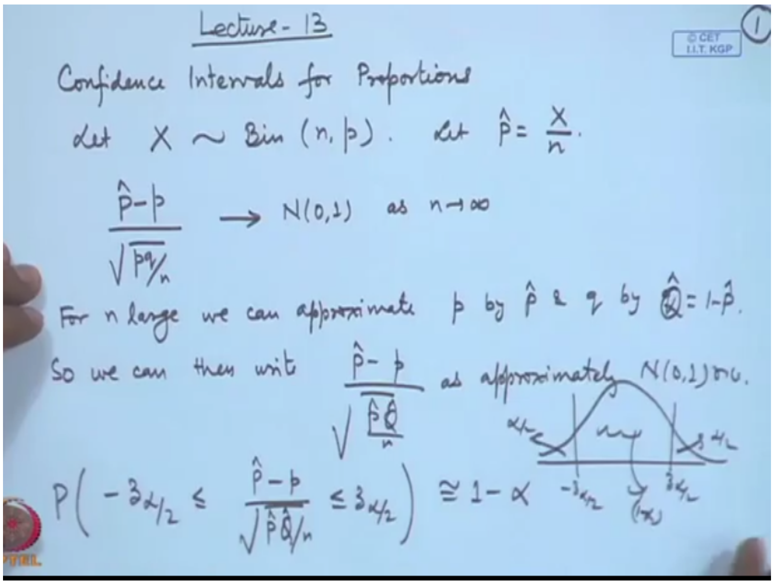
Statistical Methods for Scientists and Engineers
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology - Kharagpur

Lecture - 13
Parametric Methods - V

In the last lecture, I have introduced the concept of interval estimation and I discussed 1 method of constructing the confidence intervals with a given confidence coefficient. This method is the method of pivoting and we constructed the confidence intervals for parameters of normal populations when we have 1 sample or 2 sample problems. Today, I will also discuss briefly in the confidence intervals for proportions.

That means we are dealing with a binomial problem, for example we may have people who favor a certain proposition by the government, people who can be categorized as 1 type in a population. So if we are doing the sampling from there then to construct the confidence intervals we can use the binomial approximation to the normal distribution.

(Refer Slide Time: 01:18)

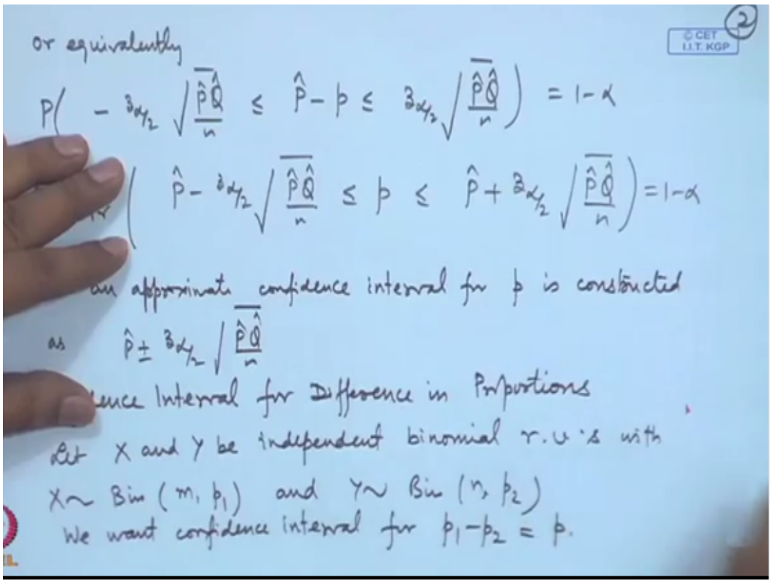


So let us consider confidence intervals for proportions. So typically we will have the data like this that we have a sample of n observations and out of that we have X number of successes. So let us define say the sample proportion as X/n. We want to construct the confidence interval for the parameter p that is the proportion of successes in a binomial population. So we can consider say P hat-small p/square root pq/n.

This converges to normal 0, 1 as n tends to infinity. This result is known. Now for n large, we can approximate p/P hat and q/Q hat=1-P hat. So we can then write P hat-p/square root of P hat Q hat/n as approximately normal 0,1 random variable. So we can use the pivoting method by considering the interval from –z alpha/2 to +z alpha/2=1-alpha because we are considering the 2 points on the standard normal curve.

This is z alpha/2 that is this probability is alpha/2 and if this probability is alpha/2 then this is –z alpha/2 so this in between probability is 1-alpha. So the probability of P hat-p/square root of P hat Q hat/n this is approximately 1-alpha. So we can construct the confidence interval from here. We can adjust the terms.

(Refer Slide Time: 04:02)



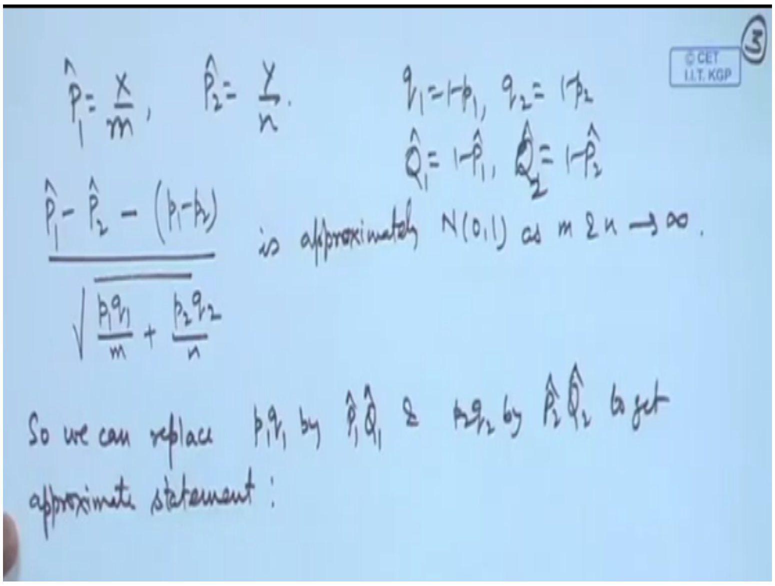
So this is equivalent to we can write –z alpha/2 square root P hat Q hat/n <= P hat-p <= z alpha/2 square root P hat Q hat/n=1-alpha or probability of P hat-z alpha/2 square root P hat Q hat/n <= p <= P hat+z alpha/2 square root P hat Q hat/n=1-alpha. So we have the confidence interval for p here.

That is from P hat-z alpha/2 square root P hat Q hat/n 2 P hat+z alpha/2 square root P hat Q hat/n. So here P hat is the sample proportion X/n. So this is an approximate so an approximate confidence interval for p is constructed. We may even consider comparing 2 binomial proportions. For example, it could be like proportion of the people who drive a certain vehicle in city A and proportion of the people who drive a certain vehicle in city B.

So the proportions may be different p1 and p2 and we may want to have a confidence interval for the difference to have an estimate whether 1 of them is less than the other or equal. So we may consider confidence interval for difference in proportions. Let us consider say let X and Y be independent binomial random variables with X following say binomial m, p1 and Y following say binomial n, p2.

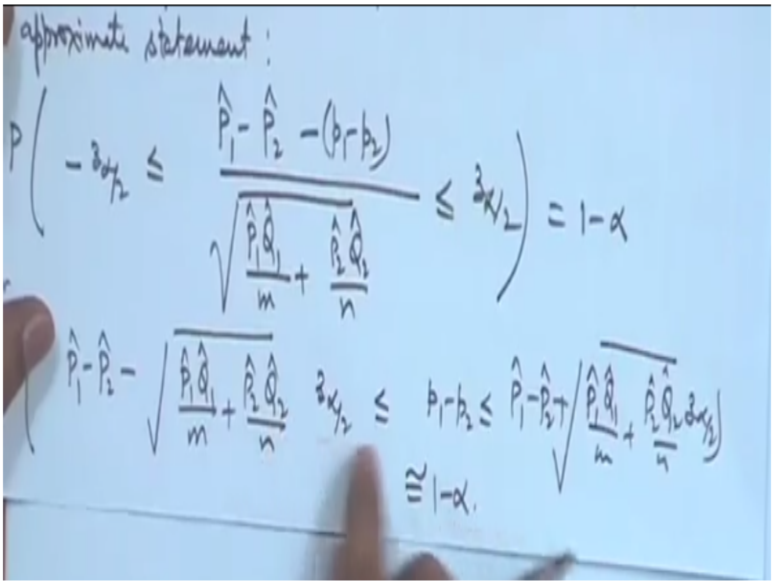
So here obviously m and n are known. We want confidence interval for p1-p2, let us say it is equal to p. Once at the end, we will make use of the approximation of binomial distribution to the normal.

(Refer Slide Time: 07:27)



So if we consider say P1 hat=say X/m, P2 hat=say Y/n. Then P1 hat-P2 hat-p1-p2/square root p1q1/m+p2q2/n where here I am using q1=1-p1 and q2=1-p2 and Q1 hat=1-P1 hat and capital Q2 hat=1-P2 hat. So this is approximately normal 0, 1 as m and n tend to infinity. So then we can write we can replace p1q1/P1 hat Q1 hat and p2q2/P2 hat Q2 hat to get approximate statement of the following nature.

(Refer Slide Time: 08:56)

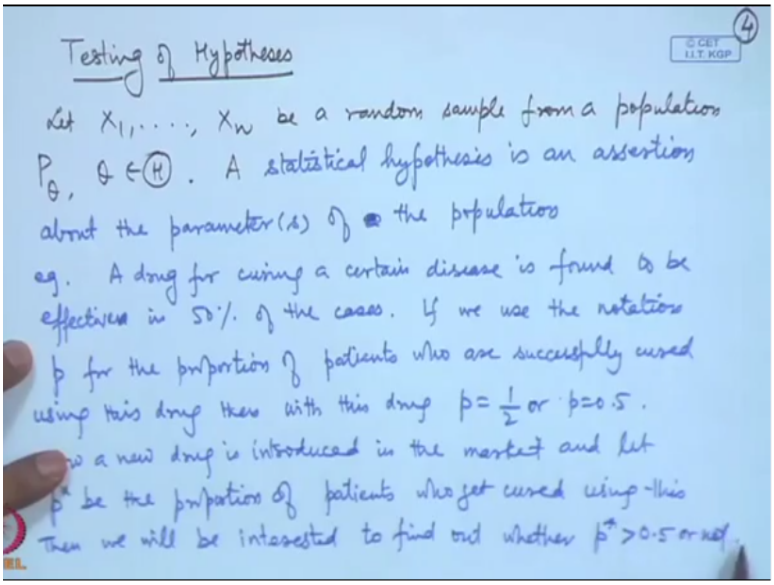


That is probability of –z alpha/2 <= P1 hat-P2 hat-p1-p2/square root of P1 hat Q1 hat/m+P2 hat Q2 hat/n. This is <= z alpha/2=1-alpha. So once again as before we can simplify, so we can write probability of P1 hat-P2 hat-square root P1 hat Q1 hat/m+P2 hat Q2 hat/n z alpha/2 <= p1-p2 <= P1 hat-P2 hat+P1 hat Q1 hat/n+P2 hat Q2 hat/n z alpha/2=1-alpha.

So we have an approximate confidence interval for p1-p2 of this form that is P1 hat-P2 hat+- square root P1 hat Q1 hat/m+P2 hat Q2 hat/n z alpha/2 where again z alpha/2 is the point on the normal distribution curve. This method of pivoting as I have explained can be used for various distributions whenever we are able to find out the pivoting quantity. Usually as we have seen it can be dependent upon the sufficient statistics.

And it is also coming from the theory of Neyman–Pearson's based test so now I will move over to the concept of the testing of hypothesis. Let us look at the basic notation and terminology for the problem of testing of hypothesis.

(Refer Slide Time: 11:04)



Let me introduce the problem first. So I have mentioned to you the problem of statistical inference that is we are considering certain population and we are looking at its characteristics. So for example we may be looking at the average heights of the say adult males in an ethnic group. We may be considering say average precision or the precision of a measuring instrument, measuring device which is used for measuring something.

We may be considering the amount of symmetry or asymmetry present in curve. We may be interested in estimating the average life of an electronic component and so on. Now in these we are making that we are having no prior knowledge about the parameter so we consider estimation, but there could be another type of thing. For example, we have a certain brand for a particular item.

Now a new brand of that item has been introduced in the market. Naturally, the manufacturer or the shopkeeper or the customer will be interested to know whether the average longevity or the average life will be more than the previous brand. Suppose there is a drug which is being used for curing a certain disease. Now a R&D division of a drug company, it introduces a new drug in the market.

It finds out or it invents a new drug. Now certainly everybody will be interested to know whether the new drug is more effective in curing the same disease than the previous. They may be looking at its efficiency in the terms of less time taken the proportion of people who are getting cured that could be more or the average cost of the medicine and so on. There can be several factors that can be used to test.

That means here we may have some information about the parameter, but we want to test. So this is called the problem of testing of hypothesis. We can roughly say that it is a statement about since we are dealing with the parametric methods we can say it is some statement about the parameters of a population. In general, a hypothesis would be any statement about the probability distribution.

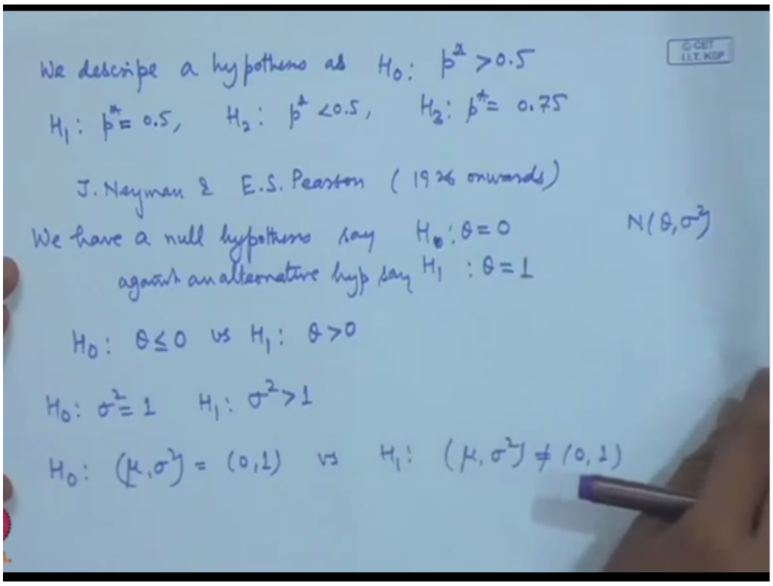
For example, you may even say that okay we want to test whether the data is coming from a normal population or the data is coming from a gamma population that could be a more general statement of the testing of hypothesis problem, but in the beginning we will restrict attention to the parametric methods that means the population is identified but we want to test something about the parameters values, whether the values equal to something or it is less than something and so on.

So we pose the problem in the following fashion. So let X1, X2, Xn. At our disposal, we have a random sample. Let X1, X2, Xn be a random sample from a population say p theta, theta belongs to parameter space theta. This theta could be a scalar or a vector. A statistical hypothesis is an assertion about the parameter of the population. So for example, a drug for curing a certain disease is found to be effective in say in 50% of the cases.

So if we use the notation say p for the proportion of patients who are successfully cured using this drug then with this drug p=1/2 or p=0.5. Now a new drug is introduced and let p star be the proportion of patients who get cured using this. Then we will be interested to find out whether p star>0.5 or not. So I have stated the problem in a very simple terms that we want to make some statement about the parameter of a population.

So here it could be like you take observation that means you can see that a sample of patients, out of that you find out how many get successfully cured and not and based on that you will conduct a statistical procedure. So let us discuss this. So firstly we will try to write down a hypothesis in this fashion.

(Refer Slide Time: 18:07)



We write a hypothesis like this. We describe a hypothesis as say H0 p star >0.5 or some we may say H1 p star=0.5 or H2 say p star<0.5 or say H3 p star=0.75 and so on. These are various statements. In each of them we are actually identifying the value of the parameter. In some cases, we are telling a range, in some cases we are exactly specifying. Now in general, hypothesis testing problems the common formulation that we give we firstly have a statement.

For example, we may like to say p star=0.5 or p star>0.5 then if we make a statement this is called a null hypothesis and then we test against another one so that is called an alternative hypothesis. Now this type of formulation for testing of hypothesis problems was developed by J. Neyman and E.S. Pearson in 1926 onwards in a series of papers where they developed this theory.

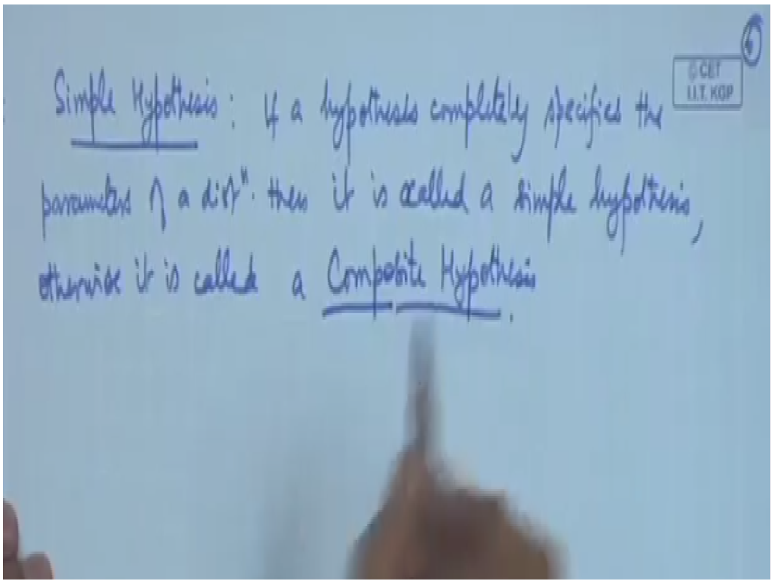
In this formulation, we have a null hypothesis say H0 so let us say if we are considering say normal distribution with parameter theta and say variant sigma square. We may like to test whether theta=0 against an alternative hypothesis say H1 theta=1. We may write in different ways also like H0 theta is <= 0 against say H1 theta is > 0.

We may like to write H0 sigma square=1 against say H1 sigma square is > 1. We may like to write H0 mu sigma square=0,1 versus H1 mu sigma square is != 0,1 and so on. So there can be various hypothesis, which may be required to be tested. Now we make a simple classification here.

When the value of the parameter specifies the distribution itself for example here in this binomial testing problem if we say p star=0.5 then the distribution is completely specified. This is called simple hypothesis and when we say p star<0.5 etc, then the distribution is not completely specified. This is known as a composite hypothesis. For example, if I write mu sigma square=0,1 then this is a simple hypothesis.

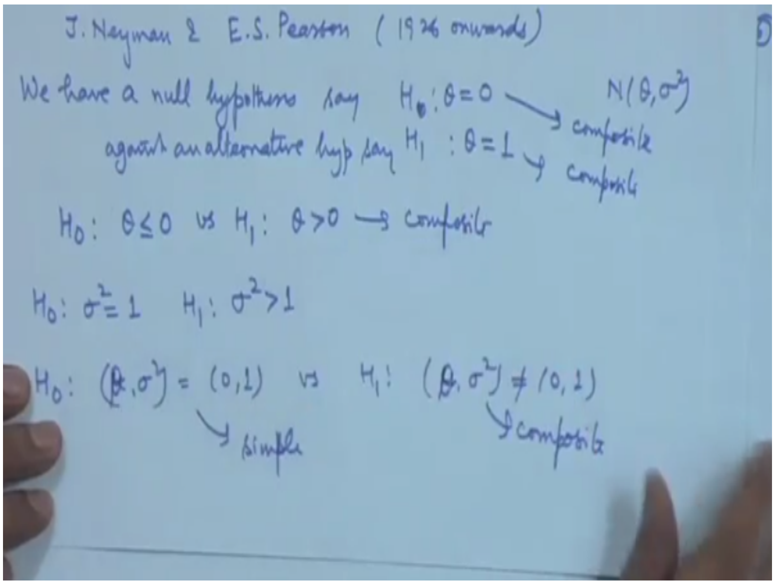
But if I say theta=0 then this is not a specified sigma square. So this is the composite hypothesis. So we have the concept of a simple hypothesis.

(Refer Slide Time: 21:35)



If a hypothesis completely specifies the parameters of a distribution then it is called a simple hypothesis, otherwise it is called a composite hypothesis.

(Refer Slide Time: 22:35)



So for example this is a composite hypothesis, this is a simple hypothesis. This hypothesis is composite because this does not specify sigma. This is composite. This is simple. Sorry this is theta here. This is composite, these are all composite. Now a statistician based on a sample will like to test the hypothesis. That mean he will give a procedure and he will decide. That procedure will try to make a decision in favor of a certain hypothesis.

For example, we may say suppose we consider a sample of 100 patients, we find that nearly 75% of the patient that is 75 patients get cured from the new drug. Then certainly we may tend to believe that p star>0.5. On the other hand, we may find that only 25 out of 100 get cured then we may say p star<0.5. Now this is something like you can say Lehman's kind of thinking that we can certainly say that if out of 100 75 get cured then it is too large than 50%.

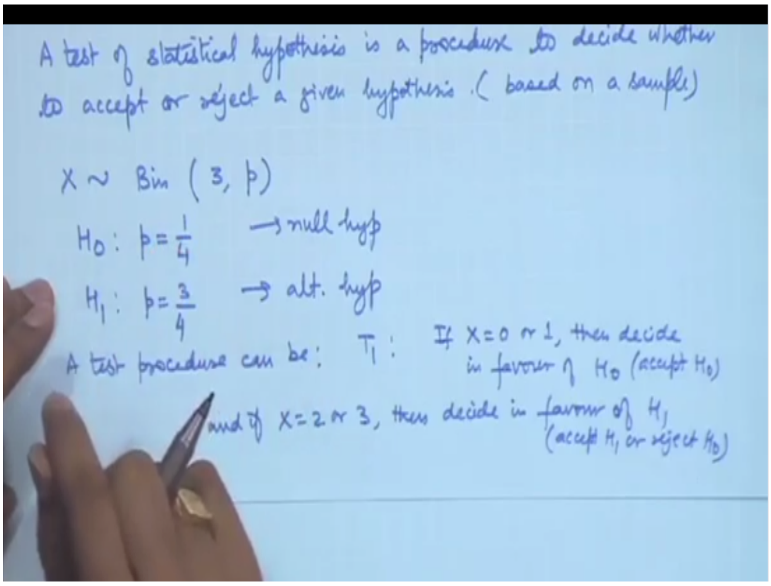
And therefore we may tend to believe that p star>0.5, but what happens suppose it is in the sampling that we have done. It turns out that out of 100 say 57 patients get cured successfully, then would we still be in favor of the statement p star>0.5 with the same convincing argument than the previous one? Can we say that it is significantly higher? The effectiveness is significantly more than p star=0.5.

Now that is the question that a statistician would like to answer in a more effective fashion. Similarly, if we are considering say the hypothesis theta=0 and theta=1, now if we consider a random sample X1, X2, Xn from the normal distribution, we may consider X bar as an estimate of theta and then you may say that okay if X bar is 0 then accept H0 and if X bar=1 then accept H1.

Now a thing is that if we are considering the sampling from the normal population then X bar is also a normal distribution with mean theta and variant sigma is square/n. So it is a continuous distribution so the probability that X bar is 0 or the probability that X bar is 1 both are equal to 0. Therefore, it does not make sense to give a test of this type and not only that. See what happens if X bar is say equal to -1?

What happens if X bar=say 1/2 or what happens if x bar=2? Therefore, in place of having a point test we may have to give a range so that we can significantly differentiate between the 2 hypothesis H0 and H1.

(Refer Slide Time: 25:58)



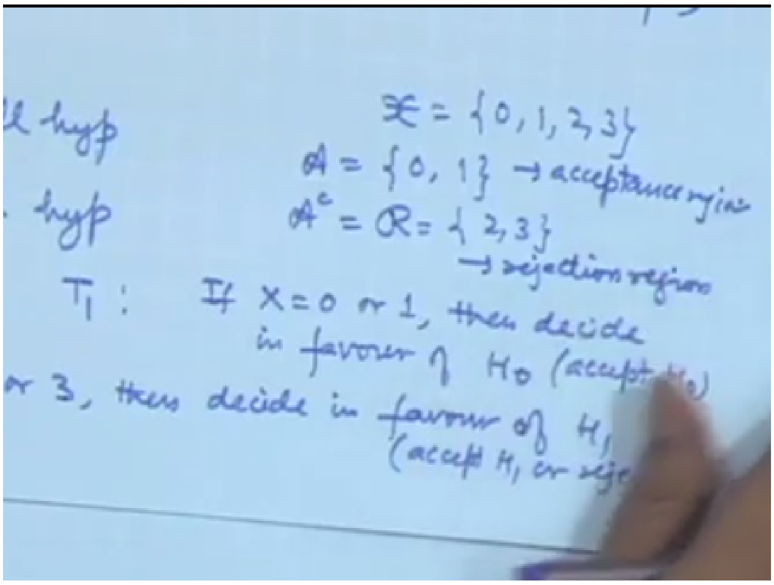
So we can say that a test of statistical hypothesis is a procedure to decide whether to accept or reject a given hypothesis. Now let us consider say and the decision will be based on sampling scheme based on sample, so let us take an example say X follows binomial say 3, p and our hypothesis is whether p=say 1/4 or H1 p=3/4. So this is that means we have considered a sample based on 3 observations out of which we say that X is the number of successes.

Now a Lehman's procedure could be that we may consider a test procedure can be let us call it T1 procedure that if X=0 or 1 then decide in favor of H0 and if X=2 or 3 then decide in favor of H1. So now you can see that this procedure is a heuristic procedure what we are saying is that if X=0 or 1 then it means that number of the proportion of the successes is smaller.

And therefore we may say that the probability of success should be smaller and therefore we go in favor of the hypothesis p=1/4. On the other hand, if out of 3 tosses or out of 3 trials you get 2 or 3 successes then you may say that the probability of success should be high and you feel that probability p=3/4 must be the correct statement and therefore we decide in favor of H1.

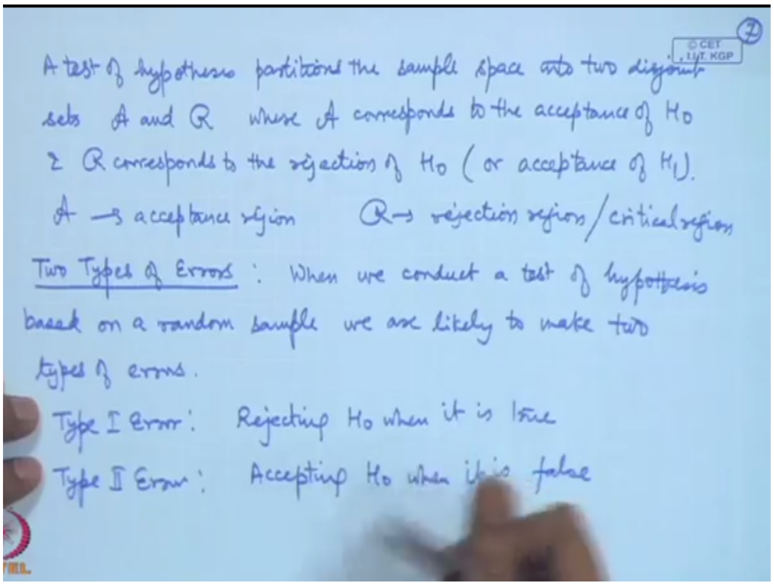
So we say we give a statement accept H0 and here we say accept H1 or we can say reject H0. Since in the original problem we write 1 hypothesis as the null hypothesis that means the initial one and another one as alternative hypothesis. We may make the statements like rejecting H0 or accepting H0 or we may say accepting H1 if we say reject H0 and so on.

(Refer Slide Time: 29:28)



Now based on this we are able to so basically what we are doing. We are having the sample space here consisting of 4 points 0, 1, 2, 3 and we are dividing it into 2 parts, we call it acceptance region that is 0,1 and the A complement that is rejection region we call it 2,3 so this is called acceptance region and this is the rejection region.

(Refer Slide Time: 30:05)

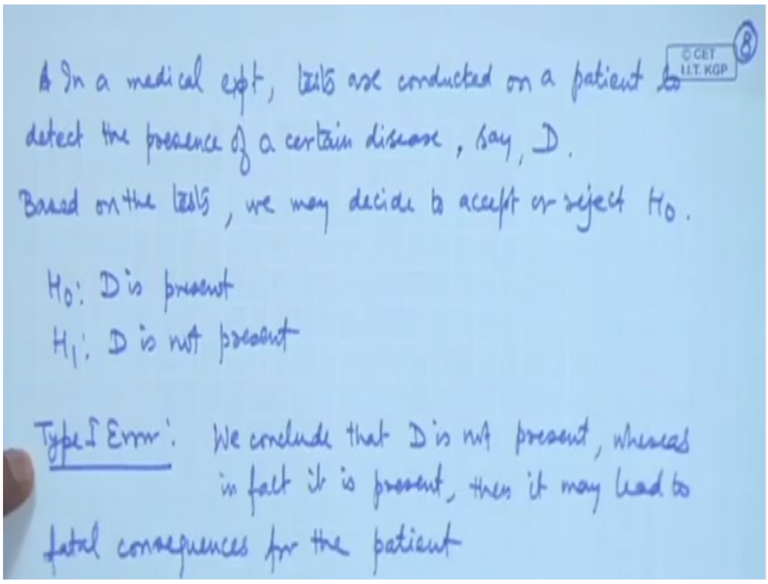


So basically a test of hypothesis partitions the sample space into 2 disjoint sets say A and R where A corresponds to the acceptance of H0 and R corresponds to the rejection of H0 or you can say acceptance of H1. So that is why this A we call to be acceptance region and R we call to be the rejection region or critical region. Since we are basing our decision on the outcome of a random experiment that means we are doing the sampling.

Therefore, certainly there is a chance of error in the form of introducing this type of error so we call it 2 types of errors. So when we conduct a test of hypothesis based on a random sample, we are likely to make 2 types of errors. So first one we call type 1 error that means rejecting H0 when it is true and second one is type 2 error that is accepting H0 when it is false.

Now the consequences of the 2 types of errors can be of various types depending upon different problems. Let us take an example related to say medicine.

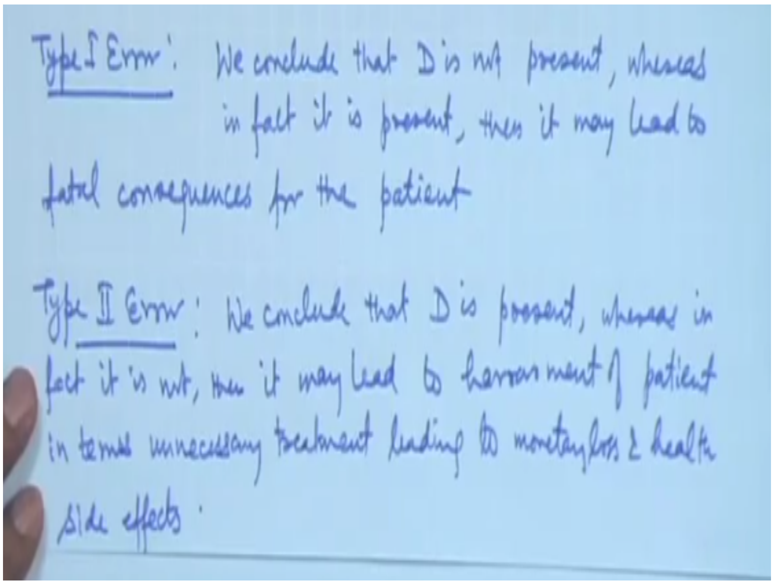
(Refer Slide Time: 33:15)



So in a medical experiment say tests are conducted on a patient to detect the presence of a certain disease say D okay. So now based on the tests we may conclude so your hypothesis is like H0 D is present that means the person has the disease or H1 D is not present. So now you see we may decide to accept or reject H0. Now what are the consequences? So if you look at type 1 error that means we are concluding that rejecting H0.

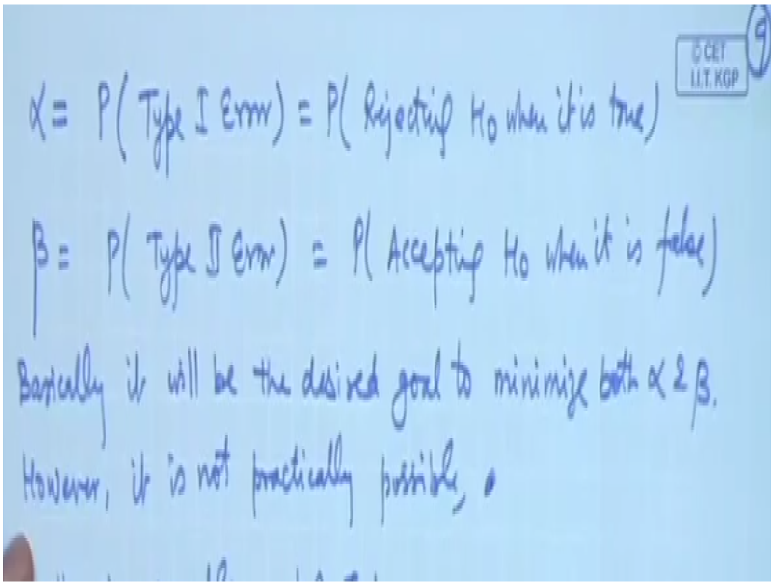
That is, we conclude that D is not present whereas in fact it is present then it may lead to fatal consequences for the patient.

(Refer Slide Time: 35:08)



If we consider say type 2 error that means you conclude that D is present whereas in fact it is not, then it may lead to harassment of the patient in terms of unnecessary treatment leading to monetary loss and health side effects. Now therefore in any given problem it is of important to control the 2 types of errors. So we give measures for these 2 types of errors.

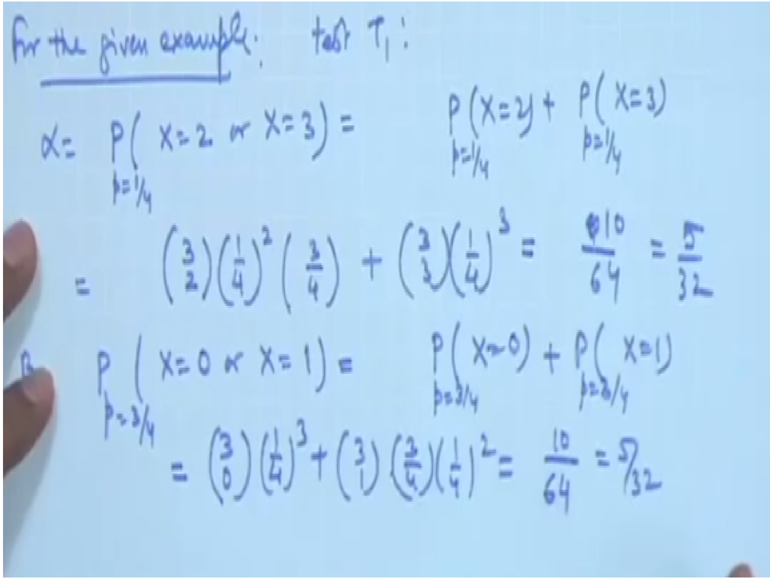
(Refer Slide Time: 36:23)



We consider say alpha=the probability of type 1 error that is the probability of rejecting H0 when it is true and similarly we consider beta that is equal to probability of type 2 error that is equal to probability of accepting H0 when it is false. So in any given problem it will be interesting or you can say it will be desirable to control both the errors alpha and beta.

Basically, we will like to have them to be a minimum. So basically it will be the goal to minimize both alpha and beta; however, it is not practically possible. The reason is that if I reduce alpha then beta will increase and if I reduce beta then alpha will increase. You can think from this example that I gave.

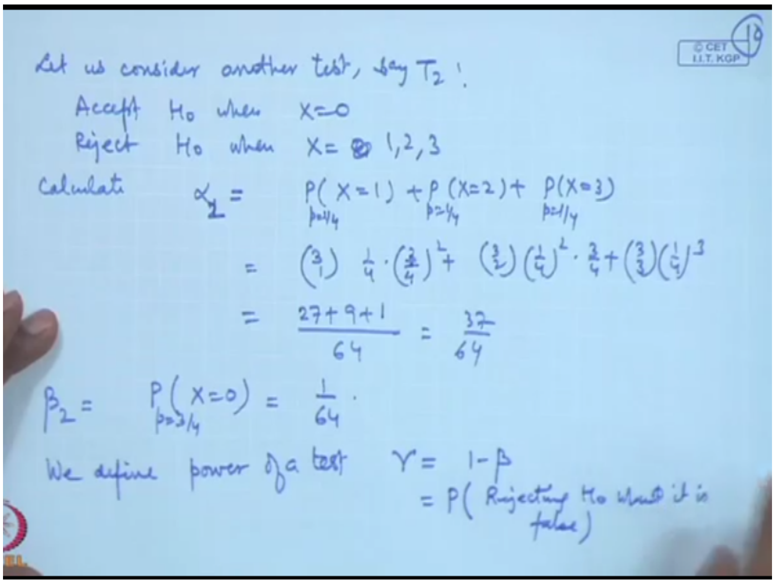
(Refer Slide Time: 37:59)



For this example, let us calculate, let us consider this test T1, what is alpha here? Alpha is the probability of rejecting H0 that means X=2 or X=3 when it is true that means when p=1/4 that means=probability of X=2 when p=1/4+probablity of X=3 when p=1/4. So that is equal to 3 C 2 1/4 square 3/4+3 C 3 1/4 cube. So you can write these values. It is equal to 9/64+1=10/64=5/32.

Let us look at beta, beta=probability of X=that is probability of accepting H0 when it is false. So we accept H0 when X=0 or X=1 when it is false that means when p=3/4. So that is equal to probability of X=0 when p=3/4+probability of X=1 when p=3/4. Once again we calculate these quantities, it turns out to be 3 C 0 1/4 cube+3 C 1 3/4*1/4 square. So once again it is equal to 10/64 which is equal to 5/32.

(Refer Slide Time: 40:05)



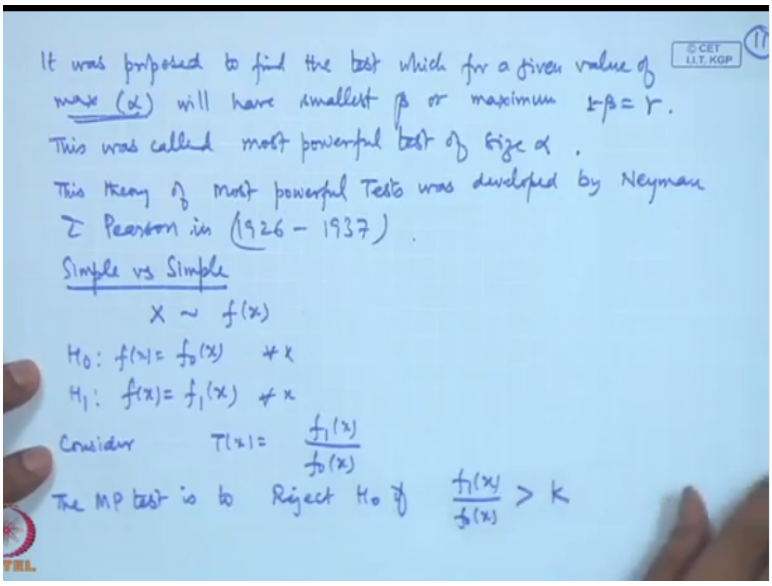
Now I design another test say let us consider another test say T2 okay. That is accept H0 when say X=0 and reject H0 when X=1, 2 and 3. For this test, let us calculate say alpha let me call it alpha 1 say for the test this one I will call alpha 1 and beta 1. Now here I will call it alpha 2 and beta 2 so that is equal to probability of X=1+probability X=2+probability X=3 when H0 is true that is p=1/4.

So that is equal to 3 C 1 1/4*3/4 square+3 C 2 1/4 square 3/4+3 C 3 1/4 cube =27+9+1/64=37/64 and let us look at say probability of type 2 error then that is becoming probability of X=0 that is probability of accepting H0 when it is false. So this is simply equal to 1/64. So you can see here that by using this particular test we have been able to reduce beta 2 from 10/64 to 1/64.

But at the same time the probability of type 1 error has increased from 10/64 to 37/64. In the same way, we can consider reduction of alpha but then beta will increase. So therefore a practical way which the Neyman and Pearson suggested was that we fix an upper level for probability of 1 type of error and then try to find out a test procedure for which the other type of error is minimum or we can say 1-the probability of the other type of error is maximum.

So as a convention it was considered we define power of a test say let us call it gamma that is equal to 1-beta that is probability of rejecting H0 when it is false.

(Refer Slide Time: 43:19)



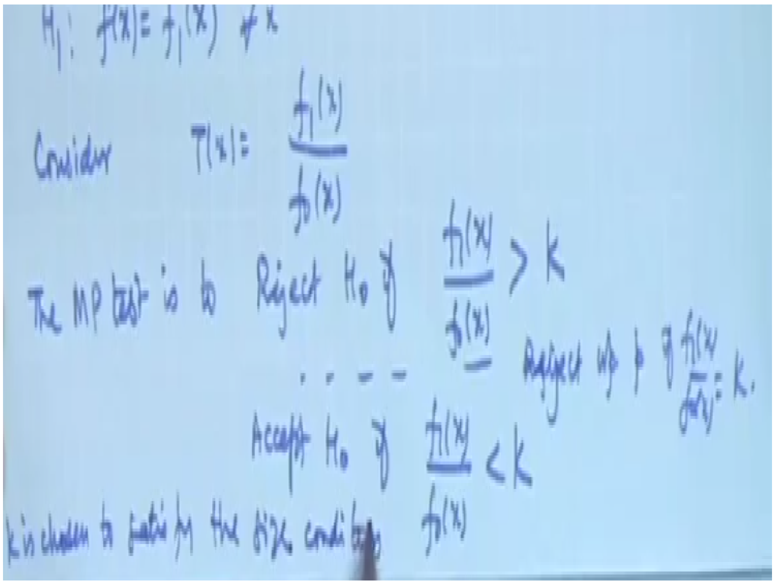
So it was proposed to find the tests which for a given value of maximum alpha will have smallest beta or maximum 1-beta that is gamma. So this was called most powerful test because maximum power most powerful test of size alpha because we put the maximum value of alpha that is called the size of the test or the level of significance. There are various names of it.

And we consider the test which will have the minimum probability of type 2 error or the maximum power that most powerful test. So the theory of most powerful test, so for simple versus simple case, a complete solution was obtained by Neyman and Pearson in 1926 and thereafter it was analyzed to the concept of uniformly most powerful test later on by the same authors and for composite hypothesis.

And also for some other situations where even uniformly most powerful test does not exist so they considered certain restricted class of test called unbiased test and among those tests they found the most powerful test. This theory of most powerful test was developed by Neyman and Pearson in 1926 to say 1937 in this period. So firstly they considered the solution for the simple versus simple case.

So suppose we have the problem let us write in terms of observation so X is a following fx okay and we make the hypothesis whether fx=f0x or H1 fx=f1x. So consider say Tx=f1x/f0x. The most powerful test is to reject H0 if f1x/f0x>k. Basically this is not the complete description we also have the range for example we may have discrete distribution.

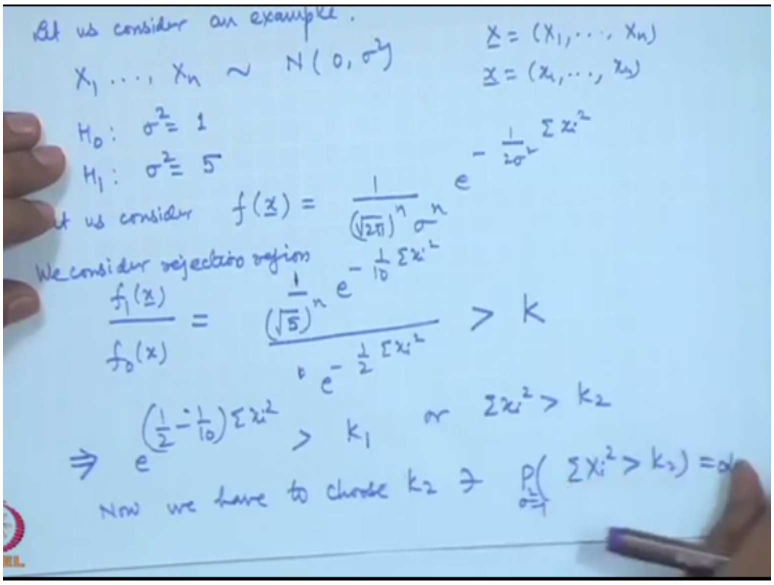
(Refer Slide Time: 46:47)



And in that case we also have reject H0 this, accept H0 if f1x/f0x<k and there was also a portion=that is reject with probability say p if f1x/f0x=k. Now this constant k is chosen to satisfy the size condition. However, even importantly to ask that it is not only sufficient it is also necessary condition for the most powerful test. So simultaneously they showed the existence of such a test, existence of the most powerful test.

And also that if there is a most powerful test, it has to be of this form. Now this turned out to be extremely useful result and let me explain through one example.

(Refer Slide Time: 48:03)



Let us consider say a simple testing problem say X1, X2, Xn say follow normal 0 sigma square. We were having the testing problem say sigma square=1 against say H1 sigma square=say 5. Now let us take let us consider the density function here of all the observations.

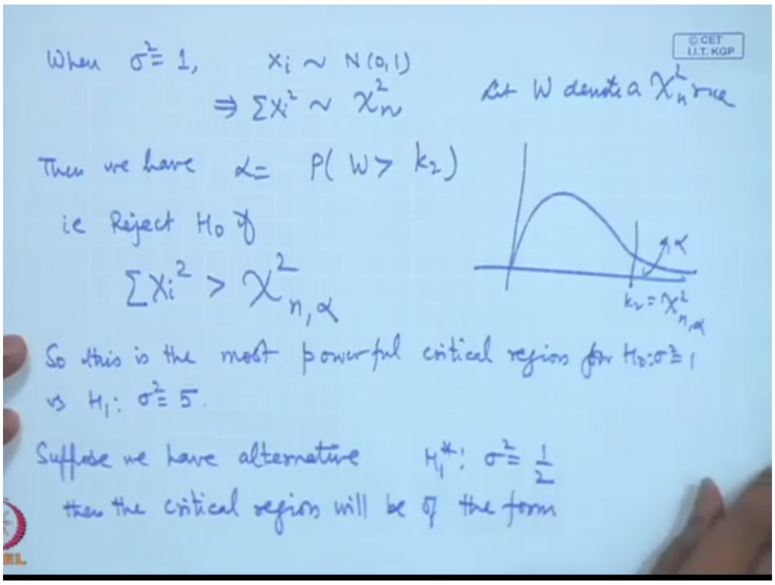
So x=x1, x2, xn where your X=X1, X2, Xn. So this is equal to 1/root 2 pi to the power n sigma to the n e to the power-1/2 sigma square sigma xi square.

So we consider f1x/f0x that means the ratio of the densities when sigma square=5 and when sigma square=1. So this will become equal to now this 1/root 2 pi to the power n will get cancelled out. You get 1/root 5 to the power n e to the power-1/10 sigma xi square/1 e to the power-1/2 sigma xi square. So we consider the rejection region. This is > some k.

Now this you can write in a modified fashion because this constant I can adjust on the right hand side and it will become e to the power 1/2-1/10 sigma xi square > some k1. I can take logarithm here so it will reduce to sigma xi square > some k2. Now we have to choose k2 such that probability of sigma xi square > k2 under sigma square=1=alpha.

So you can easily see that when I am doing the sampling from the normal distributions I can actually calculate the distribution here.

(Refer Slide Time: 51:10)



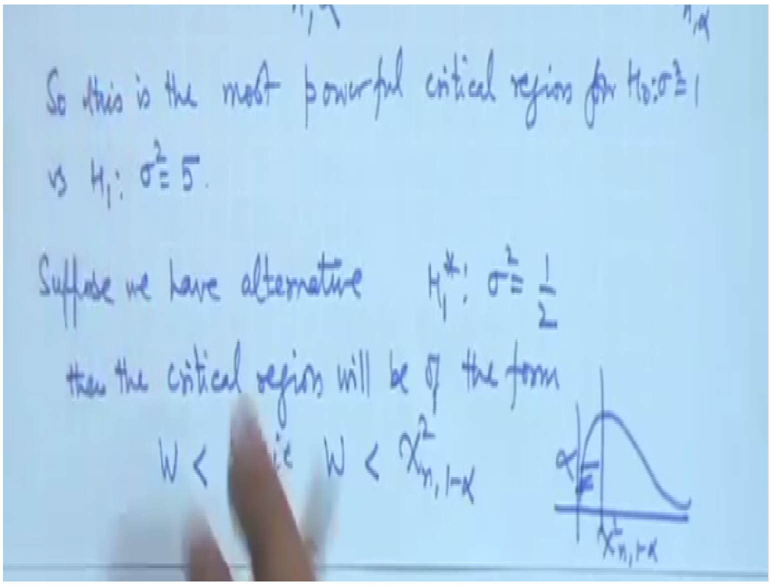
So under when sigma square=1 then I have xi is following normal 0,1 this will imply that sigma xi square will follow chi square distribution on n degrees of freedom. If that is so then this statement is reducing to let us call it say let W denote a chi square n random variable then we have alpha=probability W>k2 that means if I am considering a chi square curve on n degrees of freedom, then this k2 point is actually chi square n alpha.

That is reject H0 if sigma xi square is > chi square n alpha. So for the most powerful test the rejection region is of this form. So this is the most powerful critical region for H0 sigma square=1 against say H1 sigma square=5. Let us consider little generalization of this problem. See you notice here, here I took the null hypothesis 1 and in the alternative sigma square was 5 which was slightly bigger.

And therefore you have seen here in the denominator we had this -1/2 here and when we took the difference this becomes a positive quantity and therefore the region is in the form sigma xi square>k2. On the other hand, suppose I modify this. Suppose I consider suppose we have alternative say H1 star sigma square=say 1/2 if that is so then in this particular place we will get sigma xi square.

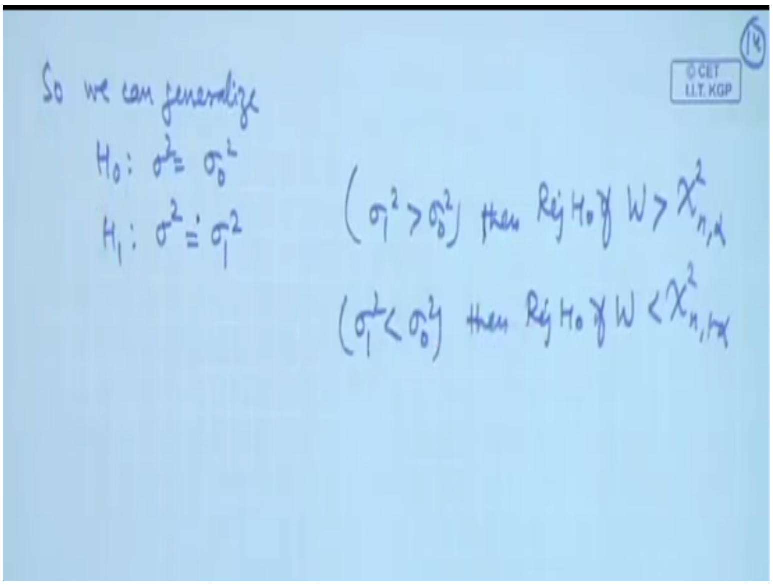
And if that is happening then you will get negative quantity here. So if I take log the region will get reversed.

(Refer Slide Time: 53:48)



Then the critical region will be of the form W < some k3. So if that is so then if you consider the region then if I want the probability alpha then this should be chi square n1-alpha that is W will be < chi square n1-alpha.

(Refer Slide Time: 54:34)



So we can generalize to this problem suppose I can see that sigma square=sigma 0 square against sigma square is equal to sigma 1 square and if sigma 1 square>sigma 0 square then reject H0 if W>chi square n alpha and if sigma 1 square<sigma 0 square then reject H0 if W<chi square n1-alpha. This is also pointing out to some important characteristic of this distribution.

Then we are considering normal 0 sigma square in the density in the exponent we are having sigma xi square as a sufficient statistics and there is a property here actually. This is called a monotone likelihood ratio property, which is satisfied here and therefore the region of rejection will be decided by the direction in which sigma xi square is taking value.

So since you can also think of it as a maximum likelihood estimator and from there also you can see that for the larger values of sigma xi square I will favor the hypothesis H1 and for the smaller values I will favor H0 and similarly in reverse fashion we will consider here for the smaller value I will favor H1 here and for larger values I will favor H0 in the other case when sigma 1 square is < sigma 0 square.

In the following lecture, I will give you the test for various hypotheses for the parameters of the normal distributions which are based on this theory. Basically, these results have been extended to the composite hypothesis. For example, I may consider here sigma square<sigma 0 square against sigma square>sigma 0 square and vice versa. When we consider those situations we have the uniformly most powerful test.

When we have 2 parameter situations then we have uniformly most powerful unbiased tests of these hypotheses. Now without mentioning these things I will be explicitly giving tests for the various normal population problems in the next lecture.