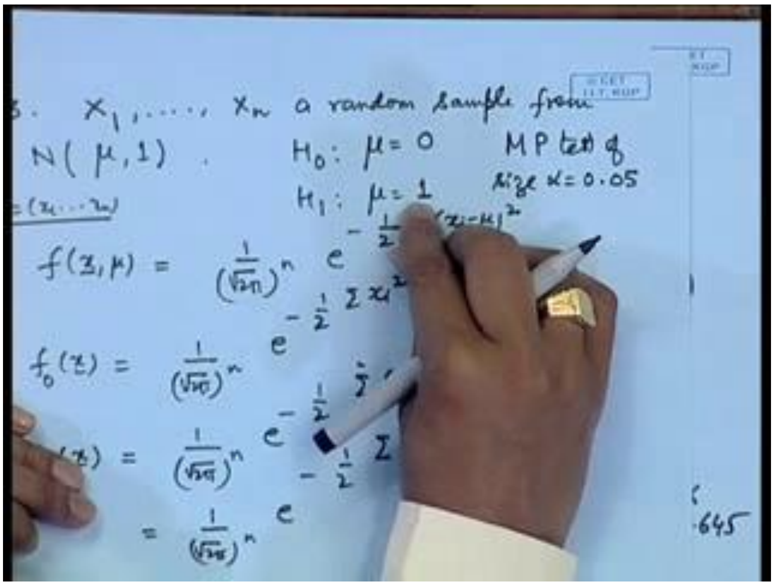
Probability and Statistics

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

Lecture - 69
Applications of N-P Lemma – II

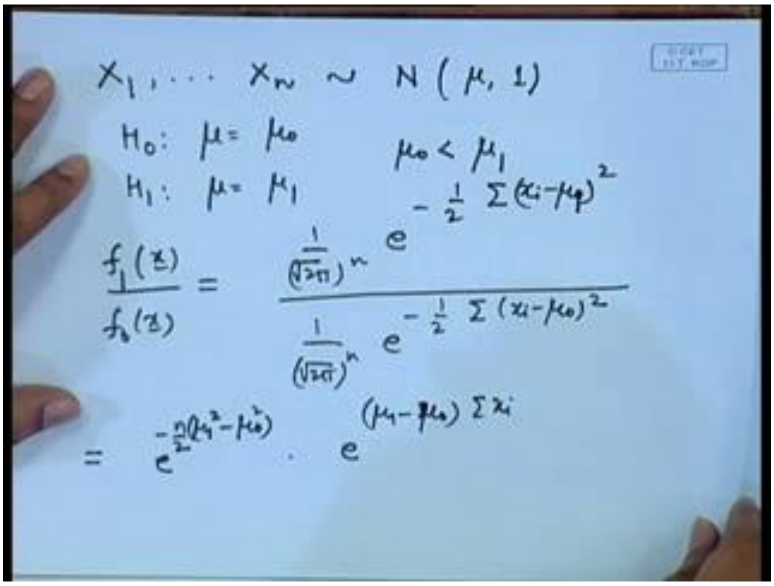
So, we may actually consider it in a slightly broader sense.

(Refer Slide Time: 00:25)



In place of mu is equal to 0 and mu is equal to 1 if we substitute say some values mu naught and mu 1 and then let us see the effect of this.

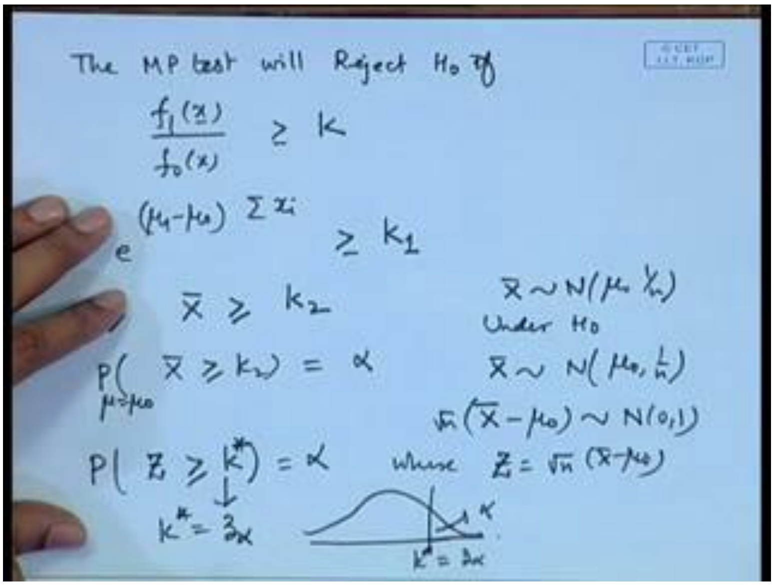
(Refer Slide Time: 00:36)



So, now let me generalize this problem X 1, X 2, X n follows normal mu 1. And we are testing the hypothesis whether mu is equal to mu naught against H 1 mu is equal to mu 1 where let me take mu naught to be less than mu 1. Now let us write down the density ratio that is f 1 x divided by f naught x.

So, it is 1 by root 2 pi to the power n e to the power minus 1 by 2 sigma xi minus mu naught square; sorry this one will be mu 1 square divided by 1 by root 2 pi to the power n e to the power minus 1 by 2 sigma xi minus mu naught square. Now you can see that these terms cancels out e to the power minus 1 by 2 if you expand this you get sigma xi square minus twice mu 1 sigma xi. So, after simplification this term becomes e to the power n mu 1 square minus mu naught square with a minus sign e to the power mu 1 minus mu naught sigma xi; remaining term gets canceled out here.

(Refer Slide Time: 02:28)



So now if we look at the most powerful test, this is reject H naught if f 1 x by f naught x is greater than or equal to k. So, if we utilize this here given mu 1 and mu naught whatever be the values this is some constant here. So, this region is reducing to e to the power mu 1 minus mu naught into sigma xi greater than or equal to say some k 1. So, if mu 1 is greater than mu naught then this region is equivalent to x bar greater than or equal to some k 2. Therefore, the test is to once again reject H naught for larger values of x bar.

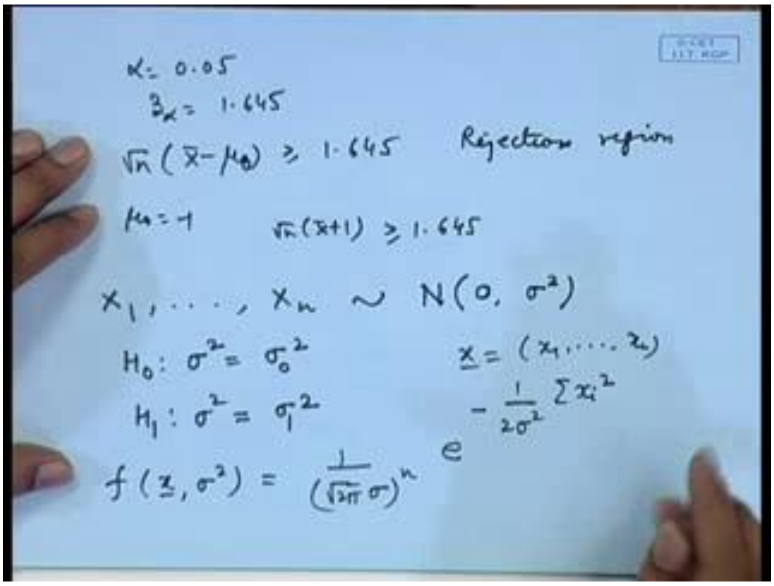
So, if you compare with the previous situation where I had taken mu naught to be 0 and mu 1 is equal to 1 then we were rejecting for larger value of x bar. So, as I mentioned here that the only deciding factor is the value of x bar, but we wanted to know the scale of x bar that on what scale we will consider x bar to be large what should be the small value that is decided by the probability of the type one error.

So, in the same way here you are seeing that if mu naught is less than mu 1 the reason is actually same, but how much it is same that will be dependent upon the probability of type one error. So, if we write here probability of x bar greater than or equal to k 2 then mu is equal to mu naught this is equal to alpha then we observe the distribution here. So, the distribution of x bar here is normal mu 1 by n. So, under H naught x bar follows normal mu naught 1 by n. So, we can do the calculations here by simplification x bar minus mu naught into root n that will follow normal 0 1 distribution.

So, under H naught this statement can be written to be equivalent to Z greater than or equal to some k is equal to alpha, in place of k let me write here k star where Z is defined to be root n x bar minus mu naught. So, this point k star becomes the upper hundred alpha percent point of the standard normal distribution; this is the point k star, this probability is alpha. Therefore, this k star is actually Z alpha point here.

So, as a practical example if we substitute different values here say mu is equal to minus 1 then there is X n region is changing the root n x bar plus 1 greater than or equal to Z alpha. We have seen the example of here.

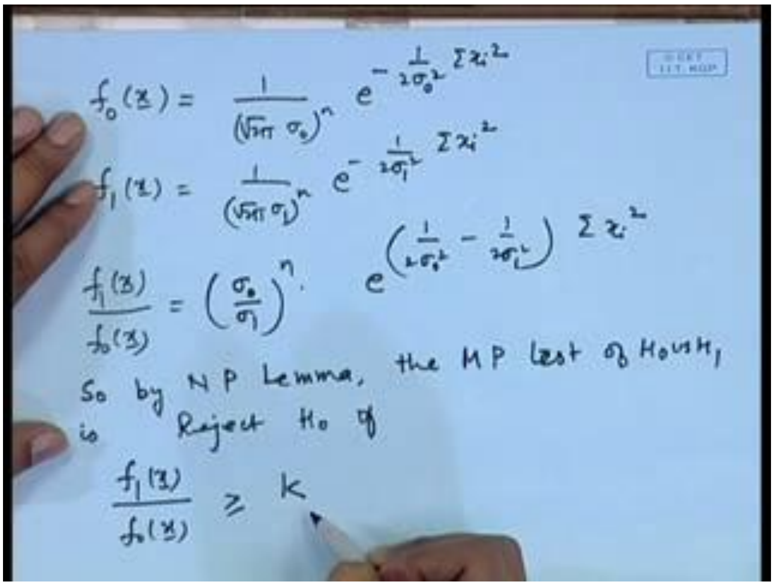
(Refer Slide Time: 06:01)



That if I am putting say alpha is equal to 0.05 then Z alpha is equal to 1.645. So, the test will become in that case root n x bar minus mu naught greater than or equal to 1.645, this is the rejection region. So, if I say take mu naught is equal to say minus 1 then this will become root n x bar plus 1 greater than or equal to 1.645. If we compare with the previous example where mu was 0 then it was root n x bar greater than or equal to 1.645. So, the magnitude of x bar which will be considered to be large depends upon the probability of the type one error. And that means, what is a value of the probability distribution point when mu is equal to mu naught.

A similar behavior is observed suppose we consider testing for the variance in normal distribution case. Let me take the case of say X 1, X 2, X n for convenience let me take the mean to be 0 and variance to be sigma square. And we are interested to make a test of hypothesis about say sigma square. Now once again let us write down the density function here; f x sigma square, so we need to write for the joint distribution of X 1, X 2, X n here so that is 1 by root 2 pi sigma to the power n e to the power minus 1 by 2 sigma square sigma xi square, since I have taken the mean of the normal distribution to be 0 so the joint distribution of X 1, X 2, X n turns out to be this one.

(Refer Slide Time: 08:09)

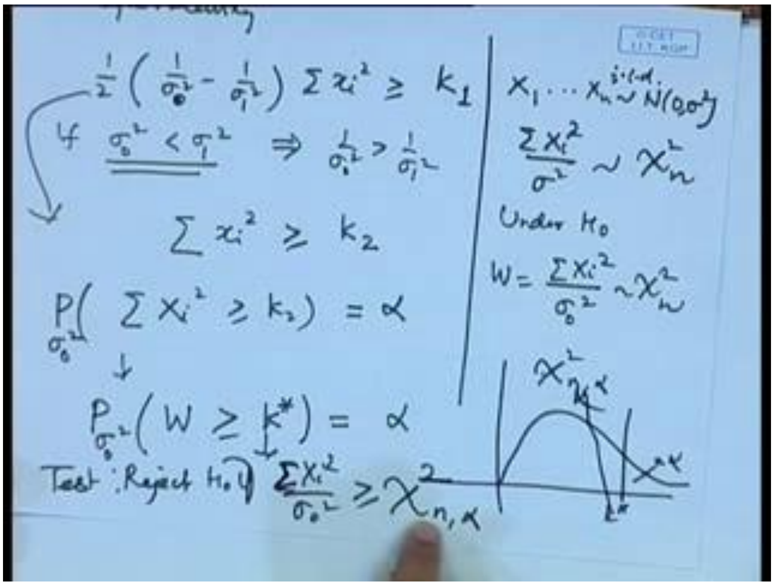


So, we write down this value corresponding to the null and the alternative hypothesis. So, when sigma square is equal to sigma naught square then this is becoming 1 by root 2 pi sigma naught to the power n e to the power minus 1 by 2 sigma naught square sigma xi square, and f 1 x in a similar way will become 1 by root 2 pi sigma 1 to the power n e to the power minus 1 by 2 sigma 1 square sigma xi square.

So, if you consider the ratio f 1 x by f naught x that is equal to. So, now this will become equal to sigma naught by sigma 1 to the power n e to the power 1 by twice sigma naught square minus 1 by 2 sigma 1 square sigma xi square. So, by Neyman-Pearson Lemma the most powerful test of H naught versus H 1 is reject H naught if f 1 x by f naught x is greater than or equal to k. Once again we notice here that this distribution of x is continuous distribution, so the distribution of the variables involved here for example; here sigma xi square is involved that is also continuous distribution. So, here without loss of generality we can write greater than or equal to, because the probability of the statement being equal to k that is f 1 by f naught equal to k that probability will be 0. So, this equality can be included here.

So now, if we look at the ratio here this is greater than or equal to k then this will reduce to because sigma naught and sigma 1 are the known constants, this condition is gone.

(Refer Slide Time: 10:20)



And if you take the logarithm then we get it as this condition is equivalent to 1 by 2 1 by sigma naught square minus 1 by sigma 1 square sigma xi square greater than or equal to some k 1.

Now, once again the relative position of sigma naught square and sigma 1 square is playing a role here. Suppose I take sigma naught square to be less than sigma 1 square, then this will be equivalent to 1 by sigma naught square greater than 1 by sigma 1 square. Therefore, the region will be equivalent to sigma xi square greater than or equal to say k 2. Where, this k 2 as to be chosen in such a way that sigma xi square greater than or equal to k 2 as a probability equal to alpha when sigma square is equal to sigma naught square. So, the condition becomes sigma xi square greater than or equal to k 2 when sigma square is equal to sigma naught square is equal to alpha.

So, in order to find out the value of k 2 we need to look at the distribution of sigma xi square when sigma square is equal to sigma naught square. So, we look at our statement here X 1, X 2, X n they followed normal 0 sigma square, and therefore if we consider sigma xi square by sigma square that will follow chi square distribution on n degrees of freedom, because we are considering this to be random sample so these are independent and identically distributed random variables. So under H naught, let me call it w that is sigma xi square by sigma naught square that follows chi square distribution on n degrees of freedom.

So, we can write down this statement as w greater than or equal to some k star is equal to alpha. Since w is following chi square and distribution the point k star becomes upper hundred alpha percent point this point is k star and this is alpha so this point is nothing but chi square n alpha. That means, the test is reject H naught if sigma xi square by sigma naught square is greater than or equal to chi square n alpha.

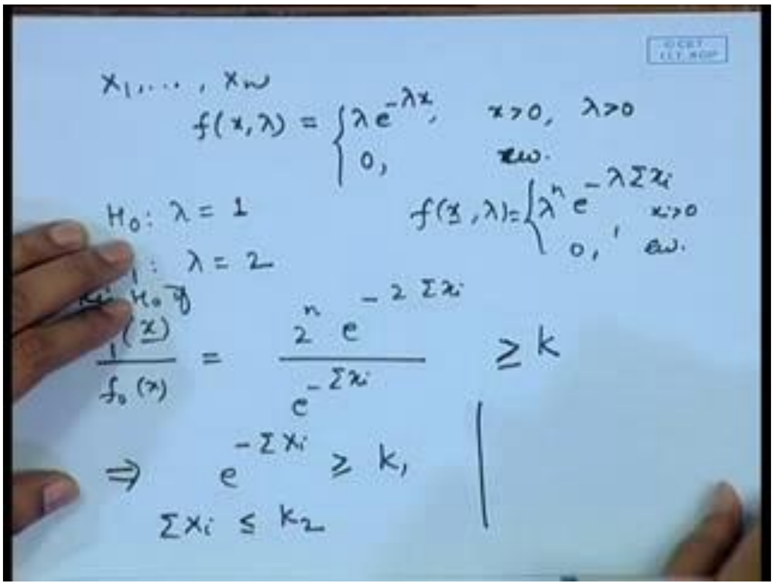
Let us interpret this test here: we wanted to test whether the variance of a normal distribution is less or more, because sigma naught square we took to be less than sigma 1 square. Now when mean is taken to be 0 sigma xi square by n is a estimator we have actually calculated the maximum likelihood estimator, so that is an estimator for sigma square. So, as a Neyman you will based your decision on the value of sigma xi square by n. That means, for a smaller value of sigma xi square by n we will tend to accept H naught and for a larger value of this we will tend to accept H 1; that is rejecting H naught.

So, now how much value of sigma xi square by n should be considered a small or large that is decided by the probability of the type one error. So, if the probability of the type one error is alpha the relative position of sigma xi square is decided by chi square and alpha, and of course the value sigma naught square also plays a role. Because if sigma naught square is much smaller compare to sigma 1 square then that value will play a role. So, the text here you can see the relative position is dependent upon the value of the parameter in the null hypothesis. And for the power function it is reverse, we are making use of the alternative hypothesis value; that is a power will increase or decrease according to the value of the parameter in the alternative hypothesis here.

Now we have seen here application of the Neyman-Pearson lemma to some continuous distribution especially normal distribution. We have seen the application to some discrete distribution such as binomial distribution. So, this is a very general result, because I can consider any distribution and if we have a simple versus simple hypothesis. In fact, it is not even necessary that we have a same form of the distribution as we have seen in the first example, where under the null hypothesis we had a uniform distribution and under the alternative hypothesis we had another distribution which was having the density 2 x.

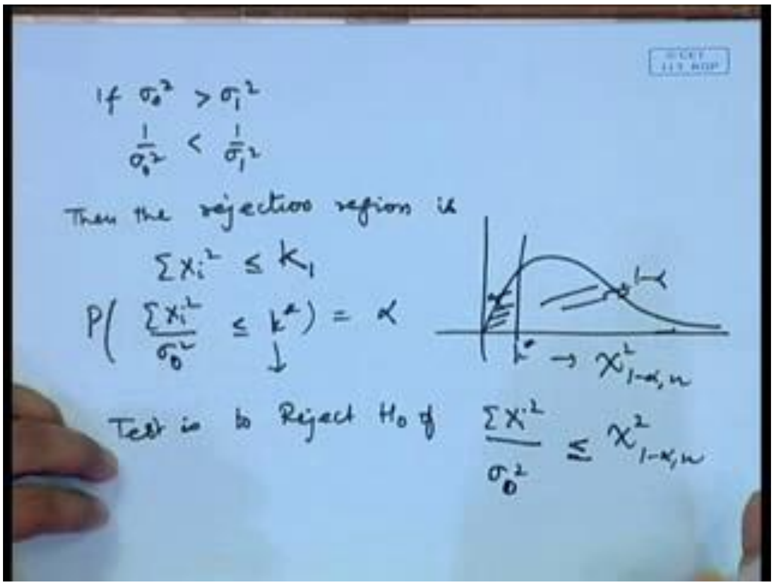
In general we are able to test whether we have this probability distribution which is completely specified or another one which is again completely specified by making use of the Neyman-Pearson fundamental lemma. Another important point that you may notice here is that in most of the situations the test function is coming in terms of the statistic which is actually a sufficient statistic. You can also say that it is coming in the terms of maximum likelihood estimator as we have seen in the example of the normal distribution; where for a mu you are in terms of x bar and for sigma square we when we are doing the test then it is coming in terms of sigma xi square.

(Refer Slide Time: 16:43)



Naturally we can check for certain other distributions also such as say let X 1, X 2, X n this follow an exponential distribution with some parameter say lambda. Before I discuss this example let me take the other case also where naught square may be greater than sigma 1 square, then let us see how we are distinguishing.

(Refer Slide Time: 17:23)



If sigma naught square was greater than sigma 1 square then 1 by sigma naught square is becoming less than 1 by sigma 1 square. Now, if you see here this value will become negative, therefore if I divide by that value the region is becoming reverse. So, we are getting the region as then the rejection region is sigma xi square is less than or equal to k say k 1 or let me put in another way in place of. So, once again we will have probability of sigma xi square by sigma naught square less than or equal to some k star is equal to alpha.

So, here it is turning out to be the left hand point we are saying this value is alpha; that means this upper value is 1 minus alpha. So, this k star in this case becomes chi square 1 minus alpha n. That means, the test is to reject H naught if sigma xi square by sigma naught square is less than or equal to chi square 1 minus alpha n.

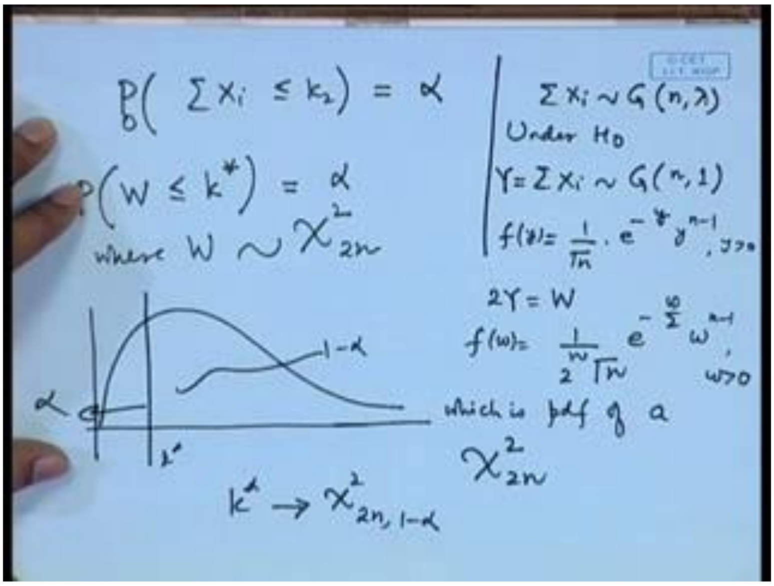
So, you can see here that the region as got reverse, why because now the null hypothesis supports a larger value of sigma square; that is sigma naught square I have taken to be bigger than sigma 1 square. So, a smaller value of sigma xi square will be favor of the alternative hypothesis which is against the previous case where higher values are supporting the alternative hypothesis. And once again that on the relative scale that how much value of sigma xi square will be considered very larger or a smaller that is determined by the probability of the type one error and the value of the parameter in the null hypothesis here.

Now, let me consider the example which I mentioned earlier that is of a exponential distribution. And we may like to test say lambda is equal to say 1 against say lambda is equal to 2. Now the question is that when we are discussing distributions which are different than the normal distribution etcetera we may get a statistic where the distribution of the statistic which is appearing in the test function may not be simple. Then we may have to use certain transformations and get the distribution of that so that one may make use of the tables of the standard distributions to find out the exact test of the hypothesis.

So, in this particular case for example, let us write down the join distribution; so f 1 x by f naught x; so here, the joint distribution f x lambda that becomes lambda to the power n e to the power minus lambda sigma xi when all the xi's are positive it is 0 elsewhere. So, f 1 will be correspond to the value of lambda is equal to 2 then this becomes 2 to the power n e to the power minus twice sigma xi divided by; when I put f naught that is corresponding to lambda is equal to 1 I will get e to the power minus sigma xi.

So, we are saying the test is reject H naught if this is greater than k. Once again we have loss of generality we may include equality here or we may delete equality, because the distribution of xi's are continuous therefore the distribution of sigma xi will also be continuous. In fact, we know the distribution of sigma xi here. Firstly, let us simply this. So, this region is equivalent to if we take this in the numerator it is reducing to sigma xi greater than or equal to some k 1, because this coefficient we can remove. And when we take logarithm this is reducing to sigma xi less than or equal to some k.

(Refer Slide Time: 22:09)

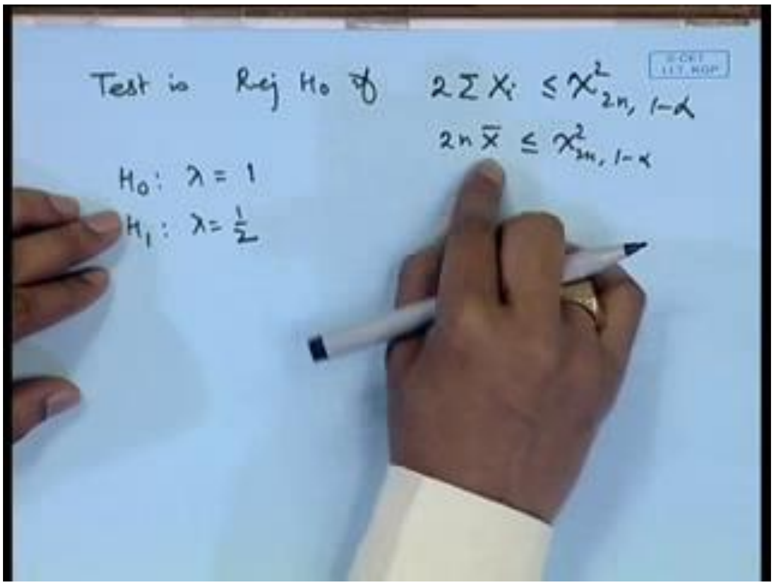


Now, we have to find out the value of k 2 such that the probability of type one error that is sigma xi less than or equal to k 2 is equal to alpha. So, we make use of the distribution theory; here as I was mentioning sigma xi will follow gamma distribution with parameters n and lambda by the additive property of the exponential distribution. The sum of independent exponential variables is a gamma variable. So, under H naught sigma xi will follow gamma distribution on n and 1 degree of freedom.

Now, what is this distribution? If we write down let me denote it by say y then the density of y is 1 by gamma n e to the power minus y y to the power n minus 1. So, if we consider say 2 y is equal to say w then the distribution of w is equal to 1 by 2 to the power n gamma n e to the power minus w by 2 w to the power n minus 1, for w greater than 0; which is nothing but probability density function of a chi square distribution on 2 n degrees of freedom.

So, now under H naught we can write this as probability of w less than or equal to some k star that is equal to alpha. So, it is w is having a chi square distribution on 2 n degrees of freedom, so this point that you are having here this is such that this probability is equal to alpha and this probability is 1 minus alpha. So, k star naturally turns out to be chi square 2 n 1 minus alpha.

(Refer Slide Time: 24:42)



So, the test function then becomes reject H naught if twice sigma xi is less than or equal to chi square 2 n 1 minus alpha. Let us again analyze this statement in a practical sense.

Here lambda is the rate of the Poisson process, so basically the mean was 1 by lambda. So, you want to test whether the mean is less or more. So in fact, the alternative hypothesis is having a higher value, because mean is 1 reciprocal of that. So, this is actually rate, so rate is less or more. Now for the rate for the average the variable or you can say the statistic would have been x bar which is of course proportional to summation of the values here. So, you may even write it in this particular form, this is actually equal to twice in x bar.

So, natural thing would be to go in favor of the null hypothesis if x bar is smaller, because if rate is larger if rate is smaller this is corresponding to the mean to be a smaller. So, the mean is represented or you can say estimated by the sample mean. So, for the sample value of the sample mean we will tend to favor H naught, whereas for the; I am sorry I just made the reverse statement. Here the null hypothesis is corresponding to lambda is equal to 1 against the alternative hypothesis lambda is equal to 2, see If you are considering say mean then 1 by lambda is 1 and 1 by lambda is equal to half here.

That means, under the null hypothesis the mean is a smaller, sorry mean is larger and in the alternative hypothesis the mean is smaller. That means, when we are using the sample mean as an estimate of that for by smaller value of the sample mean we will tend to favor the alternative hypothesis. And for the larger value we will tend to favor the null hypothesis. And the relative significance of how much is larger or how much is bigger is dependent upon the probability of the type one error, as well as the value of lambda is equal to 1 and lambda lambda is equal to 1 has been utilized here

On the other hand, if we had say H naught lambda is equal to 1 against say lambda is equal to half. Suppose just I made the change here then what will happen, in the case of the null hypothesis you will have a same value whereas in the for the alternative hypothesis this will become half and this will become e to the power minus sigma xi by 2. So, in that case in the numerator we will get e to the power sigma xi by 2 with give a positive sign. And then the test function will become sigma xi is greater than or equal to something rather than less than something.

So, when we analyze this we get here that test would be to reject for larger value of x bar, which is natural because when I say lambda is equal to half. That means, I am saying 1 by lambda is equal to 2 which is bigger than 1 by lambda is equal to 1 here. So, you can also see that in the Neyman-Pearson Lemma it the test which we are obtaining it by using the theory of most powerful test they are conforming to a Neyman approach or you can say likelihood approach for testing the hypothesis.

In the next lecture I will be discussing in the more detail how to find out the test for the composite hypothesis.