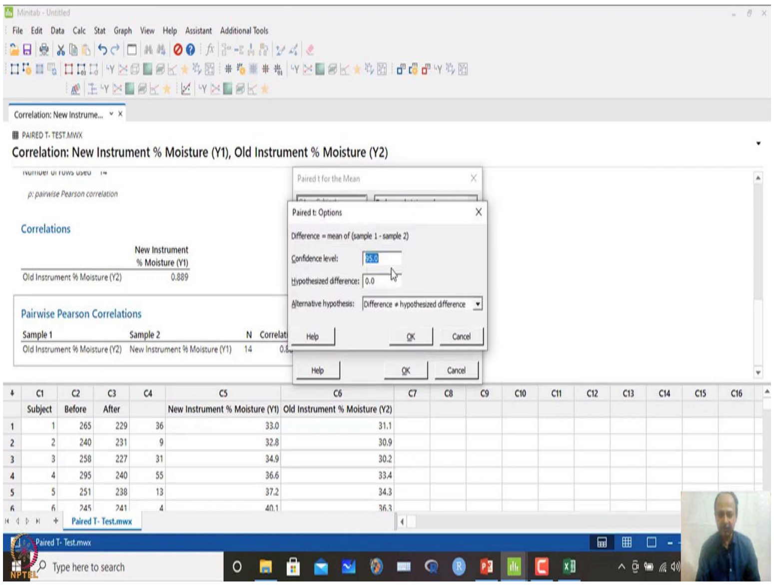


So, this you see that the p-Value is approximately equals to 0. So, in this case, what we can suggest is that; there is high amount of correlation. So, in this case what we can suggest is that go for paired t-test instead of two-sample t-test. So, what I will do is that? I will go to basic statistics and then, I will go to paired t-test. And in paired t-test, what we will give is that? New instrument and old instruments like that and I will give options over here, that same I want to see the differences better.

(Refer Slide Time: 12:41)

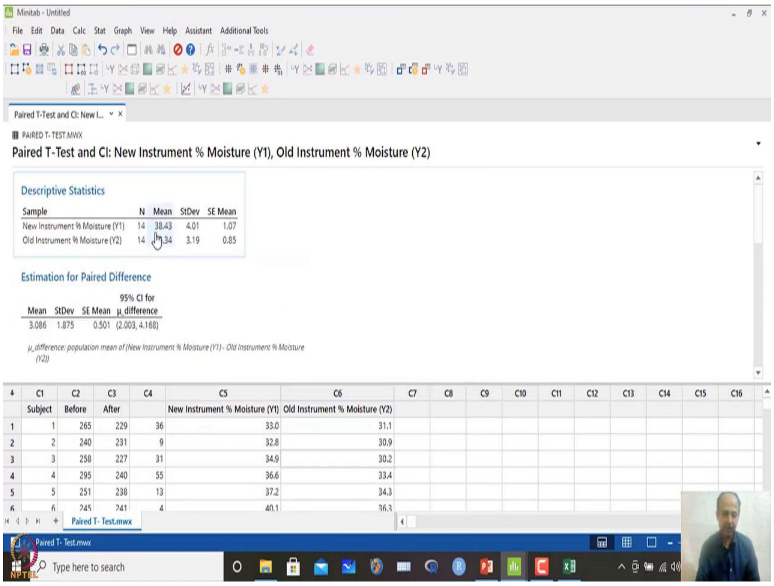


(Refer Slide Time: 12:43)



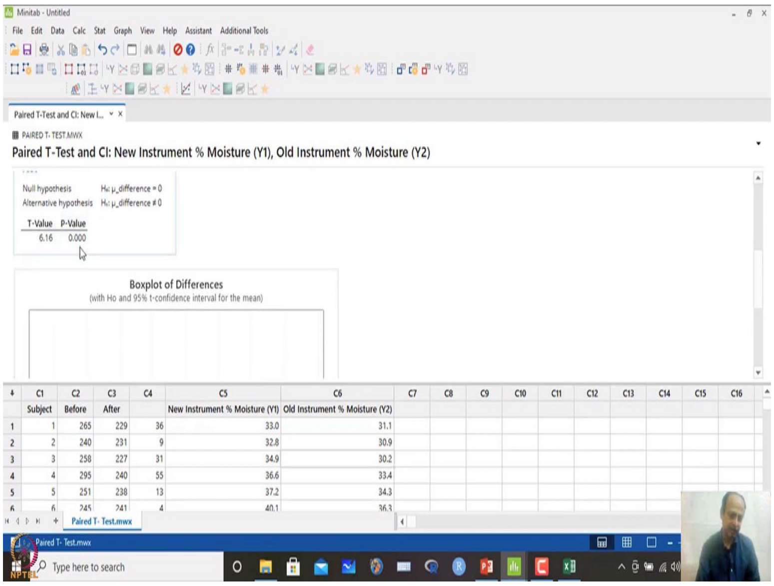
We can do one-sided also, but we are doing two-sided. So, immediately we will.

(Refer Slide Time: 12:52)



So, over here what you see is that new instrument and old instruments in this case. One is giving 38.43 reading that is new instrument on an average. So, this old instrument is giving around 35. So, there is a difference of around three units over here. And that is significant that is shown in the paired t-test that you see over here.

(Refer Slide Time: 13:05)



So, this I can copy and I can paste it over here below we can press this one and we can just enhance this and try to see. So, in this case also p-Value is less than 0.05. So, in this case what conclusion we draw this we can draw is that the readings are different the readings are different over here.

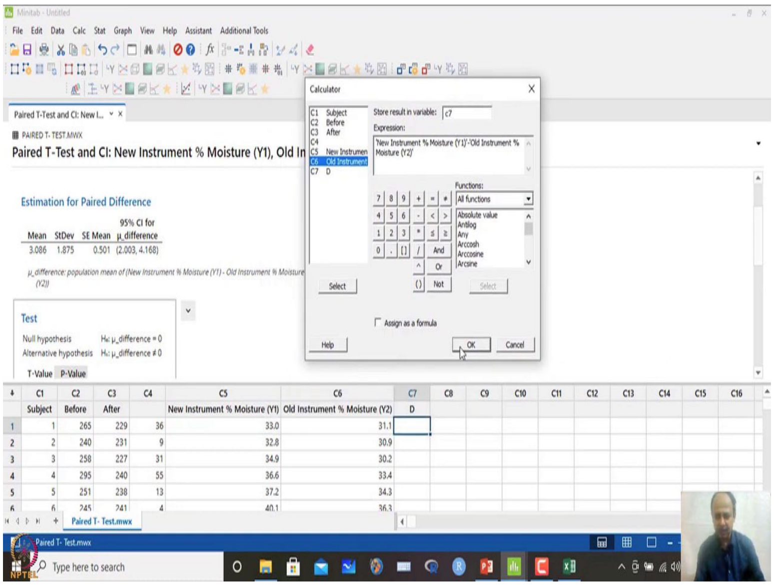
So, in this case we can just say that the two readings are different. So, before and after readings are quite different. So, we will use paired t-test whenever there is a high amount of correlation between the samples means with the same samples, I am taking two readings over here; like in manufacturing, we will find that I have hardness testing machines of different types with different tips like that.

So, whether I use different tips whether readings are different or not. So, that to confirm that one also we do paired t-test, because sample remains same. So, in that case on the same sample I am using two methods like that. I want to compare which method is giving me higher reading or their same readings like that or any of them will give the same hardness like that.

So, for that also this type of analysis is required. So, this is the all about paired t-test. So, if you want to analyze when normal distribution assumptions is not satisfactory what can be done? One-Sample Wilcoxon Test can be used.

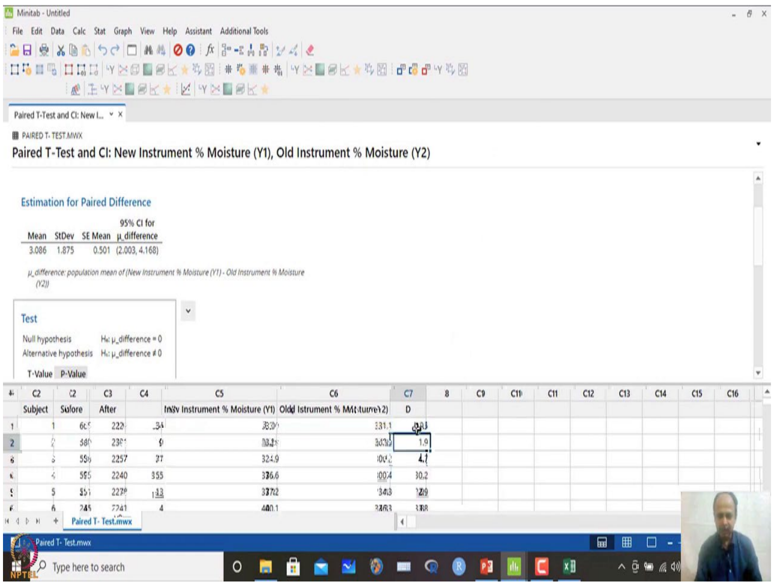
So, in this case let us say, old and new one and we are taking the difference over here let us say, D is the difference over here, we can just calculate the difference. So, calculator can be used.

(Refer Slide Time: 14:40)



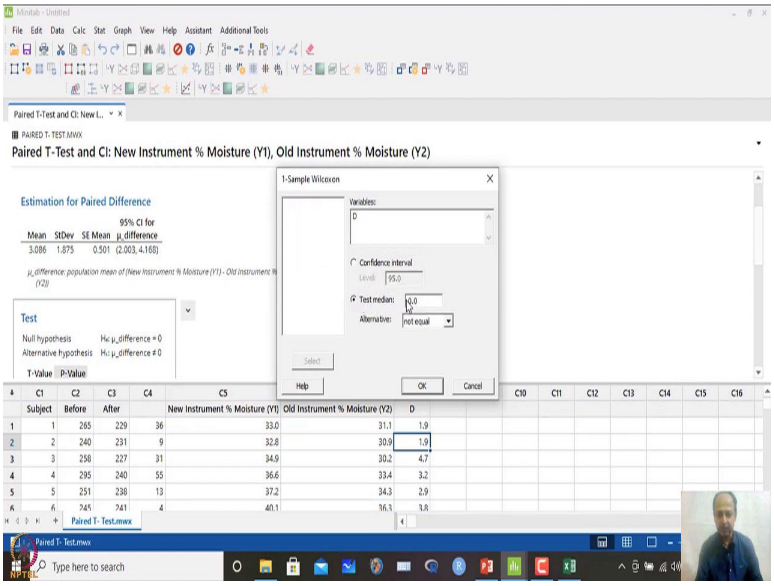
So, in this case what we can do with new instrument minus old one over here, and I click and save it in let us say, C7, I am saving it in C7 and I click ok over here the difference is recorded over here.

(Refer Slide Time: 14:51)



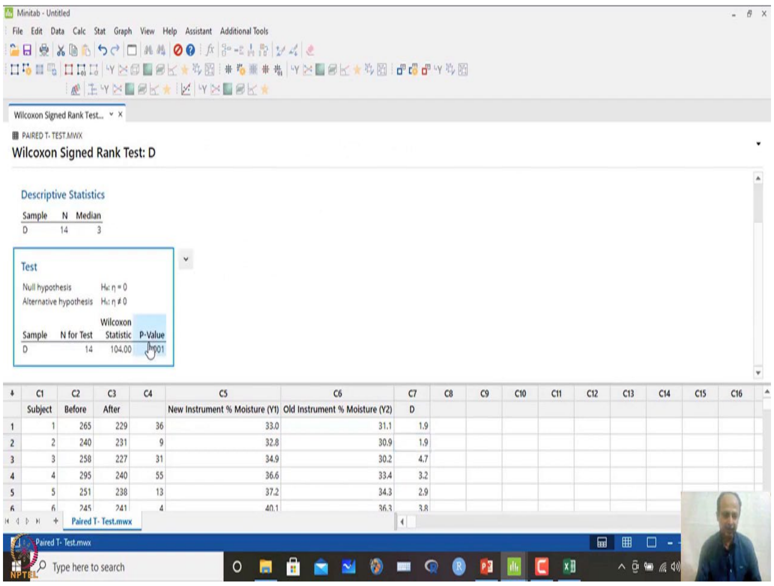
So, difference is recorded over here and maybe, it is not following normal distribution, but here it will follow normal distribution. This is a classical example taken from book. So, in this case, in case it fails, what you have to do is that? You have to go to non-parametric test, One-sample Wilcoxon test.

(Refer Slide Time: 15:11)



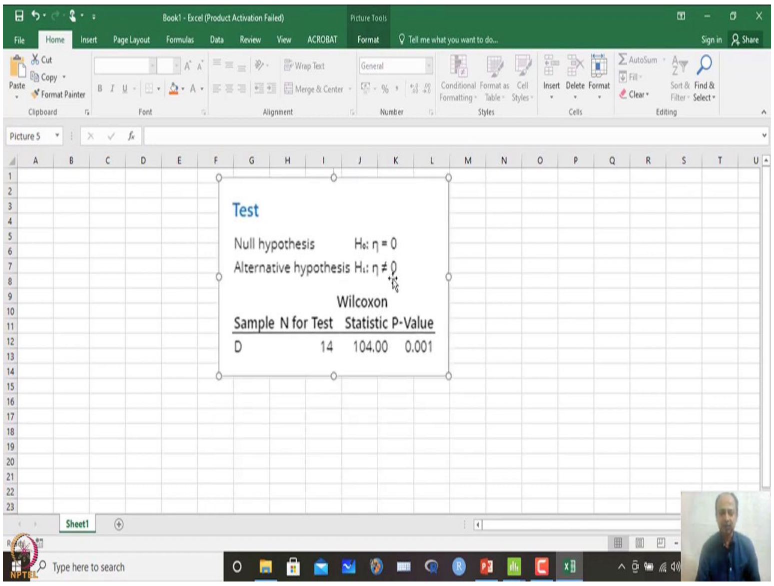
So, in this case what you can do is that difference that we want to see and median value should be equals to 0, that is our assumption difference should be close to 0.

(Refer Slide Time: 15:21)



So, in this case not equals to case we have taken. So, p-Value of this you can get.

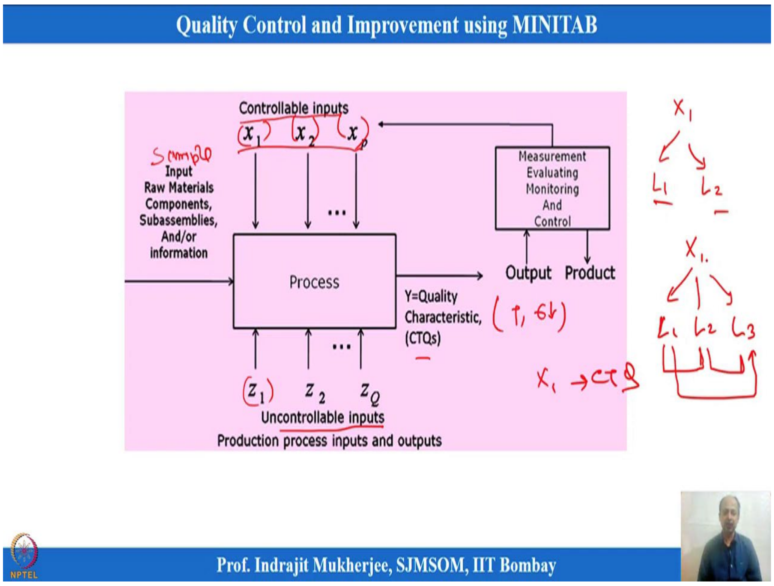
(Refer Slide Time: 15:31)



So, if I copy this image over here and I paste it like that so, this observation and I paste it what I will find is if I will get is that the difference is equal to 0 or not equals to 0, that is the condition I am testing and the p-Value approximately 0.001 which is less.

That means, there is a different significant difference between the, that difference is quite significant that is why, p-Value is less than 0.05. So, that is the alternative we have in case you are unable to satisfy the assumptions. So, in this case we will go for this non-parametric test One-Sample Wilcoxon Test. So, that is the overall idea when I am experimenting with one factor at two levels like that.

(Refer Slide Time: 16:11)



Now, we will extend this concept to an important concept, which is known as analysis of variance. So, for that Some understanding over here is required and the same diagram I am using. $x_1, x_2, ..., x_p$ are the control factors, which is in my control while doing experimentation.

So, there can be p number of factors which can be controlled by the experimenters like that, these are known as controllable factors like that ok. There will be input conditions or sample observation samples we need to go into the process like that then, these parameters will be controlled.

There will be some noise variables over here, or uncontrollable inputs like that or variables like that ok. So, in presence of this, I need to determine the setting of these conditions. So, that I get the best CTQ's or output over here which is close to target with minimum variabilities like that ok.

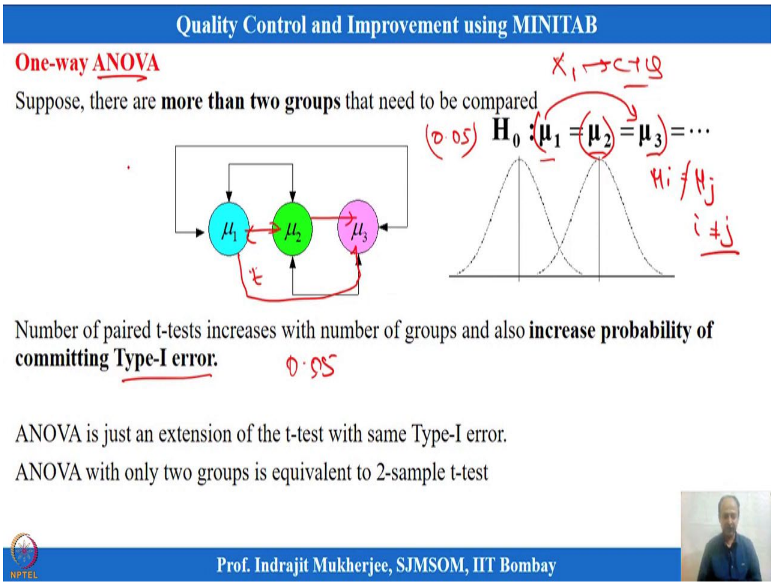
So, in this case so, for paired and two-sample t-test what we have done is that we have assumed that X_1 is a factor let us say in the process and there is only level 1 and level 2. So, if X_1 has more than one level over here like level 1 level 2 and level 3, one possibility is that I do individual assessment over here, whether level 1 and level 2 are different or level 2 is different from level 3 mean is different. Similarly level 1 with level

3 like that, I can do pair wise comparison like that. So, I can do two-sample t-testing and see which level is different from which one like that. So, why I am doing this?

Because I want to see whether this factor X_1 is influencing the CTQ. I want to screen the factors and do some preliminary analysis so that I understand that this factor can be considered as full-fledged experimentation at a later stage. So, I want to see whether one of the factor is influencing the variable CTQs, I can experiment more than two levels like that.

So, it can be two levels, but we can also experiment more than two levels for a given factor like that. So, in this case what is the best option? So, if I do pair wise comparison what will happen is that; and I have to conclude based on this all pair wise comparison. The Type-1 error generally increases over here.

(Refer Slide Time: 18:16)



So, the null hypothesis over here is that all the means are same and alternative is any of the means is different from any of the means over here. So, there is different means when $i \neq j$ when I am considering this condition like that.

So, there can be three different levels and three different means will be generated and whether they are same or whether there is a difference between the means of any two levels. So, when I am concerned about that we go for a one-way analysis of variance what I can do is that I can compare μ_1 with μ_2, μ_2 with μ_3 and μ_1 with μ_3. Like that pair-wise comparison is possible.

So, if there are three levels, we have ${}^3C_2$ combination. So, what will happen is that if I am doing this combination over here, overall judgment will be correct. So, that reduces like that.

So, the Type-1 error basically, increases over here and in each statistical book you can see why Type-1 error increases. So, in this case, what is required is that; to keep the Type-1 error as 0.05 what we have assumed for an experimentation hypothesis is over here. So, my level of significance will be always 0.05.

So, for that Fischer develop this analysis of variance and where the level of significance remains 0.05 and I can make a conclusion based on this that I will be true 95 percent of the time. So, in this case this is a extension of you can think of 2-sample t-test like that, but there are more than two levels that we want to see over here.

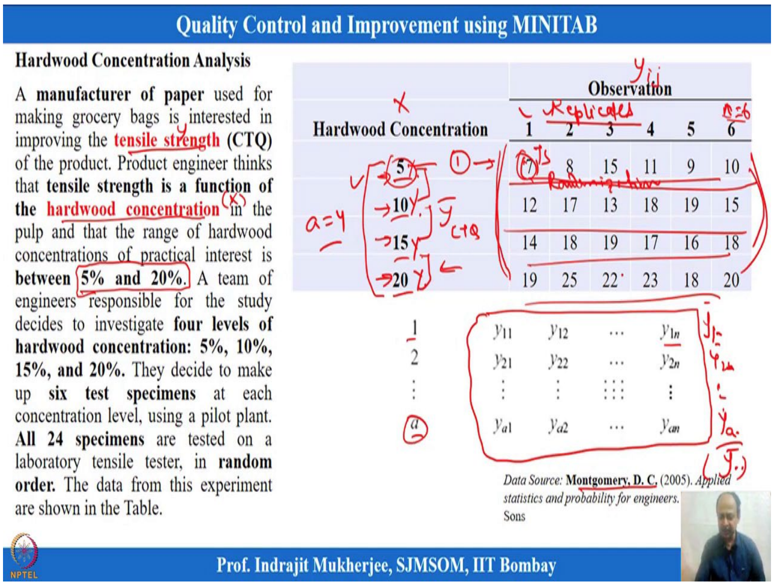
So, there are more than two levels and I want to check and, but there is only single one factor over here. So, there will be only one factor and that will have different levels and I want to check that whether when I change the levels; whether it is basically impacting the CTQ's or not.

So, it is impacting the CTQ's or not or mean CTQ's or not; that is our objective over here. I want to screen this factor and figure out that within the range if there is true or three levels of X_1. So, whether the mean of CTQ is changing or not so that is our overall objective. And if it is so, then in that case X_1 is a critical variable, which influences Y and we can take it forward to full-fledged experimentation when we go for response surface methodology or like that.

When we are going for full factorial may be one of factors to be screened. So, here experimenters can do a simple experimentation with a single factor and see whether that is influencing or not and although it is always suggested that we take all together and do the experimentation, but initial studies we can do and try to screen that one and then go for full factorial.

So, this is used for basic screening experimentation like that, but you have to remember that we cannot go for paired t or 2-sample t-test over here. So, t-test is not sufficient we have to go for a F-test analysis over here, which is given by an analysis of variance.

(Refer Slide Time: 21:19)



So, we have to consider that scenarios like that. So, let us take one example and try to understand the scenarios before we go into analysis of this. This is known as one-way analysis of variance basically ok. So, what we are doing over here is that we are analyzing variation basically, but difference between means we are trying to report over here; whether there is difference.

But the variance information is used for that ok, that is why it is analysis of variance basically. This is one example taken from again from the Montgomery's book and what you see is that tensile strength I want to maximize. So, in this case I am changing the hardwood concentration which is the only x factor that I am having and this is the only y I am monitoring over here.

And the hardness concentration the experimenters have decided that it should be between 5 and 20 that is the optimal range of x that we can take over here and it can have more than two-levels like that. So, the experiment that choose 5, 10, 15 and 20. And this is the x level that you are seeing and these are the y observations.

There are total A number of levels. So, if you write the generalized form over here what you can see is that y_{11} is the first observation. There are n number of observations n equals to 6 over here so, n number of observations. So, this one will be y_{1n} like that. So, there will be a-levels.

So, this is the way of general forms we can write. So, mathematically we can express this one in this way. And then, what we are seeing over here is that when I change the level from 5 to 10 or 10 to 15 or 15 to 20 is there any significant difference that is happening in the CTQ's voice CTQ is over here. So, mean value whether it is changing in any two-levels when I am changing over here.

So that means, if any two-levels it is changing significantly, this factor is important and can be considered for the further experimentation like that ok, but assuming that there is no other factor and I want to maximize or optimize the levels.

So, that where the hardwood concentration should be kept. So, that I have maximum tensile strength which is my objective over here. Let us assume that other factors are not prominent over here; only one factor is there I want to select the factor over here. So, this is the first step of experimentation where we want to determine the condition where we will get maximum CTQ or CTQs can be maximized, because my target is tensile strength.

So, over here tensile strength higher the better. So, if it is higher the better; that is the condition we want to achieve over here. So, one is 5 percent over. So, levels over here are more than 2. So, there are how many levels over here? Four-levels that that you can see over here.

So, a is equals to 4; that is the thing that we are assuming, n equals to 6 like subgroup size what we have seen. So, this is n equals to 6 observation and there are 24 total number of observations over here in this matrix. And all are random this is experiment is done by randomization.

So, what is randomization? So, at 5 percent some one-sample will be taken arbitrarily and 5 percent hardwood concentration will be checked and accordingly, what we will say that there is 6 specimen over here. So, in this case first specimen will be taken and what we will do is that we will run at 5 percent and observe the tensile strength.

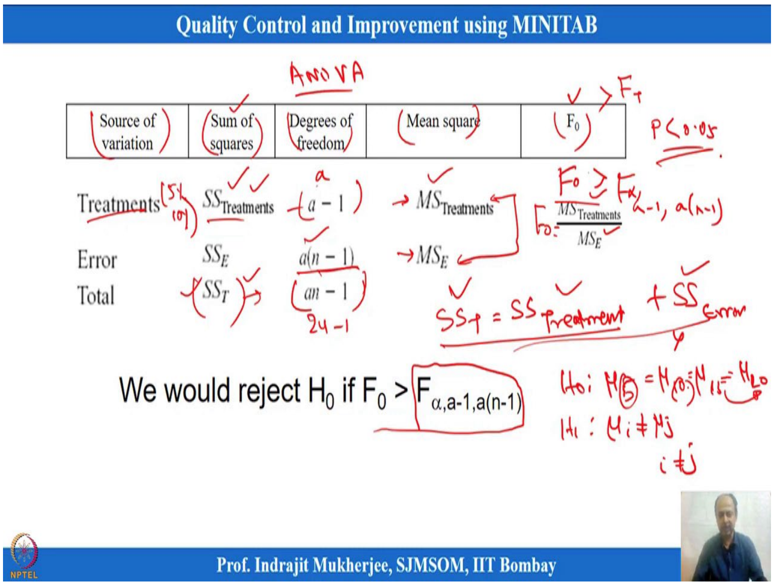
So, 7 was observed with the tensile strength like that. Then, I will randomize this means this data set is general this dataset is generated based on randomization; that means, sample will be selected. Any of the samples can be selected and any of the levels will be selected over here. So, this is known as complete randomization like that.

So this is because of some reason which we will understand afterwards. So, what I am saying is that randomization is done over here. So, levels will be selected randomly over here, samples will be selected randomly. Samples are assumed to be homogeneous over here. So, in this case so, sample to sample variation is very less only percentage variation what we are considering over here. So, at different levels we take six readings like that.

So, we will have an average of this. So, $\overline{y}_{1n}$, $\overline{y}_{2n}$ $\overline{y}_{an}$ over here and then we can make a grand average over here. So, this can be a grand average like that. So, this is symbolic notation that you will find in Montgomery's books. So, that is the symbols that I have written over here.

You can use the different symbol also. So, that is my overall objective over here is to determine the level whether it is 5, 10, 15 or 20; where the tensile strength will be maximized over here. For that I have changed intentionally. So, this is systematic induced variability that I have considered over here. I have changed the levels arbitrarily over here and then, I have recorded the tensile strength and based on this mean value over here, this and, this and, this. I want to freeze a level where my tensile strength will be maximized like that.

(Refer Slide Time: 27:06)



So, how do we do that? That is important for us. And we will discuss about that. So, that is the and what Fischer has given is that, what Fischer has done is that; Fisher has given a table over here which is known as ANOVA table like that. And MINITAB will report this ANOVA table over here where you will see source of variation, sum of square over here. Degree of freedom and mean square calculation and a F statistic that could be reported like that.

This is important F statistics over here and it will be compared with the tabulated value and if this F is greater than this one. So, in this case p will come out to be less than 0.05 ok. So, this is the process this is the process we make the interpretation like that. So, over here treatment means when I change from level-five like that 5 to 10 or something 10 percent over here.

So, treatment whenever I am changing the levels the total amount of variability that is that it creates like that. So, that can be calculated as and known as $SS_{Treatment}$ like that. So, because of a change in treatments, which is level 1 to level a what is the overall variation that is happening, because of these changes like that. So, that we can think of a sum of square measure.

So $SS_{Treatment}$ will be calculated, SS_T is the total variability of the data so, because treatment will not explain all variability. So, that is the condition that is SS_T will be

$SS_{Treatment}$ over here, and there will be some error SS_{error} over here. So, Fischer has given you two-way to calculate this one and how to calculate SS_T formula is given from the given data set that is generated over here.

So, from this matrix, we can generate these values of SS_T. We can generate $SS_{Treatment}$ and SS_{error} can be calculated. SS_{error} will be $SS_T - SS_{Treatment}$; that is a formulation we can consider and degrees of freedom will be reported also over here. So, degrees of freedom in say number of treatments is considered over here. So, a-1 is the degree of freedom and total observation is 24, 24 minus 1 this is a n minus 1 is the total degree of freedom.

And if you subtract this an-1 and then, minus this treatment combination what you get is that error degree of freedom basically ok. Now, when we have calculated this degree of freedom over here. So, $SS_{Treatment}$ divided by degree of freedom gives you mean square treatment like that.

So, this has to be calculated similarly SS_{error} as an-1; this will give me MS_{error} like that. So, then this two will be compared like that and the ratio will be taken $MS_{treatment}$ by MS_{error} over here. This will give you a F_0 values like that. This F_0 value has to be compared with tabulated value.

So, this will have a degree of freedom a-1 over here numerator degree of freedom and the denominator degree of freedom will be an-1 like this. So, F will be calculated at let us say α level of significance and this will be calculated and F_0 will be compared with this tabulated value over here and if this is greater than this in that case p will come out to be less than 0.05.

So, we are going by p method like that we are not going by tabulation and all this. So, p will indicate whether to accept or reject the null hypothesis which is null hypothesis over here. Null hypothesis is whether the mean at different levels 5 percent over here and equals to 10 percent over here equals to 15 percent over here or this is 20 percent over here.

And the alternate hypothesis over here is any of this $\mu_i \neq \mu_j$ over here for $i \neq j$, that is the condition I am checking over here ok. So, here you can see the formulation that I have given I have written like that. So, this is will be compared and if it is getting we will reject the null hypothesis basically.

So, this is the overall theoretical aspect that is covered over here; that means, how to collect the data randomization is considered over here and n will be the sub-groups what we are considering. So, this is also known as replicates over here this is also known as replicates.

In experimentation, this is an important thing that we have to consider. More and more replicates what will happen is that more and more I will be assure of the final conclusion that I am drawing. So, but number of sample size will increase over here. So, total number of samples here 24 experimentation was carried out. If I have taken only two replicates in that case, only I have to do $4 \times 2 = 8$ experimentation like that.

So, you have to find out the optimal combination of this where I need to stop. How many replicates I will consider in the experimentation, considering the cost like that. So, in this case and if you can collect this data randomize the data and collect the experimented results over here; what will happen is that then, ANOVA analysis becomes easier.

And it will show you whether when I change the levels this will show you F value will tell that there maybe, if that is significant in that case it will indicate maybe 5 is different from 10 or maybe, 10 is different from 15 the mean value of y CTQ 10 is different from 15 or 15 is different point.

At least there are two-levels when I change from that level to the other level what is happening; the mean of the CTQ is changing significancantly between the two-levels like that.

So, they are different statistically basically ok. So, that will be the conclusion, but which level ANOVA will not tell you from which level to which level this has happened basically. It will say any two labels are different. So, that will be; that will be concluded based on this F values or p-values that you generate over here. So, MINITAB will generate p-values and based on that we will conclude.

So, we will continue discussion of this ANOVA analysis further to understand more about ANOVA analysis in experimentation. So, and how to implement that immediately;

in our next session ok. So, thank you for listening and we will stop over here and we will start from the ANOVA next time again ok.

Thank you.