Probability and Statistics

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

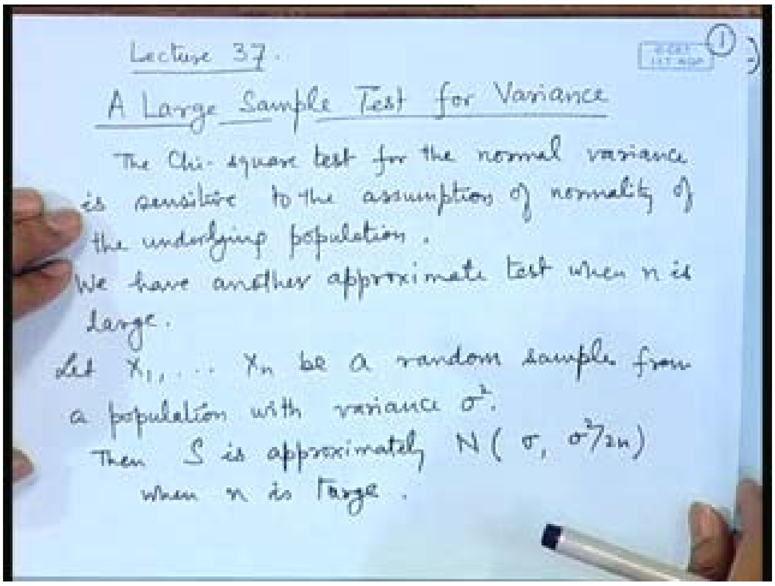
Lecture No. #37
Testing of Hypothesis-V

In the last lecture, I have explained how to test hypothesis for parameters of a normal distribution. For example, if I have available a random sample from a normal distribution with mean mu and variance sigma square, then how to conduct the tests for mean that is, mu or for sigma square? We have considered various cases that means, when we are testing the mean is equal to certain quantity, or it is less than or equal to certain quantity then whether variance is known or unknown; in the different two cases, we have two different tests- one is based on a normal distribution and another is based on a T distribution.

Similarly, when we are testing for the variance then we are assuming mean may be known or mean may be unknown and you have a two different chi square tests. We also discussed the tests for proportion that is, binomial distribution for the large n. So, we approximated it by a normal distribution and we considered the test statistic as X minus np naught divided by square root of np naught q naught, where p naught is the value which is to be tested against a general hypothesis point.

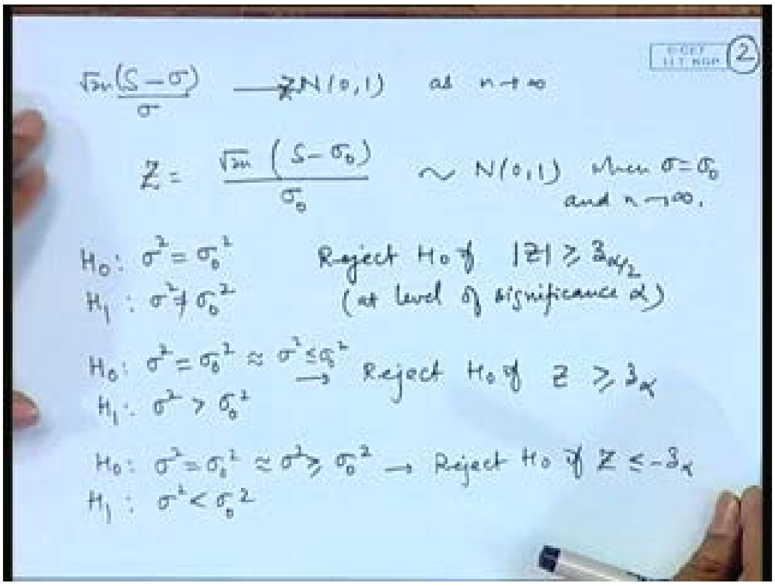
In the case of the variance testing, we used a chi square test, that test is quite sensitive to normality assumption, if the assumption of the normality is not satisfied, then the test may be somewhat bad, it may give a false result. However, if we observe the distribution of S square that is used there, which is chi square under the normality assumption; even if we do not have the normality assumption, for large distribution of S it can be approximately normal distribution.

(Refer Slide Time: 02:17)



So, the chi square test for the normal variance is sensitive to the assumption of normality of the underlying population. So, we have another approximate test when n is large. So, let X1, X2, Xn be a random sample from a population with variance sigma square. Then, S is approximately normal sigma, sigma square by 2n when n is large.

(Refer Slide Time: 04:16)



So, since S square is unbiased estimate for sigma square and S is not unbiased, but asymptotically it is unbiased for sigma therefore, we may make the following thing that S minus sigma divided by sigma by root 2n then, this is approximately normal 0, 1 as n tends to infinity, this tends to a. So, if we define our random variable Z is equal to root

2n S minus sigma naught divided by sigma naught, then this follows normal 0, 1 when sigma is equal to sigma naught and n tend to infinity.

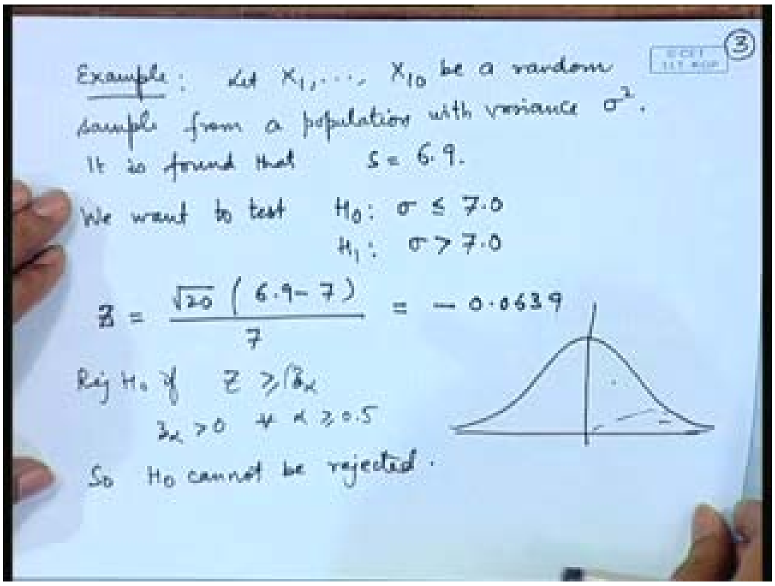
So, if we want to test about say, sigma square is equal to sigma naught square against say, sigma square is not equal to sigma naught square, we can make a test based on this Z variable. So, we can use the test as reject H naught if modulus Z is greater than or equal to Z alpha by 2- so, at level of significance alpha, the test is that you reject H naught when the absolute value of Z is greater than or equal to Z alpha by 2.

Similarly, we can write one sided test. For example, if I have H naught, sigma square is equal to sigma naught square against say, sigma square is greater than sigma naught square, then we will reject H naught if Z is greater than or equal to Z alpha of course, this hypothesis is equivalent to writing sigma square is less than or equal to sigma naught square.

Similarly, if I am considering sigma square is equal to sigma naught square against one sided alternative, sigma square less than sigma naught square, then we will reject H naught if Z is less than or equal to minus Z alpha, and again this is equivalent to sigma square greater than or equal to sigma naught square.

Let me give one example here. Also the framing of the hypothesis is important here, since we are controlling the probability of the type 1 error that is, rejecting H naught value when it is true therefore, rejecting H naught is always a considered as a strong conclusion. Therefore, the hypothesis for which we want to give an importance, we can put it in H1 of course, that may not be possible if you are testing for the equality because then inequality cannot be put in the alternative. I will give one example here where this role of null and alternative hypothesis is important and one has to take a justified decision that in what way we should write down the null and the alternative hypothesis.

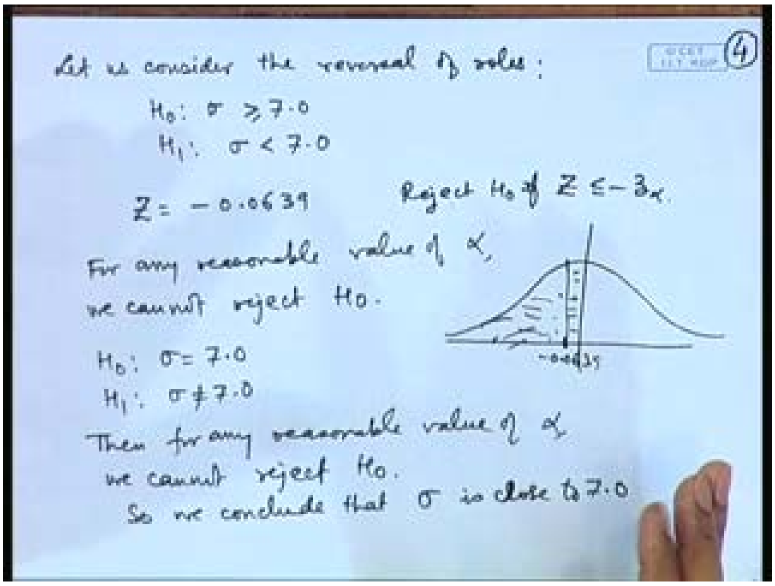
(Refer Slide Time: 07:25)



Let us consider, let X1, X2, X10 be a random sample from a population with variance sigma square, it is found that S square, S is equal to 6.9 say. Now, we want to test say, H naught, whether sigma square is less than or equal to 7 against H 1, sigma is greater than 7. So, let us create this Z variable here, that is root 2n is root 20, we have ten observations here, S is 6.9 minus 7 divided by 7- so, this can be evaluated easily, it is minus 0. 0639.

Now, if we consider say, we have to reject H naught if Z is greater than or equal to say Z alpha; now, for any level of alpha up to say 0.5, the value of Z alpha will always be non-negative because Z is a symmetric distribution about 0, so, all this points, these all are positive, so, this value will never negative, this Z alpha is positive for all alpha greater than or equal to 0.5, so H naught cannot be rejected. Now, one may feel that since we have rejected H naught, we cannot reject H naught that means, there is a strong support to the hypothesis that the variance is smaller than 49, or the standard deviation is smaller than 7, but this is slightly misnomer here, we should be careful in the choice of the null and alternative the hypothesis.

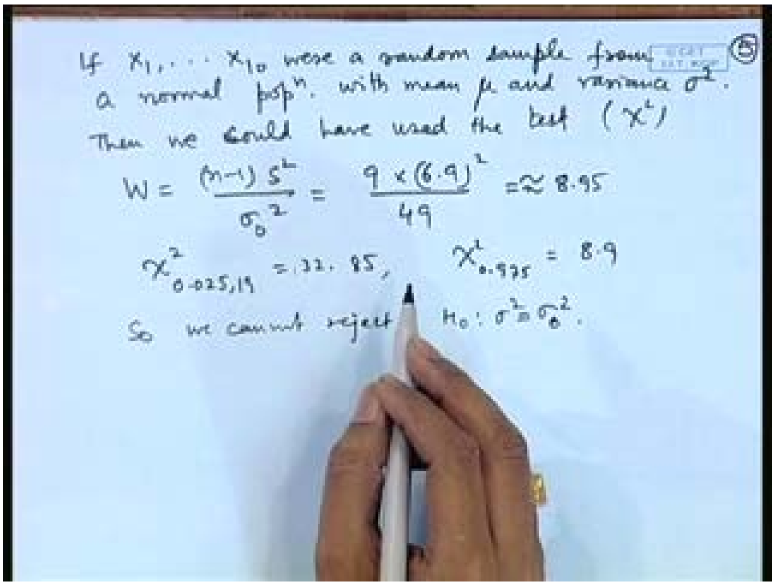
(Refer Slide Time: 10:00)



Let us consider say, the reversal of roles. Let us consider say, H naught as sigma is greater than or equal to 7.0 against H1 sigma is less than 7.0. Now, if you consider your Z value, that is minus 0.0639 and we are saying reject H naught if Z is less than or equal to minus Z alpha; now, you can see for any reasonable level of significance- because these are the negative values- so, we are saying Z is equal to minus 0.06, that is somewhere here that means, there is a very small probability about this, so, for a large value of alpha the hypothesis of H naught cannot be rejected because all these values see, this value 0.00639, this is much higher. So, for any reasonable value of alpha we cannot reject H naught. What does it mean, does it mean that sigma is greater than or equal to 7? Because in the previous test we concluded that sigma is less than or equal to 7 from here. So, I mentioned this point earlier that we frame the hypothesis in such a way that H naught is actually a weaker hypothesis, or you can say accepting H naught is a weaker conclusion therefore, we put stronger emphasis on the alternative hypothesis.

So, if we see carefully, probably, the reason is that the Z value is smaller that means, there is a reason to suspect that if we consider the hypothesis the sigma is equal to 7 against sigma is not equal to 7, then for any reasonable value of alpha we cannot reject H naught. So, we conclude that sigma is close to 7 that means, hypothesis of sigma is equal to 7 cannot be rejected that means, sigma is actually not significantly different from 7. So, this is an example of explaining that we should frame the null and the alternative hypothesis in a judicious way, as well as we should analyze the result for possible rejection of various kind of hypothesis, because we should not conclude falsely here for example, if we look at this and since this cannot be rejected, we are getting that actually sigma is greater than 7, but it is not significantly greater, because when we consider sigma less than or equal to 7 then also this hypothesis could not be rejected. So, the only conclusion that we can draw is that actually sigma is closer to 7.

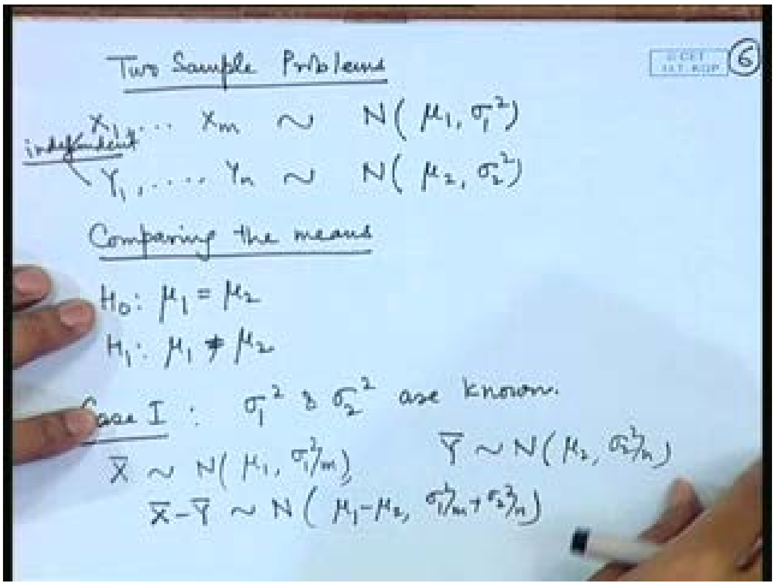
(Refer Slide Time: 13:41)



In this case suppose normality assumption was there, if the assumption was there, then we may consider if X1, X2, X10 were a random sample from a normal population with say mean mu and variance sigma square, then we could have used the chi square test that is, W is equal to n minus 1 S square by sigma naught square, that is 9 into 6.9 square by 49, that is approximately- this value can be calculated as- so, this will be closer to 9, but slightly less than 9 some 8 point may be approximately 8, let me put it here.

If we see the chi square value here form the tables, chi square say, 0.025, that is 32.8 5, chi square say, 0.975, that is 8 point something, so, this is closer to this thing 8.95 something, this is 8.9 here, so, we can see here that neither W greater than this is satisfied nor W less than this is satisfied. So, we cannot reject hypothesis sigma square is equal to sigma naught square here also. So, the main point which I was trying to make is that here the hypothesis of equality cannot be rejected under any alternative. Now, we consider two sample problems- that means, we may have to compare the means of two different distributions, the variances of two different distributions, etcetera.

(Refer Slide Time: 16:34)

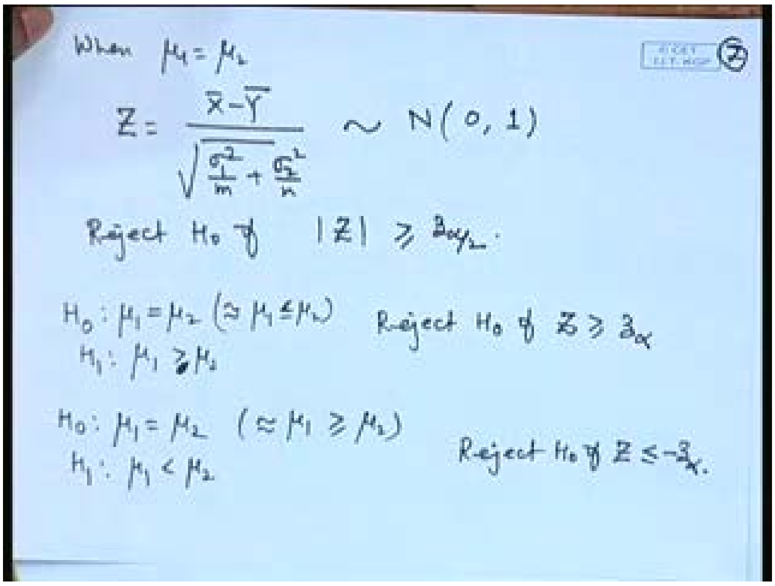


So, let us consider the case when there is a normality assumption- so, two sample problems. So, in the beginning, let us consider the normal model, let X1, X2, Xm, Xm be a random sample from normal distribution with mean mu1 and variance sigma1 square and let Y1, Y2, YN be a sample from another normal population with mean mu2 and variance sigma2 square. This type of situation occurs when, many times we are interested in comparing two different setups for example, we may have two brands of certain product and we want to compare their average life, average performance; we may have say, two ethnic populations we may like to consider their say average expenditure on health care per month; we may have two different countries, we may like to compare the average longevity or average life of people in the two different countries.

In such cases, the appropriate model is that we are considering two independent random samples, one from one normal population with certain mean and variance and another from another normal population with certain mean and variance. Of course, the assumption of normality could have been replaced by some other populations also, but we are considering certain procedures which are applicable to the normal populations. Now, suppose we want to compare the means, now, if you want to compare the means that means, we may have a hypothesis of the types say, mu1 is equal to mu2 against say, mu1 is not equal to mu2, or one sided hypothesis like mu1 is less than or equal to mu2 against mu1 greater than mu2, etcetera.

Now, there are various cases here. So, first case is we may have the information about sigma1 square and sigma2 square, this may be through some prior experiment or from past data, the values of the variances may be known; in that case, the testing problem becomes much simpler, we may consider X bar as normal mu1 sigma1 square by m and Y bar follows normal mu2 sigma2 square by n. Now, X bar and Y bar are independent because these two samples are considered independent. Therefore, if we consider the distribution of X bar minus Y bar, that can be normal mu1 minus mu2 sigma1 square by m plus sigma2 square by n. Under the null hypothesis when mu1 is equal to mu2 this mu1 minus mu2 term becomes 0.

(Refer Slide Time: 20:03)

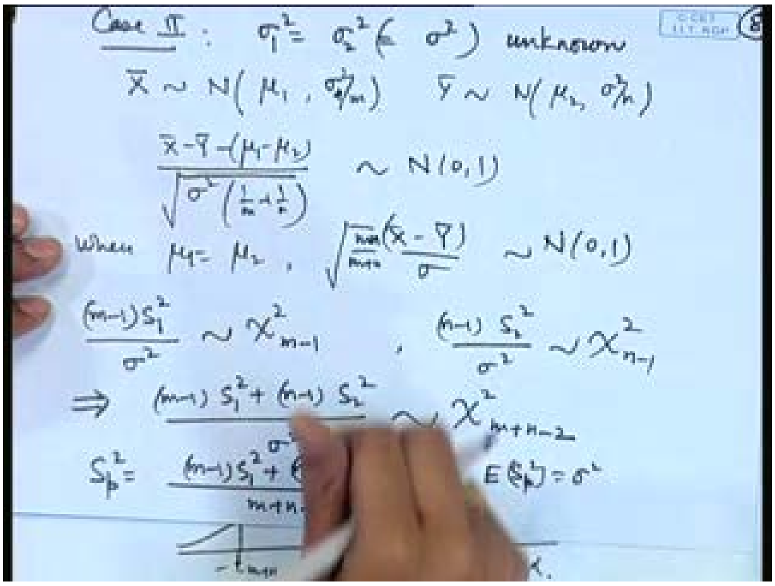


And therefore, easily we can construct the test function when mu1 is equal to mu2, the function X bar minus Y bar divided by square root of sigma1 square by m plus sigma2 square by n, this follows a normal distribution with mean 0 and variance unity. So, this is a perfect test statistic to be utilized for testing this hypothesis. So, we may reject H naught if modulus of Z is greater than or equal to say Z alpha by 2. So, this is a test for level of significance alpha. We can answer the question for the one sided hypothesis testing problems also.

For example, if I have mu1 is equal to mu2 against say, mu1 greater than, or greater than mu2 of course, this one is also equivalent to writing mu1 less than or equal to mu2 because a rejection region is dependent upon the alternative hypothesis, so, we may take the rejection region as reject H naught if Z is greater than or equal to Z alpha. Likewise, we may write for the one sided hypothesis where mu1 is less than mu2, mu1 is greater than or equal to mu2. So, here you will reject H naught if Z is less than or equal to minus Z alpha. We will consider examples of this a little later. Firstly, let us consider the other cases also.

So, if the variances sigma1 square and sigma2 square are unknown, then we cannot utilize this test statistic, in that case, we will have to put certain estimates for sigma1 square and sigma2 square.

(Refer Slide Time: 22:21)

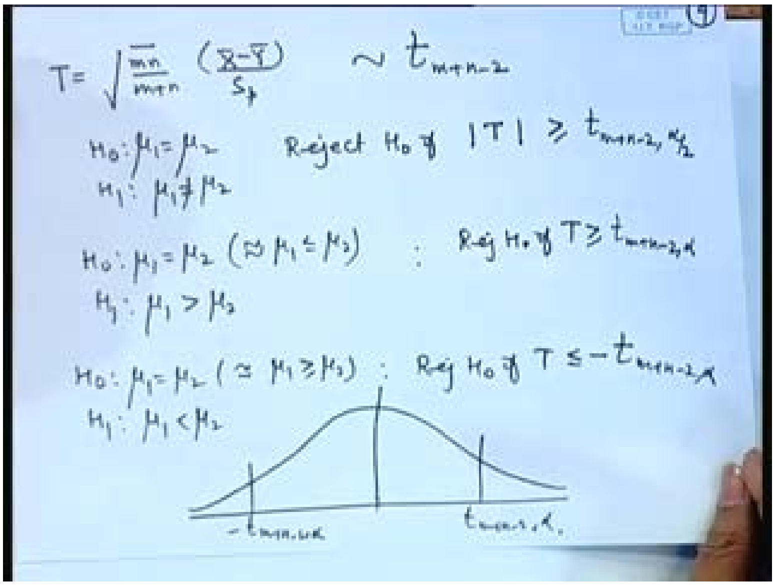


Once again, here we have two different cases. So, we have case two when sigma1 square is equal to sigma2 square is equal to say, sigma square, which is unknown. So, here we are assuming that the variances are not known, but they are known to be equal, this type of situation occurs in the cases such as, we are considering measuring two different instruments using certain device and our measuring instrument, since it is the same measuring instrument the variability in the measurements will be same, but since it is measuring two different instruments therefore, the value which is being measured that is, the means may be different- so, you may have different means, but the variances may be same. So, we may not know what is the variance, but since it is, we are using the same instrument then the variability is likely to be the same.

For example, if we have a, we conduct a test and the test paper has a certain difficulty level, therefore if we conduct on two different sets of students, the average marks may differ, but the variability may be the same because the test procedure is the same, we are using the same test. So, in such a situation, we may make use of, that X bar follows normal mu1, sigma square by m and Y bar follows normal mu2, sigma square by n. So, here X bar minus Y bar minus mu1 minus mu2 divided by sigma square 1 by m plus 1 by n square root, this follows normal 0, 1- once again, the independence is utilized here. Now, when mu1 is equal to mu2, you will get X bar minus Y bar square root mn by m plus n by sigma, this follows normal 0, 1 distribution.

Now, we will consider estimation of sigma here because we cannot utilize this sigma in the test statistic. So, we look at the estimates of sigma square from both the samples and we may merge them. So, if we look at S1 square, then S1 square is an estimate for sigma square in fact, we have m minus 1 sigma1 square S1 square by sigma square, this follows chi square distribution on m minus 1 degrees of freedom, and n minus 1 S2 square by sigma square follows chi square distribution on n minus 1 degrees of freedom. Since the samples are independent this statistic and this statistic are independent and therefore, by the additive property of the chi square distributions, we have m minus 1 S1 square plus n minus 1 S2 square by sigma square follows chi square distribution on m plus n minuS2 degrees of freedom. So, if you define Sp square as the pooled sample variance, then this follows, then expectation of Sp square, that is sigma square.

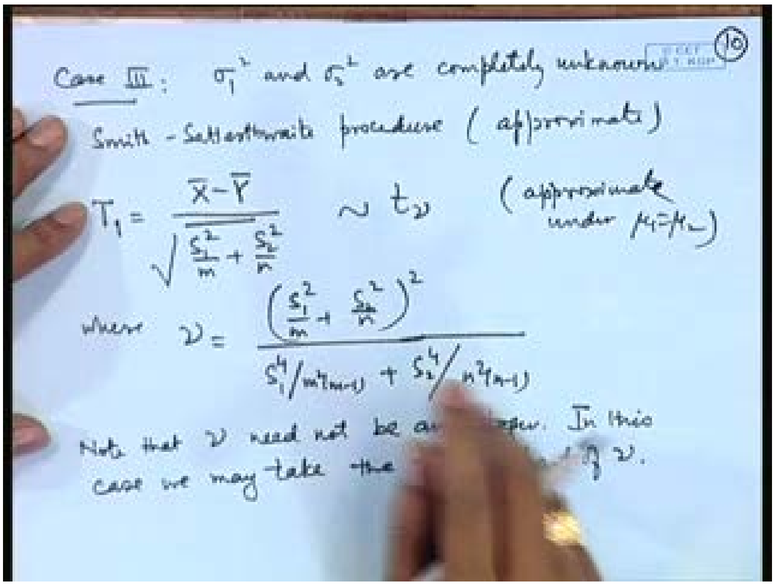
(Refer Slide Time: 26:17)



So, if we consider here the ratio, we get square root mn by m plus n X bar minus Y bar by Sp, that will follow t distribution on m plus n minus 2 degrees of freedom Let us denote this by T. So, for H naught, mu1 is equal to mu2 against mu1 is not equal to mu2. We can consider the test as reject H naught if modulus of T is greater than or equal to t m plus n minus 2 alpha by 2- so, this is the level alpha test here. If we have one sided alternatives for example, mu1 is equal to mu2 against say, mu1 is greater than mu2 of course, this null hypothesis can also be replaced by mu1 less than or equal to mu2, we will have the test reject H naught if T is greater than or equal to t m plus n minus 2 alpha. On the other hand, if we have mu1 is equal to mu2 against mu1 is less than mu2 of course, this null hypothesis is again equivalent to mu1 greater than or equal to mu2, then we can use reject H naught if T is less than or equal to minus t m plus n minus 2 alpha. Because the density of T is symmetric about 0, so, this is minus t m plus n minus 2 alpha point and this will be plus t m plus n minus 2 alpha point.

On the other hand, there may be a situation where sigma1 square and sigma2 square may not be equal, they may be unknown as well as unequal if that is so, then we may not be able to merge it here. For example, here we will have sigma1 square and here we will have sigma2 square likewise, here also we will have sigma1 square by m plus sigma2 square by n. Naturally, if we consider this ratio, the term will not cancel out; in this case it is prudent to replace the unbiased estimators of sigma1 square and sigma2 here and see whether we can do something about the distribution of that, it is observed that that has an approximate t distribution, so, in that case we use the following procedure.

(Refer Slide Time: 29:00)

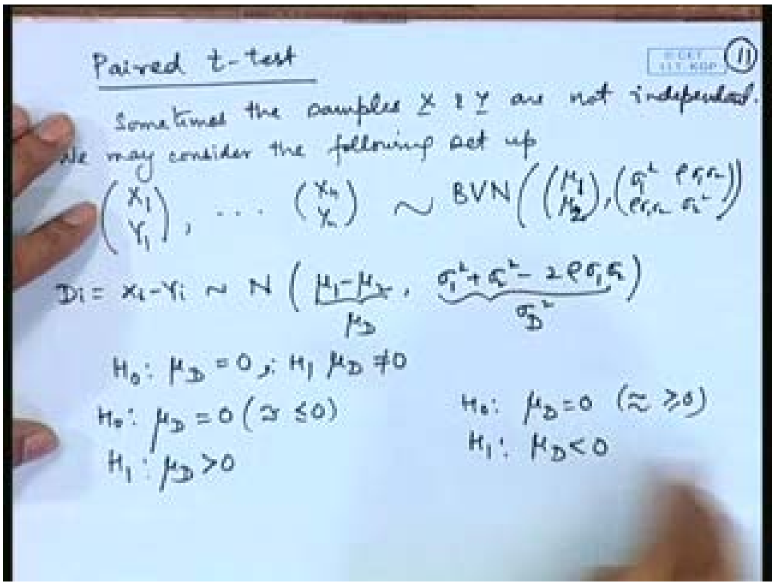


So, we have the third case, when sigma1 square and sigma2 square are completely unknown that means, we have no prior information about them. In that case, we have the following Smith-Satterthwaite procedure, so, this is approximate procedure, it is not exact because exact distribution cannot be determined here. Say, if you consider say T1 is equal to X bar Y bar square root of S1 square by m plus S2 square by n, then this has approximate t distribution on nu degrees of freedom under the assumption that mu1 is equal to mu2, where this mu value is given to be S1 square by m plus S2 square by n whole square divided by S1 to the power 4 by m square into m minus 1 plus S2 to the power 4 divided by n square in to n minus 1.

Now, note here that this nu, note that nu need not be an integer, in this case, we may take the integral part of nu. So, rather taking the rounded off value, which may be higher also, it is better to take the lower value because the power of the test will increase if we consider a lower degrees of freedom here. There is yet another situation which arises here. For example, here we have made an assumption that the two samples are taken independently and we are comparing the means here, but there may be situations where the samples may not be independent. For example, we are testing the effect of a certain medicine on some patients, so, earlier say, one set of patients is chosen they are given one medicine, now, we observe the effects of that medicines for example, it is a medicine to reduce the blood pressures, so, the average effectiveness of this medicine is recorded, now, the same set of patients is given another medicine, may be after 1 month, and then again the effect of the medicine is recorded. Now, here since the set of the patients is a same therefore, the observations Xis and Yis, suppose I call the first set as X1, X2, Xn and the second set of observations as Y1, Y2, Yn, here we cannot assume them to be independent.

Now, if we are not able to assume that they are independent, then the procedures that I describe earlier in case one, case two and case three cannot be adopted here. Because these procedures assume that the things are independent because we have made use of the additive properties of the chi square distribution, or we have used say, linearity property of the normal distribution, etcetera.

(Refer Slide Time: 32:55)

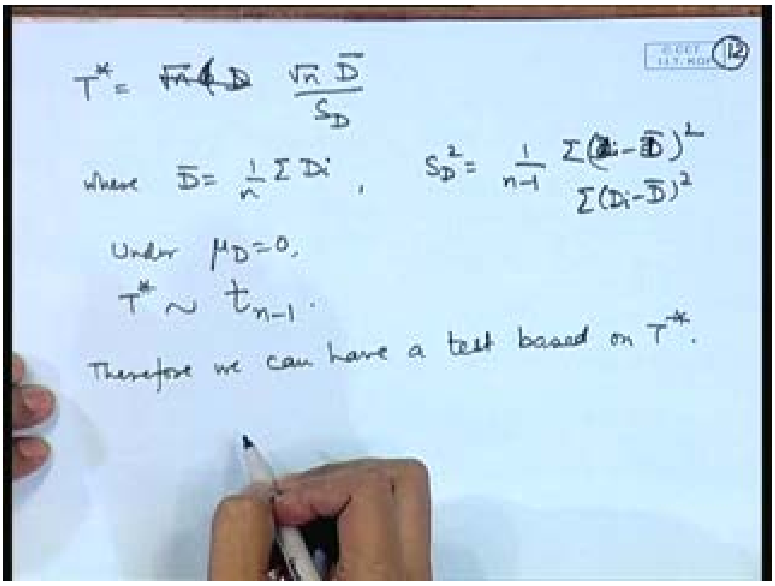


So, in this case, simplified procedure is proposed and we call it paired t test. So, sometimes the samples X and Y are not independent. So, we may consider the following set up. So, we may now consider that X1 Y 1, these are observations on the same entity, so, X1 Y 1, X 2 Y 2, X n Y n, this follows a bivariate normal distribution with mean say, mu1 mu2 and variances say sigma1 square sigma2 square and a certain correlation coefficient say rho here. So, in that case, and we have to test about the equality of mu1 and mu2, or mu1 less than or equal to mu2, etcetera.

Now, once again, you see, if we have a bivariate model, then if you consider the differences that is, Xis minus Yis, they again follow univariate normal distribution with means mu1 minus mu2 and variance will become sigma1 square plus sigma2 square minus twice rho sigma1 sigma2- we can call it say, so, we give a notation here Di and this I call muD and this is sigmaD square. So, the testing problem can be reduced to about muD, whether mu D is equal to 0 or muD is not equal to 0, or say H naught muD is equal to 0 against say, muD is greater than 0- of course, this is equivalent to less than or equal to 0 here- or muD is equal to 0 against muD less than 0- of course, this is again equivalent to greater than or equal to 0 also.

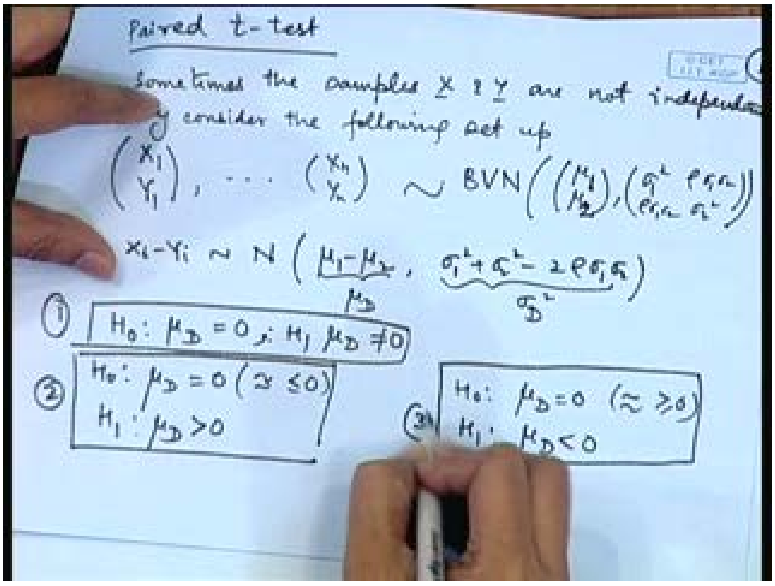
Because these hypothesis is now equivalent to for example, mu1 is equal to mu2, and this is equivalent to mu1 not equal to mu2, this is equivalent to mu1 less than or equal to mu2, this is equivalent to mu1 greater than mu2, etcetera. Now, this model again reduces to the model which we considered for one population that means, testing for the mean of a normal distribution, and we know that there is a T test for that.

(Refer Slide Time: 35:41)



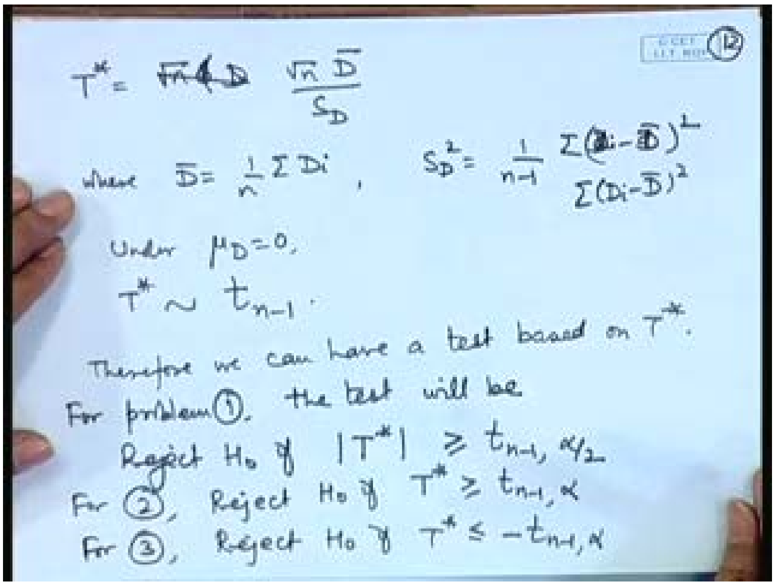
So, if we consider here the statistic defined by say T star is equal to root n D bar, so, let me write it, root n D bar divided bar SD where D bar is the mean of Dis, SD square is the sample variance of the Di minus D bar square. Then, one can see that under muD is equal to 0, the distribution of T star will be T on n minus 1 degrees of freedom- therefore, we can have a test based on T star.

(Refer Slide Time: 36:55)



So, for example, if I am considering, let me name this hypothesis here. Let me call this one as say hypothesis problem 1, this one I can consider as problem 2, this one I can consider as say, problem 3.

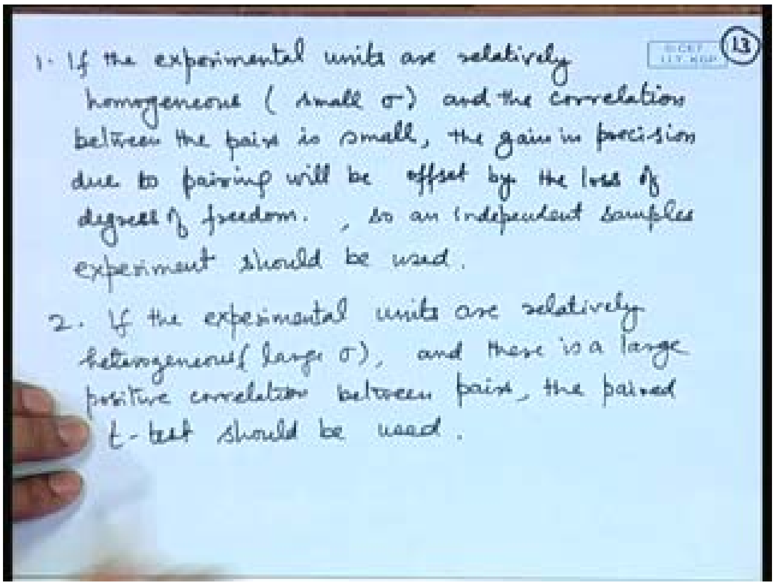
(Refer Slide Time: 37:08)



So, for hypothesis testing problem 1, the test will be reject H naught if modulus of T star is greater than or equal to tn minus 1 alpha by 2. For problem 2, it will be reject H naught if T star is greater than or equal to tn minus 1 alpha. For the testing problem 3, the test will be reject H naught if T star is less than or equal to minus tn minus 1 alpha. Once again, let me mention here that this equality or inequality in the less than or equal to greater than or equal to, it does not make any difference because we are dealing with the continuous random variables, so, the probability of equality is 0.

Now, the case that I have discussed here that we may have the dependence on the observations, but another problem that we have seen here that there is a loss of degrees of freedom here. So, if we are losing degrees of freedom here, then the power of the test is also reduced here. So, then, one has to make a very clear cut choice that in a given situation whether the data is independent or not. If the data is actually showing dependency that means, there is a high degree of correlation, then naturally the divisor becomes. For example, here if you are considering sigma D square, so, if rho is higher, then it is going to affect, because if they are independent, then the variances will be sigma1 square plus sigma2 square for Xi plus Yi or Xi minus Yi if you are taking them independent whereas, here you are making it smaller, so the estimate of that will also become smaller that means, the test is becoming more sensitive, because if this value is becoming smaller, then when you are taking the ratio, in the ratio you are dividing by smaller value, that is in this statistic, so, this value is becoming larger that means, there are more chances of rejection. If you are having more chances of rejection, then the probability of, that is the 1 minus probability of type 2 error that is, a power of the test will increase- so, the test will become slightly more powerful. On the other hand, you are losing certain degrees of freedom here. So, if the correlation is not much and still you are taking the dependence model, then you may have a loss here.

(Refer Slide Time: 40:23)

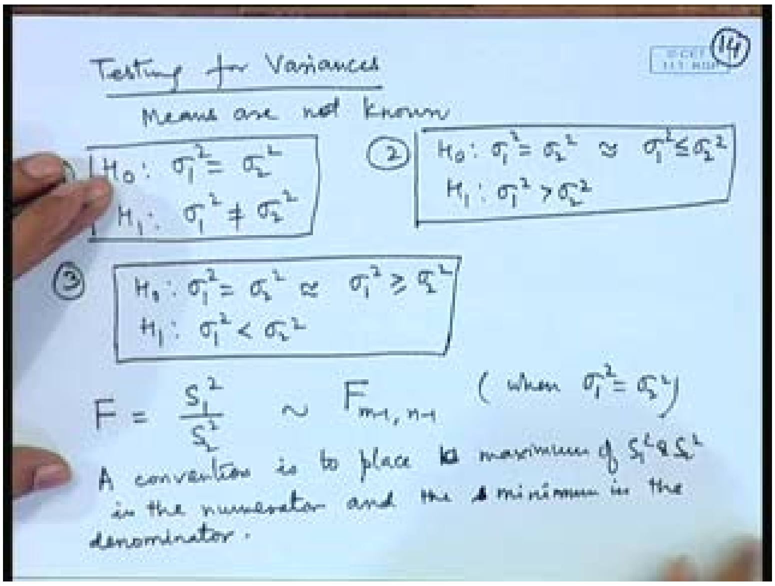


So, some general guidelines can be given here. Let me just write it here. If the experimental units are relatively homogeneous that means, variability is small and the correlation between the pairs is small and the correlation between the pairs is small, the gain in precision due to pairing will be offset by the loss of degrees of freedom, so an independent samples experiment should be used. On the other hand, if the experimental units are relatively heterogeneous that is, large sigma and there is a large positive correlation between pairs, the paired t test should be used.

So, these are some general guidelines, but of course, when a given problem is there, one has to see carefully whether the pairing is permissible or not. If it is permissible, then generally there is a better chance that you will have a better outcome, or you can say a better result if you use the pairing, or you can say paired t test.

Now, another important point is that when we are testing for the equality of the means we considered different cases and they were related to certain information about the variances that means, there is naturally question of checking the equality of the variances, because if the variances are unknown, but equal, you have another procedure, if you have them to be unequal, then you have another procedure- that means, one should carry out a test for the equality of variances before testing the equality of the means.

(Refer Slide Time: 43:55)



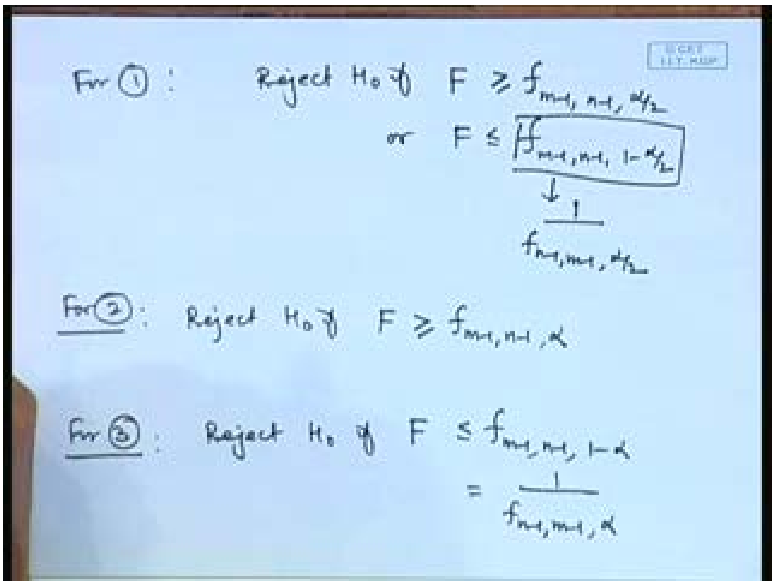
Now, we can produce a test for variances also- testing for variances. Of course, once again, we may have the case where the means are known or unknown, but it is, it will not make too much difference here in fact, we have seen in the case of testing for the normal variance that you have change in the single degrees of freedom there, so there is not too much sensitivity involved there, but in general as we know that the means may not be known here, since we are going to test for the equality of the means also.

So, let us consider the case means are not known. So, we have sigma1 square is equal to sigma2 square against say, H1, sigma1 square is not equal to sigma2 square. Once again, let me call this as hypothesis testing problem 1. Another problem would be where you have sigma1 square is equal to sigma2 square against sigma1 square greater than sigma2 square, and once again this H naught is equivalent to sigma1 square less than or equal to sigma2 square. Another case could be H naught sigma1 square is equal to sigma2 square, which is of course, equivalent to sigma1 square greater than or equal to sigma2 square against sigma1 square is less than sigma2 square.

So, let us give a test of hypothesis for all the three cases. We can consider a F test here. So, F is S1 square by S2 square. Of course, the convention is that the bigger value among S1 square and S2 square should be put in the numerator, the reason is being that if you are considering the F test, the upper 100 alpha percent points, they are bigger than one there. So, therefore, a more prudent thing would be to take the higher value that is, S2 square by S1 square in case S2 square is bigger than S1 square, or S1 square by S2 square if S1 square is bigger than S2 square. So this follows F distribution on m minus 1 n minus 1 degrees of freedom when sigma1 square is equal to sigma2 square.

So, we can give a test based on this. A convention is to place higher, or you can say maximum of S1 square and S2 square in the numerator and the minimum in the denominator.

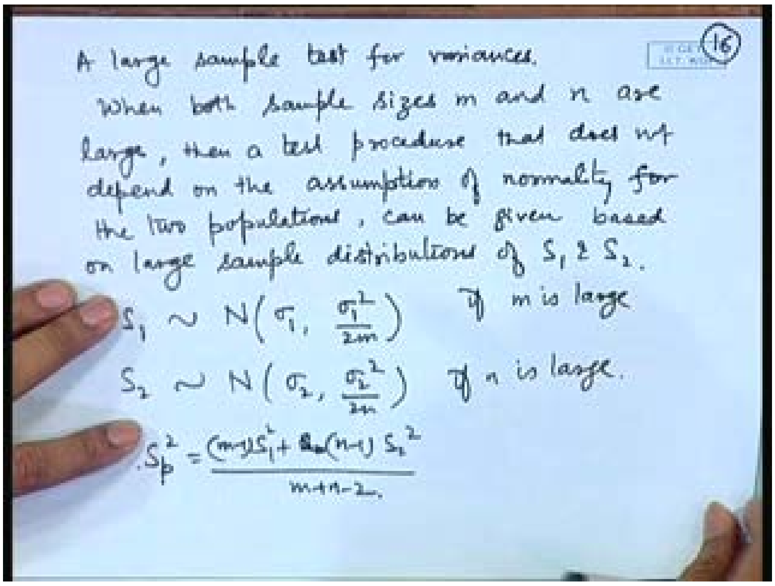
(Refer Slide Time: 47:23)



So, now, if we are making a test of the hypothesis based on this F distribution on m minus 1 n minus 1 degrees of freedom, then for 1, the test would be reject H naught if F is greater than or equal to f m minus 1 n minus 1 alpha by 2 or f less than or equal to f m minus 1 n minus 1 1 minus alpha by 2. Of course, in the f distribution because of reciprocal nature, this is equal to 1 by f n minus 1 m minus 1 alpha by 2. For the hypothesis 2, the rejection region will be reject H naught if F is greater than or equal to f m minus 1 n minus 1 alpha. For the hypothesis testing problem 3, that is sigma1 square less than sigma2 square is alternative, we will be rejecting H naught if F is less than or equal to f m minus 1 n minus 1 1 minus alpha, which is equal to 1 by f n minus 1 m minus 1 alpha.

Once again, in this situation also there may be a case when the normality assumption may not be satisfied. That means, the initial populations X1, the X1 1 sample Y1 sample, etcetera, they may not be from normal populations, in that case, we may go for large sample approximations for the S1 square and S2 square and we may use the following.

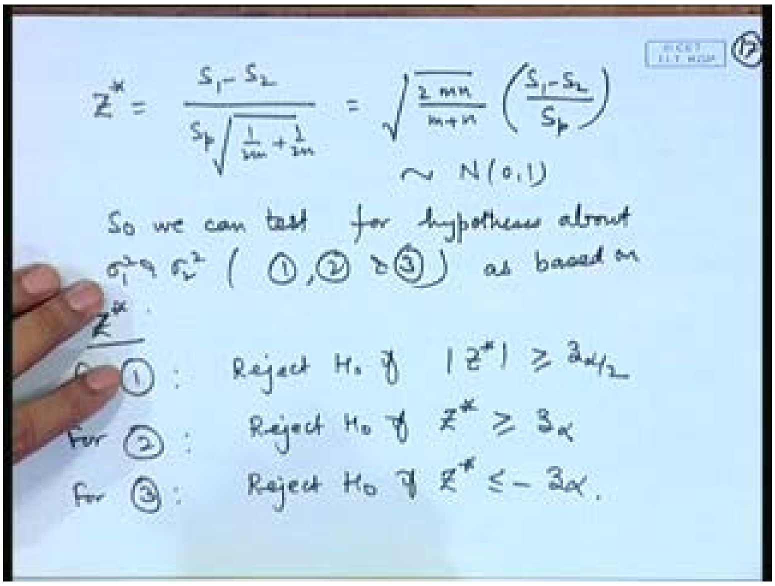
(Refer Slide Time: 49:17)



A large sample test for variances. When both sample sizes m and n are large then a test procedure that does not depend on the assumption of normality for the two populations can be given based on large sample distributions of S1 and S2. So, we can say that S1 follows normal distribution with mean sigma1 and variance sigma1 square by 2m if m is large, and S2 follows normal with mean sigma2 and variance sigma2 square by 2 n if n is large.

And once again, we can utilize the fact that S1 and S2 are independently distributed. We may also create a pooled here that is, Sp square, that is equal to S1 minus S2 under the assumption that sigma1 is equal to sigma2 we can consider like this, that is m minus 1 S1 square plus n minus 1 S2 square divided by m plus n minus 2.

(Refer Slide Time: 51:41)



And we can construct the test statistic say, Z star, that is equal to S1 minus S2 divided by Sp root 1 by 2 m plus 1 by 2 n, which is actually equal to root 2mn by m plus n S1 minus S2 divided by Sp- so, this is approximately normal 0, 1 because we are making the assumption that m and n are large. So, we can test for hypothesis about sigma1 square and sigma2 square that is, 1, 2 and 3, which I described earlier as based on Z star. So, for 1, you may say reject H naught if modulus of Z star is greater than or equal to Z alpha by 2. For the hypothesis testing problem 2, we will say reject H naught if Z star is greater than or equal to Z alpha. For the third problem, we will say reject H naught if Z star is less than or equal to minus Z alpha. So, this test can be used when the assumption of normality for the basic populations is not very strong and of course, the sample sizes are large.

In the next lecture, I will be discussing various problems where these tests can be utilized. We will also see a test for the proportions for the binomial populations when we have two binomial populations.