Advanced Engineering Mathematics

Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

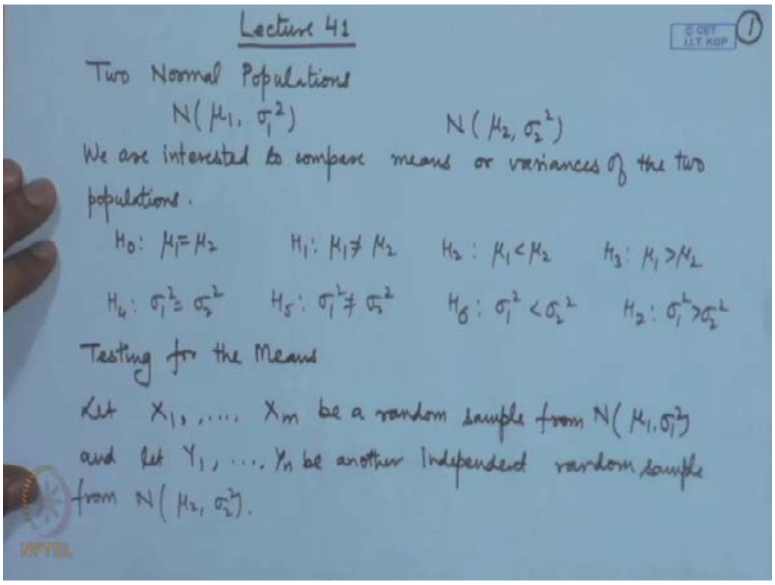
Lecture No. # 42

Tests for Normal Populations

In the last lecture, I have introduced the concepts of testing of hypothesis. What is a statistical hypothesis, and what is the method of deriving a test for a procedure to test whether a hypothesis should be accepted or should not be accepted. We have given the concept of the size of the test, and the power of the test, and based on that there is a fundamental theorem called Neumann Pearson Fundamental Lemme, which can be used to derive the test which have the maximum power for a given size or level of significance.

Now, later on I mention that this theorem has been extended to cover the cases, where we have compose it hypothesis testing problems. In particular, we consider testing for the mean, and the variance of one normal population; that is the parameters of the normal population.

(Refer Slide Time: 01:30)



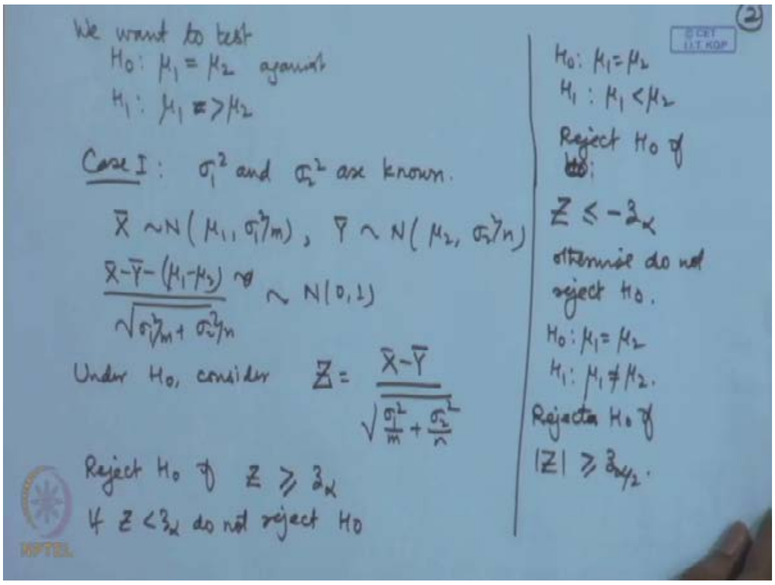
Today, I will introduce the tests for parameters of two normal populations, our situation could be like this. Let us consider two normal populations. So, we have two normal populations say normal mu 1, sigma 1 square, and another population is a normal mu 2, sigma 2 square. So, we are interested to compare we are interested to compare say means or variances of the two populations. For example, we may be dealing with the measurements which are related to say vars of implies of 2 organisation. And so they may be following normal distribution; if we are in two different groups, then one may be population normal mu 1 sigma 1 square, and another may be normal mu 2 sigma mu 2 square. We may like to check whether the average vars are the same, we may like to check whether the average variability in the vars are the same or not.

So, this leads to the problem of comparing means or variances of two normal populations. We make frame hypothesis like H naught, whether mu 1 is equal to mu 2 or mu 1 is not equal to mu 2 or say mu 1 is less than mu 2 or less than or equal to mu 2, mu 1 is greater than mu 2, etcetera. Similarly, we may have hypothesis like sigma 1 square is equal to sigma 2 square or say sigma 1 square not equal to sigma 2 square or we have sigma 1 square less than sigma 2 square or we may have say sigma 1 square greater than sigma 2 square.

So, these kind of hypothesis have to be tested. So, we need to derive the test for that. Now, as I mention that these are composite hypothesis testing problems, the method of Neumann Pearson have been extended to cover these cases; in certain cases we have in particular for one sided hypothesis testing problems, we have uniformly most powerful test. And in some two sided testing problems we do not have (( )), so we consider a class of restricted class of test; sometimes we have considering unbias test or sometimes we are considering similar test. And we find out the best there, they are called unp unbias test or unp similar test, we also have unp invariant test.

So, let me start with testing for the means. So, we consider the model say X 1, X 2, X n be a random sample, say from normal mu 1 sigma 1 square; and let say Y 1, Y 2, Y n be another independent random sample from normal mu 2 sigma 2 square.

(Refer Slide Time: 05:32)



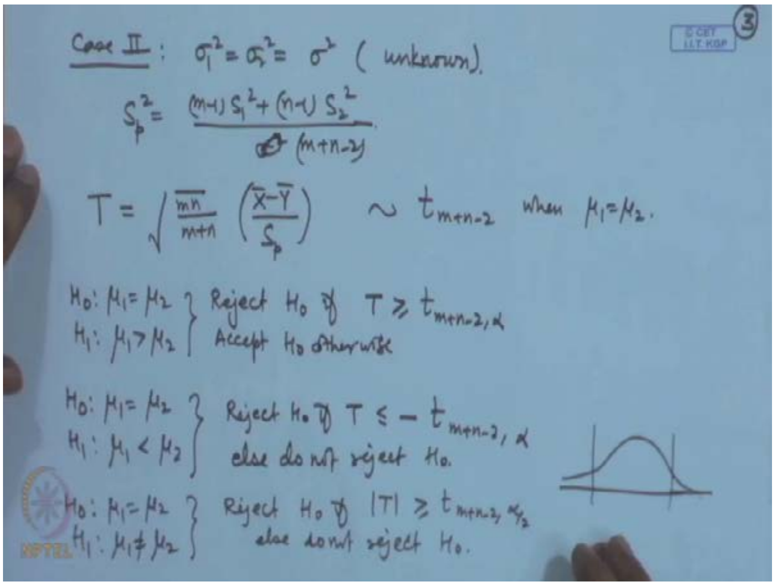
We are interested to test say mu 1 is equal to mu 2, against say H 1 say mu 1 is greater than mu 2. So, let me write various hypothesis. Firstly, let me consider this one. So, we consider the case sigma 1 square, and sigma 2 square are known, when sigma 1 square and sigma 2 square are known; let us work out the distribution theory - X bar follows normal mu 1 sigma 1 square by n, Y bar follows normal mu 2 sigma 2 square by n. Say if we consider X bar minus y bar minus mu 1 minus mu 2 divided by square root sigma 1 square by m plus square sigma 2 square by n, this follows normal 0, 1.

So, under H naught, consider let me call it Z; that is X bar minus Y bar divided by square root sigma 1 square by m plus sigma 2 square by n. So, if this value is near about 0, suddenly we will be tending to accept mu 1 is equal to mu 2, and if it is greater than certain pre specified value. Then mu 1 greater than mu 2 seem to be more possible. So, we one sided test will be a reject H naught, if Z is greater than or equal to say Z alpha, If Z is less than Z alpha, then do not reject H naught.

So, this is one sided test we also consider, here same H naught mu 1 is equal to mu 2, against say mu 1 less than mu 2. Now, in this case, if we are considering mu 1 less than mu 2, then for a smaller values of Z we will be tending to have favorable mu 2 H 1. So, in this case the test will be reject H naught, if z is less than or equal to minus z alpha, otherwise do not reject H naught. If we have two sided hypothesis say H naught mu 1 is equal to mu 2, against mu 1 not equal to mu 2. In this case the test will be two sided reject H naught, if modulus Z is greater than or equal to Z alpha by 2; and if it is less than we will be considering accepting H naught. This takes care of all the important type of hypothesis: one sided hypothesis, where one sided is on the right side, other one is the left hand rejection region, and this is the two sided rejection region.

(No audio from 09:26 to 10:02)

(Refer Slide Time: 10:10)

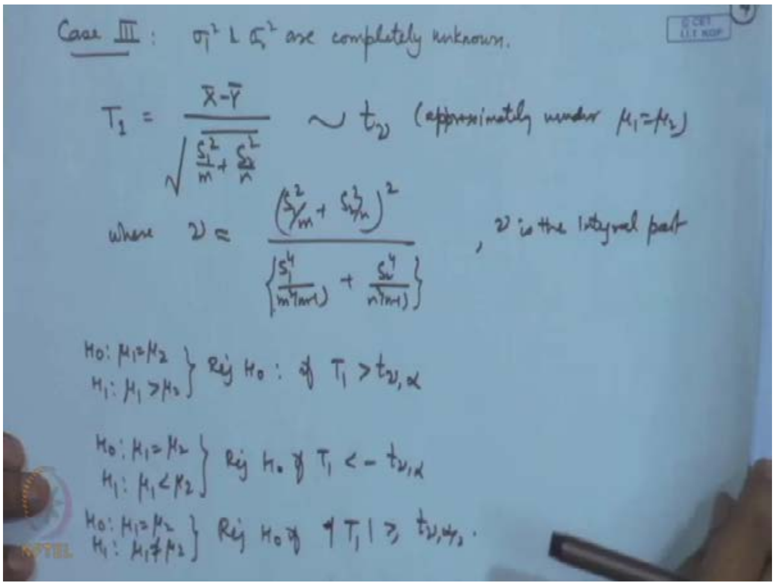


Let us consider one example here. Now, let me firstly take the case of second case: sigma 1 square is equal to sigma 2 square is equal to sigma square, but it is unknown. In this particular case, we will consider pooling; let us consider say S p square that is equal to m minus 1 S 1 square plus n minus 1 S 2 square by sigma square sorry divided my m plus n minus 2. And then, we formulate the testing statistic T, that is root m n by m plus n X bar minus Y bar divided by S p. So, this will follow T distribution on m plus n minus 2, when mu 1 is equal to mu 2.

So, once again when we are considering the hypothesis testing problem; one sided mu 1 greater than mu 2, then here we will we considering reject H naught, if T is greater than or equal to t m plus n minus 2 alpha accept H naught, otherwise… Similarly, if we consider say H naught mu 1 is equal to mu 2, against H 1 mu 1 less than mu 2, then the test will be reject H naught, if T is less than or equal to minus t m plus n minus 2 alpha else do not reject H naught. And similarly you will have two sided rejection region, when we have the two sided alternative hypothesis; in this case we will say reject H naught, if modulus T is greater than or equal to t m plus n minus 2 alpha by 2.

You can see an amazing, similarity with the procedures for finding out the confidence intervals - in the confidence intervals, we had considered the same test statistics here, and the reason is that the shortest length confidence interval for a fixed confidence co efficient are used the test statistic which is used there is also the one, which is used for deriving the best test for the corresponding testing procedure. So, there is a close association, and this was established by name in 1930s.

(Refer Slide Time: 13:11)

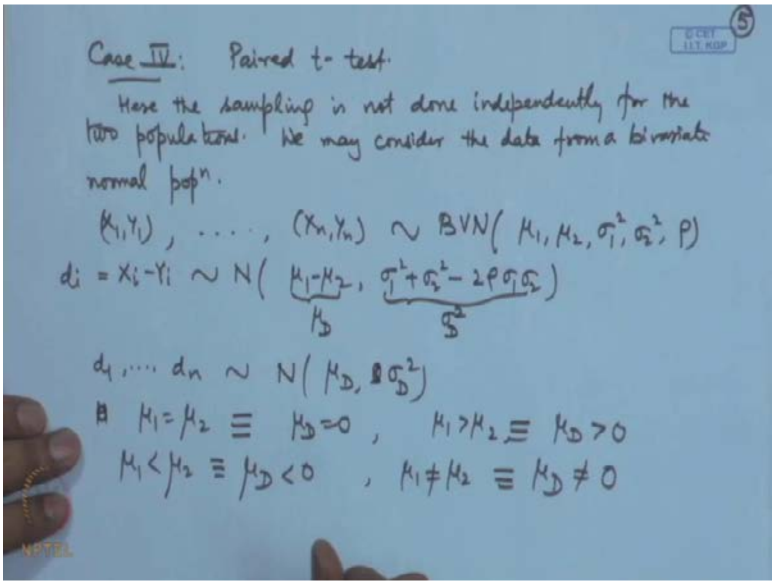


Let us consider the case, when sigma 1 square and sigma 2 square are completely unknown. In this particular case, we cannot make use of the pooling, so we consider let us say T 1 that is equal to X bar minus Y bar divided by square root S 1 square by m plus x 2 square by n. This follows T distribution on new degrees of freedom approximately under mu 1 is equal to mu 2; therefore, we can make use of this where mu is given by S 1 square by m plus S 2 square by n whole square divided by S 1 to the power 4 by n square into m minus 1 plus S 2 to the power 4 by n square into n minus 1.

And once again we take the integral part of this, so this is the actually there, the integral part of it; and we can device the T test based on this, that is for mu 1 is equal to mu 2 against mu 1 is greater than mu 2, the rejection region will be if T 1 is greater than t mu alpha. If alternative hypothesis left sided, then we have reject H naught if T 1 is less than minus t mu alpha, and for the two sided alternative hypothesis we have a two sided rejection region, if modulus of T 1 is greater than or equal t mu alpha by 2. There may be the case when the two samples are not independent, the situation may arise in the following following cases, see we may have for example we have to compare two things, but the sampling procedure may not be independent, suppose you are considering effect of certain medicine on patients.

Now firstly, a set of patients is chosen we give one medicine, and look at the effect. Then we (( )) other medicine, and on the same set of patients we give the medicine at another time, and then we observed the effect. Now, here the sampling scheme is dependent, because the same set of patients are there; this is then because it could happen that depending upon the different patients, the effect of the medicine could be different. Therefore, in order to neutralize the effect of or variability due to different patients we take the same set. Now, this is the problem of correlated data, and the previous procedure are not applicable here.

(Refer Slide Time: 16:56)

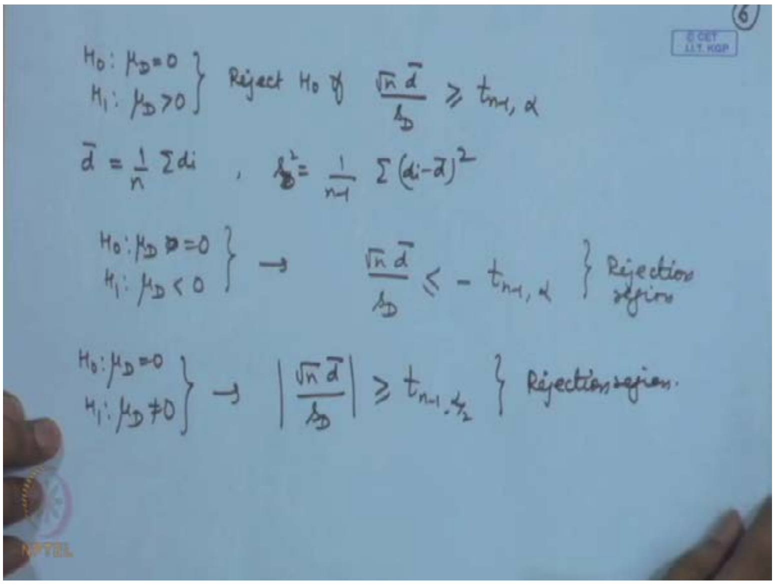


So, I will consider here paired t test, here the sampling is not done independently for the two populations. We may consider situations as the data from a bivariate normal population. so, we may consider say X 1, Y 1, X 2, Y 2, X n, Y n; this follows a bivariate normal population with means mu 1, mu 2; where ends as sigma 1 square, sigma 2 square, and a correlation coefficient row. We are (( )) testing about the mu 1 minus mu 2;

therefore, what we can do we can consider the linearity property of the bivariate normal distribution, if we consider Y X i minus Y i; this will follow a univariate normal distribution with mean mu 1 minus mu 2, and variance will become sigma 1 square plus sigma 2 square minus twice row sigma 1 sigma 2. Let us write this as say mu d, and this as say sigma square d; let me call this as d i. Then our data is become like d 1, d 2, d n follows normal mu D, S d, and sigma d square. And our hypothesis say for example, mu 1 is equal to mu 2, this is equivalent to mu D is equal to 0.

Similarly, if I say mu 1 greater than mu 2, this is equivalent to mu D greater than 0; similarly mu 1 less than mu 2 this is equivalent to mu D less than 0, if I consider mu 1 is not equal to mu 2, this is equivalent to mu D not equivalent to 0. Therefore, this problem has reduce to the testing of the mean, when we are considering one normal population that is sample from a single normal population, and this is testing about the mean mu is equal to mu not. For this the test has already been derived, let me derive for this particular situation.

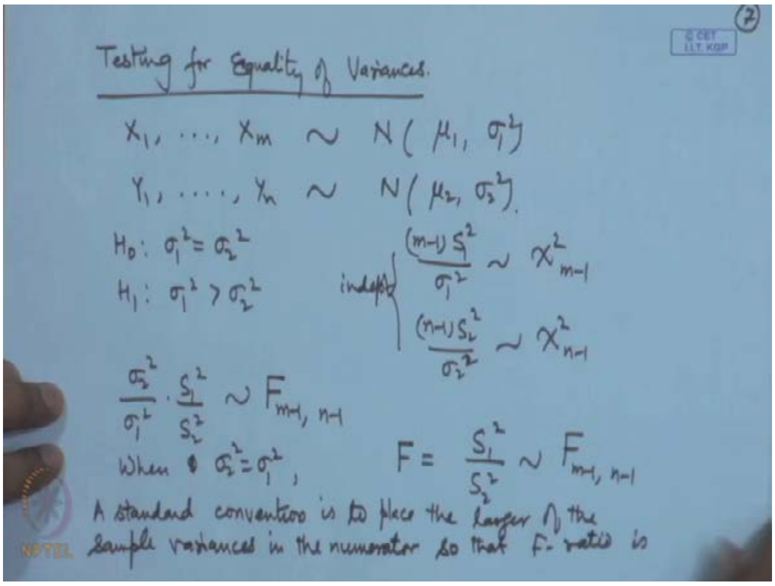
(Refer Slide Time: 19:44)



So, when we are considering H naught mu D is equal to 0, against same mu D greater than 0, then the test is reject H naught if square root n d bar divided by s D; that is greater than or equal to t n minus 1 alpha. So, what is d bar here? d bar is the mean of the d i's, and s D square is nothing but 1 by n minus 1 sigma d i minus d bar whole square.

So, if I consider say H naught mu d greater is equal to 0, against say H 1 mu D less than 0, then the rejection region will become root n d bar by s D less than or equal to minus t n minus 1 alpha, this is the rejection region. Similarly, if I am considering two sided, then that will become rejection region will become two sided; this is the rejection region.

(Refer Slide Time: 21:28)



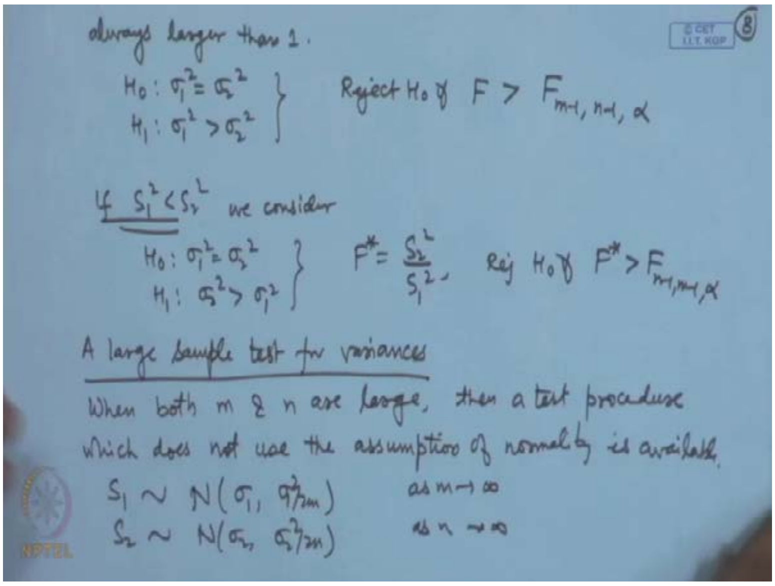
Let us consider also the testing for the variance equality of variances.

(No audio from 21:27 to 21:38)

So, we have two samples X 1, X 2, X m from normal mu 1 sigma 1 square, and Y 1, Y 2, Y n from normal mu 2 sigma 2 square. We are interested in testing say sigma 1 square is equal to sigma 2 square, against say sigma 1 square greater than sigma 2 square. Now, here we consider m minus 1 S 1 square by sigma 1 square that follows chi square on m minus 1 degrees of freedom, and n minus 1 S 2 square by sigma 2 square follows chi square on n minus 1 degrees of freedom. If we assume the independence here, then the ratio of the two chi squares divided by the degrees of freedom that will follow an F distribution.

So, we will get sigma 2 square by sigma 1 square S 1 by S 2 square that will follow F distribution on m minus 1 n minus 1 degrees of freedom. So, when sigma 1 square is equal to sigma 2 square, then we consider F that is the equal to S 1 square by S 2 square that will follow F on m minus 1 n minus 1 (( )). Now, standard convention is to keep the larger of the sample variances in the numerator. So, so that we do not have to look at the tables from both the sides, so a standard convention is to place the larger of the sample variances in the numerator. So, that F ratio is always larger than one.

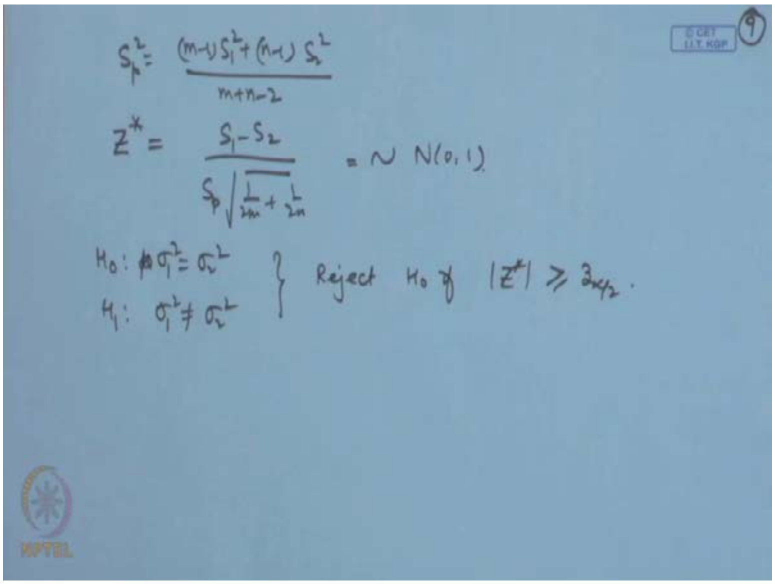
(Refer Slide Time: 23:56)



So, if we consider this hypothesis, that is H naught sigma 1 square is equal to sigma 2 square against H 1 sigma 1 square is greater than sigma 2 square. So, the test will be H naught sigma 1 square is equal to sigma 2 square, against H 1 sigma 1 square greater than sigma 2 square, we will have the test as reject H naught, if F is greater than F m minus 1 n minus 1 alpha. If S 1 square is less than S 2 square, we consider H naught sigma 1 square is equal to sigma 2 square, against sigma 2 square greater than sigma 1 square.Aand we consider F star is equal to S 2 square by S 1 square, and consider reject H naught if F star is greater than F n minus 1 m minus 1 alpha.

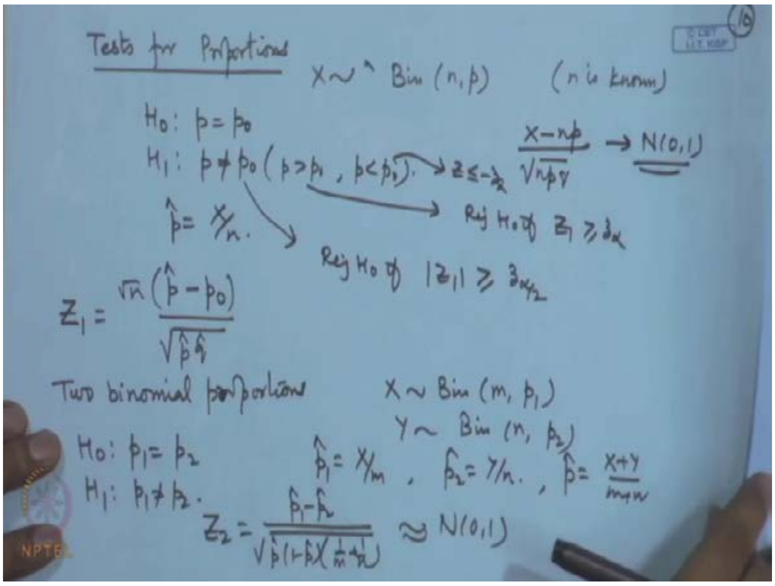
There is a large sample test also a large sample test for variances, when both m and n are large, then a test procedure which does not use the assumption of normality is available. We can consider S 1 following normal sigma 1, sigma 1 square by 2 n, as m is large and similarly S 2 follows normal sigma 2 sigma 2 square by 2 n as n is large.

(Refer Slide Time: 26:40)



We can consider here once again S p square is same as m minus 1 S 1 square plus n minus 1 S 2 square divided by m plus n minus 2, and we formulate Z star as S 1 minus S 2 divided by S p square root 1 by 2 m plus 1 by 2 n, this is approximately normal 0. So, if I am considering say mu 1 sigma 1 square is equal to sigma 2 square against say sigma 1 square not sigma 2, square then we can consider the critical region as reject H naught; if modulus z star is greater than or equal to z alpha by 2.

(Refer Slide Time: 27:51)

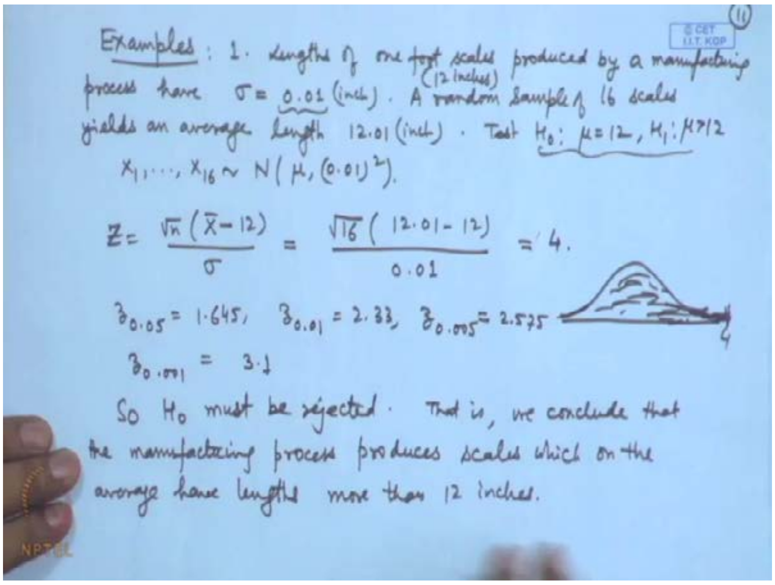


We can also consider the test for binomial proportions. Let me describe one such situations tests for proportions. So, the situation is that, we may have a binomial population. So, binomial n, p, and usually n is known; we have an observation from this binomial population, and we may have to test say p is equal to p naught against say H 1 p is not equal to p naught or say p greater than p naught or p less than p naught. In this case we make use of the normal approximation to the binomial distribution; if you remember, we have x minus n p divided by root npq; this is approximately normal 0, 1.

So, we can base our test on this we can consider. So, let us you say, let us call p hat as x by n. So, we consider p hat minus p naught divided by root. So, what we have done we have divided by n here. So, if we also consider multiplication by that. So, root n p head q head. So, we have guess test statistic here, we can call it say z 1, so we will consider say for this hypothesis reject H naught, if modulus of z 1 is greater than or equal to z alpha by 2. If I have one sided, then we can consider reject H naught, if z 1 is greater than or equal to z alpha and in this case it will become z less than or equal to minus z alpha. Similarly, we may have to compare the proportions of two normal two binomial populations, suppose we are considering two binomial proportions.

So, we may have say X following binomial m, p 1, and say Y following binomial n, p 2. We are interested to test whether p 1 and p 2 are the same or not; that means, we may have hypothesis like H naught p 1 is equal to p 2, against p 1 not equal to p 2. Let us use the notation say p 1 hat is equal to the first proportion, p 2 hat as the second proportion, and let us also define a pooled proportion X plus Y divided by m plus n p hat. Then, if we consider p 1 hat minus p 2 hat divided by square root p hat into 1 minus p hat into one by m plus one by n, then this will have approximately normal 0, 1. So, let us denote this by say Z 2, and we can give the rejection region as modulus z 2 greater than or equal to z alpha by 2.

(Refer Slide Time: 31:34)



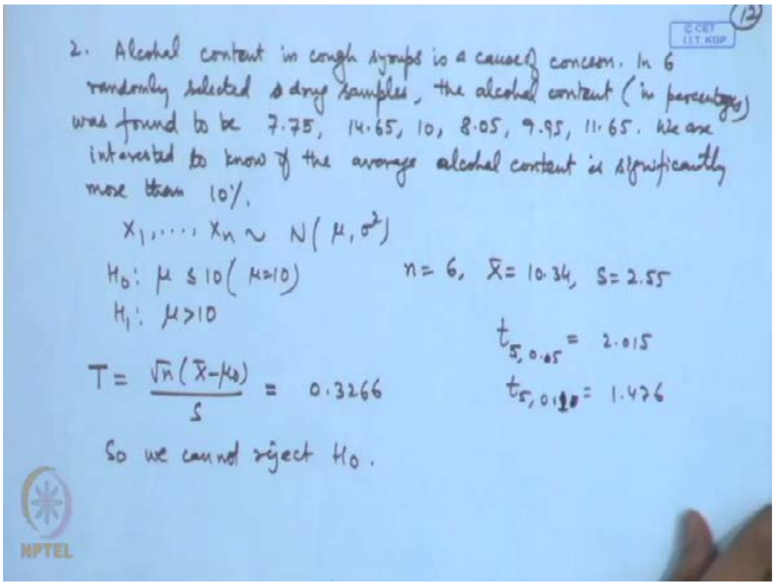
Let me consider certain examples here, on various testing problems lengths of one foot scales produced by a manufacturing process. So, one foot means basically 12 inches have sigma is equal to 0.01 inch, then we want to test whether the average length of this scale is actually equal to 12 inches or not.

So, a random sample of 16 scales yields an average length 12.01 inch. So, we want to test whether the production process is consistent or not; that means, test H naught whether mu is equal to 12, against say H 1 mu is greater than 12. That means, this actual difference of 0.01 exist significantly larger or not; actually this testing of problem testing of hypothesis problem have to be seen in a proper physical perspective, because here the practical problem for the manufacturer is whether is manufacturing process produces the one feet scales which are actually conferment to the guideline. That means, they should actually make the 12 feet. If there are in general larger than they are not good. So, it turns out this sample produces 12.01.

So, does it, is it consistent with the hypothesis whether mu is equal to 12 or not. And the variability is known that it is 0.01. So, in this case the model is that we are having X 1, X 2, X 16 following normal mu sigma S square is 0.01 square. So, we want to test this one. So, we create the test statistics root n X bar minus 12 that is mu not divided by sigma. So, here it is root 16 X bar is 0.01 minus 12 divided by 0.01. Now, this is equal to 4, now you see the normal distributions curve.

So, 4 will be come in somewhere here. So, actually all the probabilities almost all the probabilities concentrated before 4; therefore, this value is certainly large; that means, if I consider say z is equal to 0.05 that is 1.645, if I consider say z is 0.01 that is 2.33; if I consider say z is equal to 0.005 that is 2.575. If I consider say z is equal to 0.001, then that is 3.1. So, at all these levels. So, H naught must be rejected; that means, certainly the data does not support the hypothesis that mu is equal to 12, that is we conclude that the manufacturing process produces scales which on the average have lengths more than 12 inches.

(Refer Slide Time: 36:08)



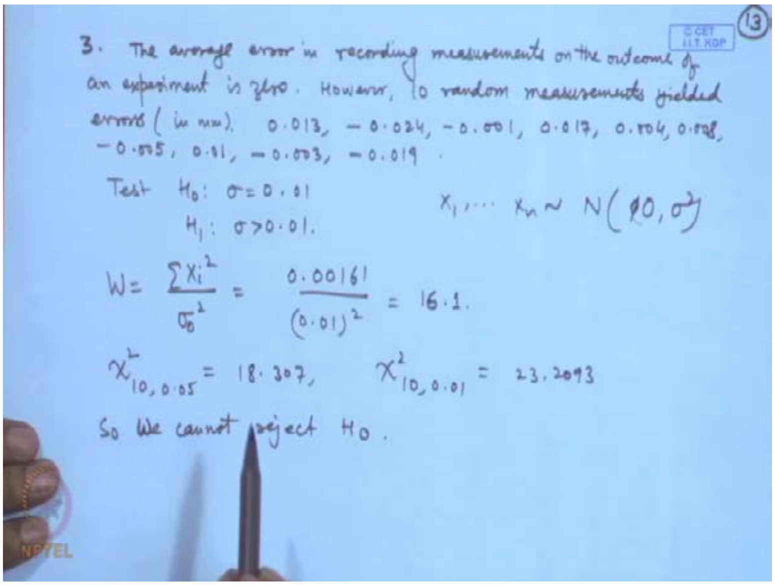
Alcohol content in cough syrups is a cause of concern, because people use it as drugs in 6 randomly selected drugs samples, the alcohol content, and it was measured in percentages was found to be 7.75, 14.65 ,10, 8.05, 9.95, 11.65. So, we are interested to know, if the average alcohol content is significantly more than 10 percent.

So, this can be consider as a testing problem, that we are having a data from normal mu, sigma square, and we want to test whether mu is less than or equal to 10 or mu is greater than 10. This hypothesis of course, we may write as mu is equal to 10 also, it does not matter, because the test procedure is dependent upon the alternative hypothesis in the Neumann Pearson theory. Here we have we will use the test statistics square root n X bar minus mu not by S. So, for this particular data set n is 6, X bar turns out to be 10.34, and S is equal to 2.55; these figures are approximated to 2 decimal digits.

And therefore, this value turns out to be 0.3266. Now, if I look at the value of the t distribution the 0.05 value on 5 degrees of freedom that is equal to 2.015, if I look at say 0.01 degrees of freedom sorry 0.1, on 5 degrees of freedom the point one point, that is 1.476. Naturally this value is smaller. So, we cannot reject H naught here on the basis of this data; that means, the drug alcohol content in the cough syrup is less than or equal to 10 percent here.

(No audio from 39:30 to 39:44)

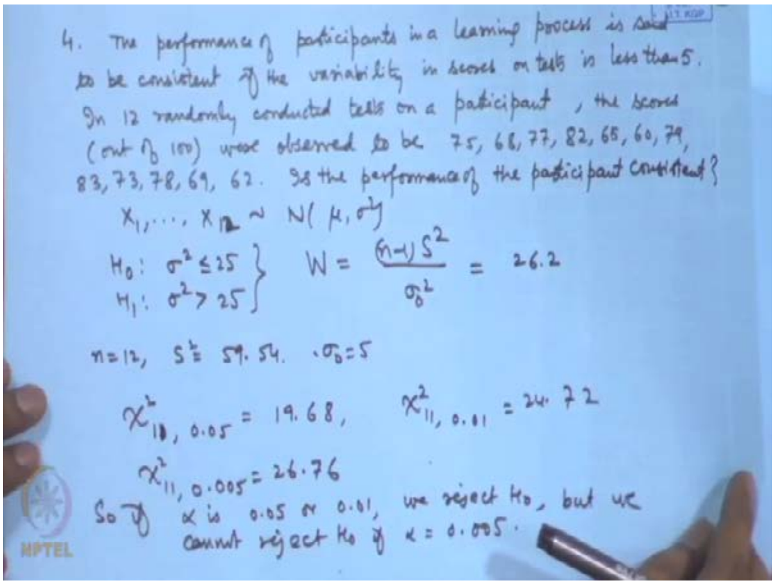
(Refer Slide Time: 39:42)



The average error in recording measurements on the outcome of an experiment is 0. However, 10 random measurements yielded errors, and these are in say millimeter 0.013 minus 0.024 minus 0.001 plus 0.017, 0.004, 0.008 minus 0.005, 0.01 minus 0.003 minus 0.019. You want to test whether the variable t is 0.01 or more. Now, this is the case when we are having the data from a normal population when mean 0, and variance sigma square. So, our test statistic is sigma X i square by sigma naught square that is equal to now for this particular data sigma X i square is 0.00161, and this is 0.01 square that is equal to 1 16.1 here. Now, if we look at chi square values on 10 degrees of freedom then 0.05 point is 18.307, if you consider chi square on 0.01 that is 23.2093, etcetera. So, we cannot reject H naught here; that means, we can claim that here the average variable t is less than or equal to 0.01.

(No audio from 42:17 to 42:30)

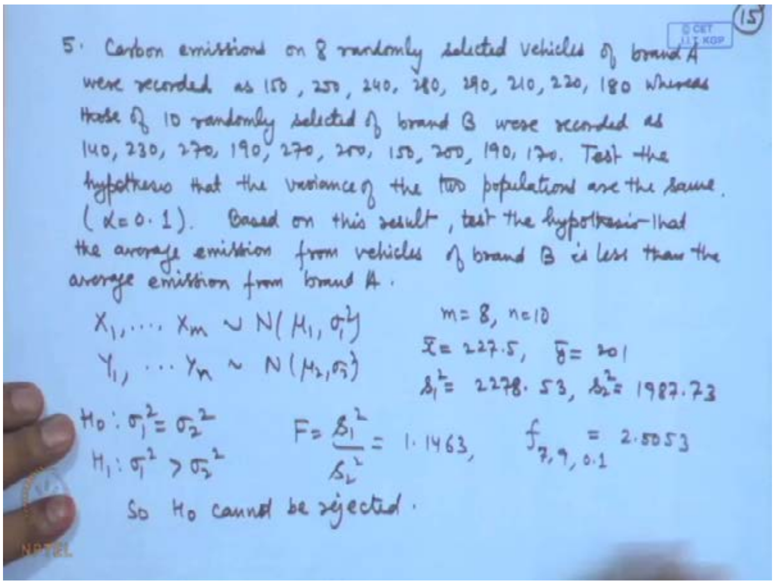
(Refer Slide Time: 42:22)



The performance of participants in a learning process is said to be consistent, if the variability in scores on tests is less than 5.In 12 randomly selected, randomly conducted tests on a participants the scores out of 100 were observed to be same 75, 68, 77, 82, 65, 60, 79, 83, 73, 78, 69, 62, is the performance of the participant consistent. So, this can be consider as the problem from that we have a data from normal mu sigma square population, and we want to test whether sigma square is less than or equal to 25 or sigma square is greater than 25. For this we construct the test statistics n minus 1 S square by sigma naught square. So, here n is equal to 12, S square we can calculate as 59.54, and sigma naught is equal to 5. So, this value terms out to be 26.2. Now, if we look at the chi square value on 11 degrees of freedom, say 0.05; this is equal to 19.68. If we consider chi square 11 on, that is 24.72; however, if I consider chi square value on 0.005 that is equal to 26.76.

So, if alpha is say 0.05 or 0.01 then we reject H naught, but we cannot reject H naught; if alpha is taken to be very very small. You can see here our decision to accept or reject H naught is dependent upon the level that we decide. So, in the significance testing we take the minimum value of alpha for which we cannot reject H naught or we reject H naught. So, that is called the p value, and in many studies we simply report p value, and it will be dependent upon the practitioner of the person who is going to use, whether that is really significant or not, so that is called significant testing, but the more about that later now.

(Refer Slide Time: 46:32)

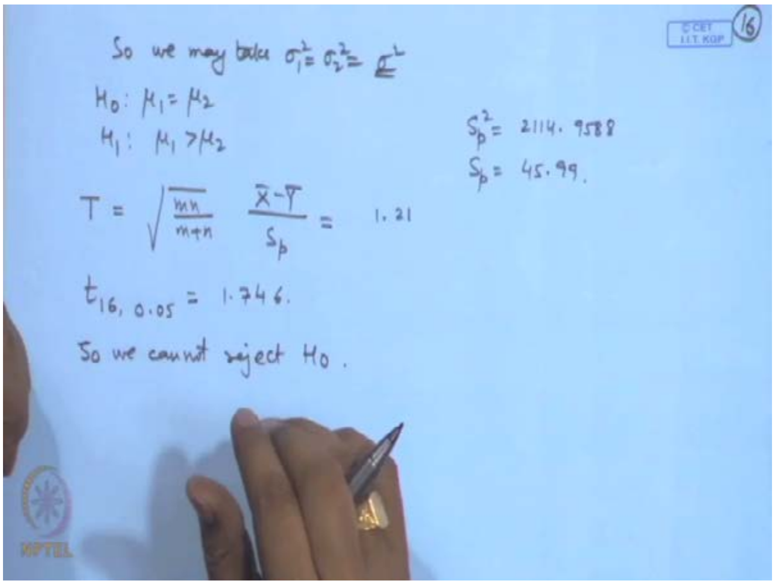


Let us consider the two sample problems, carbon emissions on 8 randomly selected vehicles of brand A were recorded as 150, 250, 280, 290, 210, 220, 180 whereas, those of 10 randomly selected of brand B were recorded as say 140, 230, 270, 190, 270, 200, 150, 200, 190, and 170. Now, first of all test the hypothesis that the variances of the two populations are the same, and for this (( )) you can takes say for example, alpha is equal to 0.1; that means, 10 percent level of significance. Now, based on this result test the hypothesis that the average emission from vehicles of brand B is less than the average emission from brand A.

Now, this is a two sample problem, we can consider the modulus one random sample from normal mu 1, sigma 1 square, and another is from normal mu 2 sigma 2 square. The two samples are consider independent; here m is equal to 8, n is equal to 10, and let us calculate the means that is 227.5 for the first sample, for the second sample it is 201, and we also calculate the sample variances from the two populations. Now, firstly we carry out f test for the equality of the variances; now notice here that we are having S 1 square larger. So, we consider the alternatives whether sigma 1 square is significantly larger than sigma 2 square.

Now, if you consider this then we take F as S 1 square by S 2 square, and that is equal to 1.1463; and the corresponding f value on m minus 1 n minus 1 degrees of freedom, and at alpha is equal to 0.1, if you see the tables of the f distribution this is 2.5053. Now, this value is smaller, so H naught cannot be rejected here.

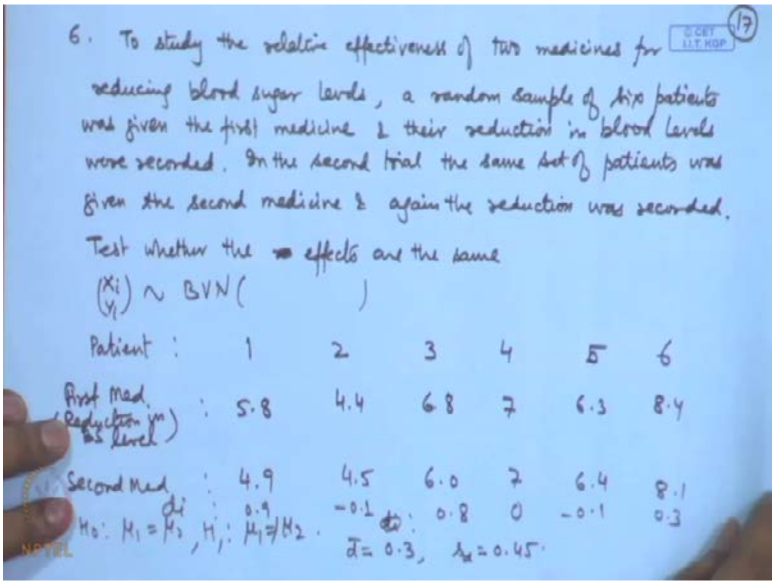
(Refer Slide Time: 51:09)



Now, if you want to do the testing for see if X naught cannot be rejected. So, we may take sigma 1 square is equal to sigma 2 square is equal to sigma square. Now, if you want to consider the test for mu 1 is equal to mu 2 against say mu 1 is greater than mu 2, then we consider the test statistics root m n by m plus n; that is for the pooling X bar minus Y bar divided by S p. So, here you calculate the pooled sample variance, that turns out to be 2114.9588, that is S p is equal to 45.99. Then this value turns out to be 1.21, if you look at the T value on m plus n minus 2 degrees of freedom as a 0.05 that is equal to 1.746.

And if I consider 0.01 etcetera, the value is going to be further larger; therefore we cannot reject H naught; that means, the carbon emission in the second vehicle average emission is not significantly smaller than the first one. Although from the values here, you can see it is 201 and here it is 227, but since the variability is quite large; therefore, we conclude here that the variability of the two of them is almost same as we concluded by f test, and then by applying the pooled sample variance test for the equality of the means, we are concluding that we cannot reject the hypothesis of the equality here.

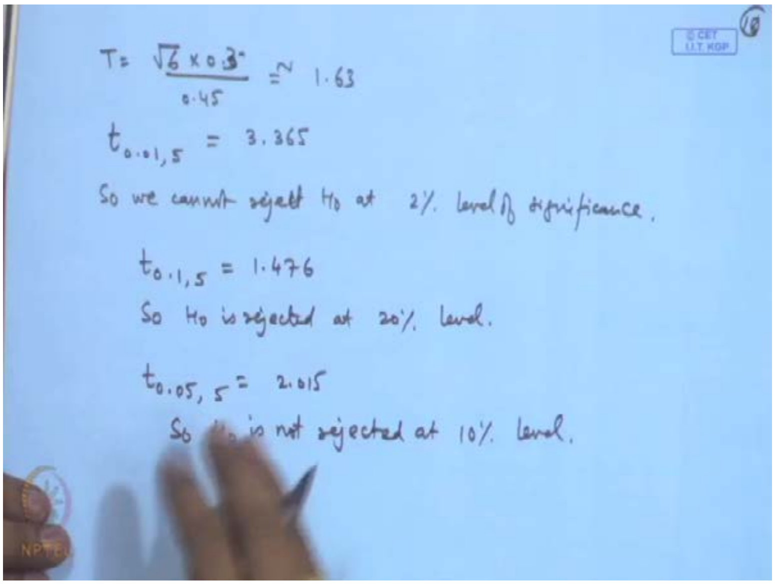
(Refer Slide Time: 53:21)



Let me give one example of the pairing here. To study the relative effectiveness of two medicines for reducing blood sugar levels, a random sample of 6 patients was given the first medicine, and their reduction in blood sugar levels were recorded. Thereafter, in the second trial the same set of patients was given the second medicine, and again the reduction was recorded. We want to test, whether the effects are the same; that means, the two medicines are equally effective or they are not.

So, here the data is paired we are considering say x i, y i following bivariate normal model. So, the data that is given here is on the 6 patients, let me name them as 1, 2, 3, 4, 5, 6. And reduction in the blood sugar level - reduction in blood sugar level that is recorded here, as say 5.8, 4.4, 6.8, 7, 6.3, 8.4; and in the second medicine it is 4.9, 4.5, 6.0, 7, 6.4, and 8.1. We want to test whether mu 1 is equal to mu 2 or mu 1 is not equal to mu 2. So, we consider the differences; the differences here, let us called d i. So, it will be 0.9 minus 0.1, 0.8, 0, minus 0.1, and 0.3. So, here d bar will be equal to 0.3, s p is equal to 0.45.

(Refer Slide Time: 56:52)



So, T is equal to root 6 point 0.3 divided by 0.45, that is equal to 1.63 approximately. If we consider the t value on say 0.01 alpha is equal to 0.01 at 5 degrees of freedom that is 3.365. So, certainly we cannot reject - we cannot reject H naught at say 2 percent level of significance; however, if we reduce the if we increase the level of significance, we can consider the t table here. Let me show you (( )) for 5 degrees of freedom; if we consider for example, if we consider for example, say 20 percent here, in place of 2 percent suppose I take t 0.1 on 5, then that value is equal to 1.476.

Now, here if you compare, so H naught is rejected then at 20 percent level. So, we need to fix up the level of significance in the given problems. However, if I consider say 10 percent. So, if I take 0.05 on 5, that value is equal to 2.015. So, H naught is not rejected at 10 percent level. Therefore, in the testing problem, it is extremely important that we carefully (( )) our level of significance, that we want we should be sure of how much level of significance we can allow in the given problem, because the related to the probability of type one error. So, this is the Neumann Pearson theory, there are other comparative theory like the significance testing dimension the minimum level of significance where the hypothesis is rejected, then there are other procedures like we have (( )) etc. So, one has to take here (( )).