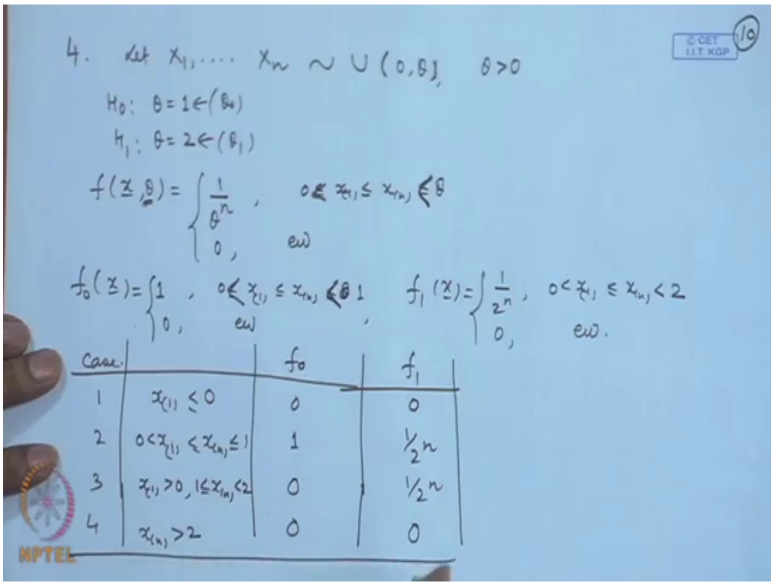
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 36
Application of NP – Lemma – II

Let me take another example, in which the range of the variable may be dependent upon the range of the parameter and let us see in that case how the Neyman Pearson lemma works.

(Refer Slide Time: 00:39)



Let us consider say X 1, X 2, X n from uniform 0 theta distribution and we consider a hypothesis testing problem say theta is equal to say 1 against say theta is equal to 2. We may also right here say theta naught and here I may right theta 1 and then I may consider the case theta naught less than theta 1 or theta not greater than theta 1. So, for convenience we have considered this is special case; theta is equal to 1 and theta is equal to 2.

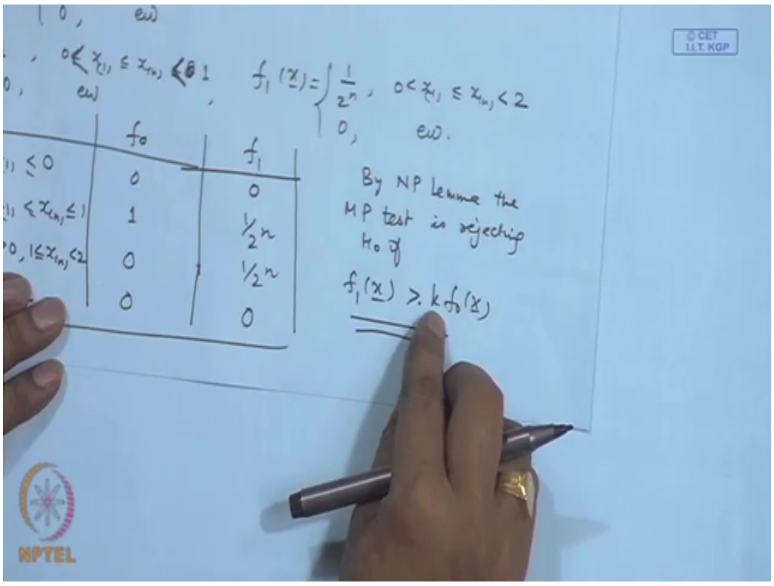
So, let us consider the most powerful test here. So, the joint distribution is 1 by theta to the power n, and here the range of the variables is from 0 to theta. So, we write it in this particular form. So, when we write for f naught and f 1 for f naught this is simply 1. So, this is simply; see we may if we write here open interval then we need not put equality here we may put it like this otherwise may put a quality. The probability of those points will be 0 so it does not make any difference.

Similarly if we consider f 1, then under f 1 theta is equal to 2. So, it will become 1 by 2 to the power n 0 less than x 1 less than or equal to x n less than 2 it is equal to 0. I was mentioning here that the range of the densities where the two densities are positive is not the same. Here you can see this density is positive for 0 to 1 and this density is positive for 0 to 2. So, let us look at the various cases of f 1 and f naught ok. So, we will make it in the form of a table, let us consider say case 1 2 3 4 like that we will write 1 2 3 4.

So, I will write all the cases which may be trivial and nontrivial cases. If we observe x 1 to be less than 0; obviously, this is not possible. So, both the densities f 1 and f naught they are 0 here. If we considered the case 0 is less than, so x 1 and x n is less than or equal to 1, in this case the first density is 1 and second density is 1 by 2 to the power n. Then we may have say x 1 of course, may be greater than 0, but x n is say beyond it is beyond 1 in that case what will happen? That this first density becomes 0; however, the second density remains 1 by 2 to the power n.

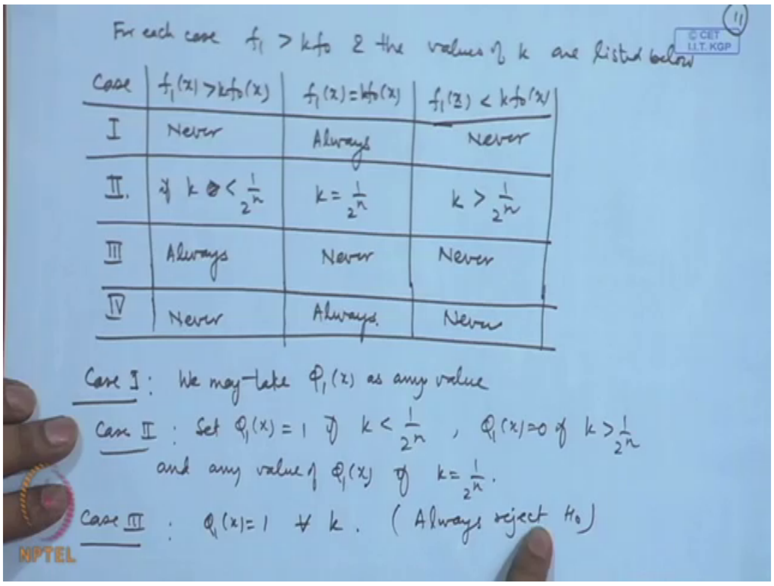
And then we may have the extreme case that is x n is greater than 2, then this is 0 and this is 0. So, broadly speaking we have to consider the f 1 greater than k f naught from the Neyman Pearson lemma under these 4 cases.

(Refer Slide Time: 04:35)



By NP lemma, the MP test is rejecting H naught if f 1 x is greater than k times f naught x. So, here the values of k we have to choose; so, let us write here.

(Refer Slide Time: 05:05)



For each case f 1 greater than k f naught and the values of k are listed below; so let us consider each case. In the case 1 if we consider f 1 x greater than if naught x now both the values are 0. So, this is never possible whatever be the value of k it is never possible; f 1 x is equal to f k times f naught x that is always true and f 1 x less than k f naught x is never possible. Let us consider second case, in the second case if we look at f 1 that is 1 by 2 to the power n greater than k times f naught.

So, this condition is true if k is greater than sorry k is less than 1 by 2 to the power n, this is true if k is equal to 1 by 2 to the power n, this is true if k is greater than 1 by 2 to the power n, let us look at the third case. In the third case f naught is 0. So, f 1 greater than k f naught is always true, and therefore this equality are less than is never possible. Let us consider the case 4 once again both of them are 0. So, inequality is never possible whereas, the equality is always true.

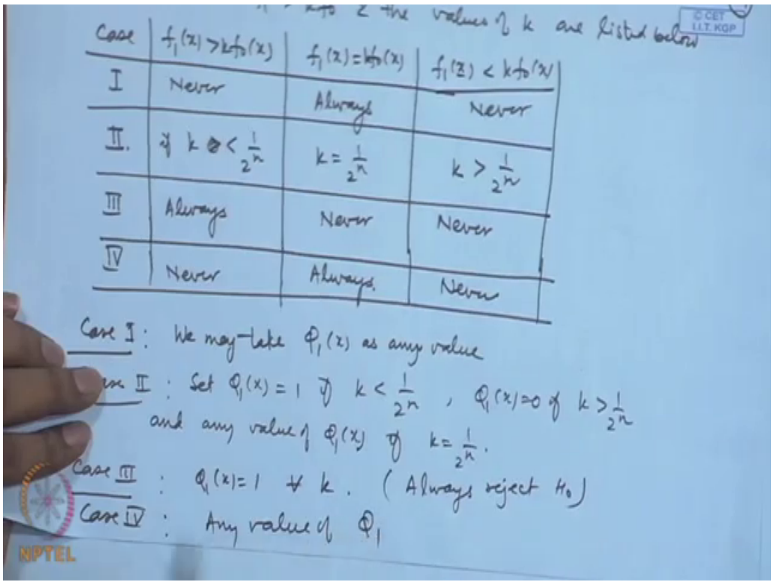
So, now based on this we should tell when to reject H naught and when to accept H naught; that means, dependent upon these 4 cases and the choices of k, we should give what is the test function. And at the same time we should also tell that whether the probability of type 1 error is equal to alpha can be achieved for a given value of alpha.

So, what we consider in case 1? Since f 1 is equal to k f naught is always possible always true therefore, whatever be the value of phi 1, it does not make any difference.

So, we may take phi 1 as any value. In case 2 if k is less than 1 by 2 to the power n, in this case f 1 is greater than k f naught that means this is the corresponding case 2 rejecting H naught. So, if case is less than 1 by 2 to the power n we will say, reject H naught and in this case we will say except h naught; that means, phi 1 is equal to 0. However, when k is equal to 1 by 2 to the power n can we may again say we may accept or reject depends upon we can assign something of course, the probability of this case will be 0.

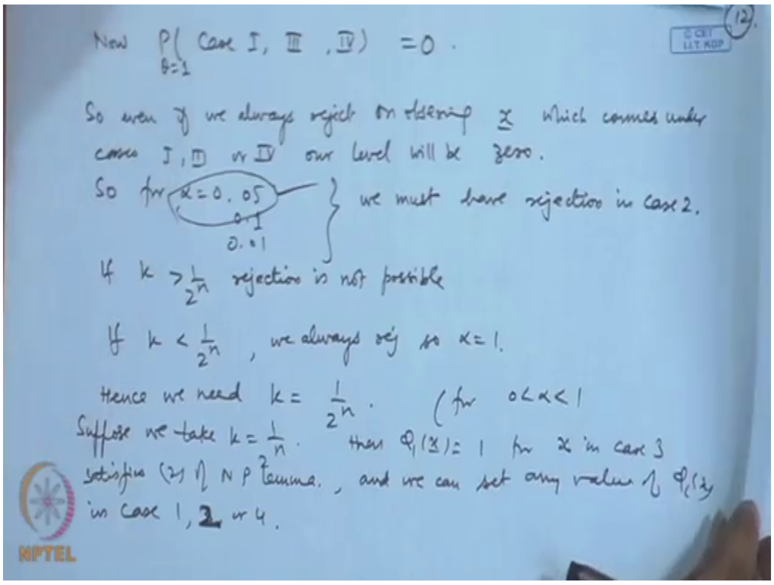
So, we can say phi 1 x that is a test function is equal to 1, if case less than 1 by 2 to the power n it is equal to 0, if k is greater than 1 by 2 to the power n and any value of phi 1 if k is equal to 1 by 2 to the power n. If you look at case 3, in the case 3 this condition is always true. So, we always reject that is phi 1 x is equal to 1 whatever v k that is always a reject H naught. Note here that these are also you restrict because what is happening? If we are getting the third case that is x n is between 1 and 2 that mean naturally the observations are from the density uniform 0 to 2, otherwise observation cannot be greater than 1 and therefore, we should definitely reject H naught and accept H 1.

(Refer Slide Time: 10:19)



And in the case 4 once again we may put any value; any value of phi 1.

(Refer Slide Time: 10:33)



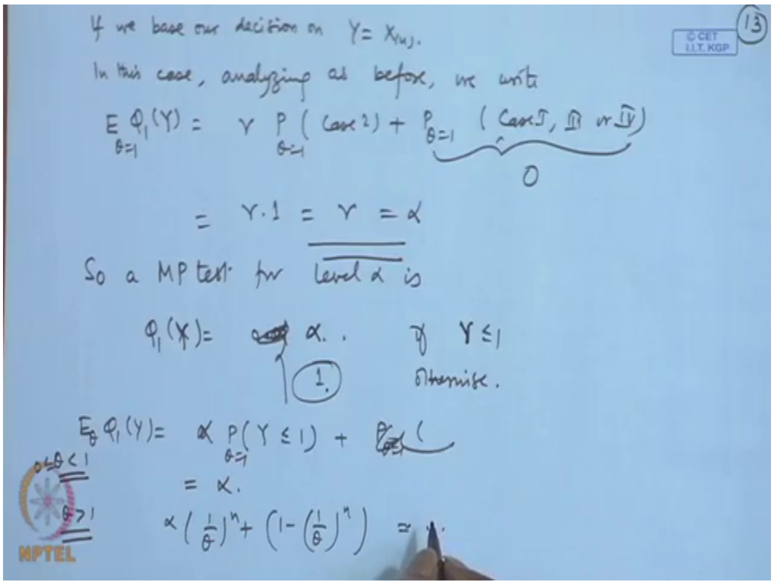
Now, probability of case 1 3 and 4 under the null hypothesis this is 0. So, even if we always reject on observing x which comes under cases I, III or IV our level will be 0. So, for alpha is equal to 0.05 at sector some given value of alpha or say 0.01 or 0.1 etcetera then what we should do? We should make the probability of rejection in case 2 to be possible.

So, we must have rejection in case 2. Now again if I take k to be greater than 1 by 2 to the power n rejection is not possible. So, the only rejection is possible for k less than 1 by 2 to the power n here rejection is becoming always true; so alpha will become 1 which is not acceptable. So, therefore, we should have this k equal to 1 by 2 to the power n as a possible value. If k is greater than 1 by 2 to the power n rejection is not possible. If k is less than 1 by 2 to the power n, we always reject.

So, alpha is equal to 1 therefore, we should have k equal to 1 by 2 to the power n for 0 less than alpha less than 1. So, suppose we take k equal to 1 by 2 to the power n, then phi 1 x is equal to 1 for x in case 3 satisfies 2 of N P lemma and we can set any value of phi 1 in case 1 2 1 3 or 4 sorry 1 2 or 4.

Let me give concrete case here when I base our decision on x n alone. You note here that when I am considering x 1 x 2 x n a random sample from you know from 0 to theta, the sufficient statistics actually x n and x n is actually playing the role here as you have already noticed here. So, let me explain that part in detail.

(Refer Slide Time: 14:23)



If we base our decision on say Y is equal to X n. So, in this case what is happening that let me write here.

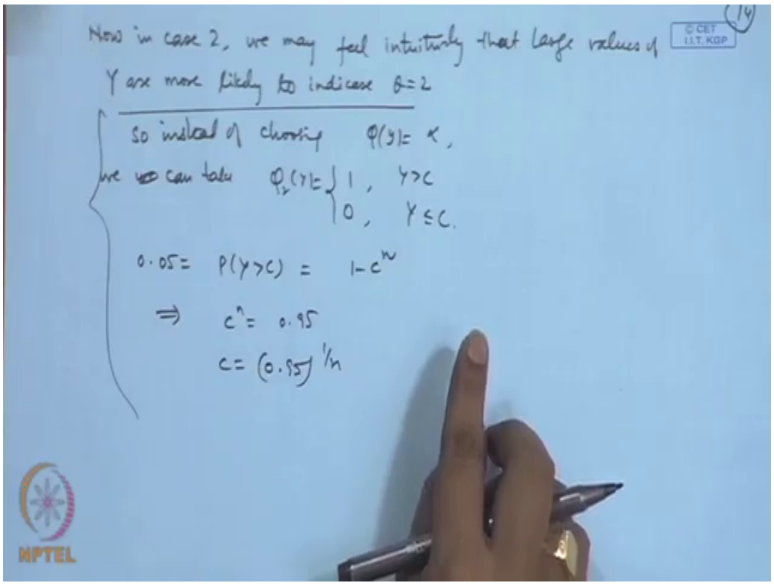
In this case analyzing as before we write expectation of phi 1 y for theta is equal to 1 as some gamma times probability of case 2, under theta is equal to 1 plus probability of theta is equal to 1 under case 1 3 or 4 now these are all 0. So, that is equal to gamma into 1 that is equal to gamma. So, we should choose gamma is equal to alpha. So, a most powerful test for level alpha is phi 1 x or phi 1 y is equal to point 0 that is alpha if Y is less than or equal to 1 it is equal to g 1 otherwise.

That means what we are saying? Reject all the time except for the case when y is less than or equal to 1. If y is less than or equal to 1 then you are rejecting with probability 0.05 and otherwise you are accepting we may also consider the power function here.

So, for example, power function here that is equal to point alpha into probability of say Y less than or equal to 1 plus probability. So, here I am taking theta is equal to 1 into. So, those cases will not occur this will have probability 0. So, this is actually equal to alpha for this is theta less than 1 and if I take theta greater than 1, then it will become equal to alpha into 1 by theta to the power n plus 1 minus 1 by theta to the power n. So, that is equal to well you can simplify this value here.

So, what we are able to do is that, we are able to provide an exact test here for testing parameters in the uniform distribution. We may also consider in a slightly different fashion let me just explain it here.

(Refer Slide Time: 17:51)



Case 2 we may feel intuitively that large values of Y are more likely to indicate theta is equal to 2.

So, instead of choosing phi y is equal to alpha, we can take say phi 2 y is equal to 1 if y is greater than c it is equal to 0 y is less than or equal to c and if you consider the probability of this. So, we are getting here say 0.05 is equal to probability of y greater than C that is equal to 1 minus C to the power n this means we can take C to the power n is equal to 0.95 or C is equal to 0.95 to the power 1 by n.

So, this is an alternative solution here of course, this is based on the heuristic consideration that large values of alpha of y are more likely 2. So, this part is not coming from NP lemma, in the NP lemma if we write exactly will take that part and the test function is of this nature that phi 1 y is equal to say alpha, if Y is less than or equal to 1 and it is 1 otherwise. So, these are the two forms that I am considering here.

Friends today we have considered in detail various applications of the Neyman Pearson fundamental lemma, how it gives exact tests for testing simple hypothesis works as a simple hypothesis. The important point that you should note here is that we need the distribution of the criteria. That means our criteria is based on certain function of the random variable which we call test a statistic, we should be able to say something about the distribution of that under the null hypothesis then only the constant k can be determined. If we are unable to determine that then we will not be able to provide the exact form of the test function.

So, in the next lecturer as I mention we will consider extension of the Neyman Pearson lemma to consider the composite hypothesis also. So, in particular we will consider the one sided composite hypothesis testing problems so, that I will be taking up in the following lecture.