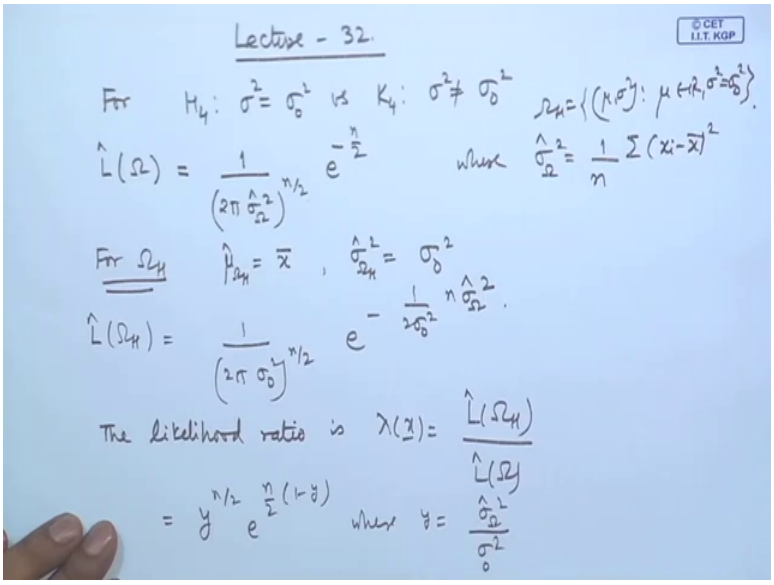
Statistical Inference

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

Lecture – 53
Likelihood Ratio Tests – III

In the last lecture, I have introduced the theory or you can say the concept of testing of hypothesis based on a heuristic consideration called Likelihood Ratio Tests. The idea was that we should consider that possibility are those that hypothesis to be more probable that gives a higher value of the maximum likelihood. Based on that, I derived the likelihood ratio test for parameters of a normal distribution. I was considering the testing for the mean and we had seen actually that the tests are almost same as the tests of derived using MN Pearson theory. For the variance testing, I derived a test for one sided null and one sided alternative hypothesis.

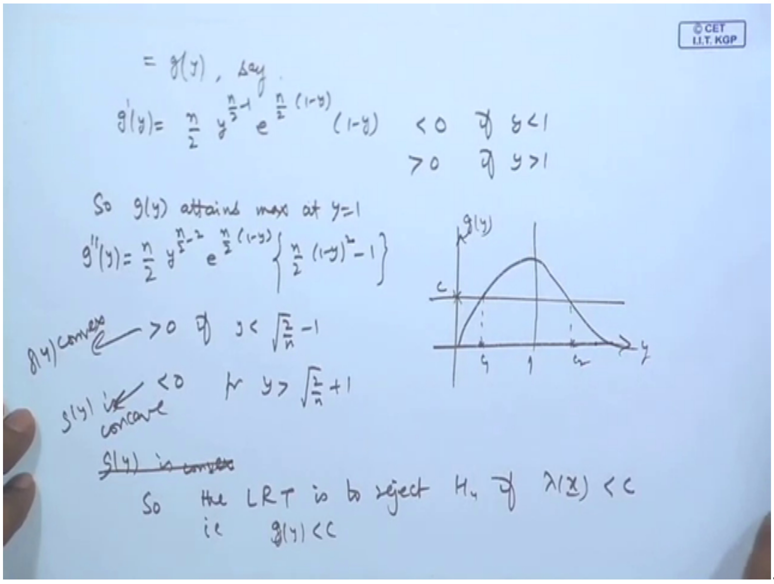
(Refer Slide Time: 01:17)



Let me also consider for H 4 that a sigma square is equal to sigma naught square against K for sigma square is not equal to sigma naught square. So, I do not have to derive L hat omega. This is derived earlier that is 1 by 2 pi sigma omega hat square to the power n by 2 e to the power minus n by 2 where sigma omega hat square is equal to 1 by n sigma xi minus x bar square. Now when we are considering omega H, then mu hat omega H remains same; however, sigma omega H hat square that is actually equal to sigma naught square because only this is the point null hypothesis. So, at this point only one value will come your omega H is actually mu sigma square mu is real and sigma square is equal to sigma naught square here.

So, L hat omega H, then this will be equal to 1 by 2 pi sigma naught square to the power n by 2 e to the power minus 1 by 2 sigma naught square and n sigma omega hat square. So, if I take the ratio the likelihood ratio is lambda x that is equal to L hat omega H by L hat omega. I can write it as same thing that is y to the power n by 2 e to the power n by 2 into 1 minus y, where y is equal to sigma omega hat square by sigma naught square. So, we have seen the behavior of this function; this is equal to say g y.

(Refer Slide Time: 03:45)

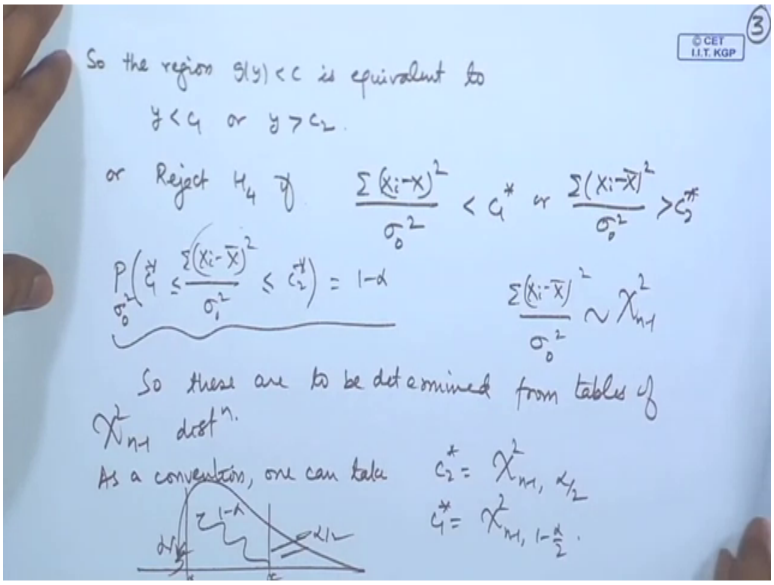


However, in the previous case y was always less than 1, but here y can take any value. And therefore, we need to look at the proper full behavior that is g prime y is actually equal to n by 2 y to the power n by 2 minus 1 e to the power n by 2 into 1 minus y into 1 minus y. So; obviously, this is less than 0; if y is less than 1, it is greater than 0 if y is greater than 1.

So, g y attains maximum at y is equal to 1. And the behavior actually we can also see what is g prime y g double prime, y g double prime y is n by 2 y to the power n by 2 minus 2 e to power n by 2 1 minus y n by 2 1 minus y square minus 1. That is greater than 0 if y is less than root 2 by n minus 1. It is less than 0 for y greater than root 2 by n plus 1.

So, it is actually that is g y is convex in this region and in this region g y is concave. So, the shape is something like this. This is one ok. So, the region LRT is to reject H 4 if lambda x is less than c. Now lambda xi have written as g y. Now g y less than c and g is a function of this nature. So now, you see here. So, suppose this is the point c and this is the curve g y against y here. So, this could be some value say c1, this is some value c 2. So, if I say g y is less than c this is equivalent to y being less than c1 or y being greater than c 2.

(Refer Slide Time: 06:29)

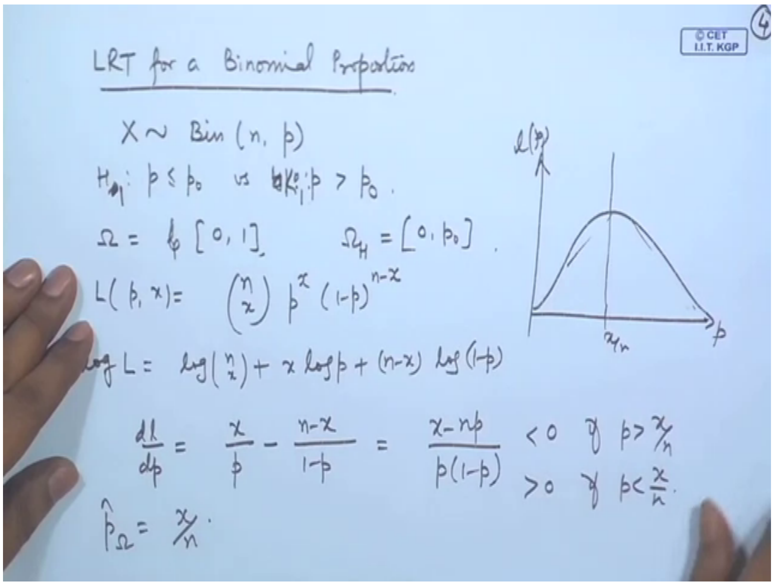


So, the region g y less than c is equivalent to y less than c1 or y greater than c 2. So, y is defined as. So, we can say reject H 4 if sigma Xi minus X bar whole square by sigma naught square is less than say c1 or sigma Xi minus X bar whole square by sigma naught square is greater than c 2. And probability of sigma Xi minus X naught X bar whole square by sigma naught square between c1 and c 2 should be equal to 1 minus alpha when sigma naught square is the true value.

Now, sigma Xi minus X bar whole square by sigma naught square follows chi square distribution on n minus 1 degrees of freedom. So, these are to be determined from tables of chi square n minus 1 distribution. As a convention one can take c 2 is equal to chi square n minus 1 alpha by 2 and c1 as chi square n minus 1 1 minus alpha by 2. Because the probability of between this suppose this is alpha by 2 probability this is alpha by 2 probability then this is 1 minus alpha.

So, this is c1 star this is c 2 star that we can take here. You can see that the form is similar to the test which is derived in the ump unbiased test. Except that in the unp and bias test we got 2 conditions there one of the conditions is the same, but one more condition was there. Here we are getting only one condition. So, there is a similarity here and. In fact, one solution corresponds to that. Before we move to two sample problems let me give an application to the distributions either which are discrete or which are not in the exponential family. So, let us consider likelihood ratio test for a binomial proportion.

(Refer Slide Time: 09:23)

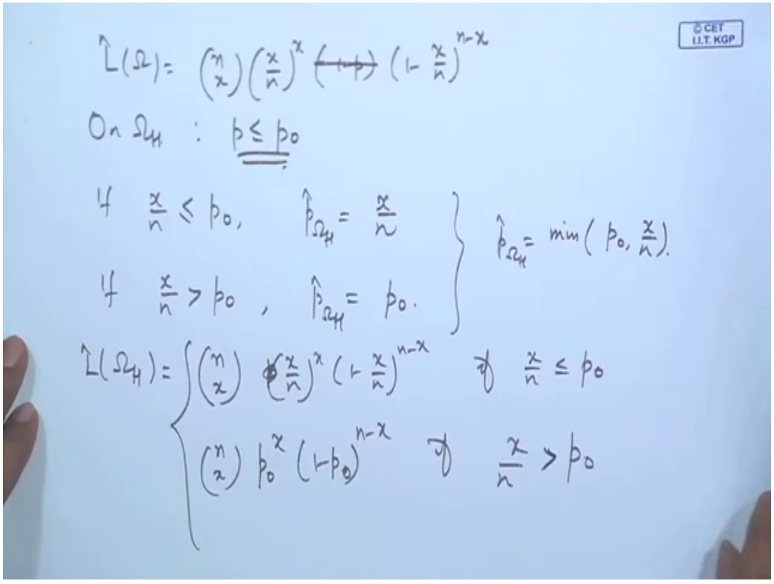


So, we have the data x following normal a binomial n p distribution. And we are considering the hypothesis testing problem say p is less than or equal to p naught against say p is greater than p naught. Let me write it as H K H 1 K 1 following our usual convention for one sided testing problem H 1 and this is K 1. Now your full parameter space is the interval 0 to 1 and omega H is 0 to p naught. Now let us write down the likelihood function. So, l p x that is the n c x p to the power x 1 minus p to the power n minus x. we have actually discussed in detail the maximum likelihood estimator for this problem.

So, derivative and putting equal to 0 all those things are the standard; however, when we consider the parameter space under the null hypothesis. Then we have to see carefully. So, I will write down the expression for that part. So, log of the likelihood is log of n c x plus x log of p plus n minus x log of 1 minus p. So, let me write it as small l. So, d l by d p that is equal to x by p minus n minus x by 1 minus p that is equal to x minus n p divided by p into 1 minus p. This is less than 0 if p is greater than x by n and it is greater than 0 if p is less than x by n.

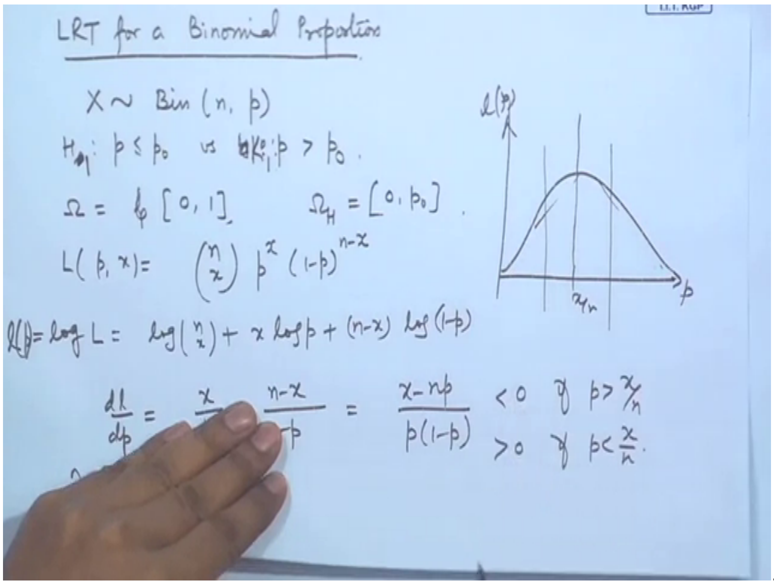
. So, if you look at the behavior of l p this side we are having l p. So, up to x by n this is increasing and thereafter it is decreasing. So, p hat omega is equal to actually x by n.

(Refer Slide Time: 12:17).



And therefore, the L hat omega that is equal to n c x x by n to the power x 1 minus p to the power sorry 1 minus x by n to the power n minus x. Now on omega H you are having p less than or equal to p naught. If I have that then there will be 2 cases.

(Refer Slide Time: 12:43)

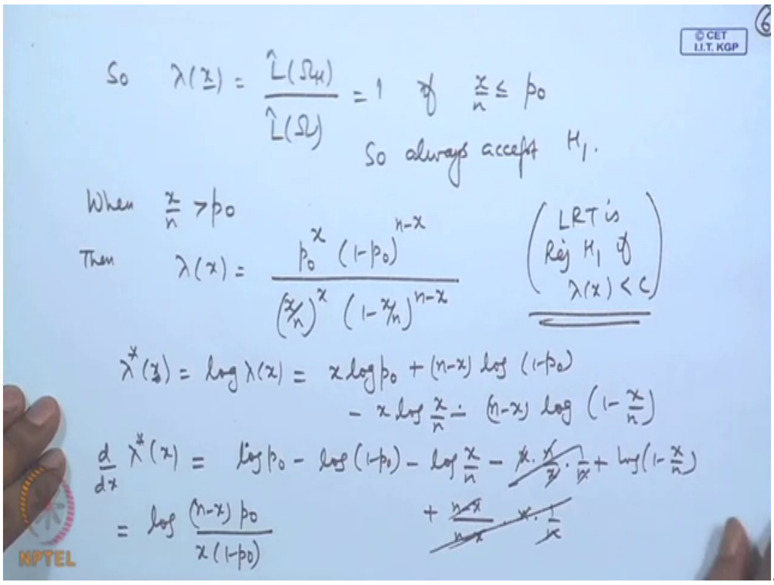


I may have p naught here and I may have p naught here. So, if x by n is less than or equal to p naught then p hat omega H that is equal to x by n; however, if x by n is greater than p naught.

In that case if this is p naught then the maximum value is attained at p naught because this is outside the region. So, then p hat omega H is actually equal to p naught; that means, we can say p hat omega H is actually equal to minimum of p naught and x by n. So, L hat omega H is equal to n c x p hat. So, x by n to the power x 1 minus x by n to the power n minus x if x by n is less than or equal to p naught and it is equal to n c x p naught to the power x 1 minus p naught to the power n minus x if x by n is greater than p naught.

So, if you compare L hat omega and L hat l hat omega H then for the case x by n less than or equal to p naught. They are the same. In that case you always accept because both are the same. So, you accept H 1.

(Refer Slide Time: 14:27)

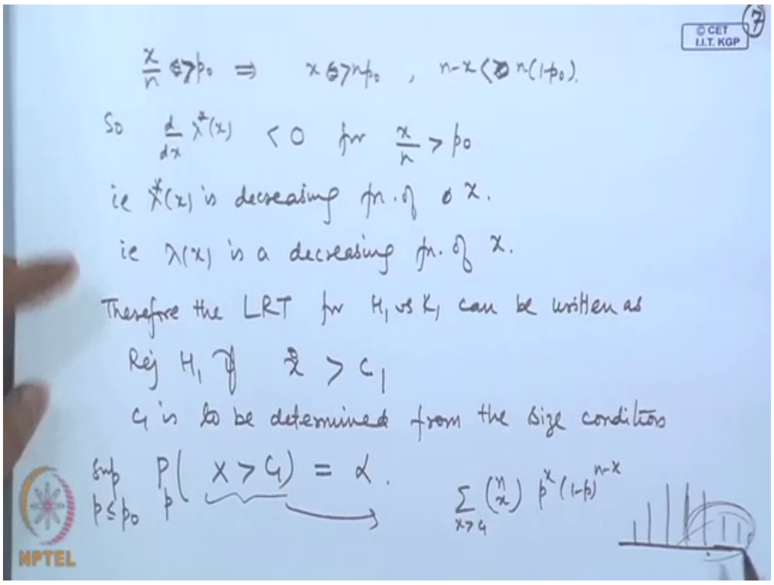


So, l lambda x that is L hat omega H by L hat omega, that is equal to 1 if x by n is less than or equal to p naught. So, always accept H 1. Now when x by n is greater than p naught in that case this lambda x that will become equal to p naught to the power x 1 minus p naught to the power n minus x divided by x by n to the power x 1 minus x by n to the power n minus x.

So, if I look at say lambda star that is equal to log of lambda x, that is equal to x log of p naught plus n minus x log of 1 minus p naught minus x log x by n minus n minus x log of 1 minus x by n. So, what is the behavior of this? Because LRT is reject H 1 if lambda x is less than c ok. Now if I write that in terms of x, what happens? So, we need to analyze the behavior of this. If I consider the derivative of this with respect to xi get log of p naught minus log of 1 minus p naught minus log of x by n minus x n by x 1 by n plus log of 1 minus x by n plus n minus x by n minus x n 1 by n. So, you can see this term gets cancelled out and this whole thing cancels out.

So, we are getting log of n minus x p naught divided by x into 1 minus p naught.

(Refer Slide Time: 17:07)

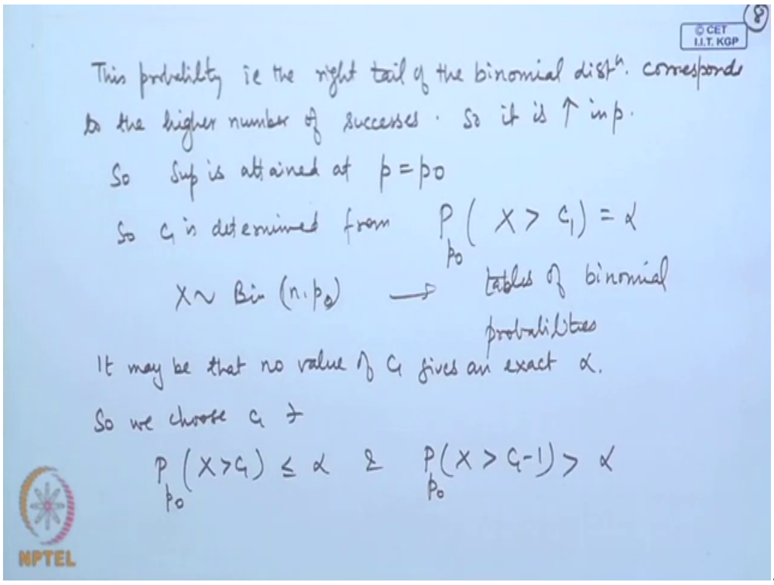


Now, we have chosen that x by n is less than or equal to p naught. This implies x is less than or equal to n p naught. And n minus x will be n minus x will be greater than or equal to n into 1 minus p naught. So, d by d x lambda star x; this will be less than 0 sorry. This is greater actually I made a mistake here x by n is greater than p naught region we are considering. So, x is greater than n p naught n minus x is less than n into 1 minus p naught.

So, what will happen that if we consider this term log of n minus x p naught divided by x into 1 minus p naught this term becomes less than one therefore, log of this becomes less than 0. So, this is less than 0 for x by n greater than p naught. That is lambda star is a decreasing function of x. That is lambda x is a decreasing function of x. So, note here I am considering the likelihood ratio that is lambda x, the region is lambda x less than c now lambda is turning out to be a decreasing function of x. Therefore, this region is equivalent to therefore, the likelihood ratio test for H 1 versus K 1 can be written as reject H 1 if x is greater than some c 1.

Now, c1 is to be determined from the size condition. That is probability of x greater than c1 for p supremum of p less than or equal to p naught. This should be equal to alpha. Now as a function of p what is this term this is actually sigma n c x p to the power x into 1 minus p to the power n minus x x is greater than c1. If you plot the binomial probabilities the behavior is like this. Now we are looking at the right tail. Now right tail corresponds to a higher number of successes.

(Refer Slide Time: 20:43)



So, this probability that is the right tail of the binomial distribution corresponds to the higher number of success. So, it is increasing in p. So, supremum is attained at p equal to p naught. So, c1 is determined from probability x greater than c1 equal to alpha for p is equal to p naught, where x follows binomial n p naught; that means, you have to see the tables of binomial probabilities. Now this being a discrete distribution, it is quite possible that there may not be any c1 for which this is equal to alpha. It may be that no value of c1 gives an exact alpha. So, we may choose c1 such that probability x greater than c1 at p naught is less than or equal to alpha, but probability of x greater than c 1 minus 1 at p naught is greater than alpha. So, this is the likelihood ratio test for the proportion of a binomial distribution.