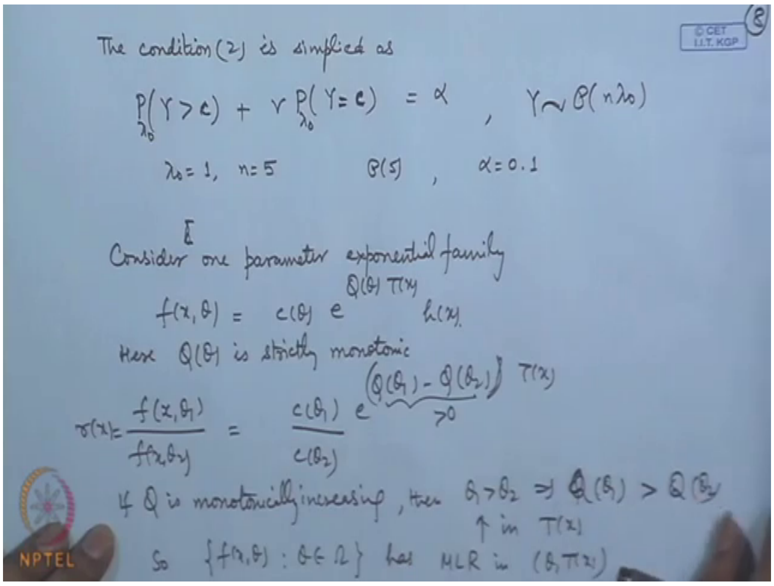
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 38
UMP Tests- II

Let me further develop this theory of the UMP Tests here.

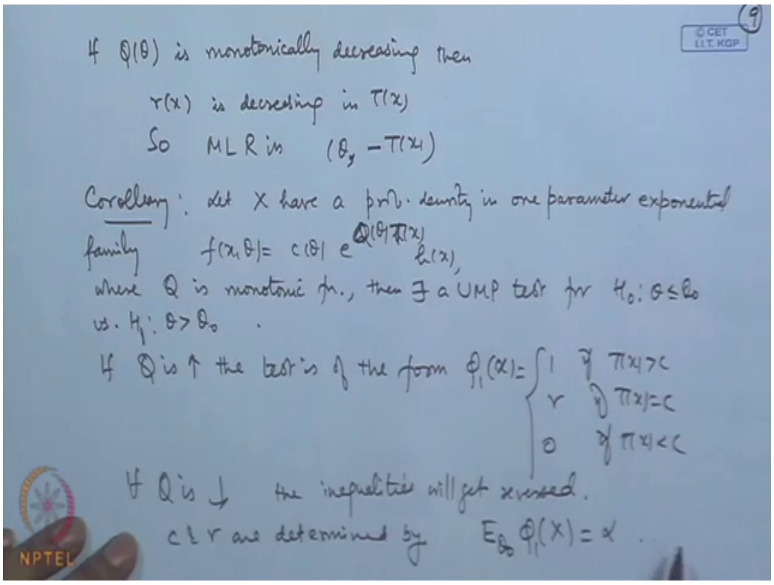
(Refer Slide Time: 00:33)



So, let us consider one parameter exponential family. So, we are considering the form of the probability mass function or the probability density function as c theta e to the power Q theta T x into h x. Here Q function is strictly monotonic function; that means, it could be monotonically increasing or monotonically decreasing. Let us write down the ratio f x theta 1 by f x theta 2. Then this is becoming c theta 1 e to the power Q theta 1 minus Q theta 2 T x and h x will get cancelled out by c theta 2.

Now, if Q is monotonically increasing then theta 1 greater than theta 2 will imply Q theta 1 greater than Q theta 2. That means, this term will become positive and you will have this as increasing function of increasing function of T x. So, this ratio becomes an increasing function of T x. So, the family effects theta this will have monotone likelihood ratio in theta T x; on the other hand if I consider say Q theta to be monotonically decreasing.

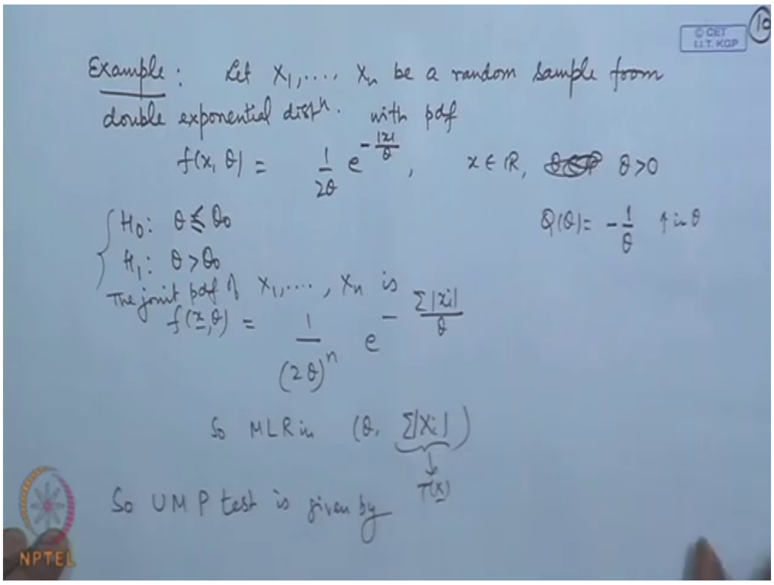
(Refer Slide Time: 02:39)



If Q theta is monotonically decreasing then this r x term what will happen here, that Q theta 1 will become less than Q theta 2 if theta 1 is greater than theta 2. Therefore, this term will become decreasing function of T x and therefore, monotone likelihood ratio will be in minus T x decreasing in T x. So, MLR will be in theta and minus T x; that means, the test function will get reversed in equality is like here we have T x greater than c it will become T x less than c.

So, as a corollary of the previous theorem we can write then let X have a probability density in one parameter exponential family, that is f x theta is equal to c theta e to the power Q theta T x into h x where Q is monotonic function. Then there exists a UMP test for H naught theta less than or equal to theta naught against theta greater than theta naught. If Q is increasing the test is of the form phi 1 x is equal to 1 if T x is greater than c. If T x is equal to c it is 0, if T x is less than c, if Q is decreasing the inequalities will get reversed. And here c and gamma are determined by expectation of theta naught phi 1 X is equal to alpha.

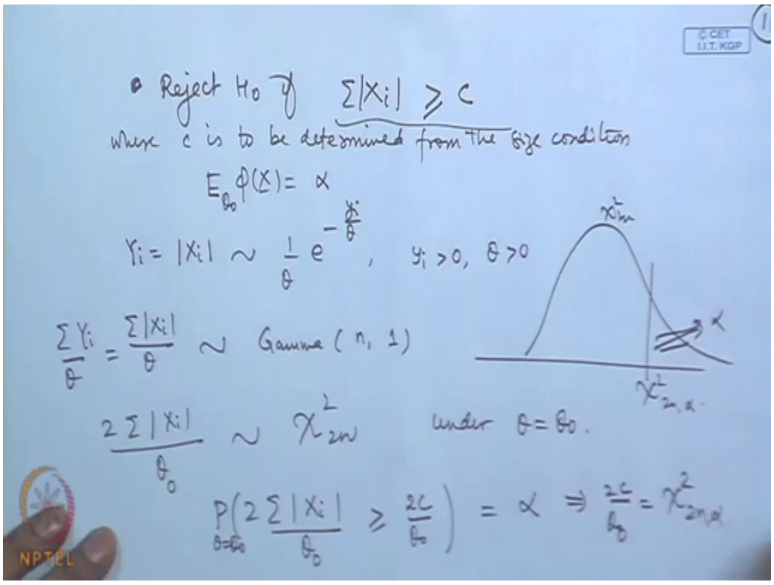
(Refer Slide Time: 05:43)



Let me consider one example let X 1 X 2 X n be a random sample from double exponential distribution with pdf given by say f x theta is equal to half e to the power minus modulus x by theta and here 1 by 2 theta. Here x is a real number and theta is any theta has to be a positive parameter here. Let us consider say theta is less than or equal to theta naught against theta greater than theta naught. You can easily see that this is a one parameter exponential family and the monotone likelihood ratio here you may consider Q theta as equal to minus 1 by theta.

So, naturally this is increasing in theta because 1 by theta is decreasing so, minus 1 by theta is increasing. So, this is strictly monotonic function. So, this suppose I do not the joint distribution of X 1 X 2 X n X 1 X 2 X n that is equal to 1 by 2 theta to the power n e to the power minus sigma. So, monotone likelihood ratio in theta and sigma X i this is T x sigma modulus of X i. Therefore, by an application of this corollary that I mentioned UMP test for this problem.

(Refer Slide Time: 08:11)

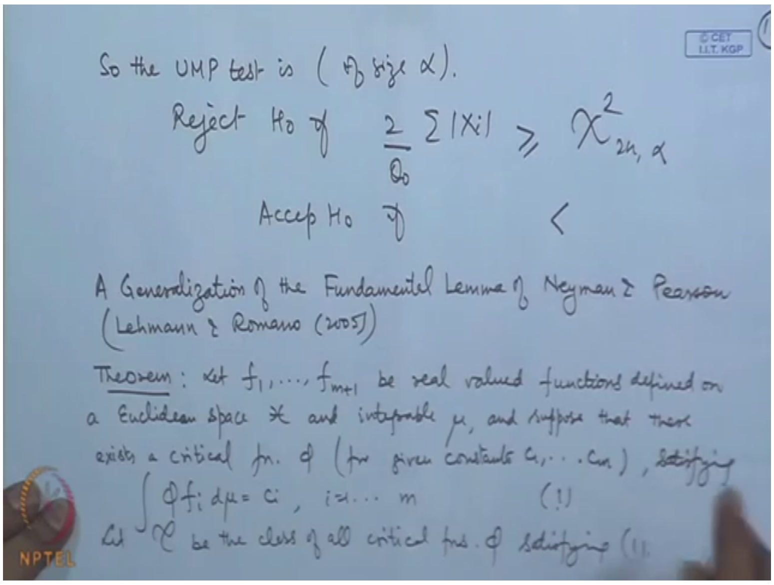


Let us say reject H naught if sigma modulus of X i is greater than c; now note here that we are dealing with the continuous distributions. So, I have written this part only, this part will be the rejection acceptance region. Now, sigma of modulus X i is equal to c we need not right this portion here because this will have probability 0. So, without loss of generality I am including the equality here. Now, what we need to do is to determine this, where c is to be determined from the size condition that is expectation of phi X is equal to alpha.

Now, let us consider say Y i is equal to modulus of X i, if X i is having double exponential distribution then modulus of x i will have simple exponential distribution that is 1 by theta e to the power minus Y i by theta. So, this will have distribution 1 by theta e to the power minus Y i by theta. So, sigma modulus of X i by theta that is Y i sigma Y i by theta that will have gamma distribution with parameters n and 1. That means, twice sigma modulus X i by theta naught that will follow chi square distribution on 2 n degrees of freedom under theta is equal to theta naught.

So, when we consider this size condition that is probability of sigma modulus X i by theta naught twice greater than or equal to some 2 c by theta naught this is equal to alpha, when theta is equal to theta naught. This implies that this 2 c by theta naught should be equal to chi square 2 n alpha; that means, n on the chi square curve with the 2 n degrees of freedom this probability is equal to alpha. So, this is chi square 2 n alpha.

(Refer Slide Time: 10:59)



So, that test is written as so, the UMP test is reject H naught if twice sigma modulus X i by theta naught is greater than or equal to chi square 2 n alpha. This is a UMP test of size alpha and of course, accept H naught if this is less. So, you can see this extension of the Neyman-Pearson theory to the families with the monotone likelihood ratio is helpful in providing the uniformly most powerful tests for one sided testing problems.

If the families have monotone likelihood ratio we are able to directly use these things here. And we are having exact test here; that means, once we have the observations and we our testing problem is clearly specified then at a given level of significance we can provide a decision whether we should accept a null hypothesis or not.

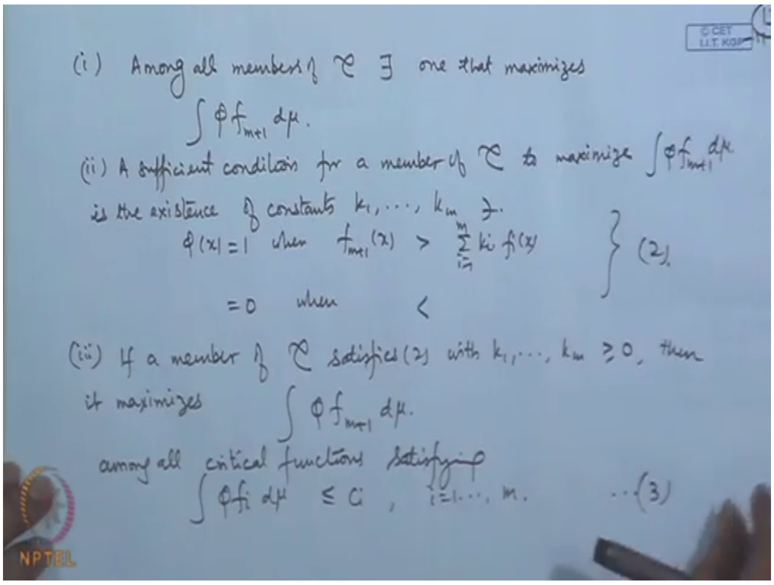
On the other hand if you do not specify alpha in advance then you can find out the p value here; now let me proceed further with this theory here. Now, for further extension of this theory of most powerful tests generalization of the Neyman-Pearson fundamental lemma was done. Let me state these results without any proof here and these results are used for solving further problem; that means, here were considering theta less than or equal to theta naught. So, it is strictly one sided and now there can be cases where we may have two sided also. For example, if I am considering se binomial proportion whether it lies in a range or it is outside a range.

Now, if I say it is within a range then it is like an interval, but if I say it is outside a range for example, I may say it is outside the interval 1 by 4 to 3 by 4. That means, I am saying the hypothesis is p less than or equal to 1 by 4 or p is greater than or equal to 3 by 4 is a two sided thing. Now, in families with the monotone likelihood ratio etcetera this Neyman-Pearson theory is applicable to this also. And then there is a another point that is regarding the determination of the constants in the test. In the one sided thing the maximum was occurring at the cut off point, that is theta naught here. When we have two sided then you will have two cut off points it will increase and then.

So, what will happen that we will consider the maximum value and at both the end points that is at both the points end points of the intervals. So, these results are proved using certain extended features of the Neyman-Pearson lemma. So, the result is known as the generalization of the fundamental lemma; let me give it here first of all. A generalisation of the fundamental lemma of Neyman and Pearson; this is statement and the proof one can find out in the book of Lehmann and Romano.

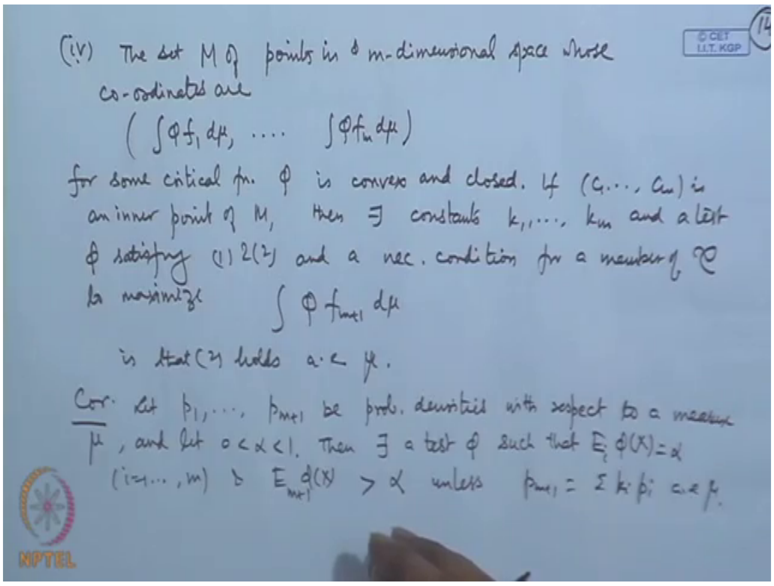
I will be skipping the details of the proof I will only give the statement here. Let f 1 f 2 f m plus 1 be a real valued functions defined on a Euclidean space x and integrable mu. And suppose that there exists a critical function phi for given constants c 1 c 2 c m satisfying integral phi f i d mu is equal to c i, i is equal to 1 to m. Let us say c is the class of all critical functions phi satisfying 1.

(Refer Slide Time: 16:37)



Then among all members of c there exists one that maximizes integral phi f m plus 1 d mu. A sufficient condition for a member of c to maximize integral phi f m plus 1 d mu is the existence of constants k 1 k 2 km such that phi x is equal to 1, when f m plus 1 x is greater than sigma k i f i x i is equal to 1 to m and it is equal to 0 when it is less. Thirdly if a member of c satisfies 2 with k 1 k 2 km greater than or equal to 0 then it maximizes integral phi f m plus 1 d mu. Among all critical functions satisfying phi f i d mu less than or equal to c i, for i is equal to 1 2 m.

(Refer Slide Time: 18:55)

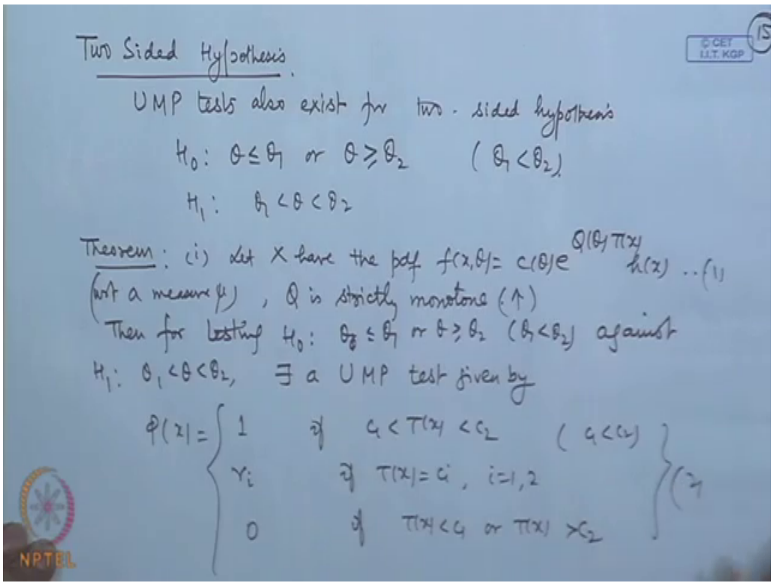


And then lastly the set M of points in the m dimensional space whose coordinates are say phi f 1 d mu and so, on phi f m d mu; for some critical function phi this is convex and closed. If c 1 c 2 c m is the inner point of M then there exist constants k 1 k 2 k m and a test phi satisfying 1 and 2. And a necessary condition for a member of c to maximize integral phi f m plus 1 d mu is that 2 holds almost everywhere.

As I mentioned I will not be giving the proof of these results, one can see the book of Lehmann. Now, this extension is helpful for solving more general testing problems, as a corollary I state the following. Let p 1 p 2 p m plus 1 be probability densities with respect to a measure mu and let 0 less than alpha less than 1. Then there exists a test phi such that, expectation of phi X is equal to alpha for i is equal to 1 2 m and expectation of phi X for m plus 1 it is greater than alpha unless of course, p m plus 1 is equal to sigma k i p i almost everywhere.

So, this will actually give the solution to more general two sided null hypothesis testing problems. So, we have the following result then that is if I am considering two sided hypothesis.

(Refer Slide Time: 22:11)



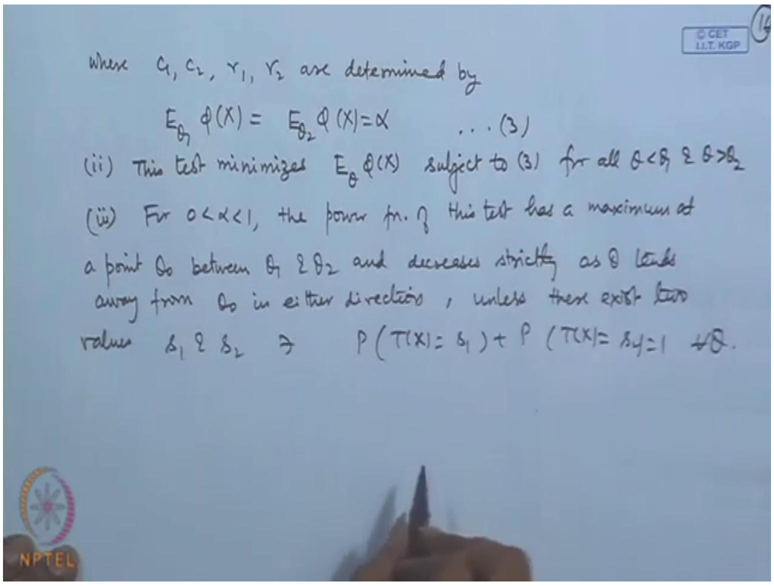
So, we can say that UMP tests also exist for certain two sided hypothesis of this nature H naught say theta less than or equal to theta 1 or theta greater than or equal to theta 2, where theta 1 is less than theta 2. So, we may like to test whether for example, say theta is the error measurements. So, we may like to check whether the error measurement lie within a certain range or they are outside a range.

It could be like we are producing certain items and say certain ball bearings are being produced and we are looking at the diameter of the ball bearings. So, whether the ball bearings diameters are within a range or it is outside the range. If it is within the range we will be accepting the product as the good item, if it is outside then will be rejecting that. So, therefore, this is a perfect case for the two sided testing hypothesis problems; we may have say H 1 as theta 1 less than theta less than theta 2. So, the result is that by the use the generalization of the Neyman-Pearson fundamental lemma we can actually give the uniformly most powerful test for these situations also.

So, we have the following theorem which I will state; let X have the probability density function with respect to a measure mu and Q is strictly monotone. Then for testing theta 1 theta less than or equal to theta 1 or theta greater than or equal to theta to where theta is less than theta 2 against the alternatives theta 1 less than theta; less than theta 2, there exists a uniformly most powerful test. Of course, here again c 1 has to be less than c 2 it is gamma i T x is equal to c i for i is equal to 1 2. So, there are two points of randomisation here and we accept if T x is less than c 1 or T x is greater than c 2.

Once again if you are considered to be strictly monotonic then the family of distributions has monotone likelihood ratio in theta T x or theta minus T x. And therefore so, here I have taken for example, increasing say because we are writing the region in the rejection region in this one. So, I am considering monotonically increasing. So, we are rejecting when the value lies between 2 ranges and we are accepting for smaller values of T x or larger value of T x. If it is decreasing then the inequalities will get reversed and at the boundary points of the interval we have done the randomisation.

(Refer Slide Time: 26:35)



Here so, let me consider this as 1 2 say where this constants c 1 c 2 gamma 1 gamma 2 they are determined by expectation theta 1 phi X is equal to expectation theta 2 phi X is equal to alpha. This test minimises expectation phi X subject to 3 for all theta than theta 1 and theta greater than theta 2. And for 0 less than alpha less than 1 the power function of this test has a maximum at a point theta naught between theta 1 and theta 2 and decreases strictly as theta tends away from theta naught in either direction.

Unless of course, there exist 2 values say s 1 and s 2 such that probability of T x is equal to s 1 plus probability of T x is equal to s 2 is equal to 1 for all theta. So, here you can see the probability of type 1 error will be maximized in the at the end points that is at theta 1 and at theta 2 that is why we are fixing that value equal to alpha. So, this is the size condition in the two sided null hypothesis problem. Then we have one sided hypothesis problem, then the maximum value is occurring at the cut off point that cut off point where the null and alternative hypothesis points are changing. But, when we have two sided then we will have 2 points; one is below and another is above. And at both the points we are having the maximum value of the probability of type 1 error, that value we are fixing has the alpha value.

In the next lecture I will be considering further amplifications of these results certain applications of this and we will also consider certain properties of this power function here which are based on actually the monotone likelihood ratio property. So, basically the properties of the expectations I will be discussing it in the next lecture.