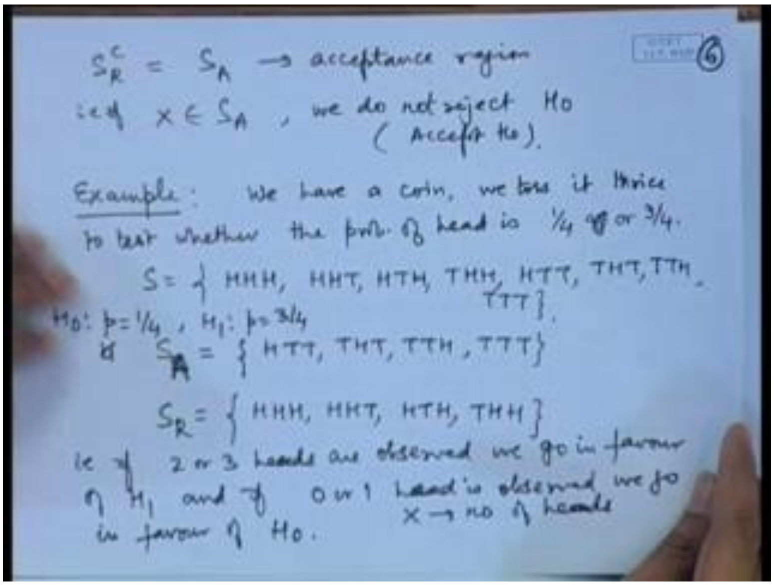
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 67
Neyman-Pearson Fundamental Lemma

So, we continue our discussion of the problem of testing of hypothesis. So, I framed it in the following terminology, we should have a null hypothesis, we should have an alternative hypothesis and then we take a random sample and we split the sample space into two portions; one portion is called the rejection region and another is called the expectance region.

As a consequence, we are likely to commit errors of two types; we call them type one error and type two error and we have the respective probabilities. I mentioned that in the case of composite hypothesis, the probabilities of type one error and two errors will be the functions of the parameters. So, the most desirable would have been to have both the type one error and type two error probabilities to be as small as possible, but as in a two dimensional decision species or you can say the two dimensional species not ordered therefore, it is not possible to minimize both of them. So, a practical approach is to keep the value helpful for to a fixed level and then find that test procedure for which beta is minimized or minus beta is maximized.

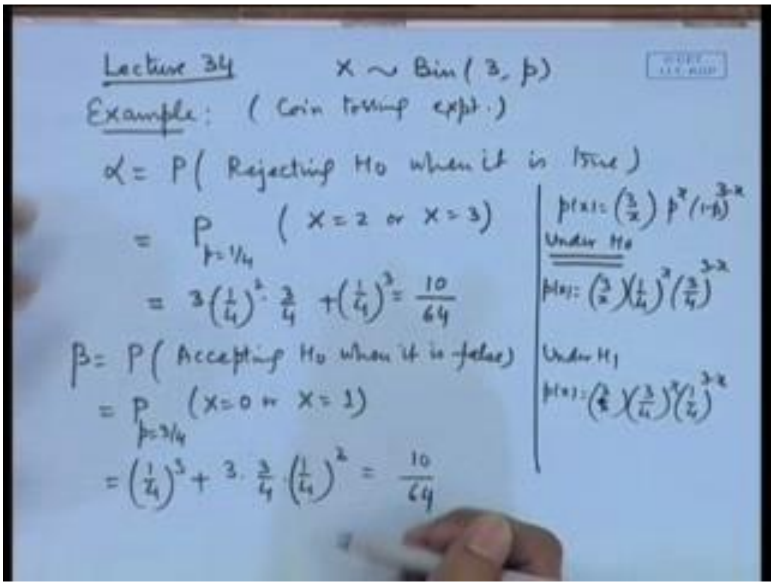
(Refer Slide Time: 01:54)



Let me explain this through one example, we discussed the problem of say checking the unbiasedness or say certain probability related to a probability of head of in a coin tossing experiment.

So, we have a coin and we tossed it thrice and we want to test the hypothesis whether p is equal to 1 by 4 against p is equal to 3 by 4. So, I have given here one region that acceptance region is that when either 0 or 1 hat is observed and we reject H naught when 2 or 3 heads are observed. Let us calculate the probabilities of type one error and type two error for this problem.

(Refer Slide Time: 02:42)

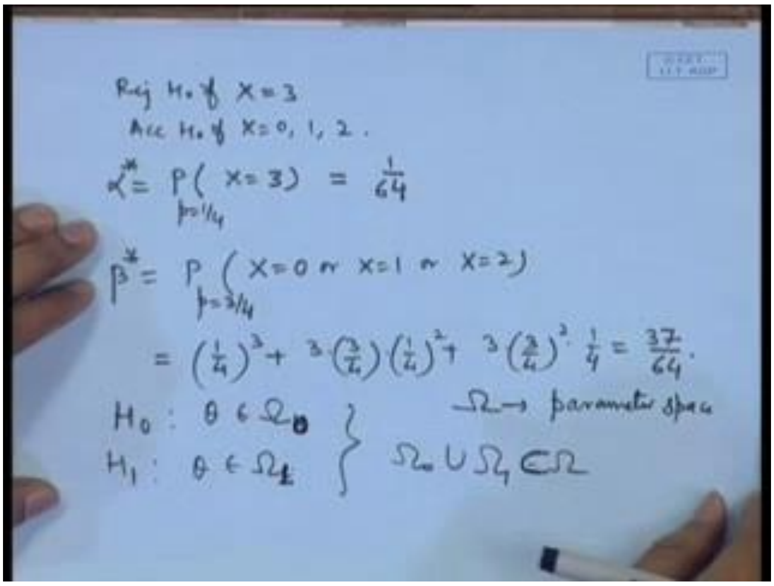


So, coin tossing experiment, so here alpha that is the probability of type one error rejecting H naught when it is true. So, we can restrict attention to the random variable X that is the number of heads. So, here X follows binomial 3 p because in the 3 tosses of the coin, you may have at the most 3 heads, so 0, 1, 2, 3. So, it is a binomial distribution the head occurs with a probability p.

So, when it is true means p is equal to 1 by 4; under this we are rejecting when H when x is either 2 or x is equal to 3. So, this is basically reducing to the probability of X equal to 2 or X equal to 3, so now this probabilities can be evaluated because we know the distribution of X that is p x is 3 c x; p to the power x 1 minus p to the power 3 minus x. Now under H naught; this p x function will be equal to 3 c x; 1 by 4 to the power x; 3 by 4 to the power 3 minus x. So, when I substitute x is equal to 2 here; I get 3, 1 by 4 is square into 3 by 4 plus, when I put x equal to 3 here this is simply reducing to 1 by 4 cube, so, that is equal to 10 by 64.

Let us look at beta that is the probability of accepting H naught when it is false; that is probability of p is equal to 3 by 4 when X is equal to 0 or X is equal to 1. Now under H 1 that is when p is equal to 3 by 4 p x is 3 c x; 3 by 4 to the power x; 1 by 4 to the power 3 minus x. So, when X is equal to 0; this value is simply 1 by 4 cube plus when X is equal to 1 it is 3 into 3 by 4 into 1 by 4 is square. So, that is equal to 10 by 64, so in this particular situation you can see alpha is 10 by 64 and beta is equal to 10 by 64; the probabilities of, now you see we suppose we try to reduce alpha; we may try to reduce alpha by taking another test.

(Refer Slide Time: 06:13)



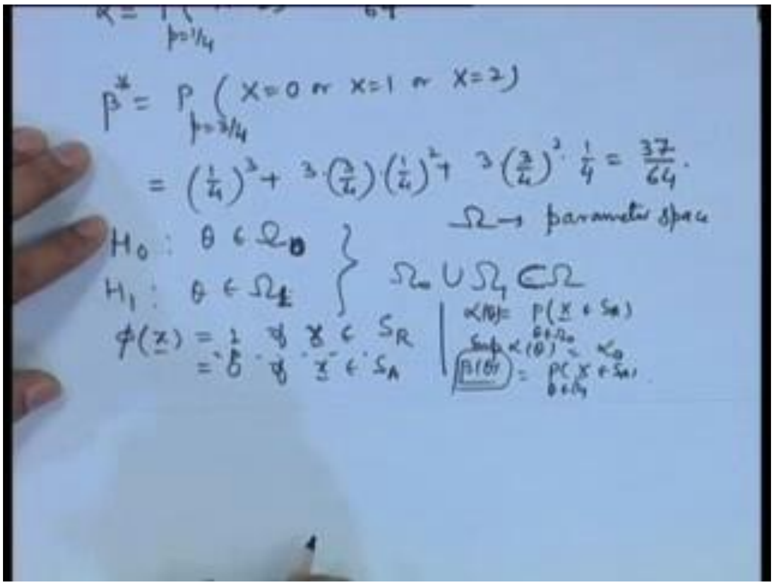
So, suppose I say reject H naught if X is equal to 3, accept H naught if X is equal to 0, 1 or 2. Now let us see what is the value of alpha; let me call it alpha is star that is the probability of X is equal to 3; when p is equal to 1 by 4. So when p is equal to 1 by 4, we noted down the distribution here 3 c x 1 by 4 to the power x 3 by 4 to the power 3 minus x. If we substitute x is equal to 3, here I get 1 by 4 cube that is 1 by 64, so naturally you can see here that this test is having alpha is equal to 10 by 64; this is having 1 by 64, so this is having a much smaller probability of type one error, but now let us see what happens to the probability of type two error; beta is star that is probability of x is equal to 0 or X is equal to 1 or X is equal to 2 under p is equal to 3 by 4.

So, when p is equal to 3 by 4, the probability distribution of x is given by 3 c x 3 by 4 to the power x 1 by 4 to the power 3 minus x. So, this will be equal to 1 by 4 cube plus 3 into 3 by 4 into 1 by 4 is square plus 3 c 2; that is 3; 3 by 4 is square into 1 by 4. So, that is equal to now you see this value turns out to be 9, 27, 27 plus 9; 36 this is becoming 37 by 64. So, compare this earlier you had the probability of type two error as 10 by 64, but as a consequence of reducing the probability of type one error, the probability of type two error has suited up, it has become 37 by 64.

So, this is the problem which I was mentioning that if we try to reduce one type of error, the other type of error increases very much. Therefore, a compromise solution is that we keep a maximum level for one type of error; that means, we say we p s sin that the probability of say type one error should not go beyond a point and then among all the other test procedures which have the same maximum level of the type one error, we choose that one which has the a smallest type two error. So, that gives us the concept of the most powerful test procedure. So, in the most general terms the theory would be represented like this that we have H naught theta belonging to say omega h. So, our parameter space is omega; the full parameter space, you have the hypothesis testing problem as theta belonging to omega H against theta belonging to omega.

So, let me put omega naught and omega 1, so here omega naught; union omega 1 may be omega or it is not necessary; it may be actually a subset also because in case we are dealing with the simple hypothesis; in that case the full parameter is space need to be necessarily this one.

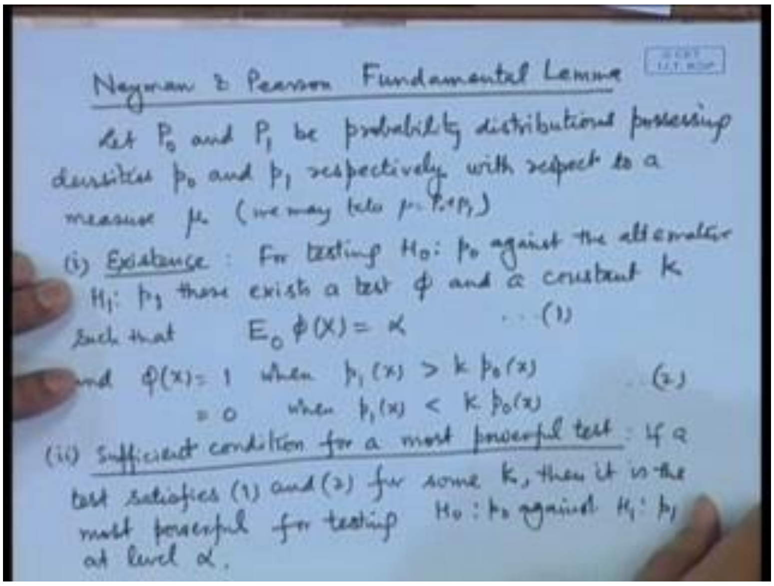
(Refer Slide Time: 10:04)



So, the procedure that we are trying to tell here is that we are devising a function phi x based on the sample. So, we are saying phi x is equal to 1; if x belongs to say S R it is equal to 0, if x belongs to S A, but in some cases as I mentioned we may go for randomization also, we may put some value p here for certain region. So, the probability of type one error that is probability that x belongs to S R when theta belongs to omega naught, so we take the maximum of this.

So, supremum of alpha theta that let us call it say alpha naught or alpha star; we choose that and then we try to minimize beta theta that is probability of x belonging to S A then theta belongs to omega 1.

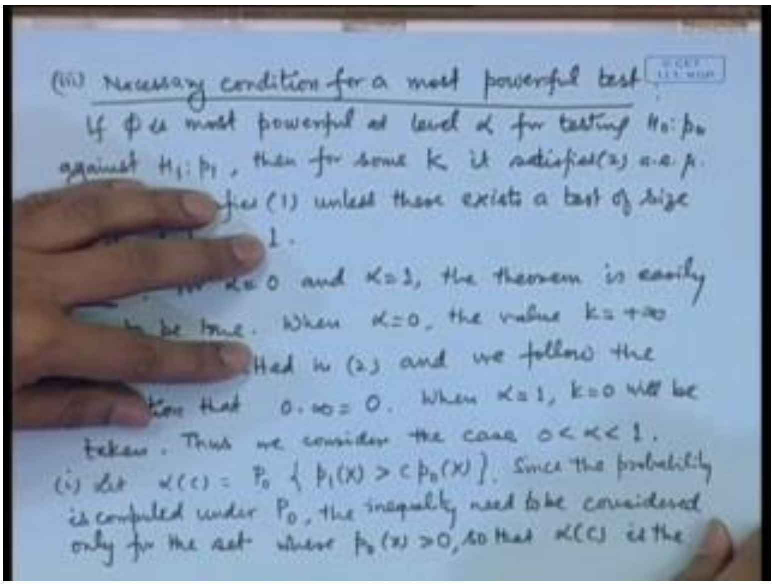
(Refer Slide Time: 11:17)



So, this optimization problem has been dealt with and the basic result in this regard is by Neyman and Pearson and the result is known as popularly Neyman and Pearson fundamental lemma also it is called n p lemma. This fundamental lemma which was given in 1927 by Statistician judge Neyman and (Refer Time: 11:50) Pearson. This initially dealt with the cases when we are having simple versus simple case, so the theorem is as follows; let P naught and P 1 be probability distributions possessing densities p naught and p 1 respectively with respect to a measure mu.

We may take say mu is equal to P naught plus P 1 also. So, the first part is existence for testing H naught that is p naught against the alternative H 1; that is p 1; there exists a test phi and a constant k such that expectation of phi X is equal to alpha and phi x is equal to 1, when p 1 x is greater than k; p naught x it is equal to 0, when p 1 x is less then k; p naught x. Second is sufficient condition for a most powerful test, if a test satisfies 1 and 2 for some k, then it is the most powerful for testing H naught; p naught against H 1, p 1 at level alpha.

(Refer Slide Time: 15:22)



The third is necessary condition for a most powerful test. If phi is most powerful at level alpha for testing H naught p naught against H 1; p 1, then for some k, it satisfies 2 almost everywhere mu. It also satisfies 1 unless there exists a test of size less then alpha and power 1. So, we see here first of all that this lemma is very powerful in the sense that if I am having a simple hypothesis versus a simple hypothesis testing problem, then the first thing it tells is that there is a test with a given size then secondly, if that test is of that form and it has that given size then it is the most powerful conversely if there is a most powerful test then that must be of this particular form.

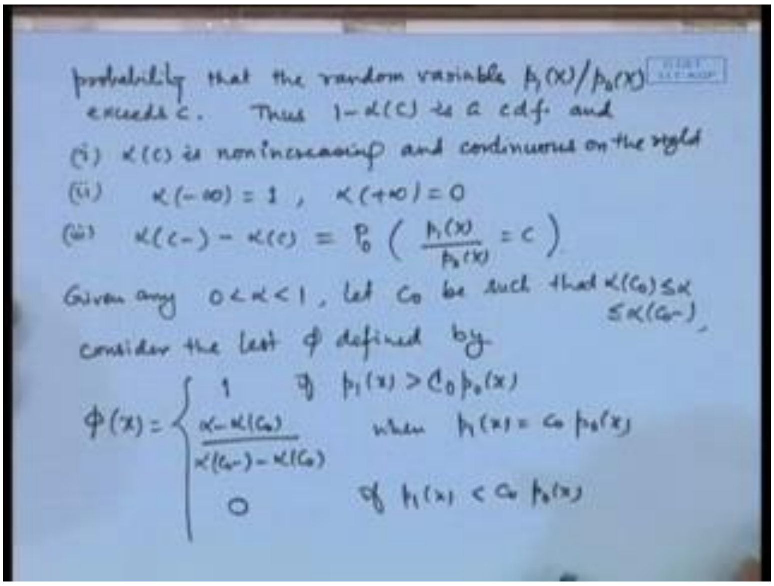
So in that sense, it is a very important result or you can say very powerful result which actually gives you the optimal solution in the case of simple versus simple hypothesis testing problems. So, let me look at the proof of this and then we will look at certain applications here. For alpha is equal to 0 and alpha is equal to 1, the theorem is easily seen to be true when alpha is equal to 0; the value k is equal to plus infinity has to be admitted in 2 and we follow the convention that 0 into infinity is equal to 0. When alpha is equal to 1, k equal to 0 will be taken. Let us look at this two choices, when alpha is equal to 0; that means, I want the probability of type one error to be 0, when will that happen; that means, probability of rejecting; that means, we should never reject, if we do not reject then this value should be infinity.

Otherwise so if this is infinity then right hand side is infinite; that means, always this condition will be true that is p 1 x is less then infinite and therefore, you will always be accepting H naught. So, the probability of type one error will become 0, so this condition is also satisfied and the whole thing is true basically because in this case, when you will look at the probability of type two error, that is probability of accepting H naught that will become 1 because you are always accepting, so the power is 1; so naturally it is the most powerful test.

Also we see the case of alpha is equal to 1; now alpha is equal to 1 will happen when I take k equal to 0 say if I take k equal to 0 this side is 0; that means, p 1 x is greater than 0 is always satisfied. Therefore you are always rejecting H naught, when you always reject H naught then the probability of type one error is 1. Now in this case what is happening to the probability of type two error, if you are always rejecting H naught; then the probability of accepting H naught will become 0 because you are never accepting that because you are always rejecting, so you are never accepting.

So, this gives you beta is equal to 0, so these are the trivial cases. Now let us look at the conventional cases when; so, let us define a function alpha c is equal to P naught that is the probability under H naught when p 1 x is greater than c p naught x. Since the probability is computed under P naught, the inequality need to be considered only for the set where p naught x is a strictly positive.

(Refer Slide Time: 22:31)



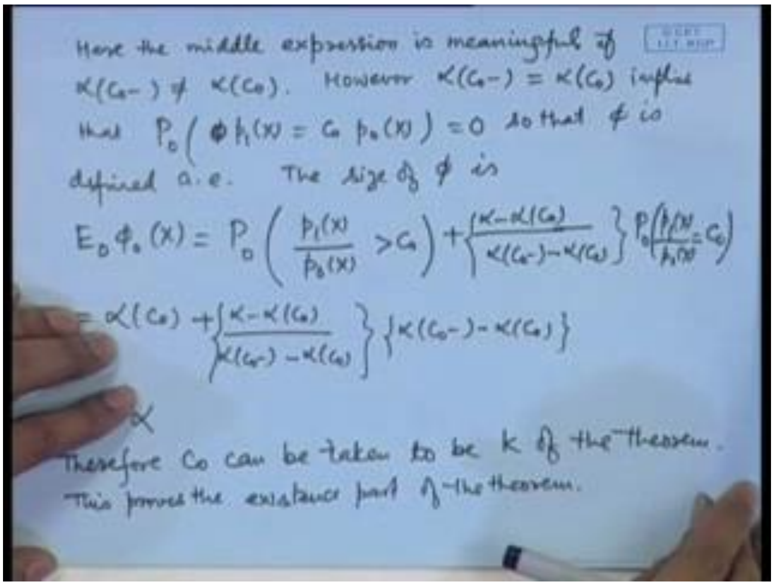
So, that alpha c is the probability that the random variable p 1 x by p naught x exceeds c, thus 1 minus alpha c is a cumulative distribution function and we have the following properties; that is alpha c is non-increasing and continuous on the right that is the pro properties of the cdf. So, if 1 minus alpha c is non-decreasing when alpha c will be non-increasing. Secondly, alpha of minus infinity will be 1; that is the limit of alpha c has c tends to minus infinity because 1 minus alpha c is cdf and alpha at plus infinity will become equal to 0. The third is that alpha c minus; minus alpha c that is the left hand limit at c minus alpha c that is the probability that p 1 x by p naught x is equal to c.

So, given any alpha such that alpha is between 0 and 1, let c naught be such that alpha c naught is less than or equal to alpha, less than or equal to alpha c naught minus.

Consider the test phi defined by, so we define phi x is equal to 1; if P 1 x is greater than c naught; p naught x and we define alpha minus alpha c naught divided by alpha c naught minus minus alpha c naught; this denotes the left hand limit at c naught; when p 1 x is equal to c naught p naught x. So, this is the randomization that I was mentioning earlier that when there is equality we put some value because finally, we want to achieve the power alpha; the size alpha and it is 0; if p 1 x is a strictly less then c naught p naught x.

Now you compare this conditions with the original function, we defined here the phi x is equal to 1 when p 1 x is greater than k p naught and it is equal to 0 when p 1 x is less then k p naught. So, if you compare this greater and less conditions are exactly matching here, so only we have introduce one quantity for equality that is the randomization point which may be required in the case of discrete distributions. So and of course, as I mentioned this is meaningful only when alpha c naught is not equal to alpha c naught minus because if it is a continuous distribution; this will be 0. So, you do not need to define this thing; that means, this is not useful because the probability of this event will be actually 0. Only in the case of discrete distribution, when the c naught is having a positive probability for the function p 1 x by p naught x; then this value will be of use.

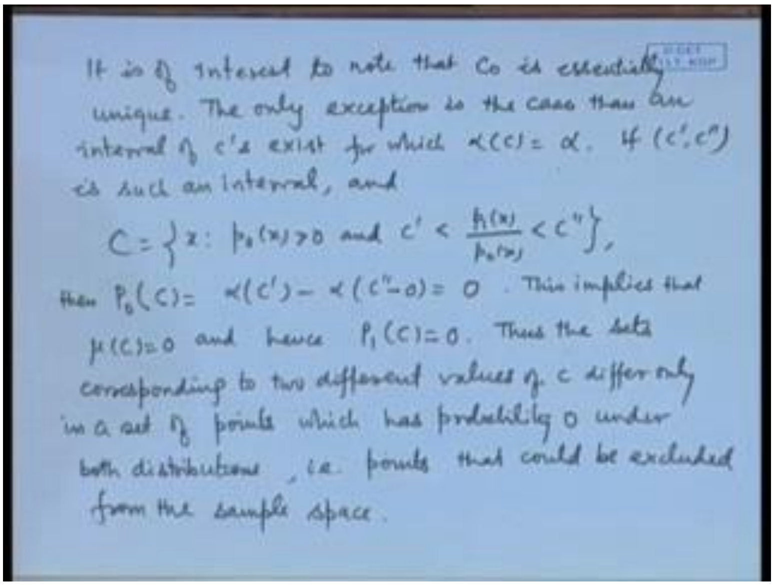
(Refer Slide Time: 26:27)



Let me write that comment here; here the middle expression is meaningful if alpha c naught minus is not equal to alpha c naught; however, alpha c naught minus is equal to alpha c naught implies that P naught phi 1; p 1 x is equal to c naught p naught x is equal to 0 so that phi is defined almost everywhere. Now let us look at the size of phi, that is the probability of rejecting when H naught is true; that is probability of p 1 x by p naught x greater then c naught plus alpha minus alpha c naught divided by alpha c naught minus minus alpha c naught into probability of p 1 x by p naught x is equal to c naught.

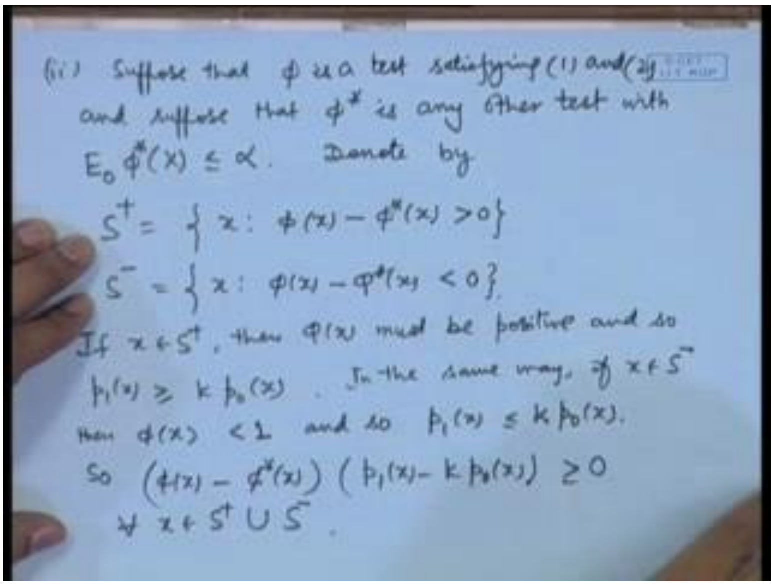
So, by the definition here this is alpha c naught plus alpha minus alpha c naught divided by alpha c naught minus; minus alpha c naught and this value is again alpha c naught minus; minus alpha c naught. So, this term cancels with this and this cancels with this, so this is actually reducing to alpha. Therefore, c naught can be taken to be k of the theorem, so this proves the existence part of the theorem because we have exhibited that there exists a test which has size equal to alpha of a given type because we fix the type also here in the existence part that there exist a test of this type. So, of course this was not complete because this not take care of the equality part, so we defined that part here and it is having this power alpha; so this k value is well defined; here this proves the existence, let me pay some attention to this value c naught here.

(Refer Slide Time: 29:37)



Now, it is of interest to note that c naught is essentially unique; the only exception is the case that an interval of c's exist for which alpha c may be equal to alpha. So, if c prime to c double prime is such an interval and c is equal to x such that p naught x is greater than 0 and c prime is less than p 1 x by p naught x is less than c double prime, then P naught c is equal to alpha c prime minus alpha c double prime minus 0 is actually equal to 0. This implies that mu c is equal to 0 and hence P 1 c is equal to 0, thus the sets corresponding to two different values of c differ only in a set of points which has probability 0 under both distributions; that is points that could be excluded from the sample space.

(Refer Slide Time: 32:25)

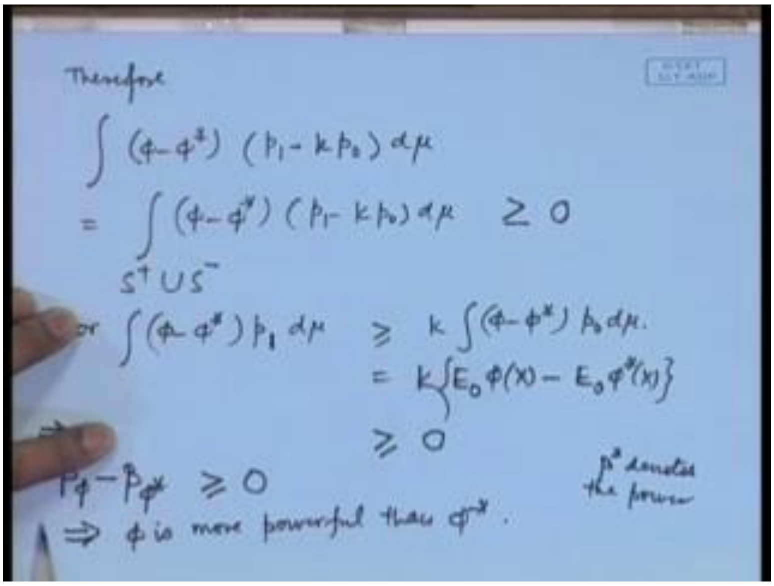


Now let us pay attention to the sufficiency part, so suppose that phi is a test satisfying 1 and 2 and suppose that phi star is any other test with say expectation of phi star less than or equal to alpha. Let us use the S plus notation for the set of those points for which phi minus phi star is greater than 0 and S minus is the set of those points for which phi x minus phi star is less than 0. Now these two are test functions, so both phi and phi star take values 0 or 1 or between 0 and 1. So, if x belongs to S plus; then what we are getting that phi x is a strictly greater than phi star, then phi x must be positive. Now if it is positive then the way we have defined our test function here, if you remember here the definition of the test function that it is positive; if it is 0 then only it is less; that means, in other cases it has to be greater than or equal to.

So, we will have this then phi x must be a strictly positive and so we will have p 1 x greater than or equal to k P naught x; let me repeat this argument. If x belongs to S plus then phi x is a strictly greater then phi star x. Now phi star x is a non negative function therefore, this phi x has to be strictly greater than 0, if phi x is strictly greater than 0 then by our definition of the test function p 1 x has to be greater than or equal to k p naught x. In the same way, if x belongs to S minus then here phi x will be strictly less then phi star x; phi star x can take values between 0 and 1. Therefore, phi x is less than 1 and so now, less than 1 condition by the definition here is satisfied for five functioned for p 1 x less than or equal to k p naught x.

So let us look at this; we are having phi x minus phi star x greater than 0 when x belongs to S plus and for that x p 1 x minus k p naught x is greater than or equal to 0. So, if I multiply these two terms, I will get non negative quantity; on the other hand if x belongs to S minus then this is negative and this is also p 1 x minus k p naught x is also negative less than or equal to 0, so the product will become greater than or equal to 0. So, what we are getting is that phi x minus; phi star x into p 1 minus k p naught x is greater than or equal to 0 for all x belonging to S plus union S minus.

(Refer Slide Time: 36:45)

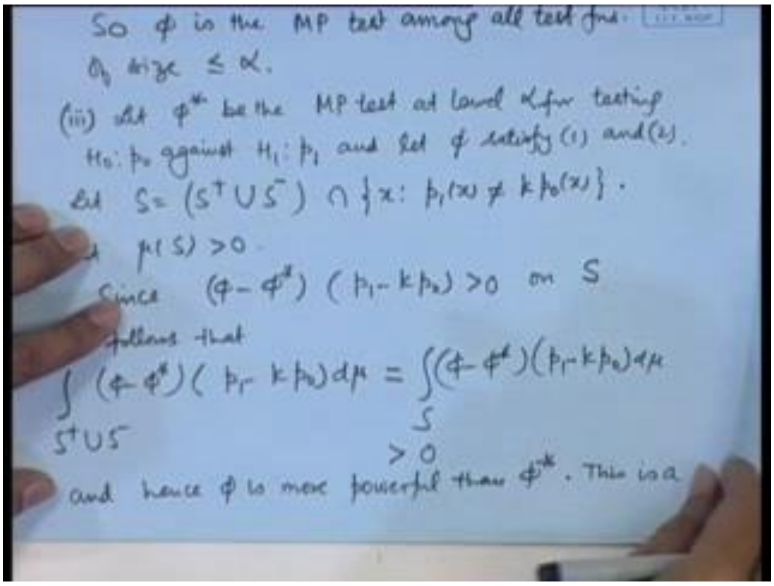


Now, let us use this; if we consider phi minus phi star, p 1 minus k p naught; d mu. So, this is a generalized term; that means, if we are dealing with the discrete distribution, this will be summation; otherwise it is a integral, so this integral will be equal to integral over the region. So, we have exhausted all the regions because over S plus; this term was positive and over S minus it is negative. So, if we go out of S plus and S minus, then this will be equal to 0. So, in that case this integral value integer and will become 0, so we can ignore that. So, we are looking at only the portion where it is nonnegative and this is greater than or equal to 0. So, this we can simplify; we can write it as phi minus phi star into p 1 d mu is greater than or equal to k times; phi minus phi star p naught; d mu.

Now, you look at the right hand side; this phi minus phi star; p naught, this value is nothing, but the expectation of phi under H naught and expectation of phi star under H naught; that is we can write it as k times expectation naught phi x minus expectation naught phi star x. Now expectation naught phi x is alpha and this value we have chosen to be less than or equal to alpha, so this is greater than or equal to 0. Now what is the right hand side; sorry what is the left hand side, this value is the probability of rejecting when H 1 is true; that means, it is the power function. So, we use the notation say beta is star for the power, so let me say beta is star denotes the power function then this is beta phi minus beta phi star; this is greater than or equal to 0.

This means that phi is more powerful than phi star, now in this one, what we did? We started with a test function phi which satisfies the conditions 1 and 2 so; that means, it has size alpha and phi star we took to be any other test function which is having size less than or equal to alpha; that means, equal to alpha case is also cover and then we are able to prove that the power of phi is more than or equal to the power of phi star, now this phi star is any arbitrarily chosen test for which the size is less than or equal to alpha; that means, among all the test functions which have size less than or equal to alpha, the power of phi is the maximum; that means, phi is the most powerful test.

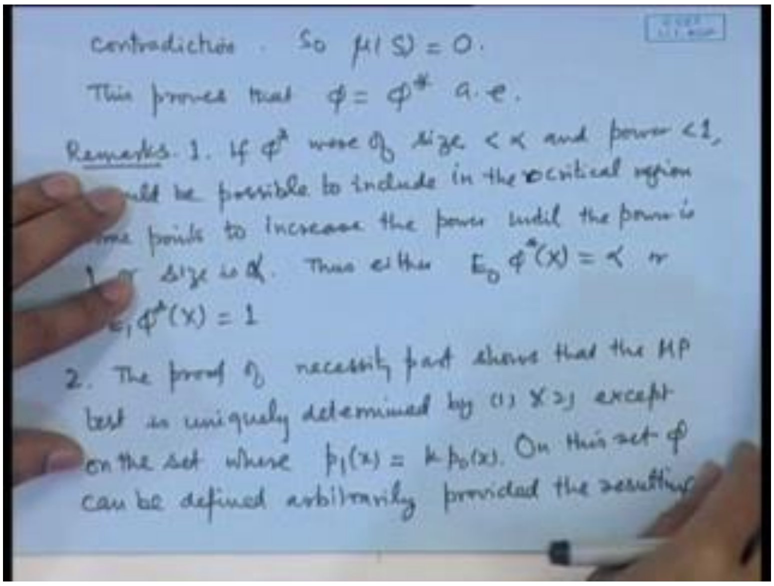
(Refer Slide Time: 40:21)



So, phi is the most powerful test among all test functions of size less than or equal to alpha. So, this theorem is very powerful in that sense that for a simple versus simple situation, it gives you a test procedure with a pre-assigned size; which is the most powerful. So, you have actually an optimal solution in this situation, but there is something more to this here; if there is a test which is most powerful then it will satisfy conditions 1 and 2. So, this is another important thing that there will not be any other test also, so in that sense it is a necessary in sufficient condition; let me prove that also.

So, let phi star be the most powerful test at level alpha for testing H naught, p naught against H 1; p 1 and let phi satisfy 1 and 2. Let us take say S is equal to S plus union S minus intersection the set of the values for which p 1 x is not equal to k p naught x. Let mu of S be positive, now we have already seen that on S plus and S minus; the quantity phi minus phi star and p 1 minus k p naught will be greater than 0. So, as already observed that phi minus phi star into p 1 minus k p naught is greater than 0 on S; it follows that S plus union S minus phi minus phi star p 1 minus k p naught d mu; that is equal to phi minus phi star p 1 minus k p naught t mu; this is a strictly greater than 0, so this means that phi is more powerful than phi star.

(Refer Slide Time: 43:46)

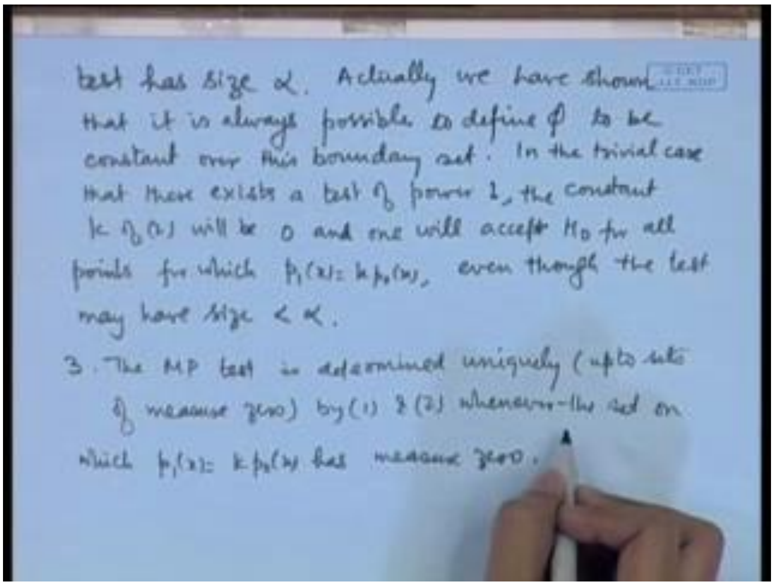


So, this is a contradiction because I started with phi star to be the most powerful, so mu s must be equal to 0; that means, the set where you have this phi minus phi star p 1 minus k p naught is actually greater than 0 that set must have measure 0. This proves that phi and phi star are same almost everywhere, so in this third part what we have done is that if there is a most powerful test; it must be the same as a test which satisfies the conditions 1 and 2; that means, and that is almost everywhere; that means, over a set of measure 0 you may modify the things here.

So, in essence this Neyman and Pearson fundamental lemma; gives you entire conditions under which you can derive a most powerful test uniquely up to almost everywhere. Let me give a few remarks here; if phi star were of size say less then alpha and power less than 1; it would be possible to include in the critical region; some points to increase the power until the power is 1 or size is 1 either of the things will happen.

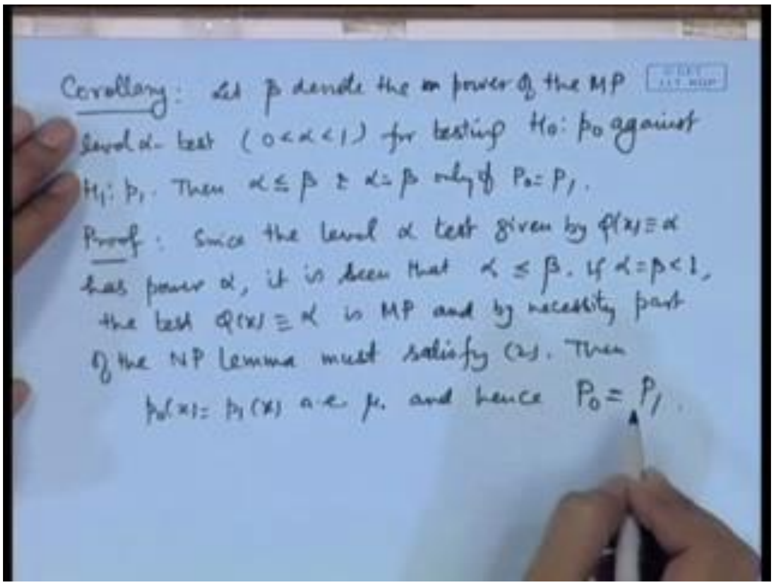
Thus either you will have expectation of phi star X is equal to alpha or expectation 1 phi star X is equal to 1; that means, either the size will become alpha or the power will become 1. The proof of necessity part shows that the most powerful test is uniquely determined by 1 and 2 except on the set where p 1 x is actually equal to k p naught x; that means, on this portion; we can define it arbitrarily, but the size has to remain alpha.

(Refer Slide Time: 47:46)



So, on this set phi can be defined arbitrarily provided the resulting test has size alpha. Actually we have shown that; it is always possible to define phi to be constant over this boundary set. In the trivial case that there exists a test of power 1, the constant k of 2 will be 0 and 1 will accept H naught for all points for which p 1; x is equal to k p naught x, even though the test may have size less than alpha. Third remark is that the most powerful test is determined uniquely up to sets of measure 0 by 1 and 2 whenever the set on which p 1 x is equal to k p naught x has measure 0.

(Refer Slide Time: 50:36)



We have a corollary here then let beta denote the power of the most powerful level alpha test, for testing H naught, p naught against H 1; p 1.

Then alpha is less than or equal to beta and alpha is equal to beta; only if P naught is equal to P 1. Let me see the proof of this; since the level alpha test given by phi x is equal to alpha; that means, throughout; this has power alpha, it is seen that alpha has to be less than or equal to beta. If alpha is equal to beta is less than 1, the test phi x is equal to alpha everywhere is MP and by necessity part of the NP lemma; it must satisfy 2, if it satisfies 2 then p naught x is equal to p 1 x almost everywhere mu and hence you must have P naught is equal to P 1; that means, basically there is no testing problem, if the null and alternative hypothesis are same; then the testing problem is disordered.

So; that means, there is no inference problem left here. So, today we have seen a powerful tool to derive, the most powerful tests for simple versus simple hypothesis testing problems. So, we will see some applications in the next lectures, this entire theory for the testing of hypothesis because in most of the other cases we will have a composite hypothesis, a simple versus composite or a composite versus composite hypothesis.

There have been extensions of this Neyman and Pearson fundamental lemma, the whole theory was developed in 1930s by Neyman and Pearson, so that will be the part of the course on statistical inference. In this particular course in the remaining portion, I will be taking of the applications of the Neyman and Pearson lemma for looking at the simple versus simple problems, as well as applications to a specific parameter testing problems in the normal distributions, the tests for the proportions in both in 1 sample and 2 samples problems and we will also look at the chi square test for goodness of it. So, that will be the coverage for the next lectures.