Six Sigma
Prof. Jithesh J Thakkar
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture - 31
Hypothesis Testing: Two Population: Minitab Application

Hello friends, we are advancing in our Six Sigma journey. And going through step by step the various phases of typical DMAIC cycle. So, we have gone through the define phase, measure phase, improvement phase and now we are discussing the analysed phase. As a part of analysed phase we had the discussion on Hypothesis Testing, inferential statistics to draw better conclusion about the present situation about the quality of my processes about the change to be implemented.

And then we can confidently say that at a given level of alpha significance to what extent my claim is true or false. So, we are now say in lecture 31. And we would like to see the application of Minitab for two population hypothesis testing.

(Refer Slide Time: 01:42)

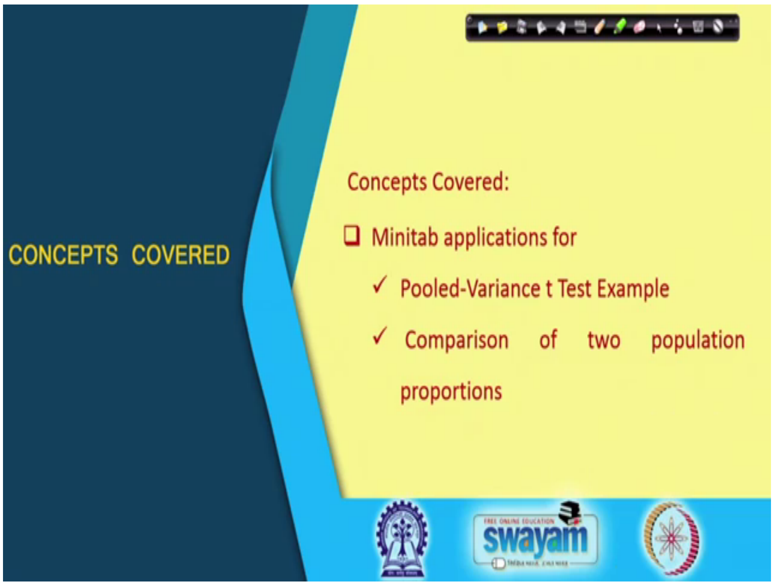
So, once again I would like to remind you that we must test the hypothesis a scientific way of checking the theory or developing the theory proposing the theory. And this is where we are constantly emphasising on our process of scientific analysis hypothesis testing. So, this is what we have seen in the last lecture that the means of two independent populations if I have to compare and they are from two different population.

Second case the proportion of two independent populations. Case number 3 we had seen the variance is of two independent population by testing the ratio of two variances and typically you call it as F statistics. And you try to say follow the procedure basic steps of hypothesis testing to see that your null hypothesis is accepted or you would reject it.

Then case number 4 dependent related samples same group before versus after treatment. So, as we have seen many examples you are trying to improve the quality, you are taking certain corrective measures. And you want to check whether the quality has really improved or not in terms of percentage defective or number of defective items produced. So, here you have a claim that I have some intervention, I have some treatment.

And before treatment and after treatment is there really a significant improvement in the process or in the number of defectives to be produced. Same thing you can apply in other context also; suppose there are patient and you want to see that a particular treatment is really effective or not. So, you would like to consider before treatment as one population after treatment as another population. And you would like to check the say test the hypothesis that whether the treatment is effective or not.

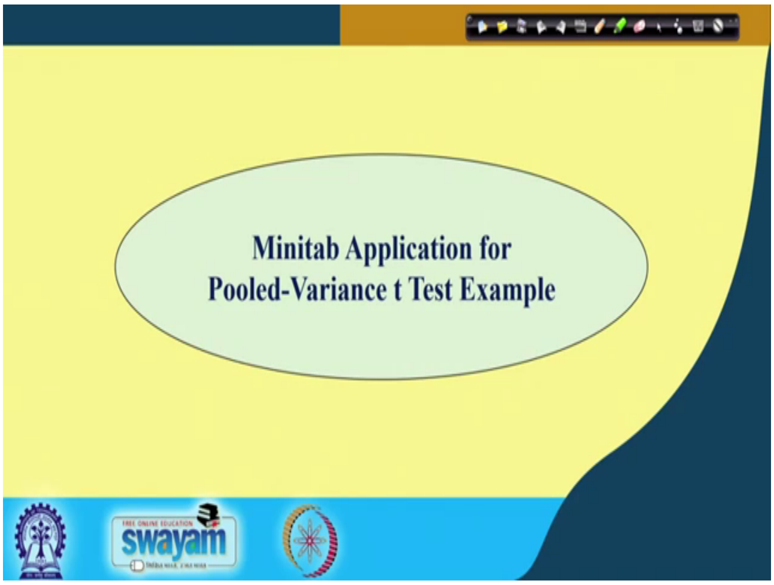
(Refer Slide Time: 03:31)



So, this particular lecture is devoted on Minitab application and here I would try to demonstrate the manual calculation as well as Minitab application. And we will compare the results for two cases; one is pooled variance t example and second is comparison of two proportions. .

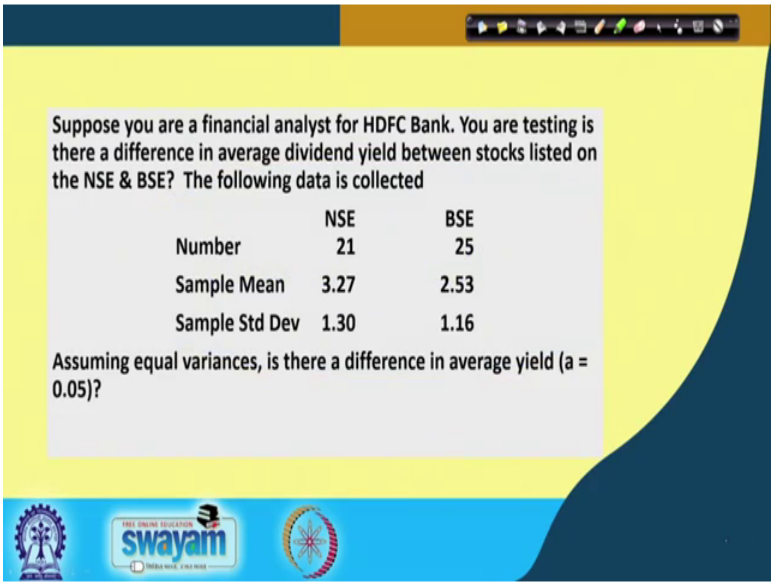
Same way you can explore the drop down menu of Minitab and whatever cases we have discussed in the last lecture you can apply the Minitab application. Here we will try to consider two cases and see that how Minitab can help us to get the necessary results in order to check my claim about the null hypothesis.

(Refer Slide Time: 04:19)



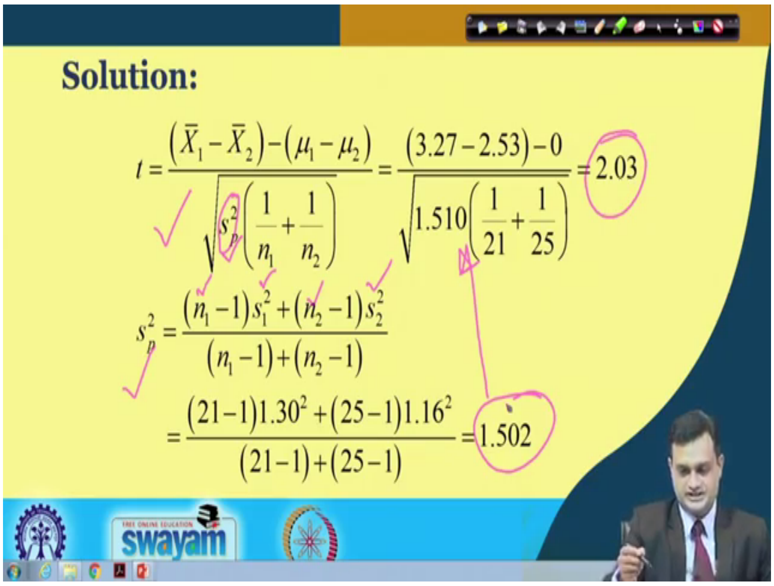
So, Minitab application for pooled variance t test example.

(Refer Slide Time: 04:22)



Once again I would refer to the same example we discussed last time. That you are a financial analyst and you want to check that the dividend yields on the two different stock exchange your NSE and BSE is it the same or there is a significant difference. I want to check my claim at alpha 0.05.

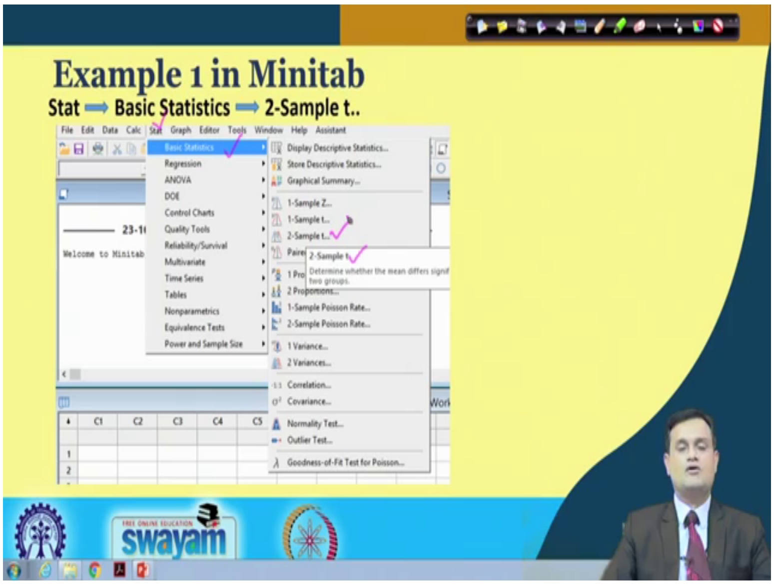
(Refer Slide Time: 04:48)



And I can compute the t statistic because I have sample size less than 30 for smaller sample size you go for the t test. And what I have computed that this is my expression for t statistic we have already seen and this comes out to be 2.03. So, I need pooled variance and pooled variance expression is already given to you.

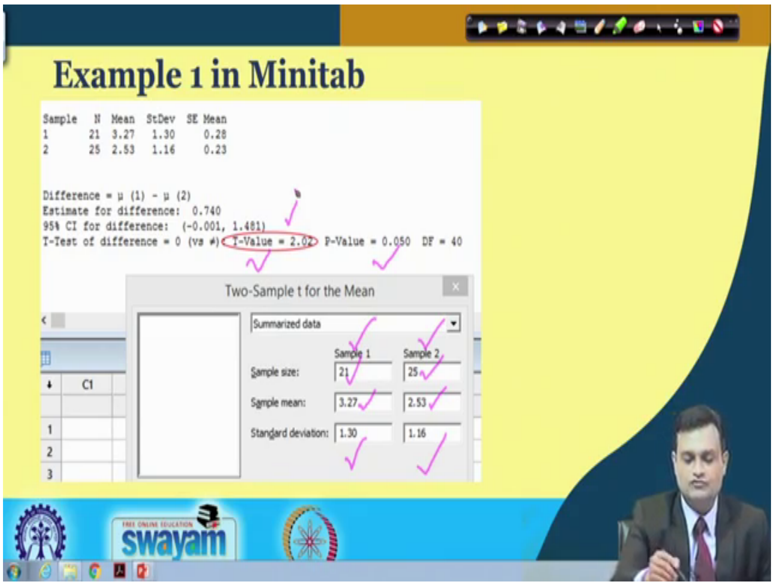
You know n 1 you know s 1 you know n 2 you know s 2. And this pooled variant is 1.502 or you can approximately take it 1.510 and you can just plug in this into. You can take this square root and plug in or take a keep this square here it is square. And just directly plug in this value into the formula of t statistics.

(Refer Slide Time: 05:45)



So, you have the computed value of t. Now let us see how this can be done in Minitab. So, Minitab it is very simple you basically go to this particular say stat. And then in this rack you go to basic statistics basic statistics you choose the 2-sample t 2-sample t. And then already you have put the data into your excel sheet you can just import the excel sheet or you can transfer the data.

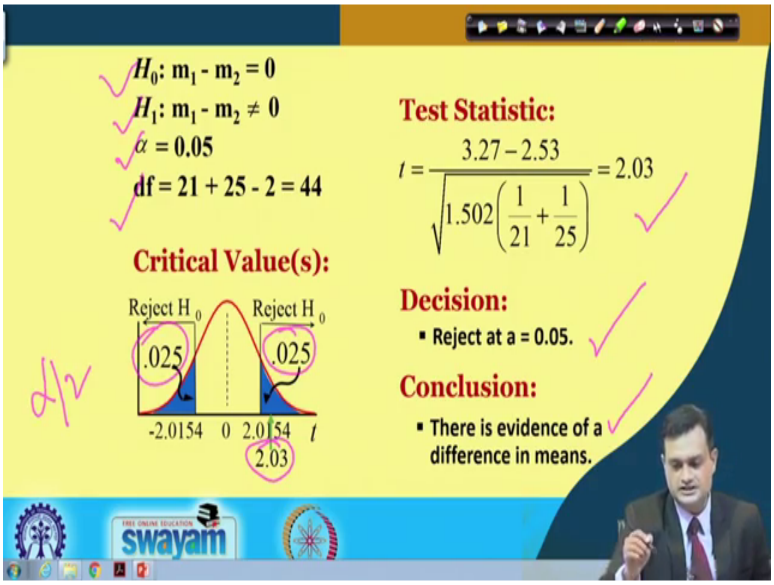
(Refer Slide Time: 06:22)



And then you can very well see in the output window of the minicab that you are getting the t value 2.02 and p value 0.05. So, I have you can see here the sample 1 I have put 21, sample 2 25 sample mean 3.27 sample 2 it is 3.53 and I have put the standard deviation. So, you have very well defined space to input insert the values.

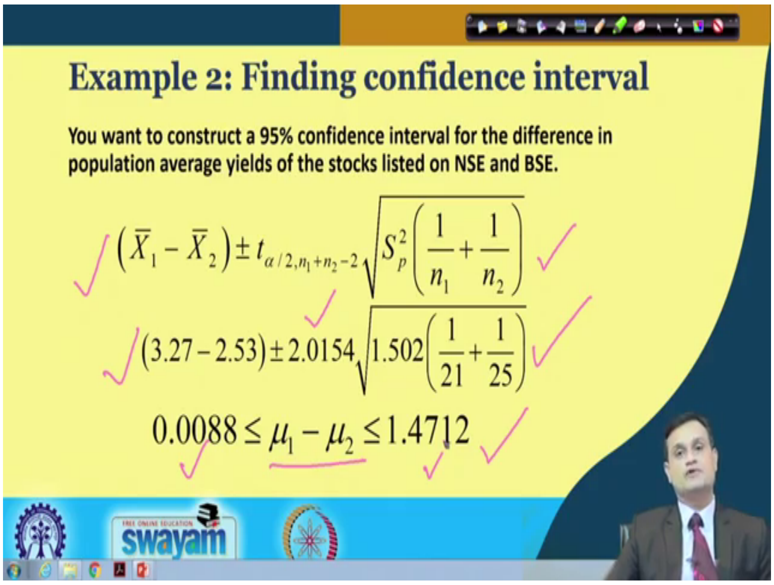
Once, you have done it this will give you the t value and p value for checking your hypothesis. And this can easily be compared with the say previous one. So, you just see that in the previous one I computed the t value and this is 2.03. Same way I can refer my Minitab outcome and Minitab outcome gives me the value which is close to 2.03 so this is 2.02. So, I can very easily compute the t statistics in the Minitab.

(Refer Slide Time: 07:33)



Now, just see that how do we check our claim. So, here my null hypothesis is m 1 minus m 2 is equal to 0.

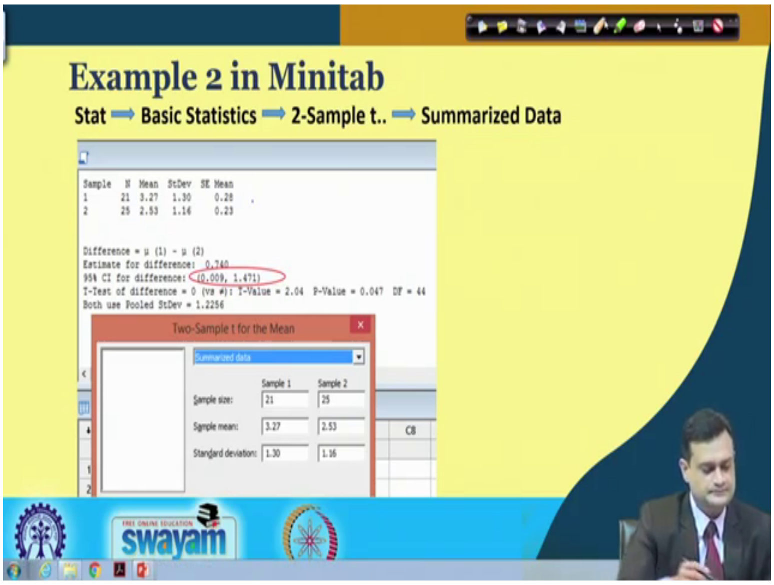
(Refer Slide Time: 07:44)



And my alternate hypothesis my null hypothesis is m 1 minus m 2 is equal to 0 and alternate hypothesis m 1 minus m 2 not equal to 0. It means null hypothesis says that there is no difference alternate hypothesis says there is a difference alpha is equal to 0.05 degree of freedom is n 1 plus n 2 minus 2 that is 42. And you can see that I am doing two tail test alpha by 2 both sides.

So, I have 0.025, 0.025 and my computed value of t either from manual calculation or Minitab it is 2.03 and this falls in the rejection region. So, I would say reject null hypothesis at a or alpha is equal to 0.05. And conclusion is that there is an evidence of a difference in means. So, this how you can very easily compute and test your hypothesis using the Minitab. I can also set the confidence interval and also I can see this in the Minitab solution.

(Refer Slide Time: 09:11)



So, confidence interval for this would be x 1 minus x 2 bar. .

(Refer Slide Time: 09:17)

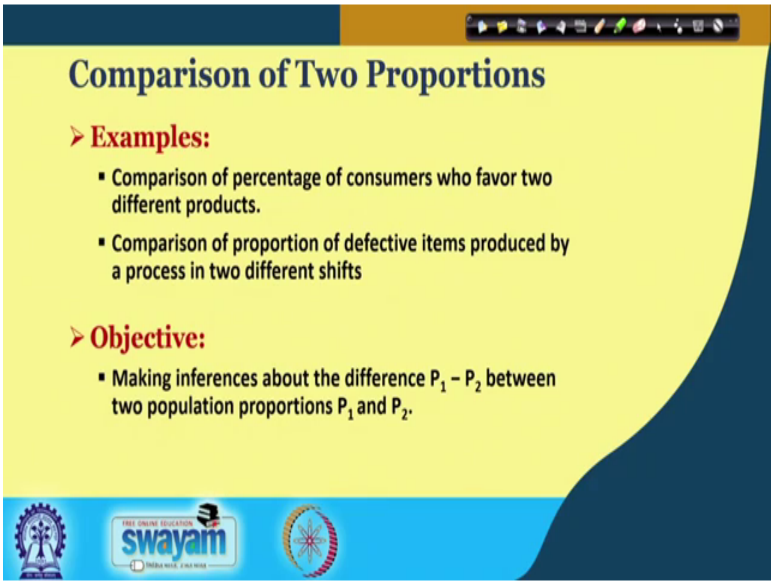
You have the confidence interval x 1 bar minus x 2 bar plus or minus t alpha by 2. And you need the degree of freedom that is n 1 plus n 2 minus 2; minus 2 and this is the quantity. So, basically it is plus or minus k sigma that we know. And you have this 3.27 minus 2.53 plus or minus 2.0154 that is the computed value of t. And this is what you have.

So, you have the confidence interval which says that your mu 1 minus mu 2 will fall say for the given level of confidence 99 percent within the range of 0.0088 to 1.4712. So, now, this is manual computation. Now if I want to check it in my Minitab you can get this that 95 percent. So, here we are referring not 99 we are referring 90 95 percent; so, 95 percent confidence interval 0.009, 1.471.

So, if we just go back then you have computed 0.0088 and 1.4712. So, this is very close to the computed value of the Minitab which is 0.0091 0.471. So, you can see that Minitab can really help us to carry out the calculations and all the t statistics z statistics various probability distribution functions their mathematical expressions are well say built in the software.

And you just need to put the necessary data and you will get the results that you should be able to interpret in terms of acceptance or rejection of the null hypothesis. So, if we just advance then let us see the case two that is Minitab application for comparison of two population proportion.

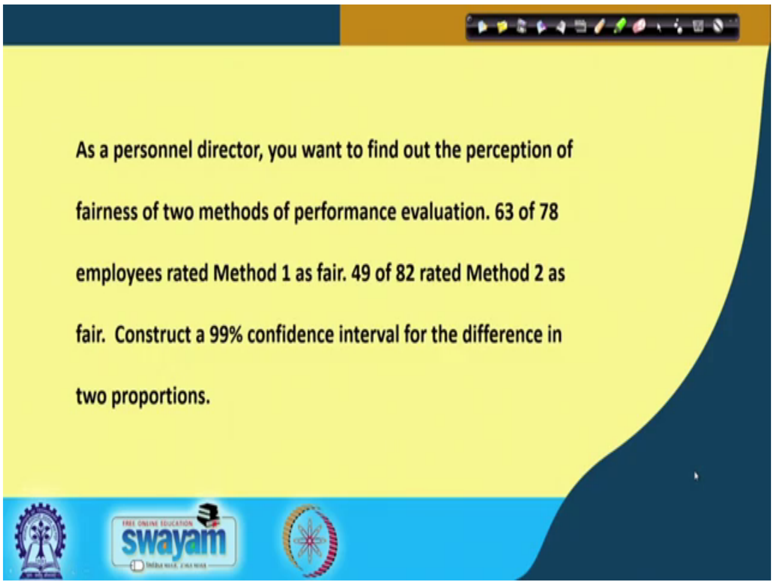
(Refer Slide Time: 11:42)



And we have seen couple of examples that what does it mean. So, here I may be interested to check that percentage of consumer who favour to different say products. So, this is the proportion of consumers that they would prefer some product. And I may be interested to see that proportion of defective items produced by a process in two different shifts.

May be morning shift may be afternoon shift and is that really a significant difference or not so that I would like to check. So, my objective when I say the proportion is to test my null hypothesis. And alternate hypothesis about the difference in p 1 minus p 2 between two population proportion p 1 and p 2 respectively.

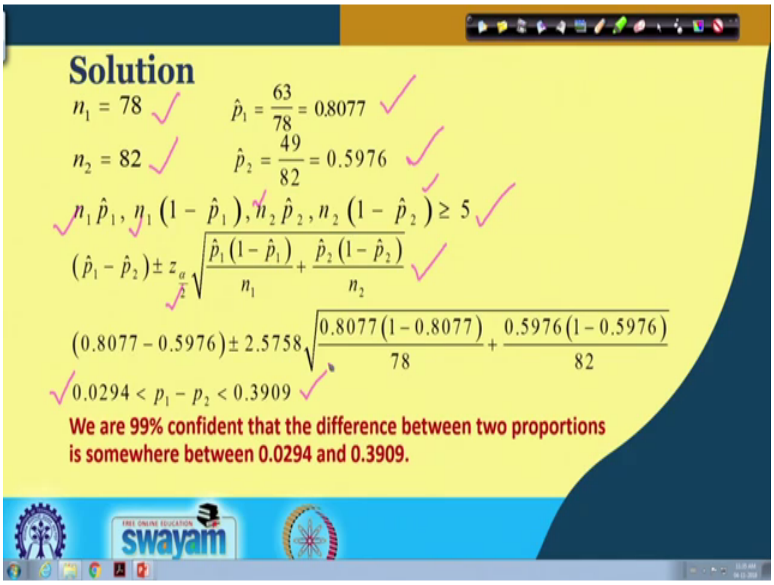
(Refer Slide Time: 12:35)



So, here let us say you have an example that; as a personnel director you want to find out the proportion of fairness of different methods of performance evaluation. So, you just took the sample 63 of 78 employees rated, say method 1 as fair in your sample and 49 of 82 rated method 2 as fair.

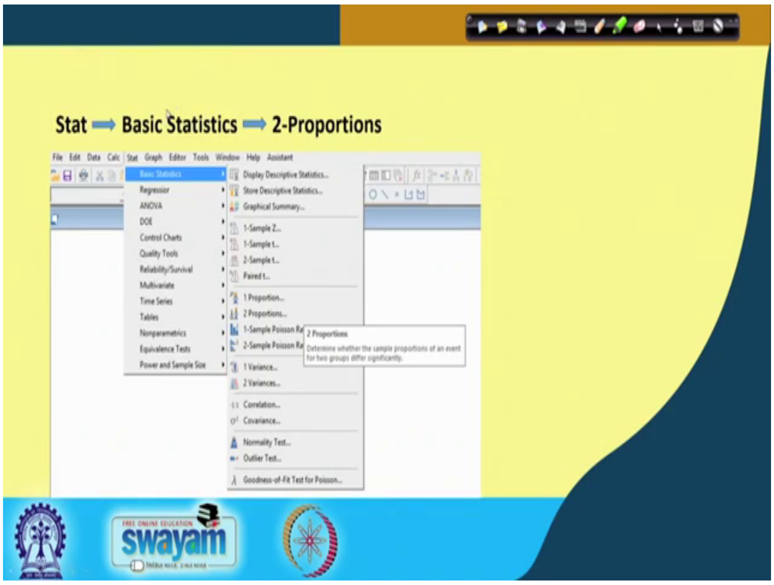
Now can we create the or construct the 99 percent confidence interval. As well as test the hypothesis for this particular problem. So, here I mean to say that whether two methods based on the proportion of reply I have received are same or there is a difference. And here I am talking about the fairness of two different methods of performance evaluation.

(Refer Slide Time: 13:28)



So, you can easily compute this n 1.

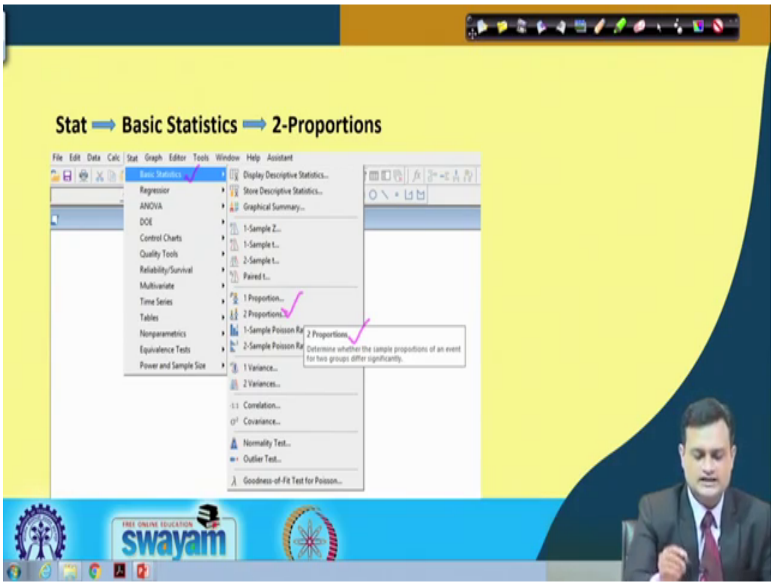
(Refer Slide Time: 13:35)



You can compute n 1 n 2 p 1 p 2 n 1 p 1 n 1 1 minus p 1 hat n 2 p 2 n 2 1 minus p 2 hat greater than equal to 5 satisfies the condition. And you can create the z alpha by 2 confidence interval. So, you will have a confidence interval says that difference in proportion p 1 minus p 2 for two different population for 99 percent confidence. It falls within 0.0294 and 0.3909. So, once again if you have a doubt I would like to clarify.

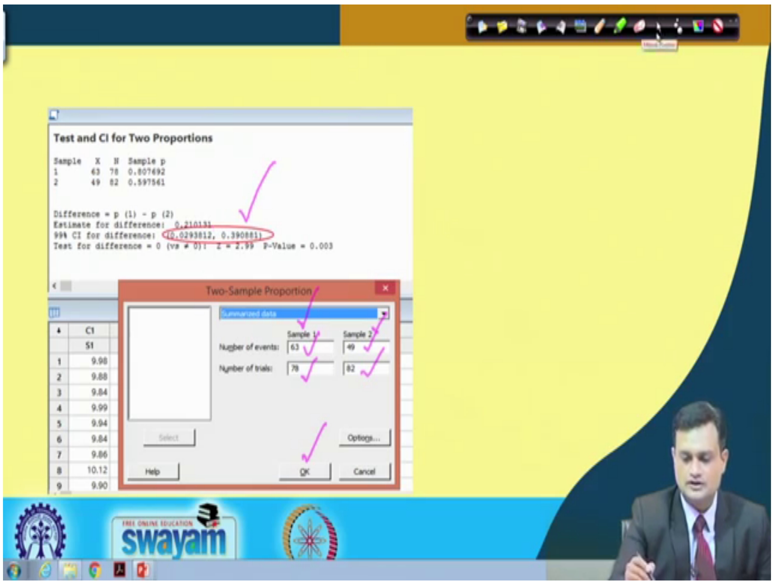
Confidence interval means that for a given level of confidence that here it is 99 percent I am sure that a particular mean or the difference will fall in a particular interval. And this interval will contain a mean or difference in mean or difference in proportion for 99 percent of time or 95 percent of the time depending upon my confidence I want to execute in this particular statistical inferential analysis. So, this is what you have done.

(Refer Slide Time: 15:00)



Now, here we want to see that how I can get the same result through Minitab. And you can do it very easily you can go to basic statistics. You can choose two proportions so you have two proportion test here. And once you have done this.

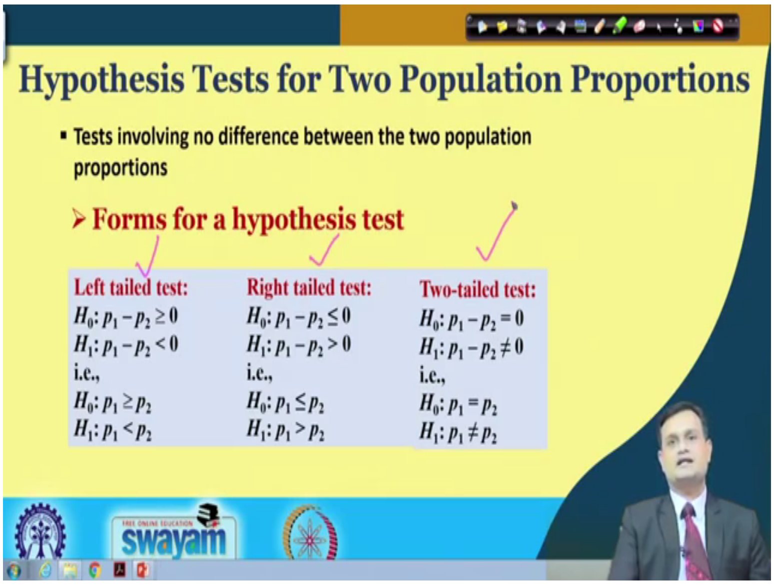
(Refer Slide Time: 15:21)



Then you can get the confidence interval here you can see 0.0293812 and 0.390881 and you have put the various values. You have given this input to Minitab sample 1 you have number of events 63, number of trials 78 and your sample 2 49 82 you press and you get this confidence interval.

So, basically let us say this is 0.029 and 0.39. So, if I compare this with my result then I will say that the calculated value through manual calculation is 0.029 and 0.39. So, both the values are matching and once again say I can do such analysis very easily say using the Minitab application of basic statistics.

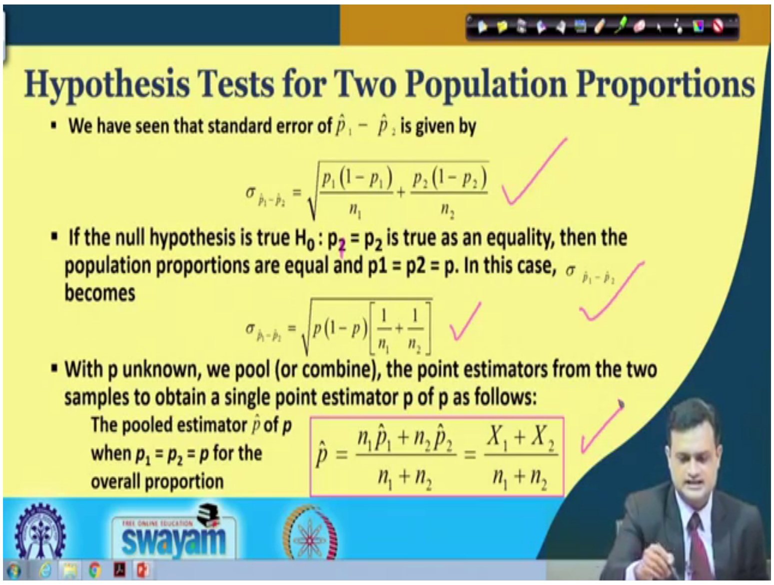
(Refer Slide Time: 16:30)



So, this is what we have seen about the second case. And you may have interest in checking the different kind of hypothesis for left tailed test. For example, you are interested to check whether p 1 minus p 2 is greater than or equal to 0. Or, p 1 minus p 2 is less than 0 right tailed test p 1 minus p 2 less than or equal to 0 or p 1 minus p 2 greater than to 0.

Or two tailed test you would like to see that p 1 minus p 2 they are same statistically or there is a difference. So, p 1 minus p 2 is equal to 0 and p 1 minus p 2 is not equal to 0. So, you would like to check different claims and this is very easily done with the Minitab.

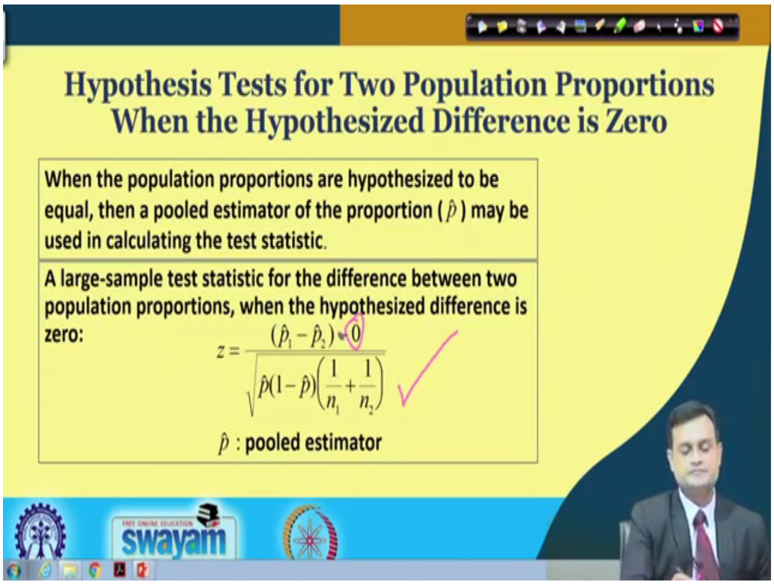
(Refer Slide Time: 17:14)



So, if you just see the expressions for testing such kind of hypothesis then you have sigma p 1 hat that is the estimated value minus p 2 hat is given by square root of p 1 1 minus p 1 divided by n 1 plus p 2 1 minus p 2 divided by n 2. And you have sigma p 1 is equal to p 2. This is p 1 is true as an equality then the population proportions are equal.

And you would like to see that what is my sigma p 1 minus p 2 and you will find the p hat that is the estimation of proportion n 1 p 1 plus n 2 p 2 divided n 1 plus n 2. So, already we have seen this is just small recap of what we have seen. In the last lecture in this lecture we are basically discussing the application of Minitab.

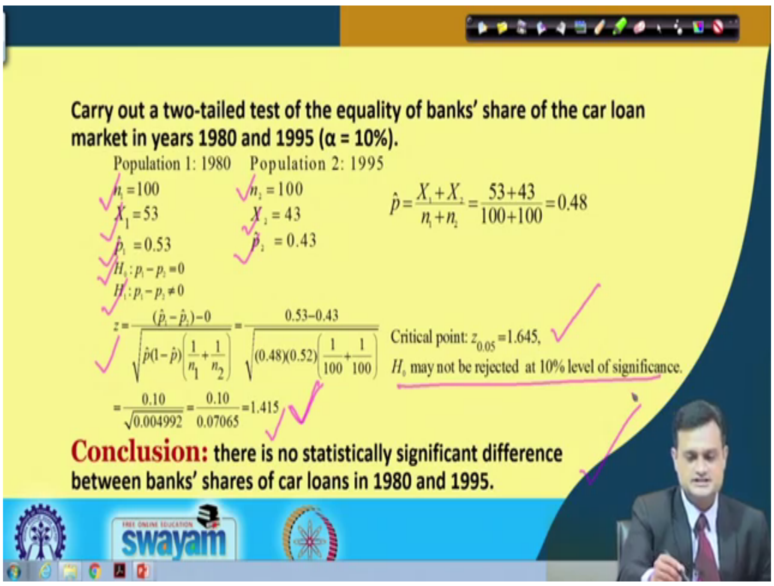
(Refer Slide Time: 18:14)



So, here my z statistics would become p 1 minus p 2 minus 0. Suppose the population proportions are hypothesized to be equal then the pooled estimator you can compute. And may be calculated the used in the calculating the test statistics. .

So, a large sampled test statistics for the difference between two population proportion. When the hypothesized difference is 0, then you will say that this component will become 0 and other expression will remain same.

(Refer Slide Time: 18:55)

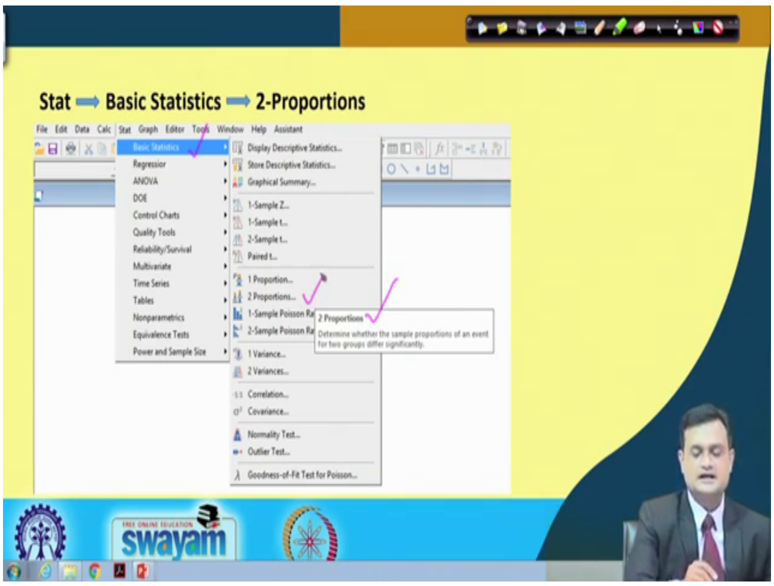


So, now let us see that for this particular situation that carry out two tailed test of equality of bank share of the car loan market in year 1980, 1995 or you can consider the recent year at the level alpha 10 percent or 0.01. So, this data you have n 1 X 1 n 2 X 2 p 1 hat p 2 hat and you want to check your hypothesis that whether p 1 minus p 2 is equal to 0 or p 1 minus p 2 that is the difference in proportion is not equal to 0. So, I am just putting these values in the standardised test statistics that is z.

And what I get is the critical value 1.415. So, my this is the calculated value 1.415 and my critical point or critical value z 0.05 is equal to 1.645. So now, you should get the visual picture of testing the hypothesis in terms of a normal distribution in the rejection region very clearly. So, you can see and that my critical value is 1.645 which is greater than 1.415. So, I compute it from 0. So, you can say that H 0 may not be rejected at 10 percent level of significance.

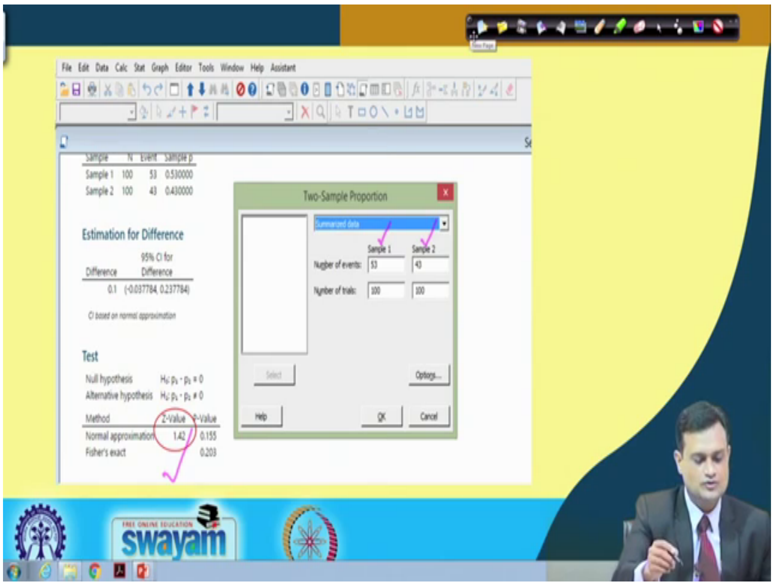
Because this value 1.415 you can very easily visualize that 0 is at the centre in case of z distribution normal distribution. And when you go to the right side this value as it is less than 1.645 will fall in the acceptance region and I will say there is no statistically significant difference between banks share of car loan in 1980 and 1995. So, this is what I conclude from this test of hypothesis. Now, if we go ahead then how to do it in the say Minitab.

(Refer Slide Time: 21:04)



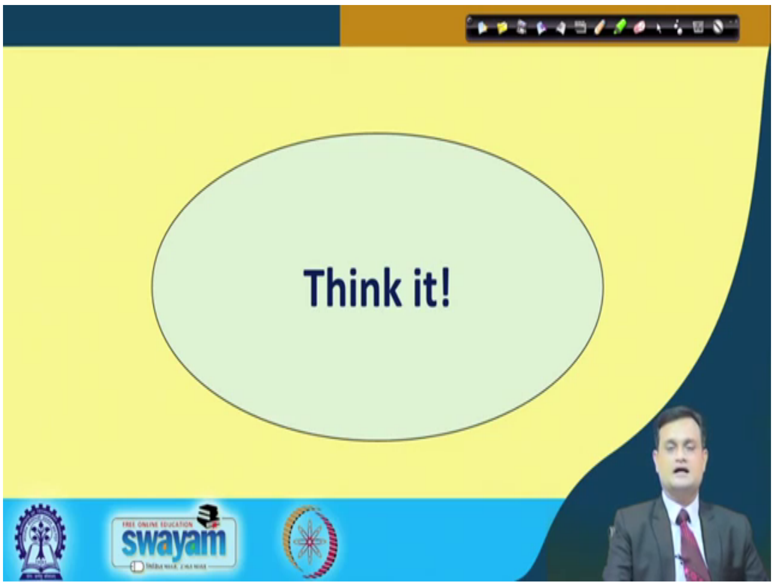
So, you can once again go to basic statistics and then you can choose two proportion. So, you will have two proportion test and this is what it is. And, once again you can put all the values into your say this Minitab input say window.

(Refer Slide Time: 21:25)



So, 5300 4300 and what you get z is 1.42 which is very much matching with the manual calculation for z for testing the difference in proportions of two different population. And this is how you can easily check the hypothesis p 1 minus p 2 is equal to 0 and p 1 minus p 2 is not is equal to 0.

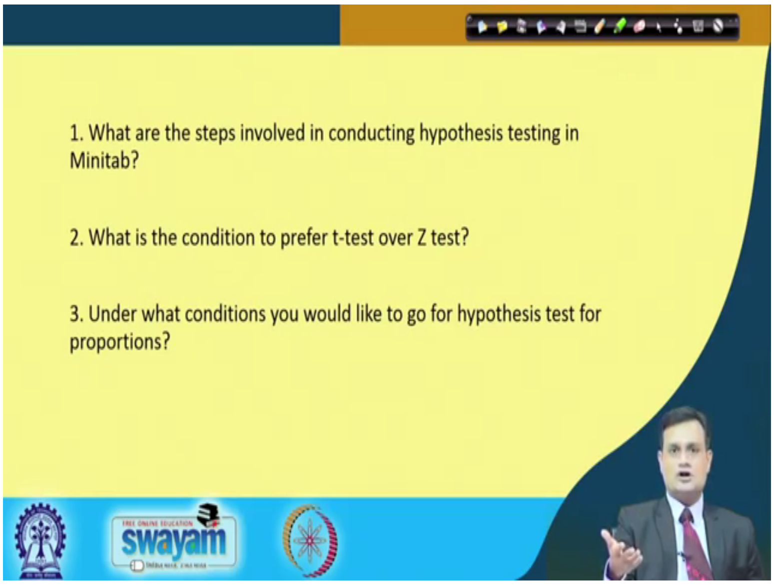
(Refer Slide Time: 22:06)



So, I hope this particular illustration must have helped you to clarify that; how hypothesis testing even can be easily done in the Minitab. This session will really relax and help the practising managers and people. Because many a times say they are in the field they do not time; they do not find the time to go for manual calculation or there is a long gap in their educational study. And now then they are learning this particular course; so, they always look for some software application and that can help them to quickly get the result just only they need to interpret it.

So, this particular lecture will give you both the population of my course one is the student regular student who are in the universities. And other is the practising manager who are working in the industry may be in the executive position, managerial position. So, this session will help both the population to appreciate the application of Minitab for hypothesis testing. And immediately they can check the validity of their claim through such kind of analysis. So, before we say end this session as a usual practice I would like to float couple of think it.

(Refer Slide Time: 23:28)



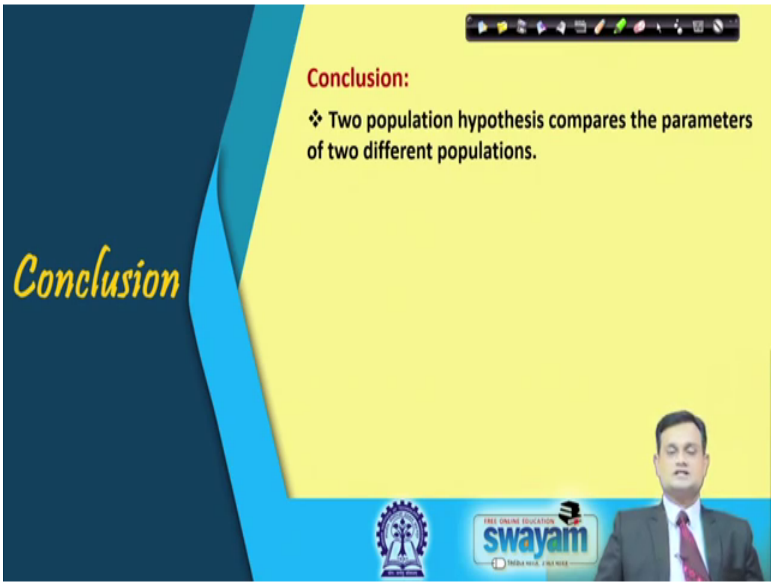
So, that you can introspect and revise the delivered content. So, what are the steps involved in conducting hypothesis testing in Minitab? In general you know the steps now if you have to execute this then what you would like to do in Minitab. What is the say condition to prefer t-test over Z-test? A little small recap and under what conditions you would like to go for hypothesis test for proportion. So, just try to think over these issues.

And this is already answered you can properly say take a note of this lecture and you can easily reply to this questions.

(Refer Slide Time: 24:10)

So, these are the references you can further use for strengthening your concept on hypothesis testing for different cases as we have discussed.

(Refer Slide Time: 24:20)



And, conclusion is that two population hypothesis compares the parameters of two different population. And, this is where I would be able to make concrete and more inferential conclusions above two different populations.

So, thank you very much for your interest in learning the Minitab application in testing the for two population. And, I hope this lecture would have given you adequate insights in operating the Minitab and not only operating the Minitab, but interpreting the results. So, please keep revising, clarifying your doubt be with me enjoy.