INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NPTEL

NPTEL ONLINE CERTIFICATION COURSE

Marketing Research

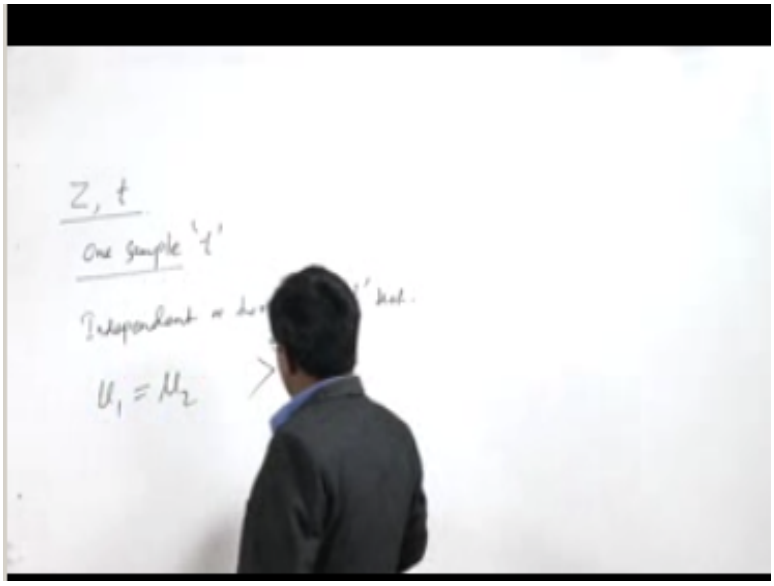
Lec-20
T, Z &F Test

Dr. Jogendra Kumar Nayak

Department Of Management Studies
Indian Institute of Technology Roorkee

Welcome friends in the session marketing research and analysis in the currently in the last session we were continuing with hypnotizes testing, previous to that we at the discussed about hypnotizes development and then we went into hypnotizes testing in which we spoke about basically the hypnotizes testing to this.

(Refer Slide Time: 00:45)



Z,t so we said they are morels similar in nature right, the only thing being that ,we use when the sumps size is very less and the population and sample and the sander relationship is unknown the population sander vision unknown, right but the basic interpretation this is nothing but is Z,

better smaller version you can understand or Z is a larger t version right that is with the way you can understand for simple terms.

So and now we explain that one sample mean, so basically what is happening now you are trying to test the hypnotizes basis on two factors either the means or the proportions right , so we discussed about the problem about the score of students GPA of education research and the normal students in which to be used the hypnotizes testing ,so one sample is one sample ‘t’test is something where you have only one sample and you were comparing the one sample these the population mean right .

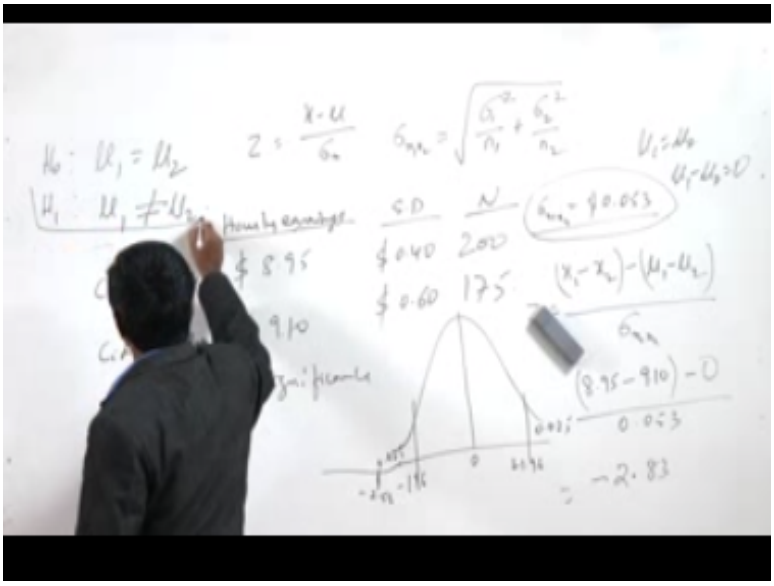
So on the other hand there are other two other things like the independent or two sample in the independent or they said two sample test, two sample do not get confused if you see these words ever ,two sample ‘t’ test right so independent sample ‘t’ test or two sample test throughout we are happening is we are trying there are two groups of samples we have pulled in two sample groups.

And now we are trying to say we whether these two samples are this coming from the same population or are they different from each other, right so assumption if you have understood by now the nil hypnotizes that would make in such sequences that the population groups from this to come are actually same right.

So U1 = U2 right is a clear suffer nil hypnotizes in such equations remember while doing a hypnotizes that increase in statics we have it going to something like there can be case of >, there can be a case of = 2 there can be case of < right the = to case is that case you can remember it forever that = to the case is the case which is used for the nil hypnotizes right then comes sometimes it could be => or = < so it could be anything right.

So but = is shown in the nil hypnosis it is okay, never get into that confusion so now the question is new one is = to new two suppose let us take a very brief example so let u take now what is happening we will compare. Now let us see a problem I have brought a problem which I have explain to this the data is given to you something like this it says a manpower development company is determining is trying to see whether the hourly wage of workers in two different cities in two different cities are the same or different what is saying there two cities basically.

(Refer Slide Time: 04:10)



City 1, city 2 so there are workers and the hourly wage is given basically dollars this is corrosion you can also find from the booker so from, the level and down book have taken. The hourly earnings given to you right pesetas this in dollars 8.95 okay, and city 2 let us city 2 gave it a name for example let us say this which is given here fx this is Eden and this is dollar 9.1 okay, what other data do you require this is a mean as given to you know what other data if you think.

So you cannot taste until unless you know how much the values are diverting from they mean which is the standard diversion okay, now let us take the standard diversion is the first case point 4 and the second case it is point 6 okay. The N the sample size taken for both the cities was 200 and 175 different sample size you should not be confused or should not be worry because end of the day when you were talking about mean or averages the sample size if it is different also does not matter okay, so now the question is what is he asking the company wants to test at 5 percent level of significant right at 0.05 level of significant okay, whether the mean or means of the two cities are different or their non-different okay.

So such a condition what you will nil hypnosis first let us understand that so my nil hypnosis I am writing my nil hypnosis here so my nil hypnosis that there is no difference that means hourly wages of both the cities are actually same and whatever differences come that is only by chance it is not a statically differences right, so significant difference. So new one is = u2 that is what I am saying so there is no difference now what is my alternate hypothesis there is a difference right very simple.

So there is difference so that means what ,if there is a difference can I know there is a difference who is bigger elder is bigger whether this city as a larger one or this city is a larger one or wise I cannot say so in such situation you will say U1 is not = U2 so what you were understand from here that means it is one directional or two directional test obviously it is two directional test because it could be lesser one could be lesser and other and we do know which one okay.

Now in such conditions how does it, how do you go about it let us say the other values also given to you so what you will do since we are saying at point 05 level of significant if I draw it in a graph or sorry let me draw it here. So how does it look it is saying this is my lesser 0 this is my Z at 95 percent level so – 1.96 this Z here +1.96.

Now my question is the Z values which I am going to calculate first of all you have understand is it a Z or this time Q I will use this is Z or Zee whatever why because the sample sizes much bigger than 30 okay, so it is more than even 100 so in that case we are using a two tale so obviously the. 025 is remaining here and .025 or the 5 percent is remaining here now I will calculate hour Z and where hour Z falls.

Okay now to do that in the arrear case you were using the formula Z = x-u Σ x okay, and here that when we were merging the Σ we were taking the sample standard variation, right we were using nut here how will calculate this Σ x how do you calculate in this case Σ x because there are not only one x here the two x right so Σ x 1x 2 okay, Σ x 1 and x2 how it will look in this case if you go by this it looks something like this √ Σ 12 / n1 right + Σ 22 / n2 now you can understand why I am doing it the arrear formula was Σ x= Σ √ n right in order to remove the √ over what I did was denominator I multiply square so it depends Σ 1 2 by n+ Σ 22 / n 2 right.

So √ now what are over to take so let us take this, calculate this and how many it is coming 9 this is coming to the Σ x 1 x 2 =0.053 dollars okay, so this estimated standard either is = 0.053 dollars okay. Now we will use this one into the formula now what is the formula for Z =early it was only x but now we have 2x so x1 –x2 –u1 –u2 divided by Σ x I hope it is clear by now so this Σ already Σ x1 and x2 so with this been calculated already from here.

Now if I take x1-x2 what is mine x1-x2, now x1 is 8.95-x2 9.1 right –u1-u2 now what is this u1-u2 we do know do we know or not is a question but if u go back to this situation where you seen

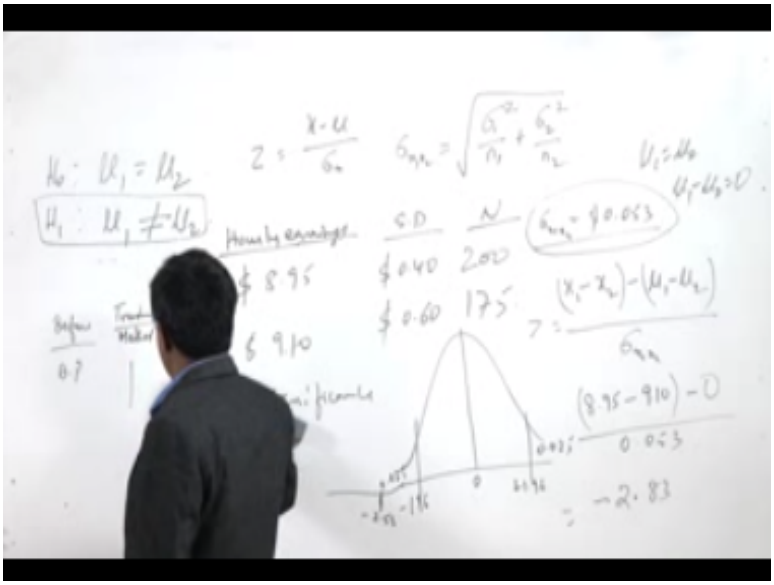
they are saying ,they always testing the position in the last session also I have said we also test that nil hypnosis okay.

So here u1 =u2 our nil hypnosis that means U1-U2 is equal to how much u1-u2 ,u1 =u2 so u1-u2 =0 right, so 0 divided by Σ x1 x2 0.053 now calculating this the value is coming to the -2.83 right, so where does -2.83 lie where did it fall -2.83 will fall obviously not to the right. So we will fall to the left and that 2 after the critical zone, the critical region. So somewhere let us say here okay, from here what is your analysis what would interpret nil obviously since it is much way from the critical zone reason so nil hypnosis to be rejected.

So what is a nil hypnosis not there is no difference between the two cities the earning or the two cities and whatever the differences it was only due to chance but that is completely wrong so that means now we will expect the alternate hypnosis it says that u1 is not =u2 and there is a significant differences between the earnings of the two cities right so this is basically what we do independent sample ‘t’ test similarly you have dependent sample ‘t’ test the difference only being that is only one sample group.

There is only one sample group but they had been taken twice now what it mean I will just explain the meaning other things remaining thus same. So in depending sample ‘t’ test what we will do basically before after case.

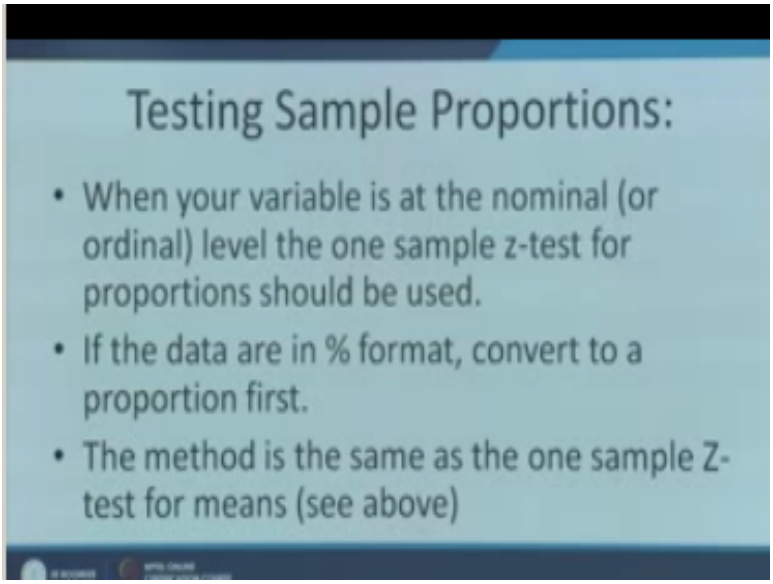
(Refer Slide Time: 12:51)



Now let us say many times the sample group is 1 many a times you want to see the effort of a medicine or training program something so we take the peoples blood pressure right before and then the medicine treatment you say the treatment is given if you remember experimental design discussed about it treatment given so what is the effort now right so when you do this right method a treatment whatever it is and then after that again you calculate BP right.

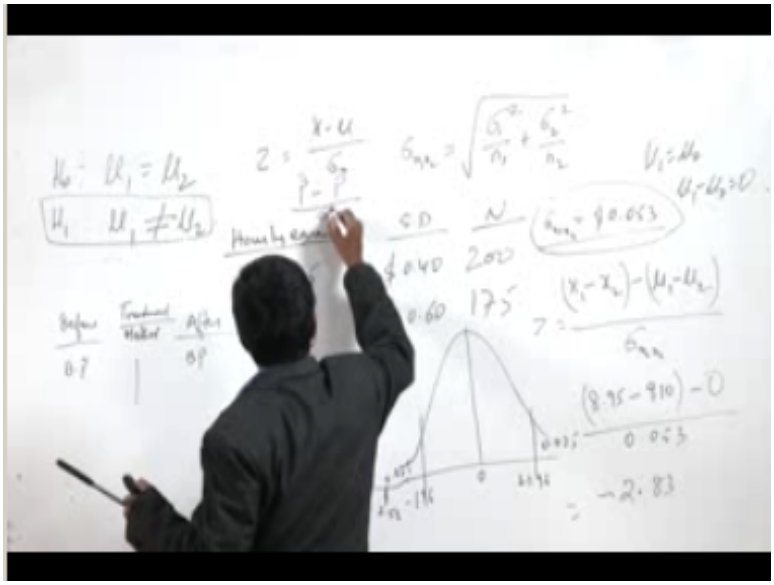
So this difference this only one sample but there are two values for us right for the same sample this is called a two sample‘t’ test right or dependent sample or paid sample’t’ test okay, but now what is I am not even started my slides. So let me go to that how do in the case of sample proportions what will you do suppose you have sample proportions how will you analysis you sample proportions when your valuable.

(Refer Slide Time: 13:54)



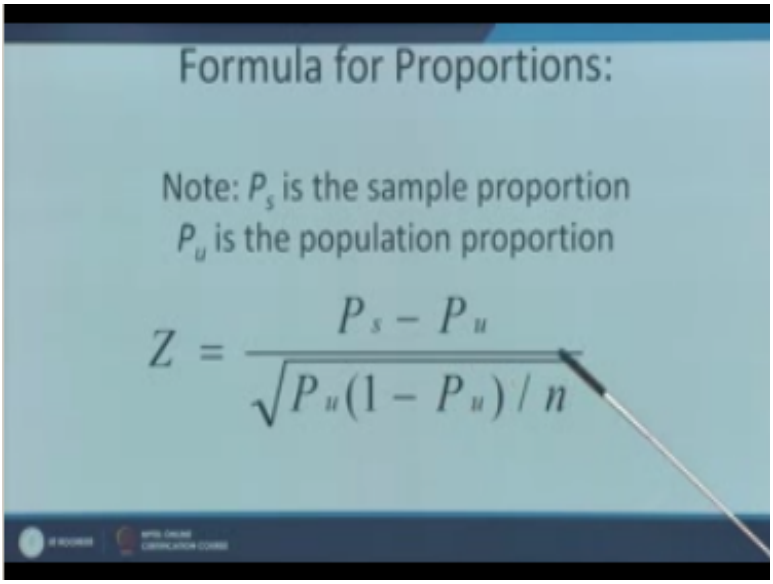
It says nominal or ordinal level the one sample Z-test for proportion should be used okay, remember proportion is nothing but beside is u form right, if the data are in % format ,convert to a proportion format let me keep it here okay, the proportion format now did I saying this method is the same is nothing absolutely no difference if you do right only thing is that.

(Refer Slide Time: 14:25)



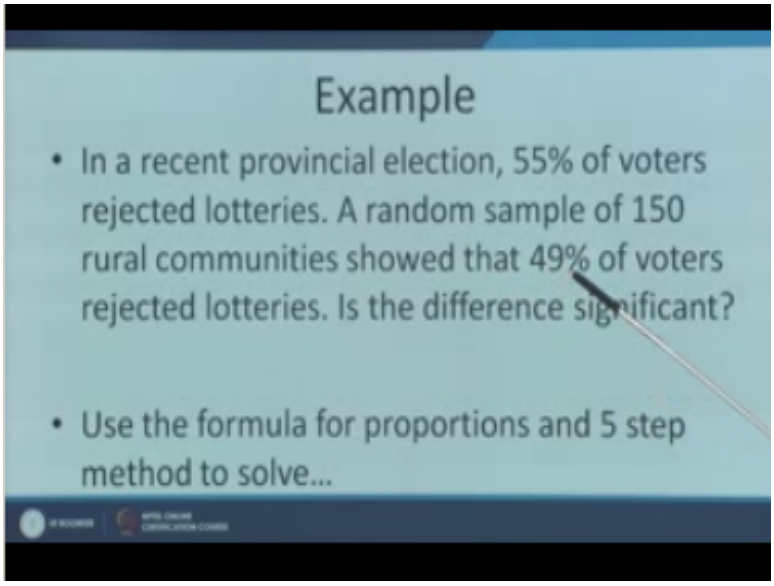
Instead of the means you will have the proportions so p- let us say you can say for example what is the population right, p group – p population so p sample –p population that is all we will require right, so other things remain the standard change to Σ x will change to Σ p so standard error of the proportions okay, so let us see this.

(Refer Slide Time: 14:54)



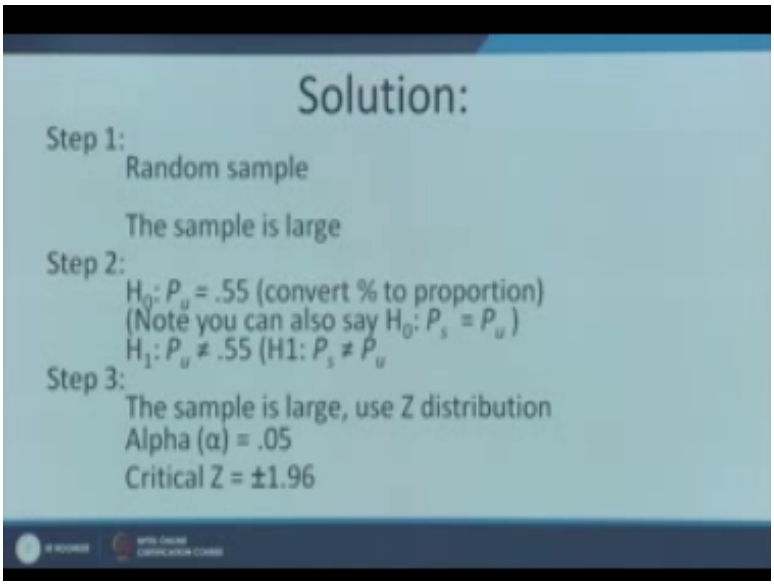
What is this saying formula for proportions Ps is the sample proportion right, P u is the population proportion which was like the arrear we were thing x is the sample mean and U was the population mean here we are saying this other things is same Ps-P u this is what I did right divided by √ of what you were doing ,you were doing Σ x= Σ √ of n right so here we were saying P u into here only thing this is changes this will become P into q that means colorant of success multiply by probity of failure divided by n okay take a problem.

(Refer Slide Time: 15:44)



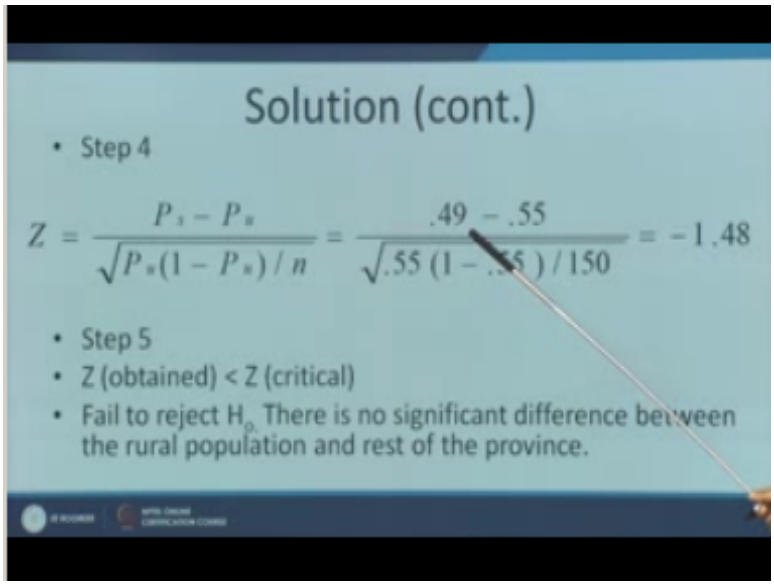
In a recent provisional election 55% of voters rejected lotteries. A random sample of 150 people showed that 49% of voters rejected lotteries are the difference significant? Now what will you do use the formula for proportions and 5 step saying steps right that means what you have write a line hypnosis this is decide the significant level decide the tale of the test all these remain the same okay.

(Refer Slide Time: 16:13)



So it is the random sample, the sample is large right, is this large yes 150 okay so what do you saying p u = .55 right that means the if you go back what you seeing 55% of voters rejected lotteries show the proportion of the voters you rejected the lotteries .55 okay H 1 says note P u is not = .55 because if you see only 49% voters are rejected so we will say we know it is not 55 % but it is less than that 49% okay. The sample is large and we use this okay.

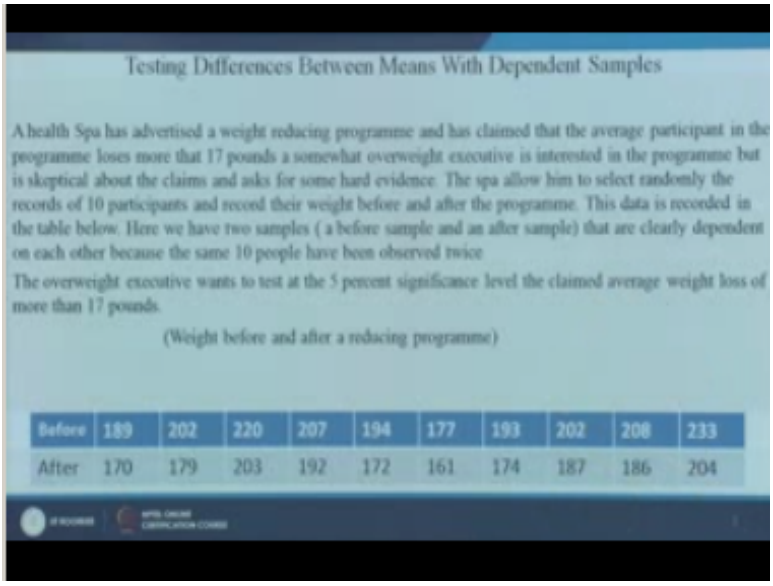
(Refer Slide Time: 16:55)



Now let us do this so .49 what are sample from the proportion and this is what the population right, and when we did it right, so we got the value of -1.48 right -1.48 now you have to check compare against the 95% confidence level is 1.96 which analysis so through value can you now hypnosis and no you cannot reject the hypnosis so you fail reject H0 there is no significant difference between that mean what there is no significant difference between the rural population and rest of the province.

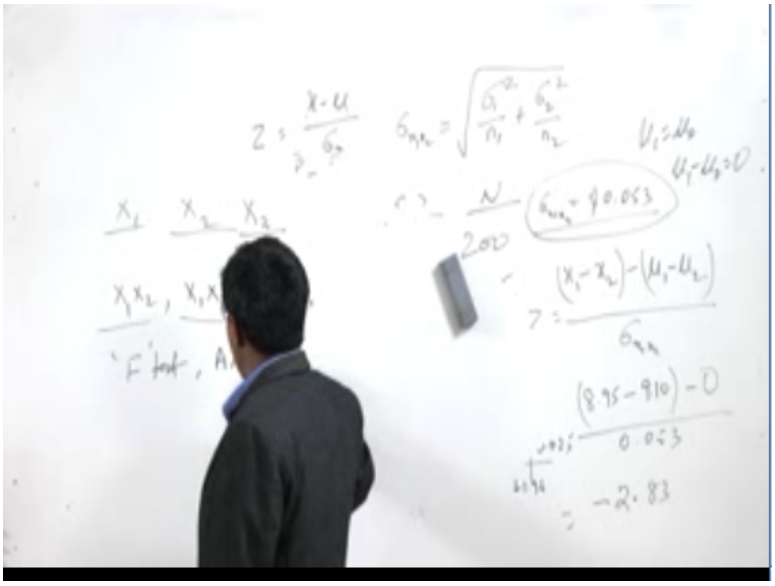
You cannot claim that there is a difference between the two population parameter two populations the rural population and the others in the province right, so this is the case of A this is the case that proportion is to be used the same methodology you can add out for two samples and dependent samples okay, you were two may be your two go through some other examples okay has is the example.

(Refer Slide Time: 18:01)



You can later on do it, this is for a weight deduction is a dependent samples okay, this is I come to know what happens I have now arrear I used to have maximum one sample.

(Refer Slide Time: 18:25)



Let us say x1 or I had maximum two samples okay, sample groups but the question is suppose let us say a person wants to test something in which mirror more than two levels or more than two groups right so in that case what I will do suppose I have x1 x2 I was doing a independent sample ‘t’ test if suppose there is an x3 coming in right if the x3 comes into this how will I do the analysis right what will I do the analysis to do this there are several possibilities I can may be do number of independent sample ‘t’ test like x1 x2 right or x1 x3 then x2 x3.

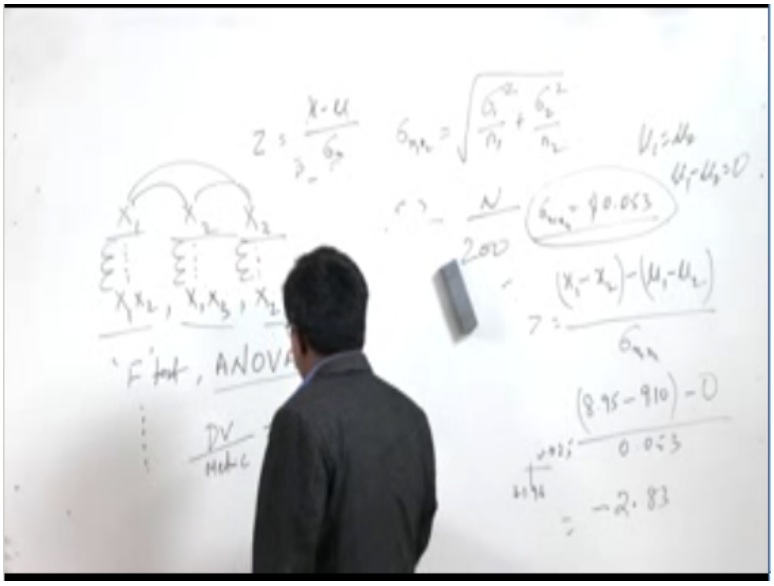
I can do this right but the question is this is not justified it should not be doing it why you should not doing it if you do this than every time you are taking an alpha of .05 each time so the error will get inflated three times if you do it is going to be inflated and it is very comparison it is confusing if you have three but suppose you would have eight groups so how many would you make combinations .So you have to evaluate the situation right you may have to think of some other way of doing it. So what is the other way the other way is we adopt it as which is called the f test f test or we also called is it as.

It is very popularly known as variants the analysis f variants so why it is called analysis of variants it is measuring the variants of the two things in account that is the groups all the group within the groups that means if there are several groups you say there are let’s say X1 X2 X3 right and each had got several samples let us say right so what is the difference between the groups that is one which is you say between

(Refer Slide Time: 20:44)

The groups and what is the kind of the variants within the group so analysis of variants helps you to basically measure the test that is hypothesis by taking it into account two ways basically variations or variants one is between the groups and the other is the within the groups as I said.
So this is a few simply if you understand on what to understand on other terms also so what happens is if you want to go through a dependent or independent variable the a nova is one where the dependent variable is basically in a continuous or metric scale.

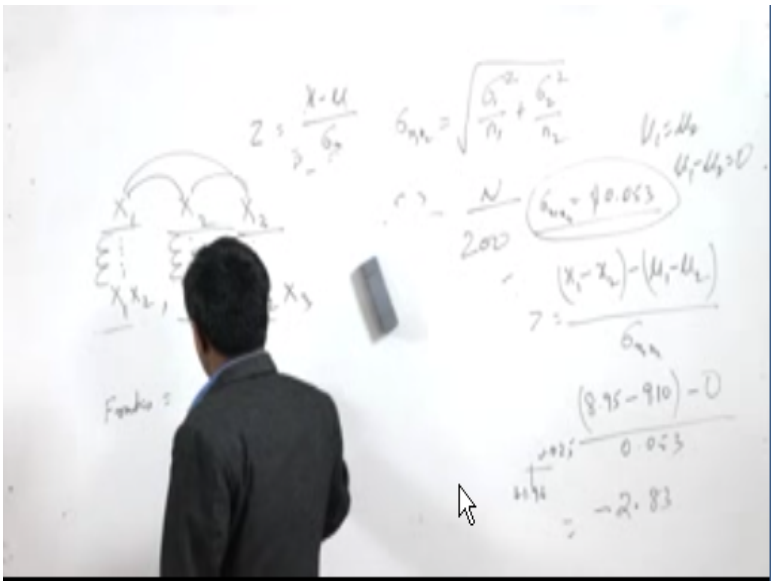
(Refer Slide Time: 21:34)



The independent variables are no metric or categorical in nature okay this is what you have to remember right if this is one dependent variable then it is a case where we are saying it is an a nova right if you get into multiple when we will have a new technique which is called me nova multiple analysis of variants okay. So now let us get into first the anova we will think we talk about it later things but maybe we will see okay.

So now what is happening in the case of analysis so variants as I said so of ratio it is calculated right now this f ratio is nothing but it says

(Refer Slide Time: 22:25)



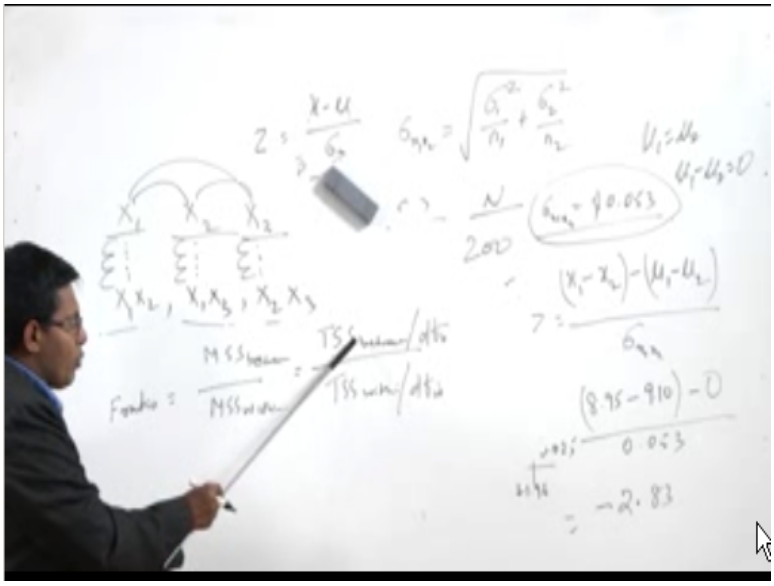
Between divided by means sum of square within no what do I mean by this between the groups and this is within the group that means there is such a test and now looking at this is basically like a you know getting slowly into a purely an experimental study a nova is used in mostly all kinds of studies in the experimental research basically where people are studying on in a chemical lab or on allergy or anything for that right.

So basically you are using this technique it is the most one of the most rudest and Avery very powerful technique where the assumptions of the normal distribution are very important for anova that is to be checked so what is the means of sum of square between sum of square within it is something which is again

(Refer Slide Time: 23:24)

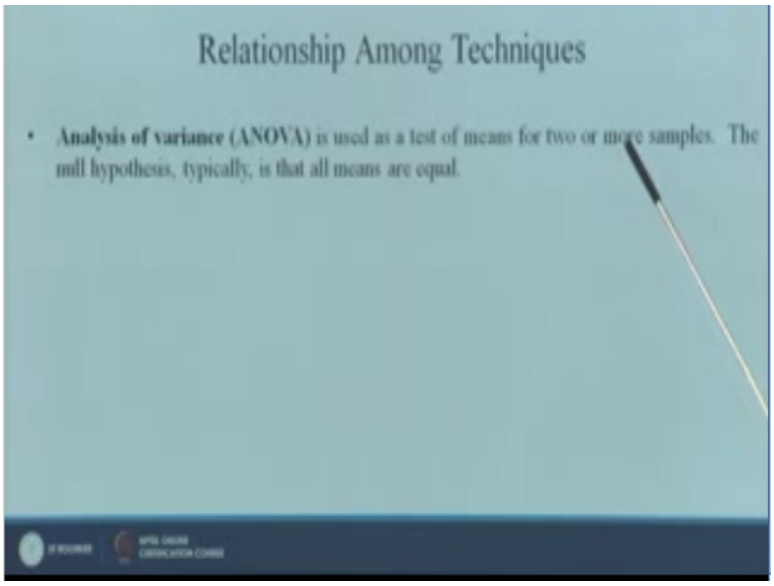
Write it like this it looks like totally sum of square between right divided by the total sum of square within right degree of freedom this is by the degree of freedom now degree of freedom this is between degree of freedom within now what is the mean sum of square that means go by this that means sum of square of which is the finally need for to calculate the f ratio is either you start from this that is

(Refer Slide Time: 23:55)



Divided by overall within the groups by the degree of freedom within the groups which is nothing but is equal to mean sum of square between the groups divided by means of sum of square it is the same thing okay this is the same thing right. So this ratio basically tells us keep so much what is the f ratio now once you have the f ratio value right other things are the same other things remain the same okay let us just have a look at it is the test.

(Refer Slide Time: 24:38)

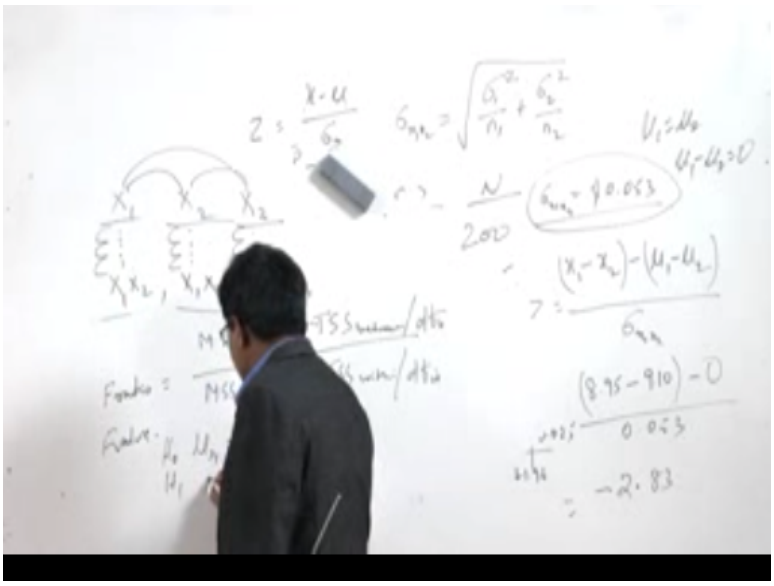


Means two or more samples hypothesis is typically that means unequal define we have three in a group let us say so what will I say new X1 is equal to X2 right is equal to μ is X3 right so when I am having three groups I am saying the means of all this three groups in my null is same there is no difference right. Means of all is three groups in my null is same there is no difference right but other is such as do I want that they should be same no obviously not so what will I what do I this is

(Refer Slide Time: 25:15)

This is my null hypothesis what my other altered hypothesis is now at least there is one difference between the groups is.

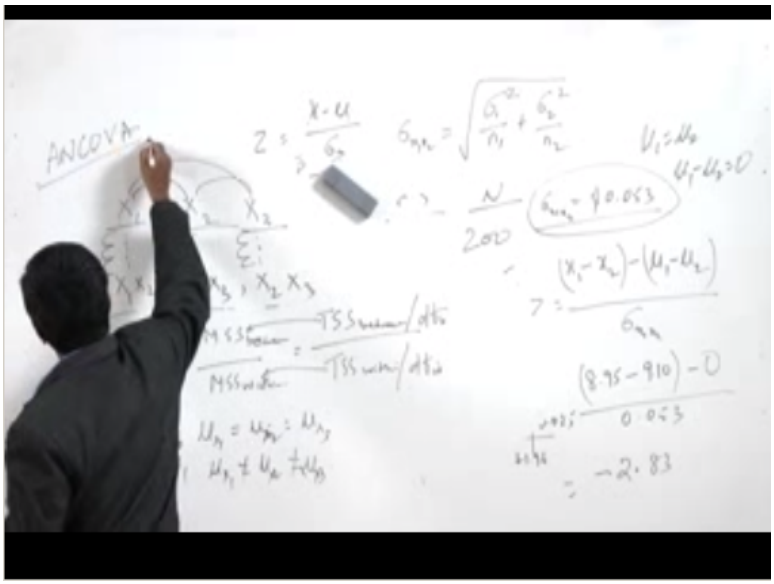
(Refer Slide Time: 25:24)



Either μx1 is not equal to μ2 or there are not equal to μ3 or μ2 is not equal to μ3 so whatever it is right so there is at least one difference between that one thing that is not similar is not equal to is a case you would say that the null hypothesis has been rejected and the alternative hypothesis is to be now excepted okay.

But let us see what how would you go ahead okay let us see some theory behind it the analysis variable is must have a dependent variable that is specific as I said and the independent variable must be non metric suppose the independent variable is let's say having also some metric let us say which we sometimes call co variants then it is new technique called anova.

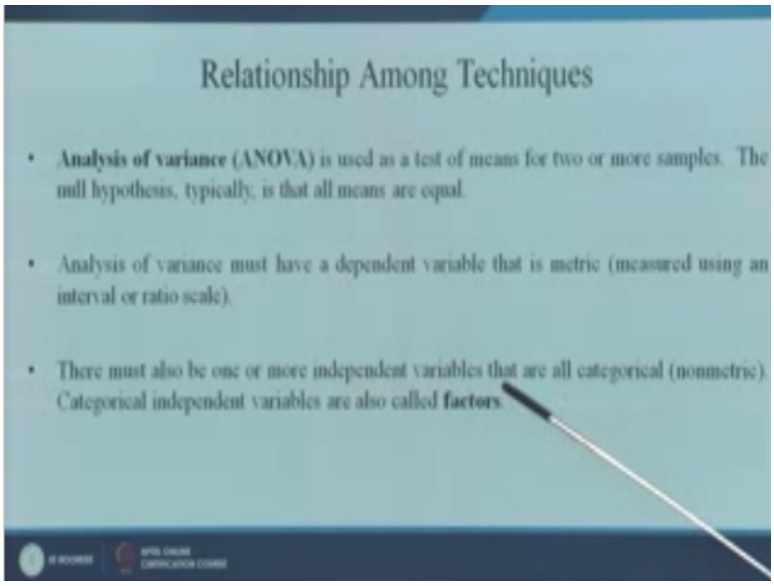
(Refer Slide Time: 26:19)



Anovas is a technique where we are saying is called analysis of co variants where the independent variables are metric as well as non metric right so it is one new method in which gives you a which makes you know provides you an opportunity to take the co variants which are basically metric in scale right so this is basically to you know to address the blocking effect basically that you do in the next segmental group.

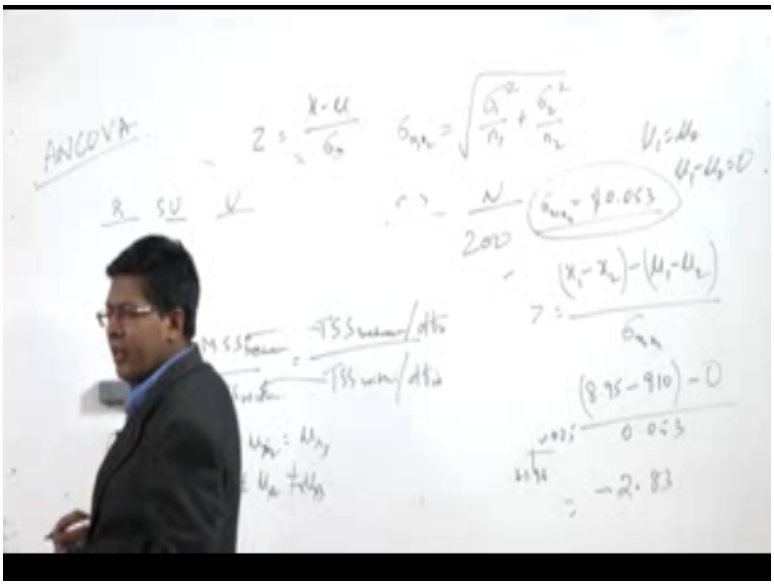
When you block a group and check and treat a group and there see the difference that is where is basically mostly used.

(Refer Slide Time: 27:00)



There must also be one or more independent variables that are called categorical and these are called the factors so let me tell you if you go through any book or something you see one factor or one way ANOVAs one way is nothing but one factor that means we have only one factor one let us say one factor that is let's say I want to see let me rub of this one.

(Refer Slide Time: 27:26)

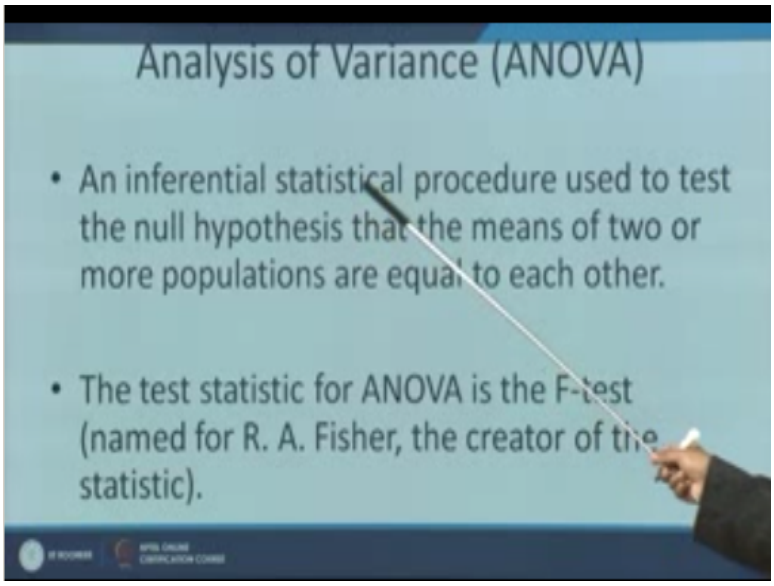


I just want to say that I have I want to seek you whether the states of no goods in a retail store depends on the size of the store or depends on the type of the city now what I have done now I have taken three types of cities well urban and let's say semi urban and urban okay three now the question is when I am measuring this three well semi urban and urban they are all basically on the basis only right so no places whatever . so I am actually measuring one factor is there which has got three levels right so in this case what did he say.

The categorical independent variables are also called factors so you know factorial design do you remember we have said if you go back to the last few sessions in a factorial session there might not be one factor there might be two factors there might be three factors they might be more than that so that what is it called now when you have two factors we say in two way ANOVAs when you have more than three factors then you say three way ANOVAs when you have five factors you say five ANOVAs.

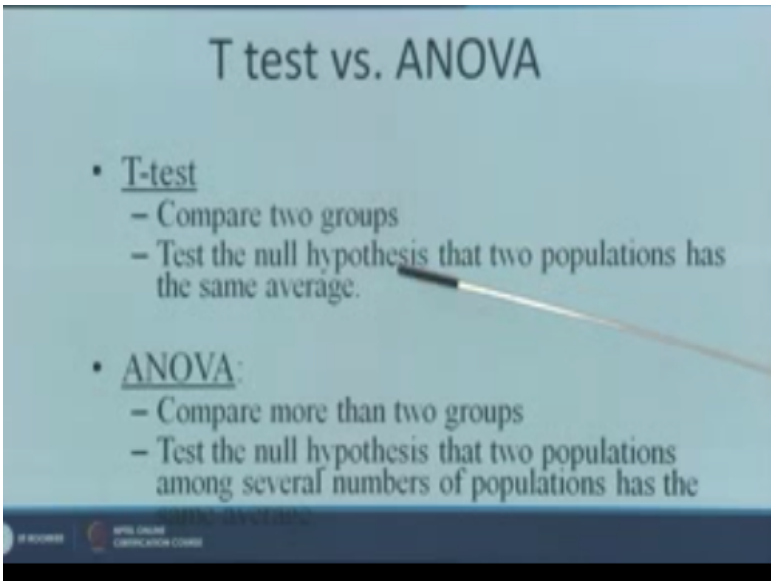
So n factors end the ANOVAs so the analysis of this variants basically it is nothing but a this is a simple factorial design that you are using in a experimental study okay so what is saying the

(Refer Slide Time: 28:58)



In the statically position you should test the null hypothesis of two or more populations or equal to each other they are equal so it is why it is named after the gentle man he was a scientist mostly done works on the agricultural fields R. A. Fisher who has developed his you know the concept of analysis of variants at their statistic test and you know all this credit goes to this gentleman okay

(Refer Slide Time: 29:32)



Just here I rewind my session here so it compares two groups test the null hypothesis that two populations has the same average or same mean but in ANOVAs compares more than two groups and tests the null hypothesis that two populations among several number of populations has the same average that means if I can re design you can say more than two groups or three groups my null hypothesis is still remaining in the same or all the means are at the same they are equal.

In that case we say it that there is no difference between the means of the groups or the samples and they are all same that is my null hypothesis in the case of ANOVAs. In the next session what I will do is will explain you how to calculate a ANOVAs right take a case of ANOVAs in there we will see key how to measure ANOVAs basically we will look at a problem where we will have a we will calculate the problem in the sum of square and then find out the f ratio and finally check a hypothesis right.

So today in the session let me recap we have explained about a independent sample it is what it talk about what should we do in case of proposition and then we next moved on to the next level in a group of you know in a case of multiple groups how would you what would you do. Which is an f test so in any marketing research I will tell a problem in the next class through you know in the marketing problem where marketers companies are using this for analysis understand one particular group let us say you know brand of cars are you say category of car is better than the ratio say Marathi has got several cars right.

Now it can say whether the states or popularity of the cars are same or it is there is the difference between the popularity of swift verses zen verses alto right it can do that so in such conditions marketers use it profusely or in very high level the other test hypothesis to prove that there is actually a different between the groups okay well this is all for today's session thank you so much.

For Further Details Contact

Coordinator. Educational Technology Cell
Indian Institute of Technology Roorkee
Roorkee 247 667
E-Mail Etcellitrke@gmail.com etcell@itr.ernet.in
Website: www.itr.ac.in/centers/ETC. www.nptel.ac.in

Production Team
Sarath Koovery
Mohan Raj. S
Jithin. K
Pankaj saini
Graphics
Binoy. V.P
Camera
Arun. S

Online Editing
Arun.S
Video Editing
Arun.S

NPTEL Coordinator
Prof B.K Gandhi

An Educational Technology Cell
IIT Roorkee Production