Business Analytics for Management Decision
Prof. Rudra P Pradhan
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

Lecture – 22
Inferential Analytics Part – 2 (Contd.)

Hello everybody this is Rudra Pradhan here, and welcome to BMD lecture series. Today our discussion is on inferential analytics 2 and that to we are in the second lecture of this particular unit. In the last lectures we have discussed details about this particular you know analytics and that to in a kind of you know multivariate frameworks.

So, here we are actually addressing some of the business problems where at least you know 2 variables or 2 samples must be in the systems, having 2 variables or 2 samples in the systems we like to check, what is the difference between these 2 sample mean and whether the sample means 2 different sample means are actually is significantly different from the population means.

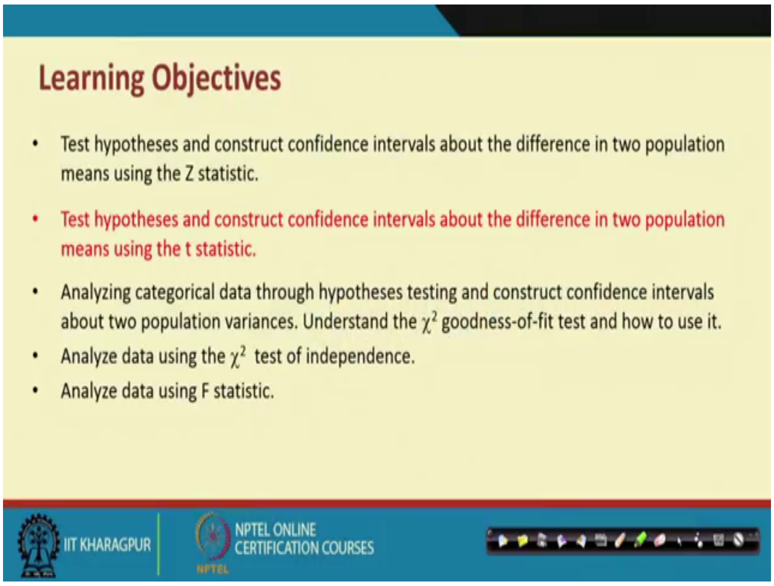
So; that means, actually, the kind of you know comparison is again to see how the sample difference you know giving the kind of you know inference to about the population mean difference and in the earlier case we have discussed this same issue through Z statistics where the kind of you know structure is little bit different, but in this case the particular you know case is through t distributions where the structure is a slightly you know again different.

So, in the first case in the class which you have already discussed that is actually through Z statistic where the condition or the requirement is your sample size should be high; that means, if there are 2 samples then the first sample and second sample both are actually should be greater than to 30 and if the sample size for both the samples are not you know greater than to 30 at least one should not be greater than to 30 then actually we cannot solve the problem through the Z statistics or you know Z distribution then the same problems we can solve through t statistics.

So, in this lectures we dissolve some of the problems by connecting t statistics. So, here the typical problem structure will be sample size will be slightly lowers that is less than to 30 and again this the structure is more or less same we are drawing 2 samples from the

2 populations and then checking the mean of this you knows 2 samples are what is the you know structure through the difference about the population mean.

(Refer Slide Time: 02:50)

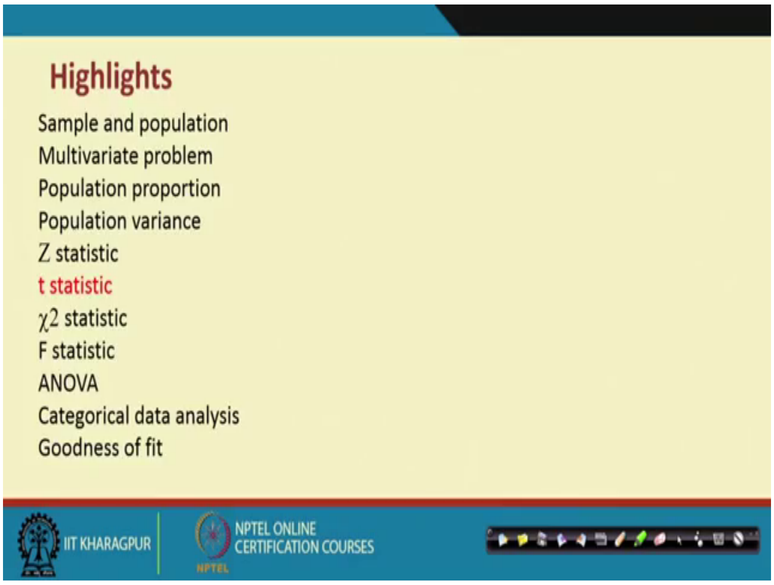


So, now accordingly we are discussing the particular structure through you know commenting about the population parameter; that means, the mean difference of 2 samples to mean the difference you know population mean difference of 2 population mean.

So, now we like to check or you know like you know say x 1 bar is the population mean sorry sample mean of the first sample and x 2 bar is the sample mean of the second sample and then the difference will be x 1 bar minus x 2 bar and against the population parameters or population mean for the first one is mu 1 and the second one is mu 2. So, now, we like to see the difference between sample mean and the difference of the population means whether they are you know same or there is a kind of you knows difference.

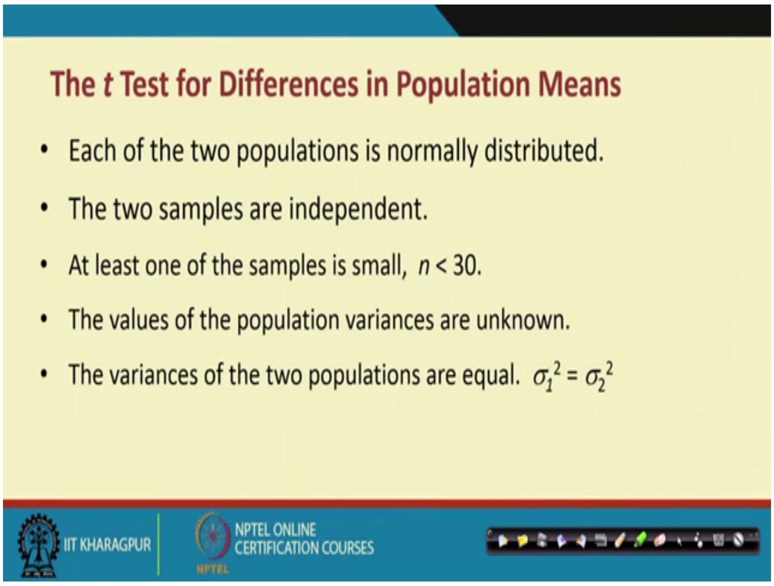
If there is a difference whether difference whether the difference is statistically significant or not and that to we are testing this through t statistics where the means it is the case where sample size will be slightly lower to 30.

(Refer Slide Time: 03:48)



So, then we will see the situation the same structure which we are you know discussing in the last lecture.

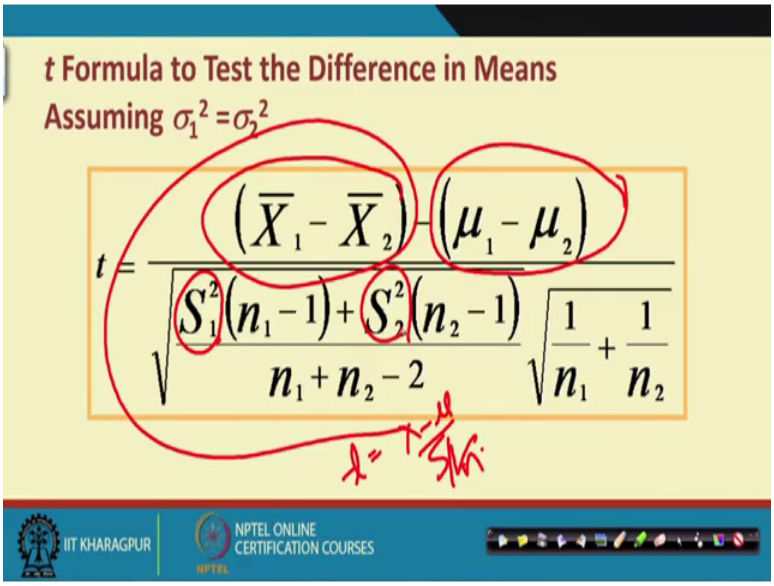
(Refer Slide Time: 03:52)



But we are actually repeating the typical case with you know different situation and in the particular structures. So, the problem structure will be like this in each of the 2 population means both the population should be normally distributed and 2 samples are independent and at least one of the samples should be less than to 30 the then the values of the population variances are unknown and the variance of the 2 populations are you know equal. So, now, corresponding to Z distribution or Z statistics where the mean of these 2 particular you know structure will be is same and then we are testing through the sample mean.

So, now here the particular structure is a slightly different with respect to sample size, with respect to with respect to the kind of you know population variance and with respect to the kind of you know structure about the population parameter right.

(Refer Slide Time: 04:53)

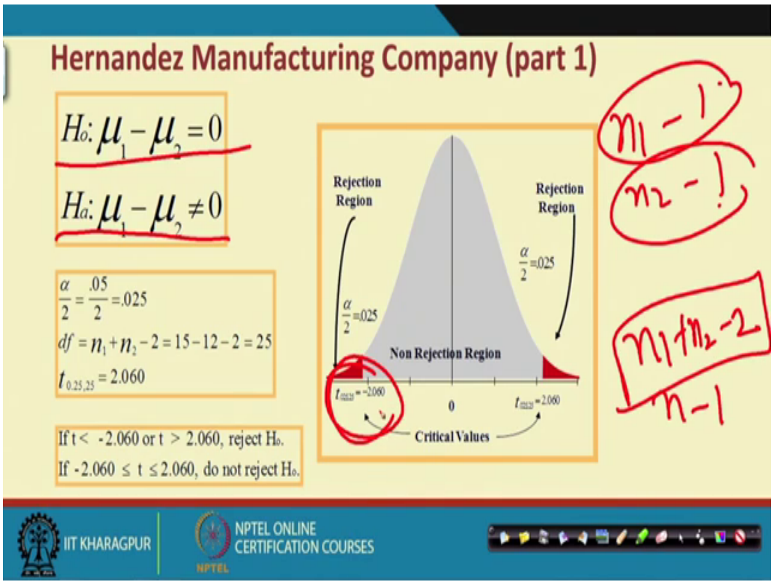


So, corresponding to this, you know testing procedurally follow these same structures like you know we have discussed earlier in the case of you know one sample case.

In the 1 sample case your t statistic will be X minus mu by S by root n. So, here we have a 2 different samples, as a result, in place of X we are putting the difference between X 1 bar minus X 2 bar and in the case of you know mu we are putting the difference between mu 1 minus mu 2 and then this is the standard error and which is actually calculated with the help of you know variants of ones for samples and variance of the second samples and this is the adjustment factors and once you get all this value then you can actually easily try to find out the difference and then we try to check whether the difference is statistically different and the same structure we have to follow.

So, you have to create a confidence interval and then we would like to compare the t critical with t calculated if the t calculated will overtake the t critical then we are actually going to reject the true null hypothesis otherwise we will accept the true null hypothesis accordingly we will actually analyze the particular case let us discuss this problem with the with you know examples and then we can get to know the kind of you know inference.

(Refer Slide Time: 06:15)



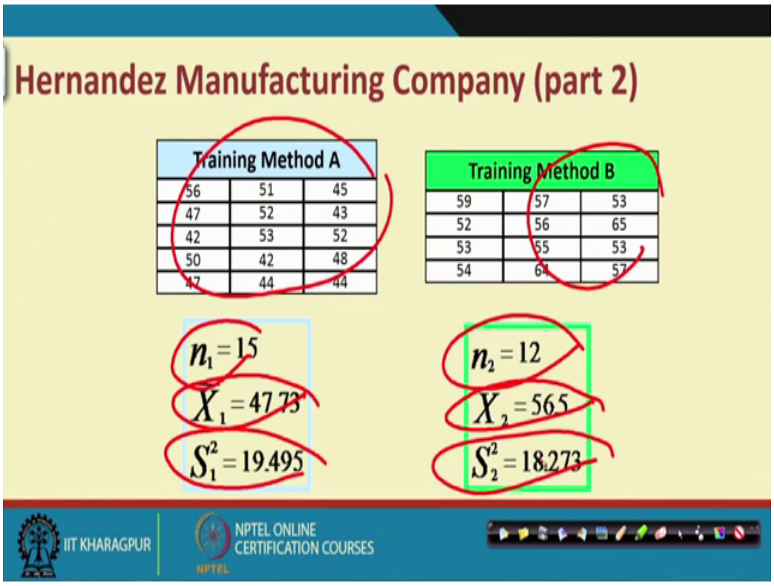
So, now it is the same structures, we are checking the case where you know there is you know equal variance and then there is a difference between you know 2 means right. So; that means, with the assumption that you know the null hypothesis that you know there is no difference between the 2 populations means as a result. So, mu 1 minus mu 2 equal to 0 like the case of you know Z statistics. So, this same is same you know hypothesis here and then the corresponding alternative hypothesis will be the difference between the not equal to 0 and then we are fixing the alpha first. So, alpha you know 5 percent and that to we are targeting 2 different sites.

So, as a result, the left hand side will be alpha by 2 that is 0.025 and the right side will be alpha by 2 again 0.025 and to know the t criticals, we need 2 things, the alpha value and the corresponding degree of freedom. So, in the case of you know of 1 sample case your degree of freedom is nothing, but n minus 1 and since here there is a 2 sample case. So, then the degree of difference will be n 1 plus n 2 minus 2. So, for instance there are 2 samples n 1 minus 1 and n 2 minus 1. So, this is the second case and this is the first case and as a result the overall degree of freedom for this particular structure will be n 1 plus n 2 minus 2.

So, now with the help of you know alpha and degree of freedom we can get to know the critical value for instance in this case having actually n 1 equal to 15 and n1 n 2 equal to 12 and fixing alpha equal to you know 0.5 in that is 0.025 for lower limit and the 0.025 for you know for upper limit then accordingly you are getting actually the critical value minus 2.06 so; that means, actually it is the total case and having alpha value and then you know degree of freedom. So, you will get the t critical right, as a result, your t critical is here 2.06.

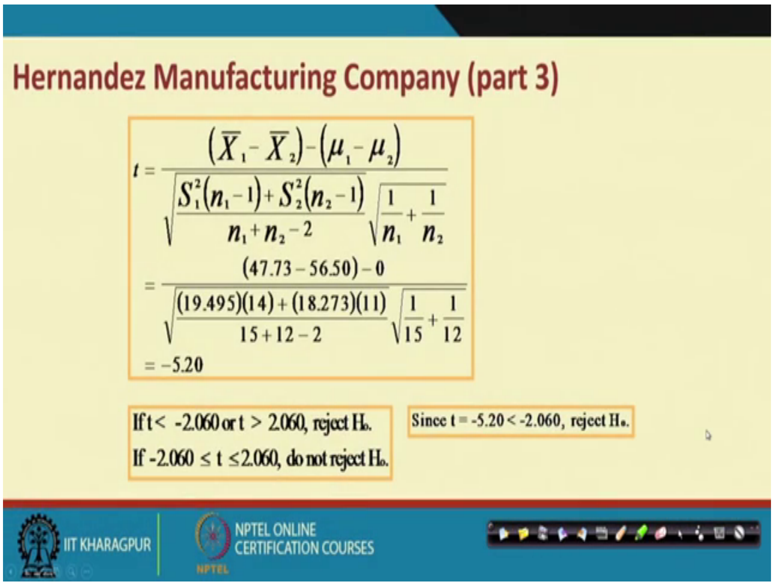
If you are coming to the left side then by default it will be minus 2.06 if you are coming to the right side then it will be; obviously, 2.06. Now, depending upon your t calculated, you have to take the decisions. So, for instance if it is actually t value is coming a positive then we can you know reject true null hypothesis provided the t value will be over take 2.06 and similarly we can reject the true null hypothesis if the t value will be lower to minus 2.06.

(Refer Slide Time: 09:07)



We are taking 2 different samples this is sample 1 and this is sample 2 and then in this samples our sample size is actually 15 and in this case your sample size is 12 and then we are calculating mean of this particular series and variance of this particular series, again mean of this particular series and variance of this particular series once you get all these informations then by default you go to this particular you know a structure.

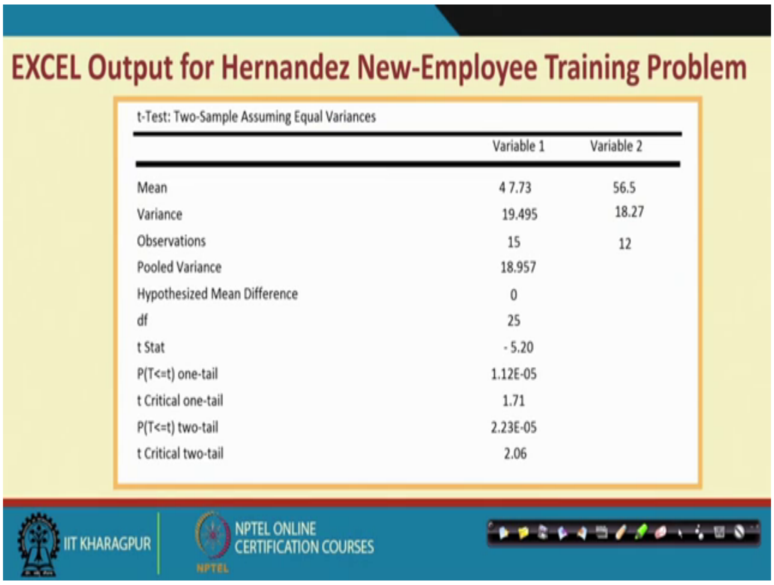
(Refer Slide Time: 09:36)



Then you just report these values and then simplify and after simplifications you will get the value you know you not value the means the t value is nothing, but actually coming minus 5.20 and earlier through the you know alpha and degree of freedom we have already reported the t criticals is nothing, but actually minus 2.06 and since it is actually coming minus 5.20, as a result, it is lower to the lower limit of you know minus 2.06.

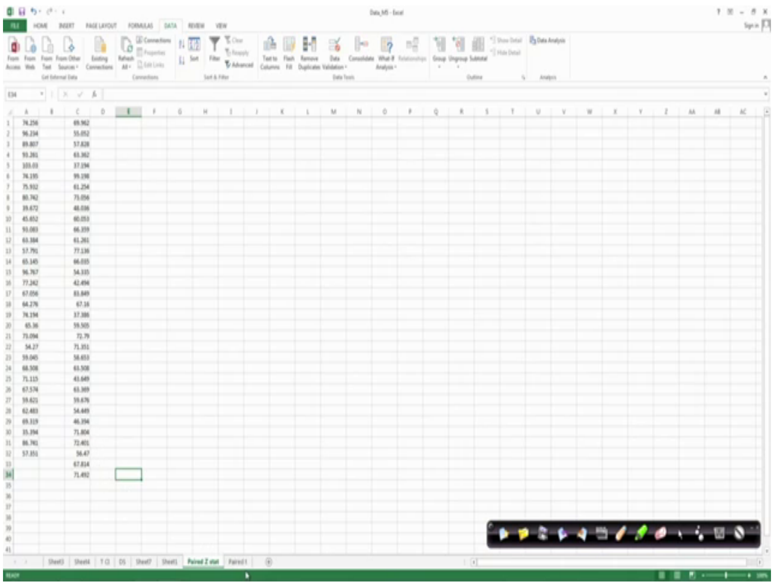
So; obviously, we are in a position to reject the true null hypothesis; that means,. So, the mean difference of these 2 means in the sample mean difference is not exactly same with you know population mean difference. So, there is a difference between the mean sample mean and you know population mean. So, sample mean difference and population mean difference and these differences are actually statistically significant. So, accordingly the business needs to be actually in you know discuss and then you know the kind of an inference can be drawn accordingly. So, with this actually we can actually move to a particular situation.

(Refer Slide Time: 10:53)



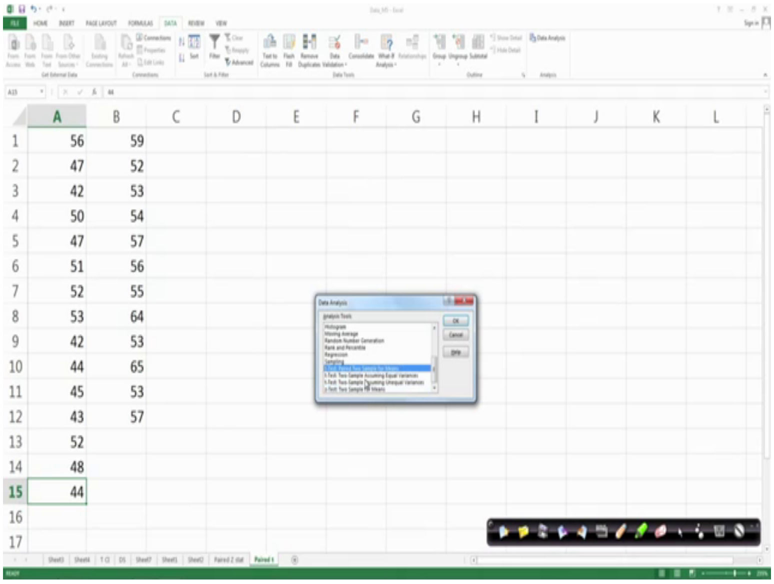
So, this is how the particular structure and then I will take you to actually the excel problems you know spreadsheet then I will highlight how actually how these things can be solved very easily.

(Refer Slide Time: 11:04)



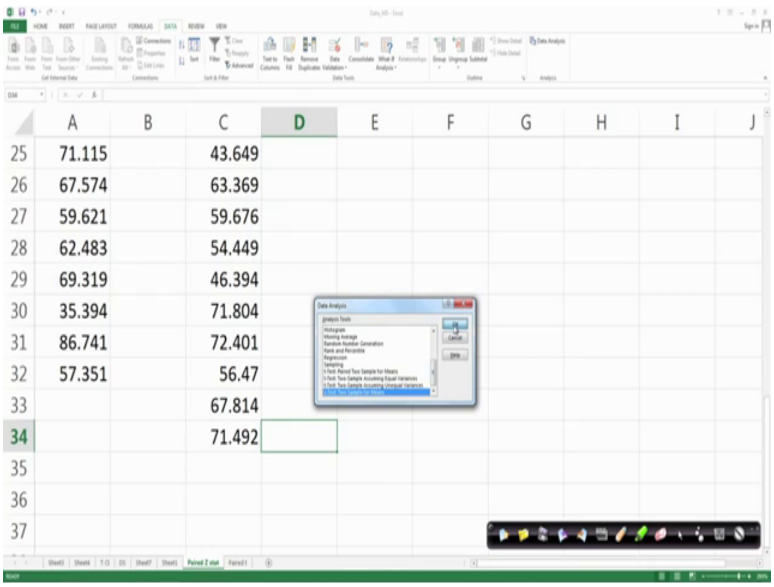
And let us assume that you know this is actually 2 sample case and these are the sample you know points and in the last lecture which you have already discussed you know through Z statistics and then here actually we are discussing the same thing through t statistics.

(Refer Slide Time: 11:27)



So, now since it is actually big samples and this can be applied through Z statistics and then against this is a kind of you know problems which is actually it can be solved through t statistic because the sample size is less than to 30 in this case your sample information is 32 and this is 34 so; that means, it crossing it is crossing in the limit of you know n 30. So, now, I will just show you here. So, let us actually see here is the extent this one. So, this is actually 2 sample case and we have a first sample 32 observation and a second sample 34 observations.

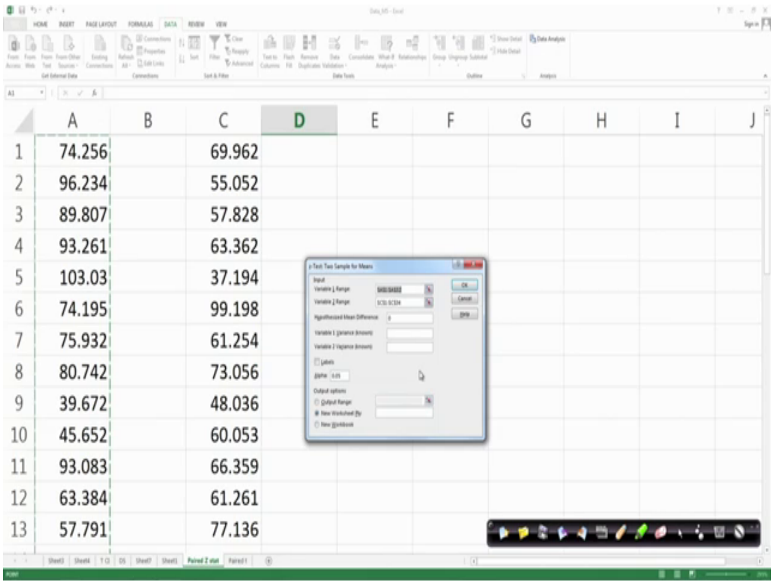
(Refer Slide Time: 12:04)



Now, if you go to the actually data analysis package then you know I will just show you know the previous example case that is through Z statistic and again, I will connect through t statistics.

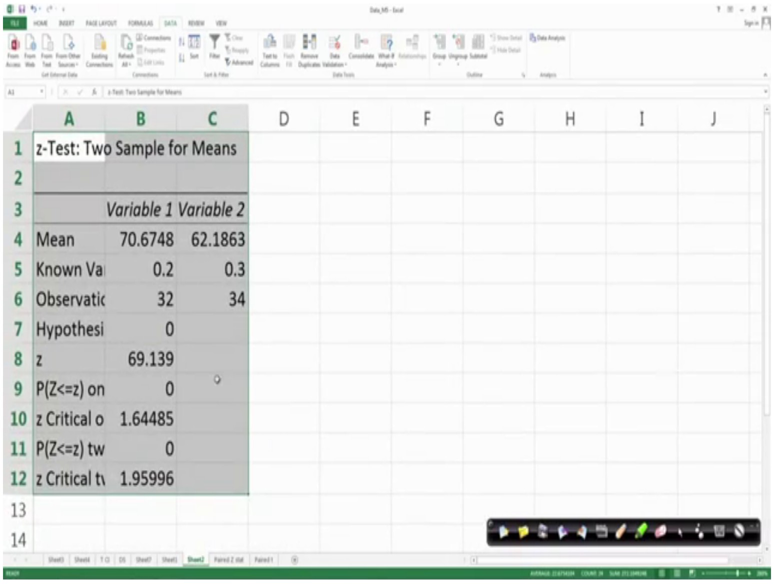
Now in this case we can strictly apply Z statistics you know because the sample size is actually crossing 30. So, if you go to the data analysis package then this is the option for you know Z test and then it would just and this will show you the kind of you know difference.

(Refer Slide Time: 12:30)



So, what will it do, it will be asking you to put the variable range for 1 and variable range for 2, as a result, you just click the variable range here. So, specify whatever variables you have. So, we have actually 32 for the first ones then against you have to specify the second one. So, go to the second tables. So, against we have to highlight for the second one. So, this is coming again up to say 34. So, now, the hypothesis mean equal to 0 then you just put. So, it will be asking that you know very there should be you know reporting of the you know variance you assume that there is a variance factors, put 0.2 some like this and 0.3 like this.

(Refer Slide Time: 13:27)

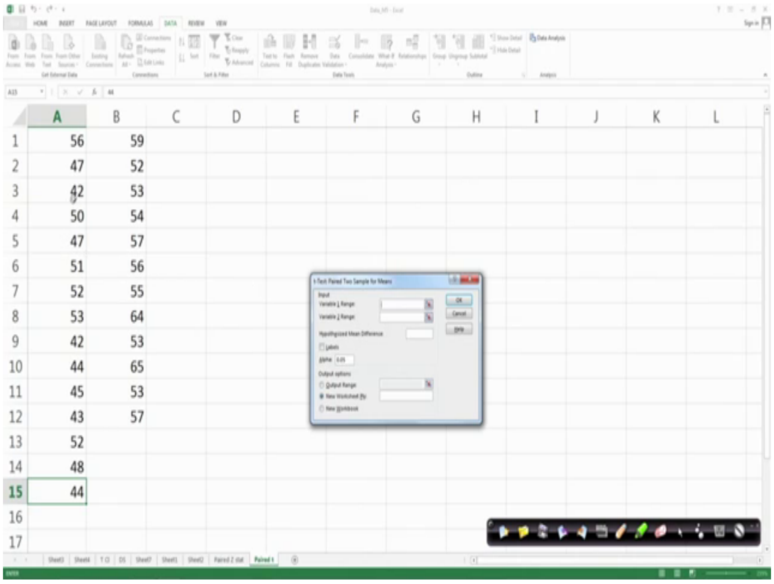


So, then you could just equal right. So, then you will find the kind of you know results is there. So, there is difference and the Z value is coming to the significantly high and it is also statistically significant so; that means, actually with the help of you know 2 samples and you know you can find out the difference and then we can check the difference whether they are statistically you know different to population difference.

Now in this in this you know discussions we are actually dealing with it is you know t statistics and the similar kind of you know problem, but here the structure you know we can use you know t statistic provided your sample size is less and corresponding to earlier problems. So, the sample size is it 32 and 34 here we have a sample size for 15 and another variable sample size is actually 12 so; that means, actually you know 2 different samples and the first sample is having 15 points and then second sample is having 12 points.

So, now again you go to the data analysis package, then you have actually here the t test you know different options. So, the first actually t test where 2 sample of means you just put ok and then you specify the range.

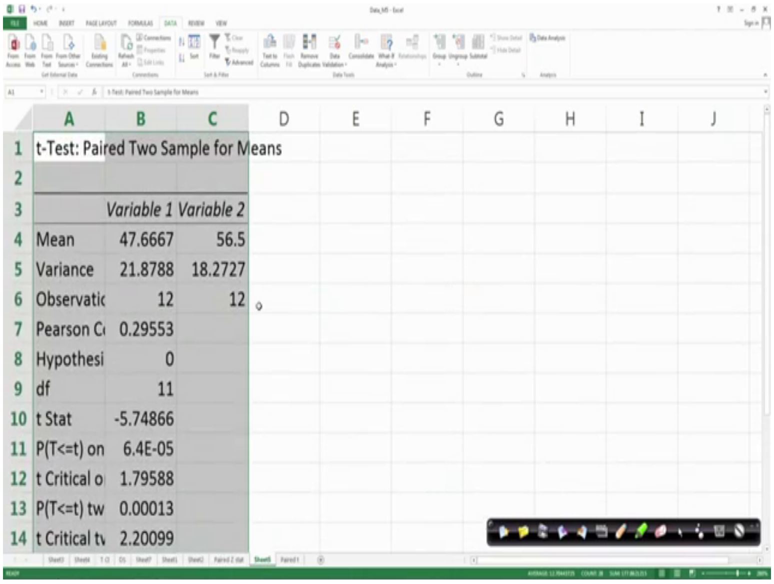
(Refer Slide Time: 14:39)



So, starting with the 1 to f 15 and then again you go to the second variables and specify the range B 1 to B 2 well depending upon the availability. So, then you just click see here is the message is coming that you know variables range must have the same number of data points so; that means, t statistic has some kind of you know restrictions.

So, what will it do again, to solve this you know you little bit you know restructure the particular process so; that means, you can test this process up to actually 12 samples which is actually uniform for both samples now you know you hypers mean difference you put actually 0 here then you put. So, this is how the result is you know or about right.

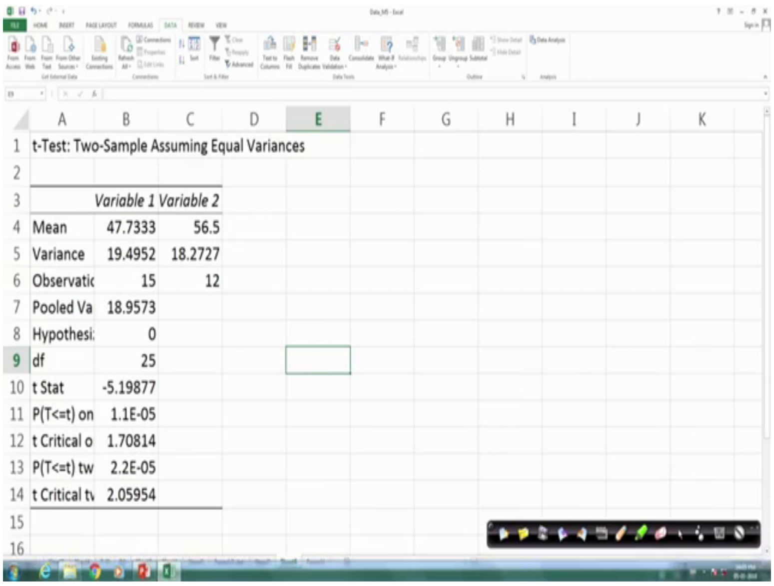
(Refer Slide Time: 15:27)



	A	B	C
1	t-Test: Paired Two Sample for Means		
2			
3		Variable 1	Variable 2
4	Mean	47.6667	56.5
5	Variance	21.8788	18.2727
6	Observatic	12	12
7	Pearson Cı	0.29553	
8	Hypothesi	0	
9	df	11	
10	t Stat	-5.74866	
11	P(T<=t) on	6.4E-05	
12	t Critical o	1.79588	
13	P(T<=t) tw	0.00013	
14	t Critical tv	2.20099	

So; that means, you know having actually 2 sample informations and that 2 sample size of you know less than 30 then you can apply t statistics and then you say check the difference of these 2 sample means and then you can give comment about the population parameters and difference.

(Refer Slide Time: 15:38)



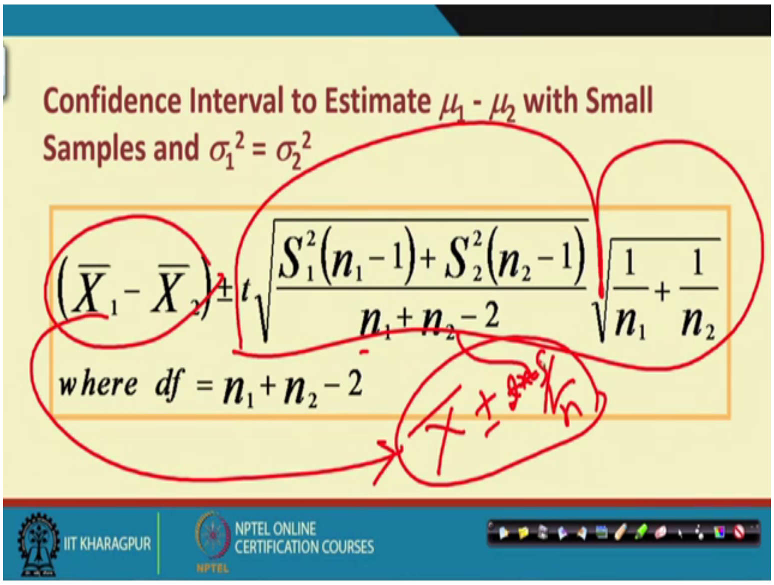
	A	B	C
1	t-Test: Two-Sample Assuming Equal Variances		
2			
3		Variable 1	Variable 2
4	Mean	47.7333	56.5
5	Variance	19.4952	18.2727
6	Observatic	15	12
7	Pooled Va	18.9573	
8	Hypothesi	0	
9	df	25	
10	t Stat	-5.19877	
11	P(T<=t) on	1.1E-05	
12	t Critical o	1.70814	
13	P(T<=t) tw	2.2E-05	
14	t Critical tv	2.05954	

So, here the differences between the 2 are again statistically coming actually significant. So; that means, it is actually minus 5.74 and which we have already got in this particular you know structure, as a result 5.74. So, it is coming actually here then it is yet a statistically significant and accordingly you will find the particular you know difference, again you go to this particular spreadsheet again and you know see here. So, what will you do here t statistic we have a different options altogether and then again you put actually t statistic 2 samples which you know equal variance again you specify the range and then again you go for second samples and specify the range and then you put and then you will find there is a kind of you know difference and the difference again it is coming statistically significant.

So, likewise actually you can actually solve some of the problems where we have a 2 samples and that to your sample size will be a less than to 30 at least one sample size should be less than to 30 and then you can you know analyze and you can compare accordingly and the idea is here to check a whether you know sample mean difference or you know converse to population mean difference if there is a difference whether the difference is coming statistically significant or not against by doing this process you can also create a confidence intervals for you know the difference of these population parameters so; that means, technically we like to know.

So, what is the difference you know confidence interval for population parameters. So, now this is what the discussion corresponding to the previous problem.

(Refer Slide Time: 17:38)



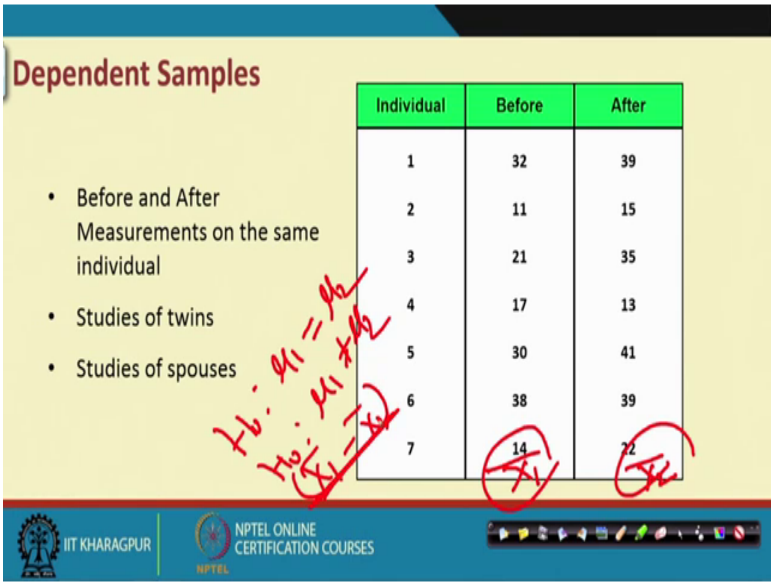
And then like the previous discussion which you have actually done through Z statistics and that to create confidence intervals when there are you know 2 samples you know to sample case you know there. So, now, in this case also by using t star you can also create a confidence interval for the population parameter.

So; that means, technically in the 1 sample case the kind of you know confidence interval will be X bar plus minus t upon 1 S by root n. So, this is what actually we have discussed earlier in the case of you know inferential analytics 1 and in this case we have actually 2 samples and the sample mean difference is nothing, but X 1 minus X 2 bar. So, as a result, this particular value will be coming in place of you know X bar and then plus minus again in the case of you know t upon s by root n.

So, this much of you know value this much value will be coming under you know this place and then again you have to create a kind of in a kind of you know confidence interval. So, again, here what will it do you to first fix the kind of you know alpha value and once you get the alpha value then you calculate actually t value corresponding to alpha and you know degree of freedom then with the adjustment of you know variance. So, you can able to calculate the kind of you know confidence intervals.

So, let us take an example and then we will highlight this particular issue. So, the examples you know again, for showing this one is.

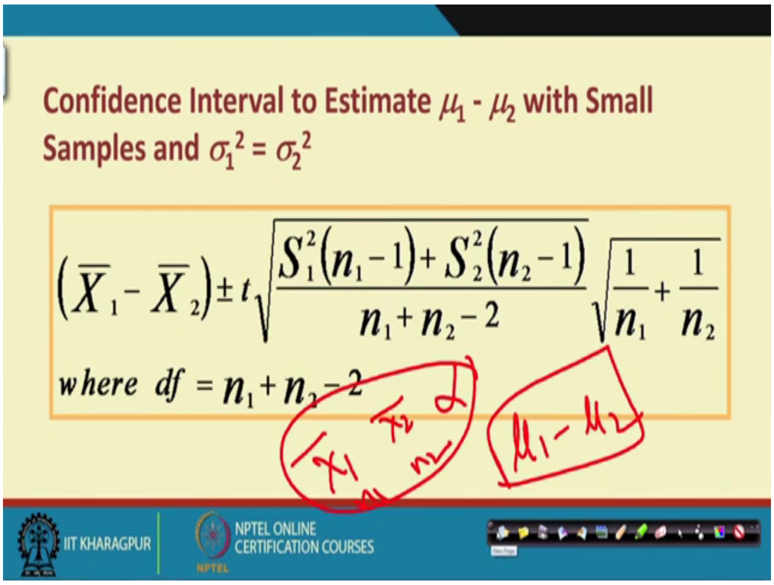
(Refer Slide Time: 19:16)



So, we will take actually all right the same examples which you can have here. So, you can use these informations like you know. So, these information you can use and then you can actually predict the confidence interval for the you know population parameters right and in this case so.

So, you know by using this information you can create a confidence interval that is for you know population parameter.

(Refer Slide Time: 19:53)



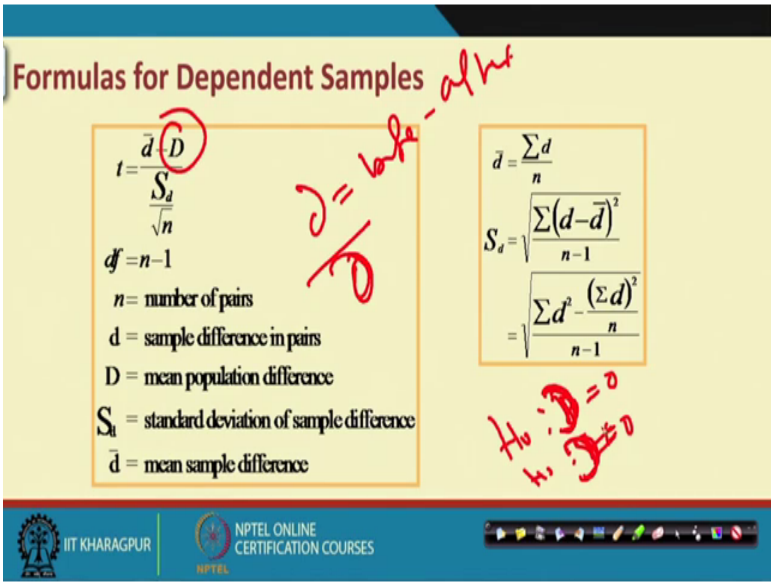
And that to know the difference between mu 1 minus mu 2 corresponding to the information about X 1 bar X 2 bar n 1 and n 2 and the kind of you know sample you know alpha value. So, once you have all these informations then you can in a position to create a confidence interval for you know for 2 population parameter the difference between 2 population parameter and a the same problems against the interesting thing is that here in the inferential analytics to solve this problem differently a now I will show you the particular case.

So, here actually 2 sample case and these 2 sample case represents here actually before and after measurements of the same individual it is a kind of you know some testing and before after. So, we like to check whether there is a difference between you know before and after situations. So, for instance if you if you solve this problem and you know you know before and after so; that means, this will be X 1 and this will be X 2.

So, as a result you to find out X 1 bar and X 2 bar and assuming that you know population will be mu 1 for this once and mu 2 for this after case then assuming that there is no difference. So, as a result all hypothesis will be mu 1 not equal to mu 2 and then alternative hypothesis will be mu 1 not equal to mu 2 and then you find out the mean and then you find out the difference X 1 bar minus X mu 2 equal to 0 ultimately we are actually checking the difference between 2 sample mean only subject to adjustment with you know standard error.

So, again you know since it is a small sample case you have to apply the t statistic and solve this problem, but the same problems can be transferred into 1 sample case and then solve through simple t star t you know t statistic you know structures. Let us see how it can be done. So, now, I having this particularly information you can solve this problem by this formula, but this formula and means you know particularly using this formula you can actually solve this problem, but what will you do this same problems can be solved through you know different kind of you know approach and that to by using a 1 sample case so; that means, what will it do here.

(Refer Slide Time: 22:25)



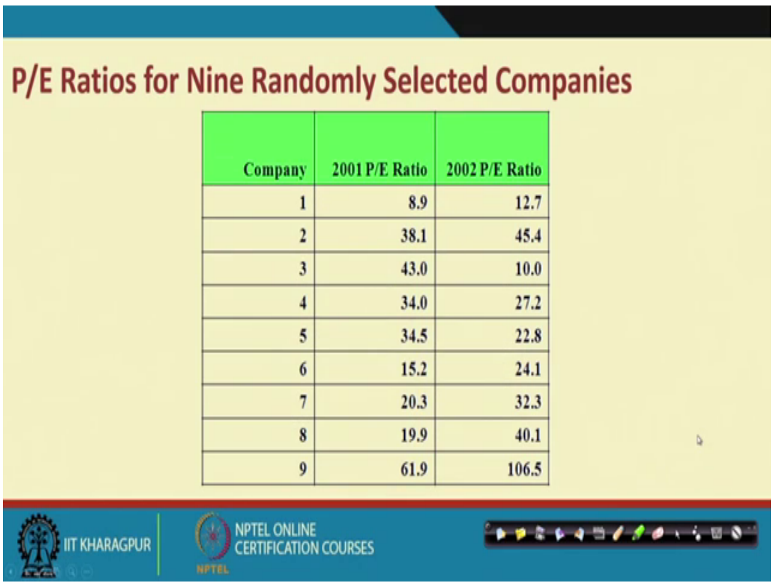
So, since it is a kind of an experimental process you have a before and after situation. So, we like to create a new variable which is called as actually D we need to create a new variable called as you know d and then the D is nothing, but the difference between a the before and after situation and then we are assuming that you know there is a the difference is nothing, but is sample single sample and that would be drawn from a particular you know populations.

So, as a result, once you calculate a you know D a then you know d j kind of you know 1 samples and then you have to calculate the mean of the dj series now and you have to calculate the mean of this particular d series and that to represented as a d bar and then that need to be tested against so; that means, the null hypothesis here that you know d bar equal to 0 alternative hypothesis d bar not equal to 0 so; that means, actually d bar is a population you know sample parameter. So, corresponding to the sample the population would be actually represented as a capital D here in this particular formula and then, d j not equal to 0.

So, now on the kind of you know the structure will be to calculate the first sample mean and that is the difference between before and after situation and then you have to calculate the mean of this particular difference and that is what actually called as you know sample mean and that to now it is the case of you know 1 sample case.

So, now, this same structure we have to follow, assuming that you know d equal to 0. So, ultimately, mean of the d need to be tested. So, mean of the D divided by standard error of this particular you know structure and ultimately you will get the kind of you know structure called as you know t calculated, again find out the you know critical value by knowing the degree of freedom and the alpha value. So, accordingly you can actually proceed here and then you check what is the difference is coming.

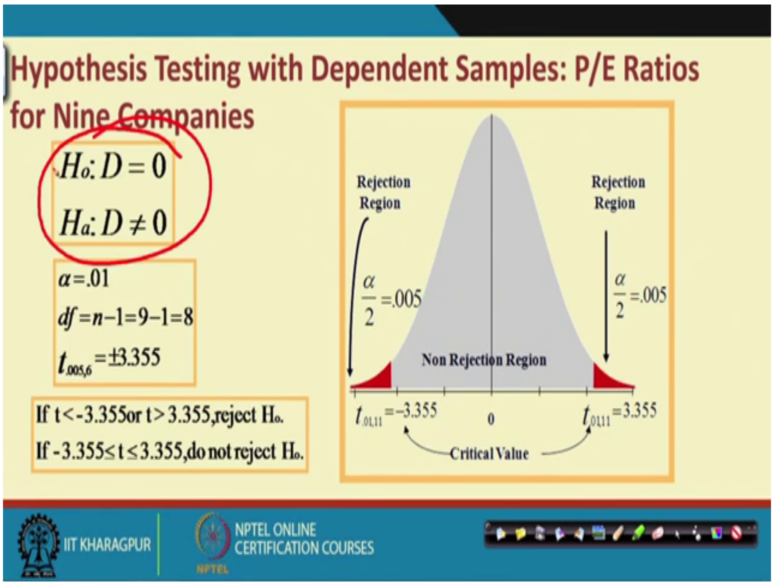
(Refer Slide Time: 24:40)


P/E Ratios for Nine Randomly Selected Companies

Company	2001 P/E Ratio	2002 P/E Ratio
1	8.9	12.7
2	38.1	45.4
3	43.0	10.0
4	34.0	27.2
5	34.5	22.8
6	15.2	24.1
7	20.3	32.3
8	19.9	40.1
9	61.9	106.5

Let us assume that this is the another kind of an example and this is a 2001 case, in 2002 case that is price earning ratio and we like to know whether there is a difference between these 2 different years the same things what will it do.

(Refer Slide Time: 24:59)



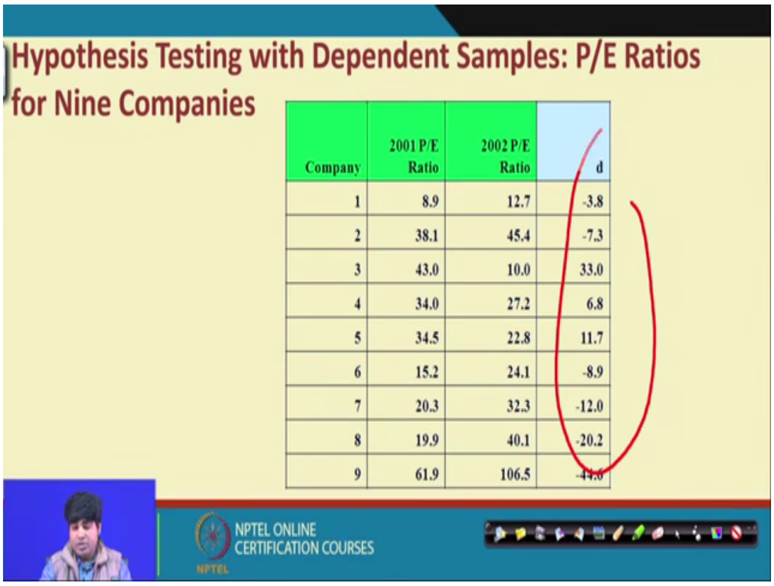
So, we will find out the kind of you know the transfer these 2 sample to 1 sample by subtracting from you know 2001 to 2002 or 2002 to 2001 and you will get a particular series and that series will be represented by your populations and then accordingly we will specify that you know null hypothesis and alternative hypothesis will be like this. So, capital D equal to 0 and capital D not equal to 0 and corresponding to this you know degree of freedom.

So, since actually in this case if you go through and you know the particular problem you will be find there are you know 9 observations. So, as a result, degree of freedom will be 8 and now fixing alpha at 1 percent your critical value at alpha 1 and degree of freedom 8 is nothing, but coming actually 3.355. So, that is what actually since actually it can be both the sides. So, now, if it is total then this value will be coming in the left side and this will be coming in the right side.

Now we will check what is the calculated figures accordingly the calculate figure will be placed whether it will be left hand kind of in our critical zones or right hand side of the critical zones and then accordingly we will take a decision to reject the true null hypothesis.

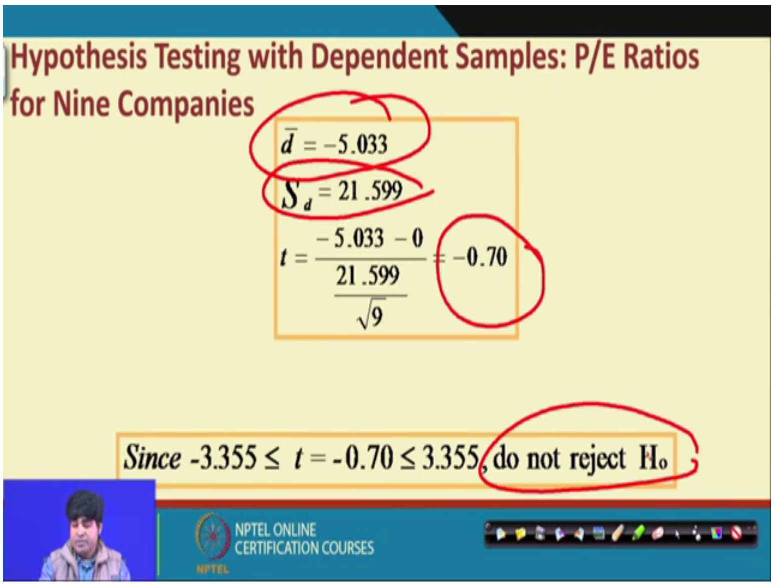
So, accordingly, if you move to this particular you know problem.

(Refer Slide Time: 26:32)



So, you will find the particular structure is coming this is how the difference and this is what the you know small d and that is the difference between the 2 different samples and after knowing this d. So, you can move to the current kind of you know calculations.

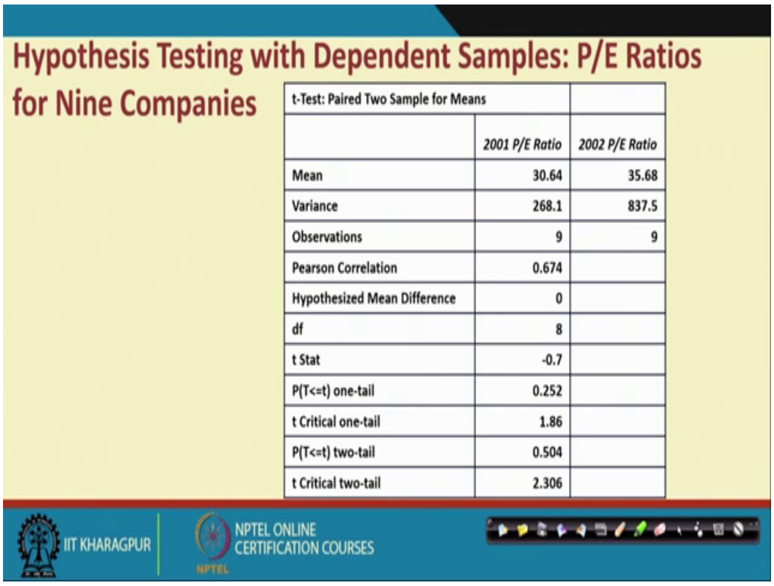
(Refer Slide Time: 26:48)



So, now this is the d is nothing, but actually called as you know sample mean and that is derived from this particular you know series and this is the sample you know variance that is standard deviations and then you have to calculate the t statistic which is nothing, but d bar minus capital d which is assumed to be 0 here that is the population mean and then you calculate actually minus 0.70 which is actually t calculated.

But our critical t is showing actually minus 3.355 in the left hand side and as a result since minus 0.70 is not actually overtaking this value. So, we are not in a position to reject the true null hypothesis; that means, the difference between the 2 years you know price earning ratio are not actually is most significant. So, accordingly we can actually highlight the particular case and then address the problem as per the particular you know requirement.

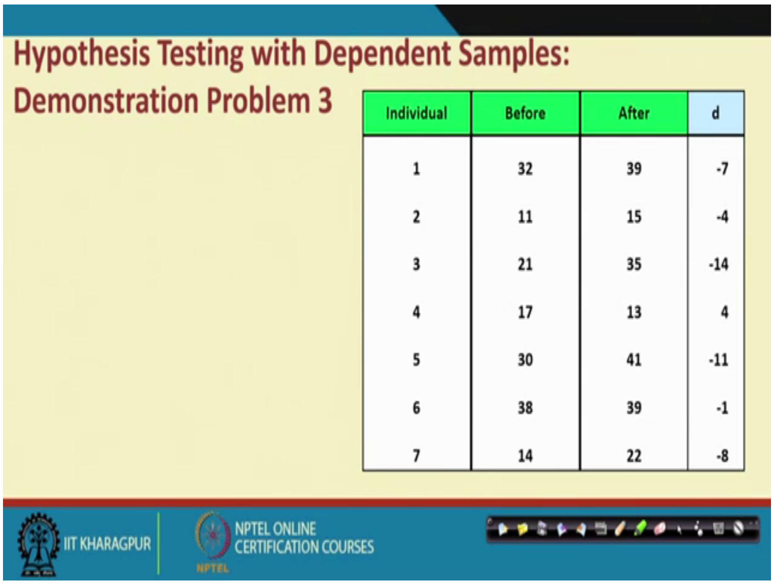
(Refer Slide Time: 27:48)



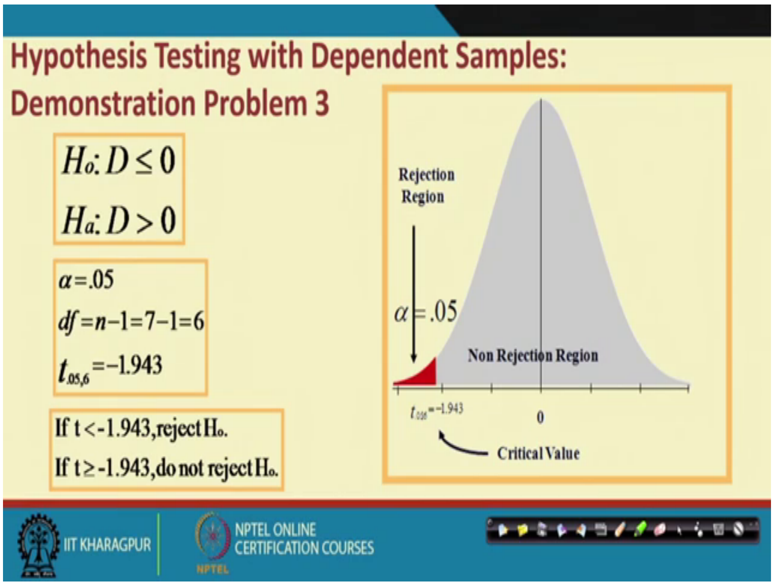
t-Test: Paired Two Sample for Means		
	2001 P/E Ratio	2002 P/E Ratio
Mean	30.64	35.68
Variance	268.1	837.5
Observations	9	9
Pearson Correlation	0.674	
Hypothesized Mean Difference	0	
df	8	
t Stat	-0.7	
P(T<=t) one-tail	0.252	
t Critical one-tail	1.86	
P(T<=t) two-tail	0.504	
t Critical two-tail	2.306	

And this is what actually the summary of this particular you know you go to the data analysis package again and then for you know the 2 series can convert into one series and against you find out the mean of the series and then standard deviation of the series and applying the t statistic you define t value is coming this much and then with the come you know if you once you compare with you know t critical and we find that you know this is not actually coming statically significant.

(Refer Slide Time: 28:21)

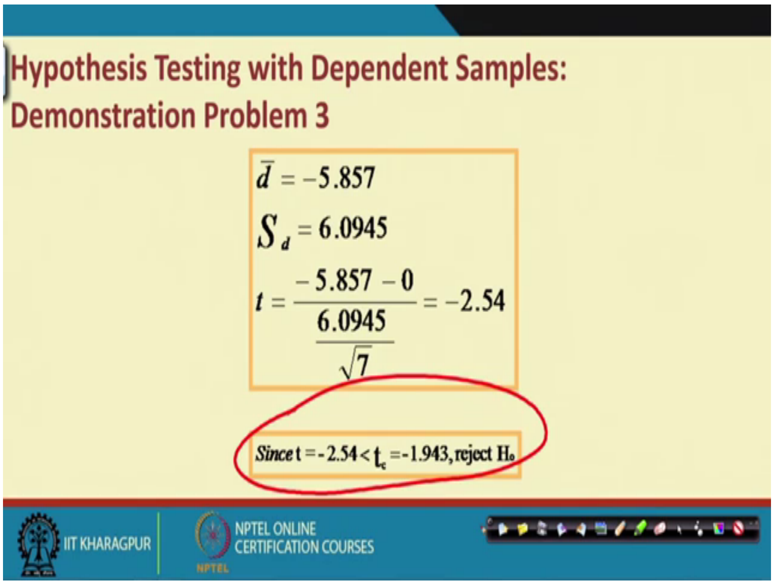


(Refer Slide Time: 28:22)



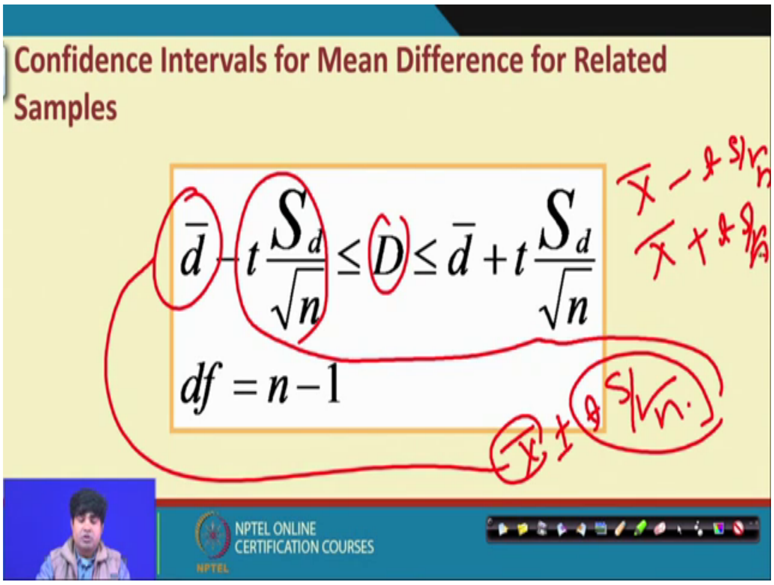
So, this is another case which you have already highlighted and same thing actually it is a fixing actually corresponding to this problem and you know d value and then here actually degree of freedom is 6 and a fixing alpha at 5 percent then critical value is coming minus you know 1.943 and we are testing this one for you know one tail test that is why D greater than 0 so; that means, this side appearing. So, as a result that means, since it is coming actually minus 1.943, we are targeting these ones. Then you know if t critical is coming lesser to this particular zone; that means, t calculate is coming lesser to this particular zone then; obviously, we are in a position to reject the true null hypothesis.

(Refer Slide Time: 29:11)



So, accordingly, let us see what is the you know result in this particular case and this is how it is coming actually you know after doing the analysis it is coming t equal to minus 2.54 and which is actually lesser to the kind of you know critical value and as a result we are in a position to reject the true null hypothesis and accordingly conclude that you know there is you know there is actually difference between the you know before and after situations.

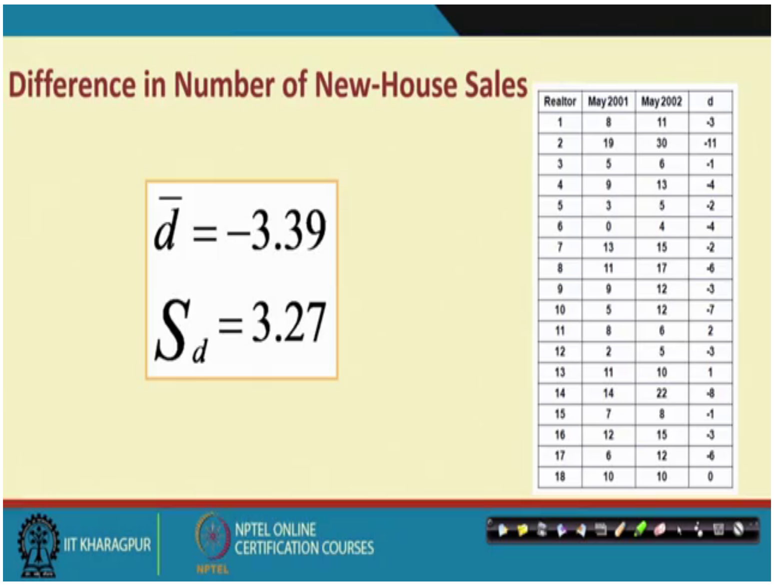
(Refer Slide Time: 29:43)



So, like the previous case we can also create here a confidence interval and the same structures since actually after doing the kind of you know conversions means to 2 sample to 1 sample case then you know standard rule is actually like you know earlier case we are putting X bar plus minus t standard deviations by root n. So, now, in place of you know X bar is just to be (Refer Time:30.08) actually at d bar and then in place of t S d root n and then you have to put this particular items right. So, as a result you are in a position to know the lower limit and the upper limit that is what the confidence interval for this population parameters and that to capital D right.

So, the lower limit will be X bar minus t of 1 S by root n and the upper value will be X plus t S by root n. So, like this we can actually calculate the population parameters right.

(Refer Slide Time: 30:47)



Realtor	May 2001	May 2002	d
1	8	11	-3
2	19	30	-11
3	5	6	-1
4	9	13	-4
5	3	5	-2
6	0	4	-4
7	13	15	-2
8	11	17	-6
9	9	12	-3
10	5	12	-7
11	8	6	2
12	2	5	-3
13	11	10	1
14	14	22	-8
15	7	8	-1
16	12	15	-3
17	6	12	-6
18	10	10	0

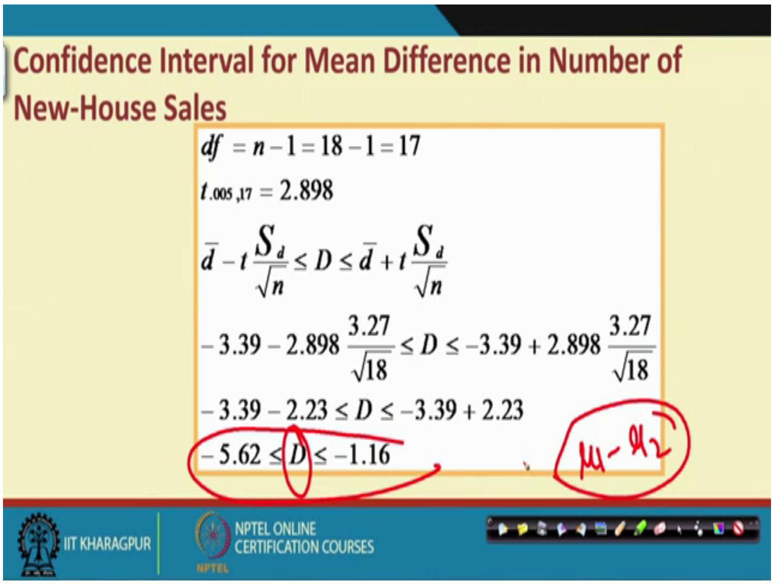
This is another kind of you know a problem same problems with you know different kind of you know examples it is a sample size is 8 and by default we can apply the distributions, but it is a 2 sample case either you will go by the using 2 different samples or transport the 2 sample to 1 sample and then you solve through one samples you know structures and that is the actually the beauty of this particular you know t distributions and what is happening here some of the problem the difference is a difference can give some kind of you know meaningful representation for instance you know before and after situation reinforced crisis before crisis after, then a year after year.

So, these are the kind of you know link or logically you should confident convinced that you know the difference can bring some kind of you know interpretations or meaningful kind of inference. So, if the difference will give some kind of in a meaningful interference it means inference and then accordingly you can actually a transfer the 2 sample case to one sample case and then you apply the t statistic for one sample case otherwise. So, the same problem can be highlighted in 2 sample case by taking the mean difference of you know 2 sample and then test with you know population difference.

But if not then you know you have to transfer the 2 sample case to 1 sample case and then you trace the 1 sample case actually a population you know mean difference and population mean for single sample is not is not an issue because in the case of 2 sample case you know we are assuming that you know mu 1 equal to mu 2 by default the difference equal to 0 and similarly in the 1 sample case we are putting mu equal to 0. So, ultimately it is the mean need to be tested or mean difference need to be tested only thing the adjustment is with respect to standard error on day.

If there is feasibility then you can do that if not then you can actually analyze the same problems which actually in with you know 2 different samples and accordingly you can structure the process and get the kind of inference. So, ultimately whatever you know whether you solve the problem with it 2 sample case or you solve the problem with 1 sample case if there is a difference and the difference is statistically same then in both the cases the results will be more or less same it will not you know having big difference it is only transferring the multivariate case to univariate case or you know or something like that only to solve the particular problem and then it gets some kind of inference as per the management requirement.

(Refer Slide Time: 33:34)



Again similar kind of you know confidence interval you can create and depending upon the mean value of the difference and the kind of you know alpha value and the kind of the degree of freedom so.

So, this is what actually the kind of you know the confidence interval for the population parameter dinner D and if it is 2 sample case then it is the we need to defined out the confidence interval for mu 1 minus mu 2. So, mu 1 minus mu 2 is deep kind of in a proxy here for actually the d. So, likewise you can actually solve the problem through 2 sample case and then you can solve this problem through 1 sample case after doing the kind of you know transformation right. So, with this actually will resolve you know we have already actually highlighted the particular case and to discuss the issue about the 1 sample case and 2 simple case and then through Z statistic and through t statistics.

So; that means, typically in inferential analytics 2 we are addressing some of the business problems connecting to 2 sample case or you know multiple sample cases; that means, typically when the problem is actually in a kind of in a multivariate framework then what are the ways you have to test the particular structure and get the kind of inference and we have already discussed you know the kind of you know structure are in the kind of unit testing and to Z statistic and through t statistics right. So, ultimately in any case whether it is you 1 sample case or 2 sample case whether through Z statistic or t statistic ultimately our basic objective is you know to check actually whether there is a kind of you know difference between sample mean and population mean.

And if whether the difference is statistically significant or not because we need actually sample means should be conversed to population mean or mean of the mean difference of the 2 sample means should converse to population mean difference and that is one of the strongest objective behind this inferential analytics and besides we can also in a position to create a confidence interval for a population parameter or difference of you know 2 population parameters and then we can also in a position to find out the size of the optimum size of the samples corresponding to a particular you know distributions.

So, these are the items which you can actually easily address in the case of inferential analytics and with this you know really we will stop here and in the next lectures we will discuss in details about the multiple case for chi square test statistics and f statistics that will be more interesting and let us see how is the kind of in a dispersal.

Thank you very much, Have a nice time.