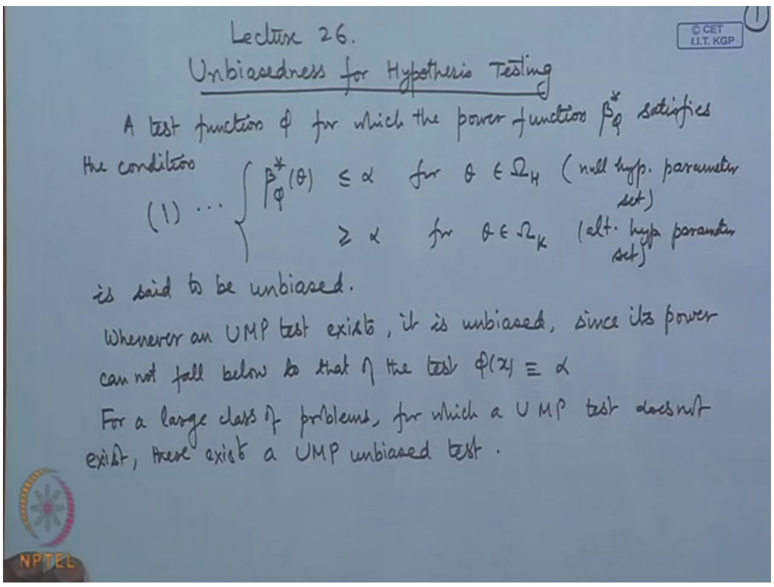
**Statistical Inference**
**Prof. Somesh Kumar**
**Department of Mathematics**
**Indian Institute of Technology, Kharagpur**

**Module No. # 01**
**Lecture No. # 26**
**UMP Unbiased Tests**

So, we have demonstrated in the last lecture that there are several problems. For example, when the distributions are monotone likelihood ratio or if the distributions are in the one parameter exponential family there are one sided testing problems are certain two sided testing problems for which U M P unbiased tests U M P test can be derived. However, we also demonstrated that there are certain cases where U M P test cannot be found in fact, we have demonstrated that they do not exist. Now, if that happens then, we can simply impose an additional condition which is called a condition of unbiasedness and then, we may try to find out the U M P test among the unbiased tests so, these are called U M P unbiased test.

(Refer Slide Time: 01:18)

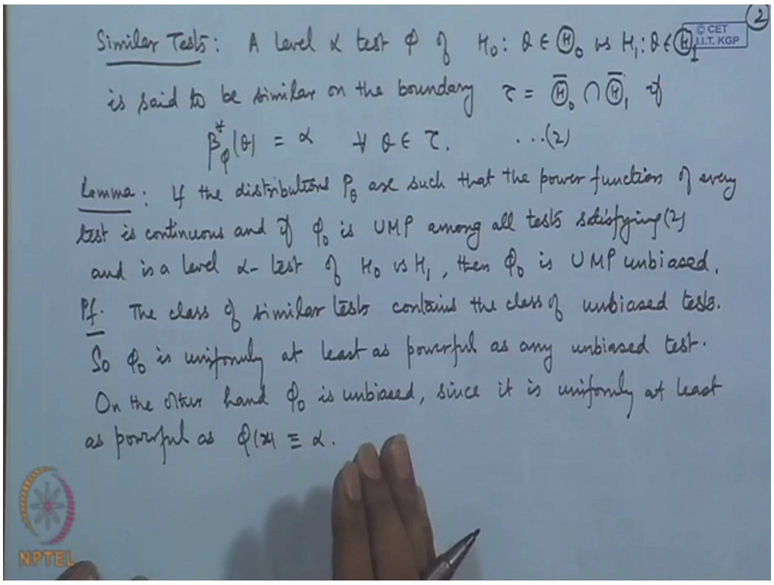



So, we may say this condition a test function phi for which the power function beta star phi satisfies the condition beta phi star theta is less than or equal to alpha for theta belonging to omega H that is the null hypothesis parameter set. And it is greater than or equal to alpha for theta belonging to omega k that it is the alternative hypothesis parameter set.

Then, this set then this test function is said to be unbiased let me call this condition one. Now, let us see what is this for example, we have considered theorem two when the family this is the monotone likelihood ratio, where there and for the testing problem for theta less than or equal to theta naught against theta greater than theta naught. So, what we showed there that beta phi theta is actually increasing function. So, at theta is equal theta naught the probability of type one error is the maximum and therefore, beta phi star theta was less than or equal to alpha for theta belonging to omega H that is theta less than or equal to theta naught and thereafter it was greater than or equal to. So, it was an example of an unbiased test.

So, whenever an U M P test exists it is unbiased since it's power cannot fall below to that of the test say phi X is equal to alpha that is the that is we always reject with probability alpha whatever be X and we accept with probability one minus alpha. And this test has expectation equal to alpha that is the power function is actually equal to alpha. So, for a large class of problems for which a U M P test does not exist there exists a U M P unbiased test. So, we may consider theta less than or equal to theta naught theta is equal to theta naught against theta naught equal to theta naught and also the cases of the new sense parameters in many of these cases, we will be actually demonstrating the existence of the U M P unbiased tests.

(Refer Slide Time: 05:08)

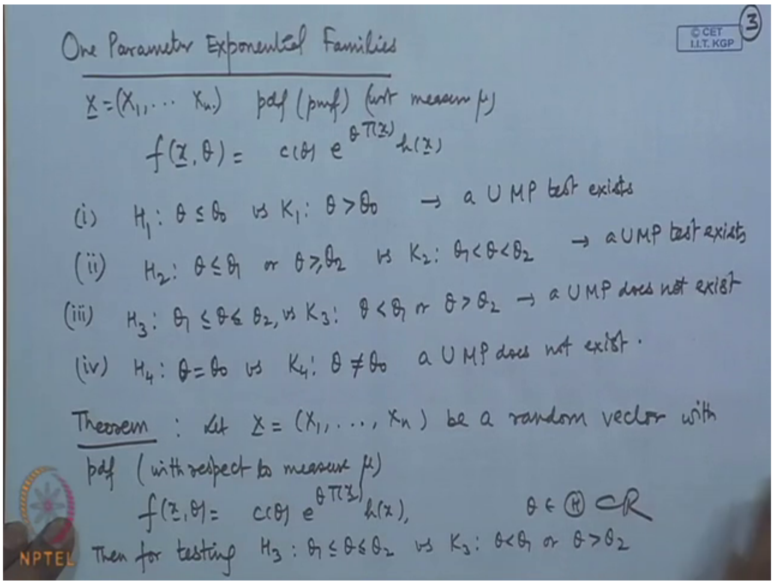



So, we give a definition here, which is called similar tests so, these are helpful in deriving U M P unbiased test. So, we call what is known as similar tests what are similar tests so, a level alpha test phi of the hypothesis testing problem H naught theta belonging to theta naught against H one theta belonging to say theta one is said to be similar on the boundary. So, boundary of this theta naught and theta one let us denote it by some say tau that is equal to theta naught closer intersection theta one closer, if beta phi theta star is equal to alpha for all theta belonging to tau. So, this is called a similar test.

We have the following lemma here, if the distributions p theta are such that the power function of every test is continuous and if phi naught is U M P among all tests which satisfy to this condition of similarity and is a level alpha test of H naught versus H one then phi naught is U M P unbiased. So, this is a very very useful result it is of a similarity is a very useful concept. So, first of all what we are saying if you look at this one can be needed here that the power function should be less than or equal to the level for the values, that is the probability of type one error basically. And the power should be greater than or equal to alpha that is what we are actually saying for the unbiasedness condition.

So, in order to achieve this we are imposing a condition of the continuity on the test function and a similarity condition that is on the boundary the value alpha should be achieved. So, then what we are saying is that, in these cases a U M P test will actually and a U M P test if it is similar test then certainly it will be U M P unbiased. So, let us look at the proof of this, which is quite simple the class of similar tests contains the class of unbiased tests. So, phi naught is uniformly at least as powerful as any unbiased test. On the other hand phi naught is unbiased since it is uniformly at least as powerful as phi X is equal to alpha.

(Refer Slide Time: 10:02)



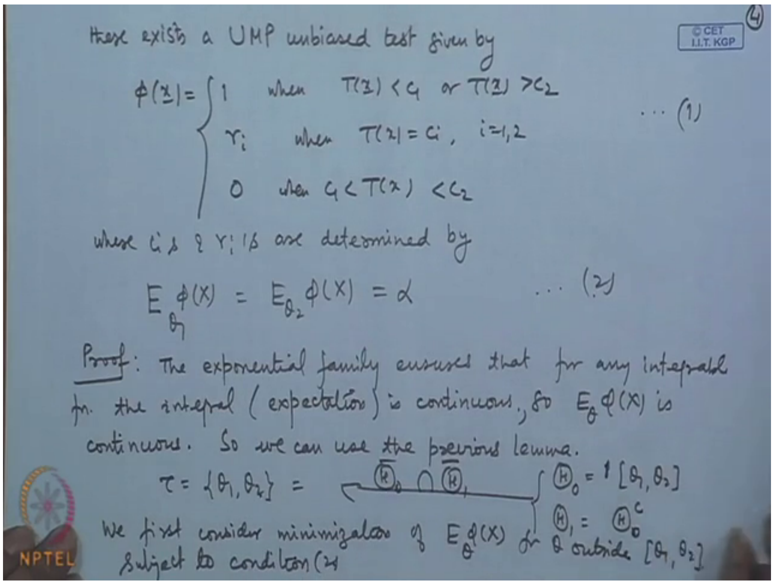

So, this proves that it is a U M P unbiased test this class of similar test contains the class of unbiased test. So, phi naught is uniformly at least as powerful as any unbiased test on the other hand phi naught is unbiased since it is uniformly at least as powerful as phi X is equal to alpha therefore, phi naught will be U M P unbiased. Now, let us consider applications to one parameter exponential family is here so, one parameter exponential families. So, as usual we are considering the X one X two X n as a random sample. So, we write as x so, p d f or p m f with respect to a measure mu we are writing down f X theta is equal to c theta E to the power say theta T X H x. So, we are having the following situation, if I am considering the hypothesis testing problem I will name this as say H 1 theta less than or equal to theta naught against say k 1 theta greater than theta naught the situation here is the U M P test exists.

Because q theta that is theta here is a strictly monotone so, the situation here is that a given p test exists. Let us consider say theta less than or equal to theta 1 or theta greater than or equal to theta 2 against say theta 1 less than theta less than theta 2 in this case also a U M P test exists. So, in the lecture twenty four we have given the form of this test and in the previous lecture twenty five I have given several applications of these two tests and the form of the U M P test has been derived. Let us consider say theta one less than or equal to theta less than or equal to theta 2 against k three theta less than theta 1 or theta greater than theta 2.

This three is actually the dual of two here a U M P test does not exist and if we consider theta is equal to theta naught against, this is also alternative is two sided here a U M P test does not exist. In the previous lecture I have demonstrated through the double exponential distribution that for such a problem a U M P test does not exist. We have shown that a U M P test which is having the maximum power for theta greater than or equal to theta naught has a smaller power than another test for theta greater than theta naught and vice versa therefore, U M P test does not exist.

So, we have the following result that is U M P unbiased tests do exist here. So, this is stated in the following theorem for detailed proofs you may look at the book of lehman or rohatgi let x. So, when we consider a random sample we generally write the line density. So, I am calling it as a random vector with probability density with respect to some measure mu. So, f X theta is equal to c theta E to the power theta T X H X where, theta belongs to say theta subset of r then for testing H 3 that is this particular four third case theta 1 less than or equal to theta less than or equal to theta 2 versus k three theta less than theta 1 or theta greater than theta 2.

(Refer Slide Time: 14:42)

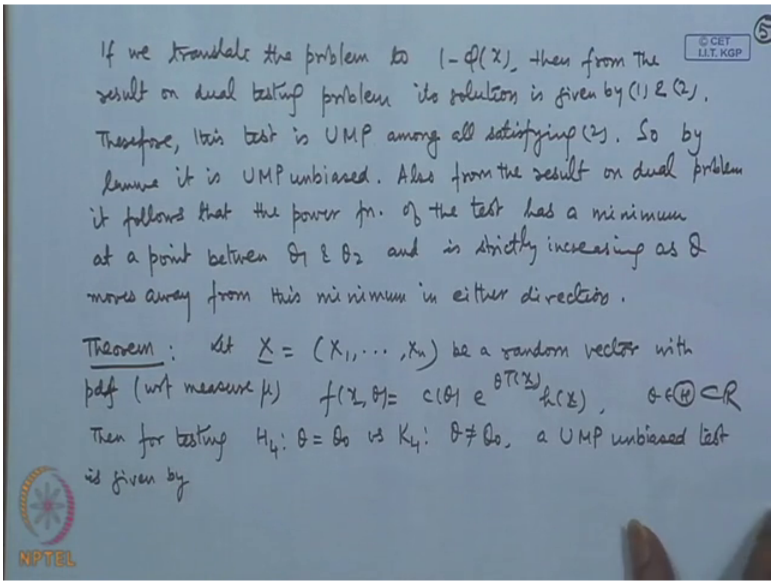


There exists a U M P unbiased test given by phi X is equal to 1 when T X is less than c 1 r T X is greater than c 2 that is equal to gamma I when T X is equal to c I for I is equal to one two and that is equal to 0, if c 1 is less than T X less than c 2. Where c is and gamma is are determined by expectation of phi X under theta one and under theta two this should be equal to alpha. Let us have a comparison with the test which I gave in the lecture number twenty four just to appreciate the dual problem. Here we had considered this q is a strictly monotone so, you had monotone likelihood ratio and T X which is satisfied here for theta less than or equal to theta one or theta greater than or equal to theta two.

Which I am actually, describing as H two a U M P test was given here now, in H three we are having simply the dual of this in H three we are having the dual of this H two here however, the test is U M P unbiased here. Let me briefly sketch the proof of this although I have been skipping the proofs of the theorems. In fact, for the detailed proofs one can look at the book of the Lehman and Romano are Rohatgi and Salih etcetera however, a few of the proofs I will just simply sketch here. So, here the distribution is in the exponential family, in the exponential family if I have an integrable function the integral of the expectation you can say it is continuous function. So, we can use this thing that is because to apply the lemma what we wanted here is that, power function of every test is continuous so, power function here will be expectation of phi X here.

So, this should be continuous now, in order to have that what we can do is we can use the condition that we are having the exponential family here. So, the exponential family ensures that for any integrable function the integral or expectation is continuous so, expectation of theta phi X is continuous. So, we can apply the previous lemma. So, according to the previous lemma we will let us consider the boundary what was the boundary? Boundary was consisting of theta 1 and theta 2 that is you had theta actually; the notation that I have used here is theta naught and theta one for the null and alternative hypothesis set.

(Refer Slide Time: 20:16)



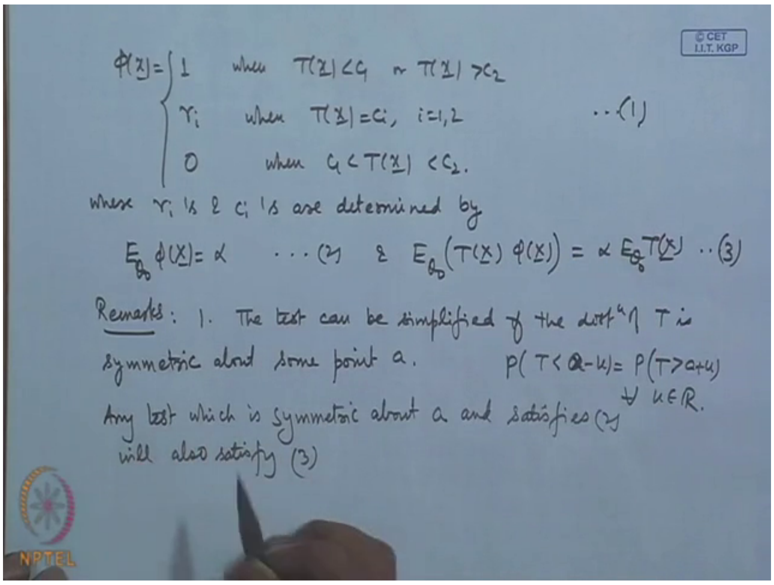

So, here theta naught was in this particular problem, the theta 1 to theta 2 and theta 1 is actually compliment of theta naught. So, if I consider this tau that is theta naught closer intersection theta 1 closer then this is going to be equal to theta 1 theta 2. So, now, let us consider we first consider minimization of expectation phi X for theta outside the interval theta 1theta 2 subject to the condition 2. Now, if we translate the problem to 1 minus phi X function, then from the result on dual testing problem it is solution is given by one and two.

Therefore, this test is U M P among all satisfying 2 so by it is U M P unbiased. Also, from the result and dual problem it follows that the power function of the test has a minimum at a point between theta 1 and theta 2 and is a strictly increasing as theta moves away from this minimum in either direction. So, basically what you have seen here is that this result is actually following a straight forwardly, from the concept of similarity and the dual problem that we consider where the null hypothesis was two sided here the alternative is two sided. So, U M P does not exist, but U M P unbiased can be found here.

Now, I will also consider and of course, you can see here the test functions forms exactly one minus the form of the test for the dual problem here condition is also the same here. Now, let us consider the point null hypothesis and the alternative is two sided now, in that case one of the conditions gets modified let me state it in the following theorem. Let

X is equal to X 1 X 2 X n so, again the same conditions are there let me just for convenience I am restating it be a random vector with probability function with respect to some measure mu that is f X theta is equal to c theta E to the power theta T X H X and theta is of course, lying in a parameter space which is a subset of the real line.

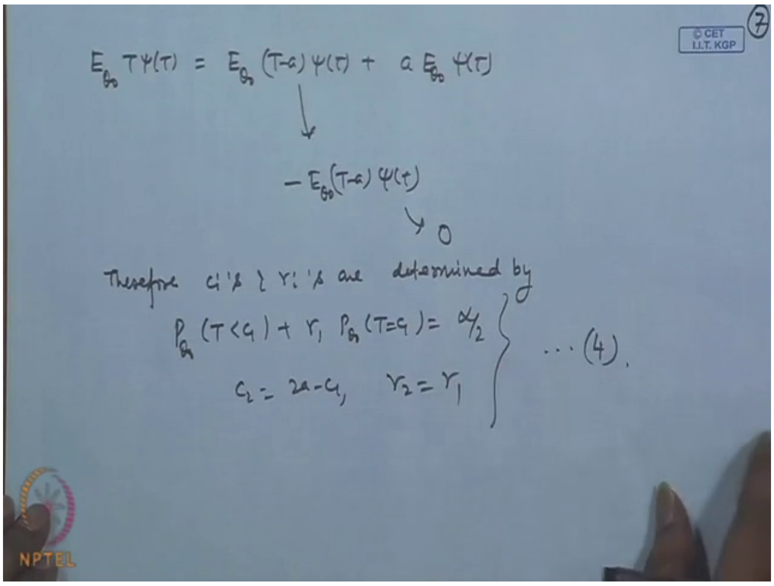
(Refer Slide Time: 24:45)




Then for testing H 4 theta is equal to theta naught against k 4 theta is not equal to theta naught a U M P unbiased test is given by you note here that actually, the form is given by the test which I were stated for the previous problem; however, the size condition will get modified. So, let me give it here it is phi X is equal to one when T X is less than c 1 or T X is greater than c 2 it is equal to gamma I when T X is equal to c i i is equal to one two it is equal to 0 when c 1 is less than T X is less than c 2. Where gamma is and c is are determined by expectation of theta naught phi X is equal to alpha and expectation of theta naught T X phi X is equal to alpha times expectation of theta naught T X. Note here that here we have a condition in which the statistic T is also involved.

So, this was not there in any of the previous results here. So, for the proof I refer to the book of Lehman and also the book of rohatgi I am not going to discuss the proof in detail here however, let me give certain comments here. The test can be simplified if the distribution of T is symmetrical about some point a. For example, so, if I say the symmetrical about the point a then actually you will have probability of T less than some X minus u is equal to probability of sorry a minus u is equal to probability of T greater than a plus u for all u on the real line. So, any test which is symmetric about a and satisfies 2 will also satisfy 3. So, automatically this condition will be satisfied and therefore, we do not have to consider two conditions in these cases.

(Refer Slide Time: 28:07)



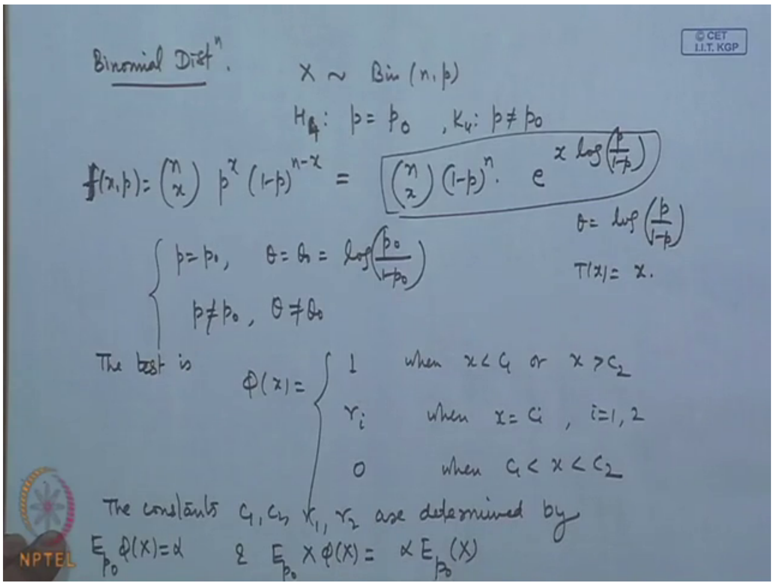

So, let me just give it here for example, you may consider here.

(No audio from 28:01 to 28:21)

So, we can actually show that this value is equal to minus expectation theta naught T minus a psi t that means, this is actually equal to zero. So, if this is actually equal to zero then so, this will become equal to zero that means, this condition becomes this condition so, it is automatically becoming true. Therefore, c is and gamma is are determined by alpha by two c 2 is equal to 2 a minus c 1 and gamma two is equal to gamma 1. So, the conditions get actually modified in place of writing down this two conditions we can actually reduce to these two conditions here.

Another important point which I would like to mention here is that, the tests that we have stated in the two previous theorems for the two sided alternative hypothesis testing problems. These are U M P unbiased they are actually is strictly unbiased what is the meaning of a strictly unbiased that as soon as we move away from the point theta naught, then the if we are going to the alternative hypothesis set it is becoming a strictly greater than alpha, if we are going to the null hypothesis set it is becoming a strictly less than alpha.

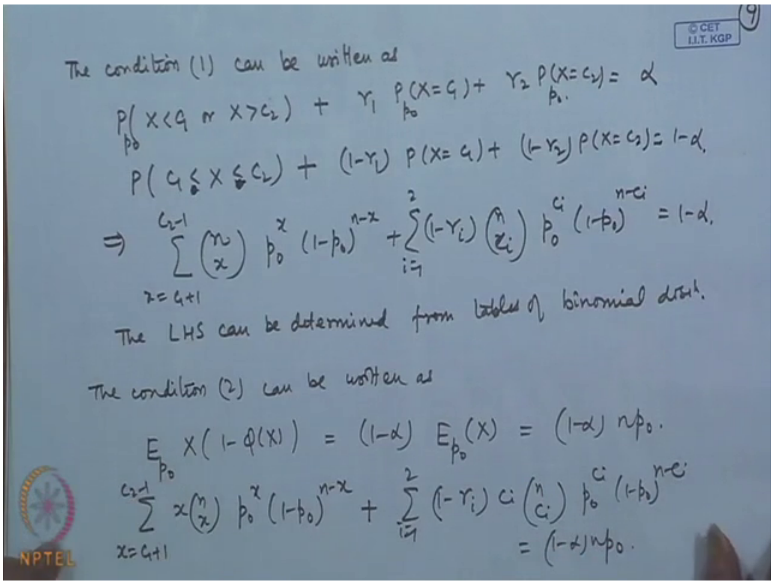
(Refer Slide Time: 30:15)




Let us consider example here, binomial distribution so, let us consider say X follows binomial n p you are considering the testing problem for H 4. So, I will use this notation H 1 H 2 H 3 H 4 reference is clear here, the alternative is p is not equal to p naught. So, here you have the distribution function the form of the density function n c X p to the power X one minus p to the power n minus X. If we are writing in the form of the exponential family we are writing it as one minus p to the power n E to the power X log p by one minus p. So, here this is X T X is equal to X and theta is equal to log p by 1 minus p and T X is equal to x. So, if I say p is equal to p naught then theta is equal to theta naught that is equal to log of p naught by one minus p naught and theta naught equal to theta naught then, it is equivalent to p not equal to p naught.

So, the hypothesis testing problem is restated in the form of this where we are getting the one parameter exponential family. So, the test is phi X is equal to 1 when X is less than c 1 or X is greater than c 2 it is equal to gamma i when X is equal to c i for i is equal to 1 2 it is equal to 0 when c 1 is less than X is less than c 2.

(Refer Slide Time: 33:07)



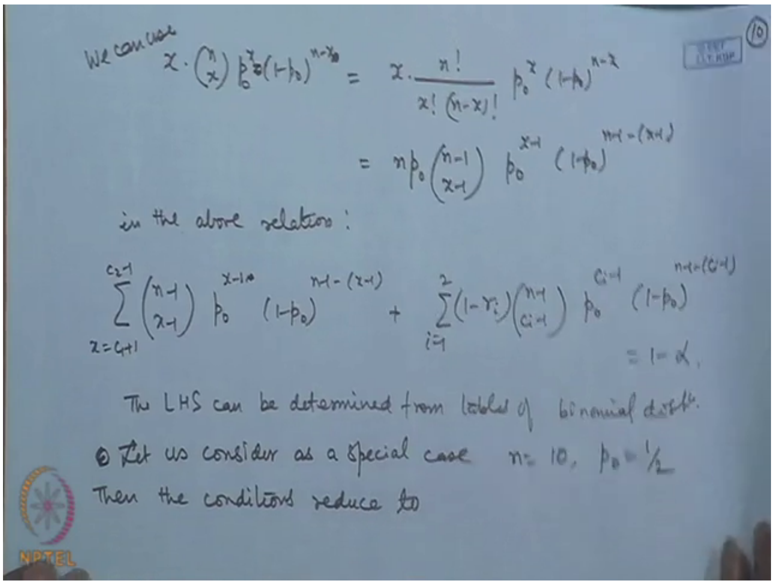

The constants c 1 c 2 gamma 1 gamma 2 are determined that is expectation of p naught phi X is equal to alpha and expectation of X phi X is equal to alpha expectation p naught X and of course, expectation of X in the binomial distribution will be n p so, here it will become n p naught. So, let me call this conditions 1 and 2 here, let me write down this conditions in detail, the condition 1can be written as this is actually probability of X less than c 1 or X greater than c 2 plus gamma 1 probability of X is equal to c 1 plus gamma 2 probability of X is equal to c 2 is equal to alpha.

I may consider the reverse of this also, I may consider probability of c 1 less than or equal to X less than or equal to c 2 under p naught of course, here. So, this I n removing from here and we can put it as one minus gamma one probability of X is equal to c 1 plus 1 minus gamma two probability of X equal to c 2 is equal to 1 minus alpha. So, this condition is then sigma n c X p naught to the power X 1 minus p naught to the power n minus X where X is from c 1 plus 1to c 2 minus one plus 1 minus gamma I n c X n c c i p naught to the power c i 1minus p naught to the power n minus c i i is equal to 1 2 that is equal to 1 minus alpha.

Now, giving the value of p naught and n one can do certain interpolation and determine. So, the left hand side can be determined from tables of binomial distribution. Now, what is the second condition? The second condition so, again we can see there one 1 so, X into 1 minus phi X that is equal to 1 minus alpha expectation of X of course, you may ask a question. That why we are taking 1 minus this is for the simplicity otherwise here we have to write two regions here we have to write only one region if I am taking 1 minus therefore, it is convenient to consider one minus in both of these places.

Now, this is equal to 1 minus alpha into n p naught here now, the left hand side if you look at once again this is becoming sigma X into n c X p naught to the power X 1 minus p naught to the power n minus X for X is equal to c 1 plus 1 to c 2 minus 1 plus sigma 1 minus gamma i into n c i into c I, because we are considering multiplication of x. So, here will be c i n c i p naught to the power c i 1 minus p naught to the power n minus c i i is equal to 1 2 that is equal to 1 minus alpha n p naught.

(Refer Slide Time: 37:15)

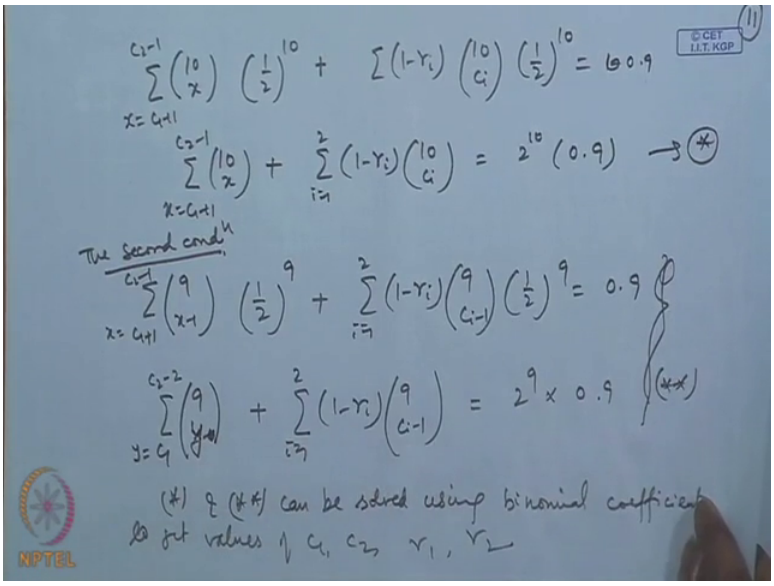



Now, here we do some sort of simplification, see this X into n c X that is equal to and of course, p naught to the power X into 1minus p naught to the power n minus X naught n minus x. So, this is X into n factorial divided by X factorial into n minus X factorial p naught to the power X into 1 minus p naught to the power n minus X. This we can write as n into n minus 1 c X minus 1 and then p naught here p naught to the power X minus 1 1 minus p naught to the power n minus 1 minus X minus 1 here.

So, obviously, we can use this in the above relation so, here we have n c X into X into this term. So, n p naught will come out in the second term it is c i into n minus c i and again here we can use this thing n p naught can be written outside and this will become n minus one c c i minus one. And this will become minus one here here this will become minus one and n p naught n p naught will get cancelled out. So, this will give me simply n minus 1 c X minus 1 X is equal to c 1 plus 1to c 2 minus 1 p naught to the power X minus 1 and 1 minus p naught to the power n minus 1 minus X minus 1 plus 1 minus gamma i n minus 1 c i minus 1 p naught to the power c I minus 1 1 minus p naught to the power n minus 1 minus c i minus 1 i is equal to 1 2 that is equal to 1 minus alpha.

And if you see this condition there is slight difference from this conditions here, here you are having n c X here this is n minus one c X minus one. So, once again you can see this left hand side can be determined from tables of binomial distribution. So, one can actually evaluate now, you will be given in a given problem n will be given to you p naught will be given to you. So, after substituting these values this is reducing to looking at the values from the so, for example, suppose I say p naught is equal to half if I take p naught is equal to half then this condition become extremely simple let me just show you it here.

(Refer Slide Time: 41:04)



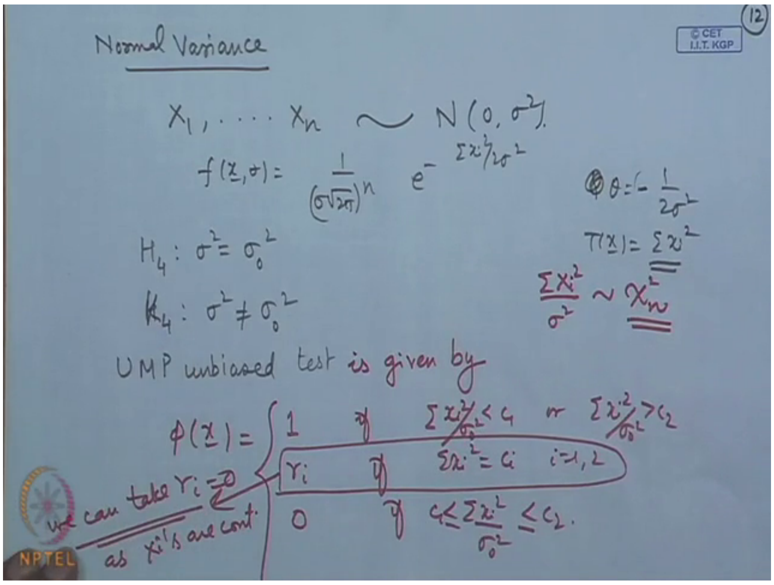

If I take say p naught is equal to half then this is becoming simply half to the power n let me just demonstrate. Let us consider as a special case say n is equal to say some value say ten and p naught is equal to half then what it is reducing to, then the conditions reduced to let me write it here one by one. So, the first condition you can see here, it is n c x so, this is becoming sigma ten c X x is equal to c 1 plus 1 to c 2 minus 1 this is becoming half to the power n that is half to the power 10 plus sigma 1 minus gamma i ten c c i and once again half to the power ten is equal to 1 minus alpha.

Suppose, I say alpha is equal to point 1 so, this is becoming point 9 say suppose I am taking alpha is equal to point one then immediately you can see you can multiply by this two to the power ten on this sides. So, this is becoming ten c X x is equal to c 1 plus 1 to c 2 minus 1 plus sigma 1 minus gamma i i is equal to 1 to 2 10 c c i is equal to 2 to the power 10 into 0.9 and similarly, if you look at the second condition. The second condition will become the second condition will become sigma 9 c X minus 1 X is equal to c 1 minus 1 c 1 plus 2 to c 2 minus 1 half to the power nine plus sigma 1 minus gamma i i is equal to 1 2 9 c c i half to the power 9 is equal to 0.9.

So, this will become equal to sigma 9 c X minus one we can put it in the terms of y say y is equal to I am putting X minus one is equal to y so, this is from c 1 to c 2 minus 2 plus sigma i is equal to 1 2 1 minus gamma i 9 c and of course, this would have become minus one here. So, that is c i minus 1 that is equal to 2 to the power 9 into 0.9. So, these two conditions condition is star and double star can be solved using binomial coefficients to get values of c 1 c 2 gamma 1 and gamma 2. So, in this problem I have demonstrated that a uniformly most powerful unbiased test exists. So, we have considered one parameter exponential family here.

So, for the point null hypothesis and the alternative is actually two sided we are able to exactly obtain the test function of course, the constants c 1 c 2 gamma one gamma two have to be further determined we have to actually carry out a little exercise here. We have also demonstrated that for continuous distribution such as normal for testing for mean or testing for the variance, again we have shown that the two sided alternative U M P unbiased test will exist here so, let us look at that case.

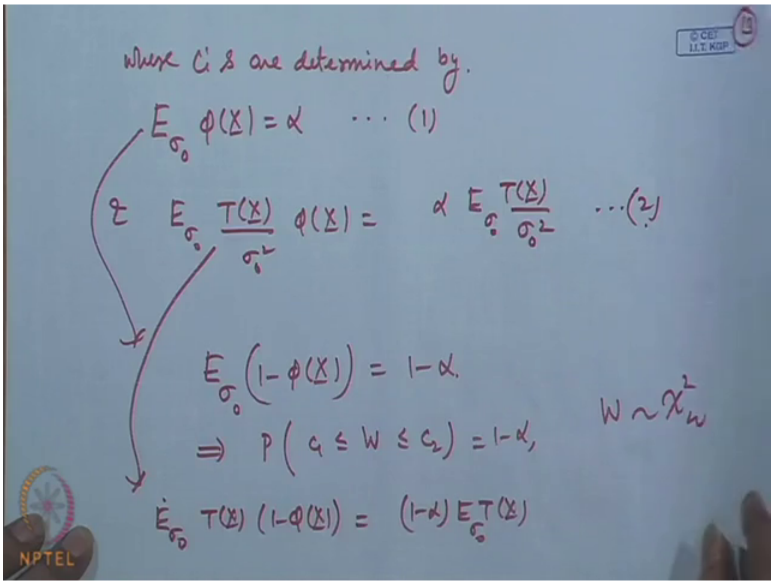
(Refer Slide Time: 45:02)

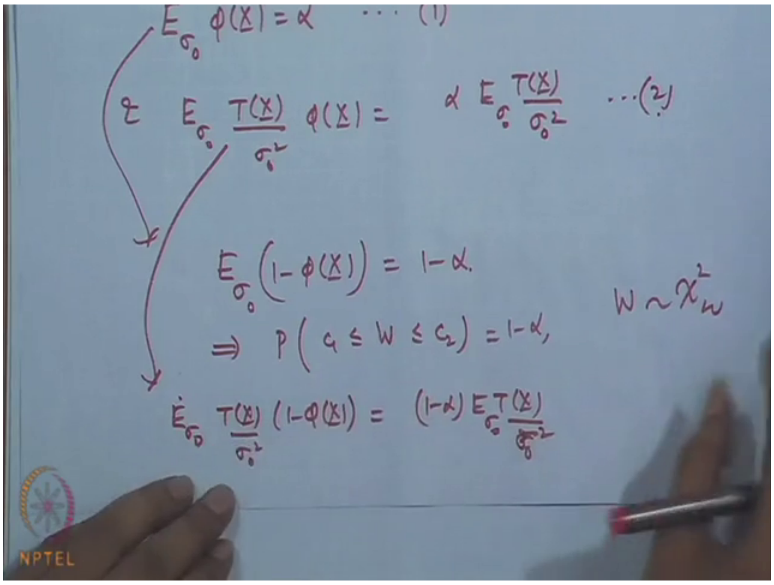



So, applications to testing for normal variance we have a random sample from normal 0 sigma square distribution. So, we need to write down the joint distribution to look at the formulation here of T X e to the power minus sigma X i square by 2 sigma square. So, here your parameter theta is equal to minus 1 by 2 sigma square of course, you can see this is an increasing function here and T X is equal to sigma X i square. So, if I consider the testing problem sigma square is equal to sigma naught square against, sigma square is not equal to sigma naught square.

Then the U M P unbiased test is given by phi X is equal to 1, if sigma X I square is less than c 1 or sigma X I square is greater than c 2 it is equal to gamma I, if sigma X I square is equal to c I i is equal to 1 2 it is equal to 0. If sigma X I square lies between c 1 to 2 and now this part we can take we can take gamma i to be 0 as X is are continuous. So, this part can be then included here also, we can note here that sigma X i square by sigma square that will have chi square distribution on n degrees of freedom.

(Refer Slide Time: 47:41)



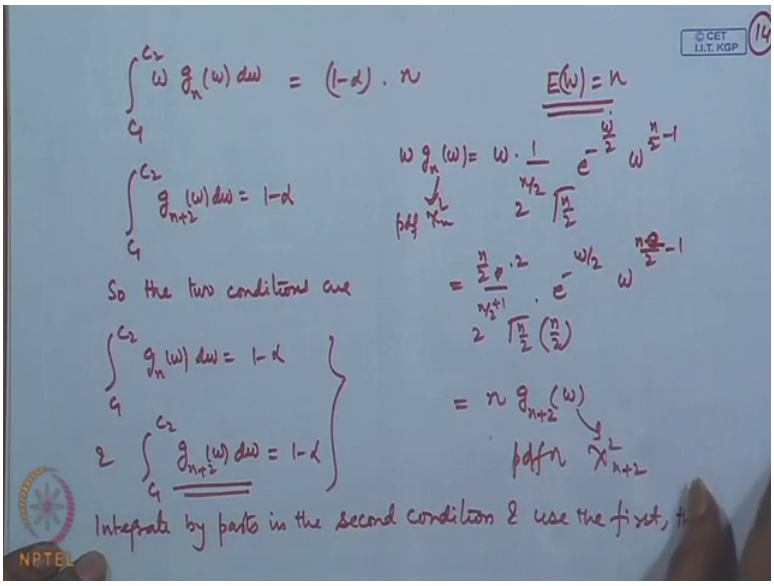




So, if I am considering under the null hypothesis I can consider deviation by sigma naught square here. So, that the distributional properties will be easy to study then, we will have say expectation of where c i s are determined by expectation of sigma naught phi X is equal to alpha and expectation of T X divided by sigma naught square phi X is equal to alpha times expectation of T X by sigma naught square. So, we will determine the constants c 1 and c 2 using both of these conditions. Since this condition can be simply written as expectation of sigma naught 1 minus phi X is equal to 1 minus alpha.

So, 1 minus phi X means this is becoming probability of c 1 less than or equal to w less than or equal to c 2 is equal to 1 minus alpha where w follows chi square n distribution.

So, this is simply to be determined from the tables of chi square distribution on n degrees of freedom however, let us look at this condition two also. What is this condition two? In the condition two you are getting once again if I write 1 minus so, this thing will become expectation of sigma naught T X 1 minus phi X that is equal to 1 minus alpha times expectation of T x.

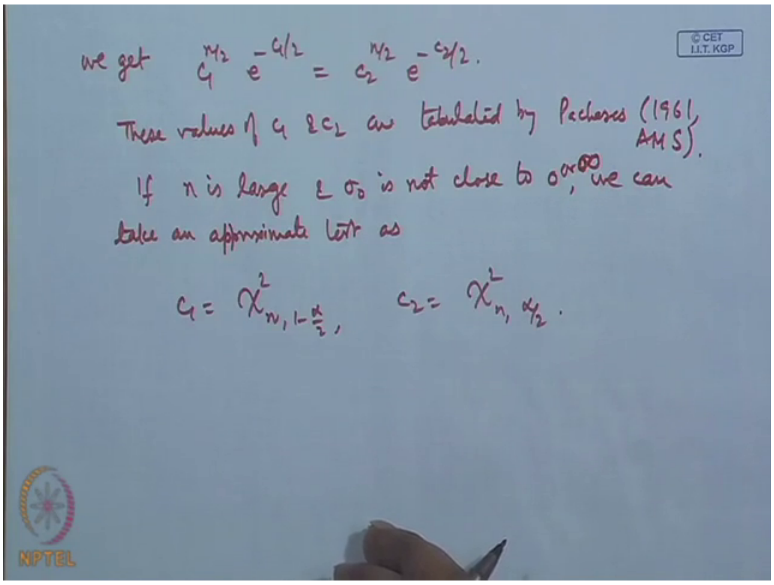
(Refer Slide Time: 49:52)




Because expectation T X gets cancelled out on both the sides so, this is then reducing to integral w and of course, we are also considering division by sigma naught square. So, this is w here so, this is w into the density function of a chi square value belonging to n degrees of freedom from c 1 to c 2 this is equal to 1 minus alpha and this is n, because expectation of w is n. So, if we consider the density of chi square on n degrees of freedom that is 1 by 2 to the power n by 2 gamma n by 2 E to the power minus w by 2 w to the power n by 2 minus 1.

Then this is equal to 1 by 2 to the power n by 2 gamma n by 2 E to the power minus w by 2 w to the power now one power is getting added so, this will become n plus 1 by 2 minus 1 if that is so, then here I had the coefficient here, I write here gamma this is n by 2 here and then n by 2 here and then I add 1 power here and I add 1 power here. So, this is becoming n times g of n plus so, this will become n plus two here rather than n plus. So, this is that means, this is density of p d f of chi square n plus 2 if this is p d f of chi square n then this is p d f of chi square n plus 2.

(Refer Slide Time: 52:34)




So, this condition then can be written as c 1 to c 2 g n plus 2 w d w is equal to 1 minus alpha. So, the two conditions are then c 1 to c 2 g n w d w is equal to 1 minus alpha and c 1 to c 2 g n plus 2 w d w is equal to 1 minus alpha. Now, if we consider integration by parts in this one integrate by parts in the second condition and use the first, then we get c 1 to the power n by 2 E to the power minus c 1 by 2 is equal to c 2 to the power n by 2 E to the power minus c 2 by 2. These tables of c 1 and c 2 are tabulated by pachares in 1961 in a paper in annals of mathematical statistics.

And another thing is that if n is large then we can consider the two conditions to be exactly the same, if n is large and sigma naught is not close to 0, we can take zero or infinity we can take an approximate test as c 1 is equal to chi square n 1minus alpha by two and c 2 is equal to chi square n alpha by 2. So, today I have demonstrated that applications of a two sided alternative hypothesis testing problem, we have U M P unbiased test and I have given applications to a binomial problem to a normal problem. And of course, here you notice that the application is limited to one parameter exponential family and in the one parameter exponential family I have taken a specific form theta into T x.

So, for such a problem then we are able to derive the U M P unbiased test. Now, in the next lecture I will be considering the completeness aspect also and we will introduce the test with the neyman structure now, using this structure we will show that we will be able to provide the, solutions when we are having the new sense parameters. The simplest case for example, we have a normal distribution with the two parameters then what happens we will actually show that in all these cases U M P unbiased test can be derived. So, as an application we will consider all the normal distribution problems for testing for the mean testing for the variance in the presence of the other parameter.

We will also look at certain applications to distribution such as exponential and others here of course, there is a concept of the an series statistic that will be utilized and completeness and the basu's theorem will also be helpful. These things we did it in the first part of this course namely, when we were discussing the point estimation we have introduced this concept. So, we need to look at those things again and we will be using these thing. So, I will continue the concept of unbiasedness in the testing problems in the following lecture.