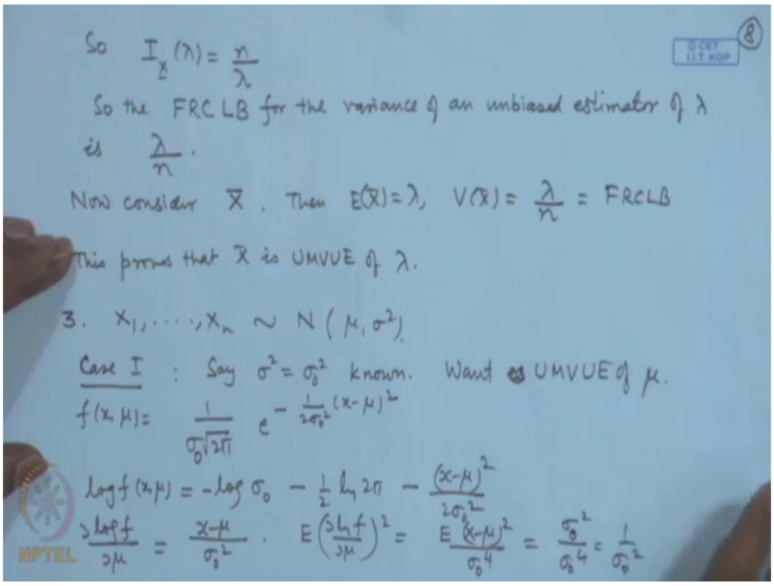
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 16
Lower Bounds for Variance - II

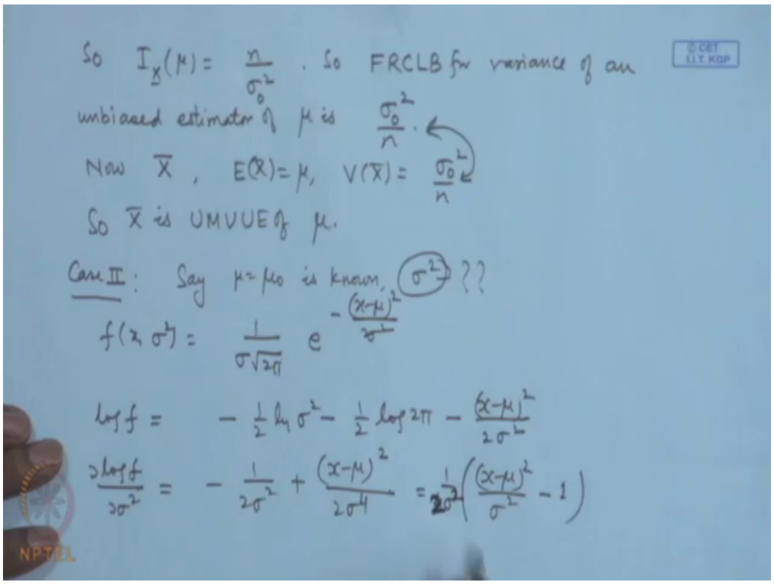
(Refer Slide Time: 00:21)



Let us take another popular example that is the normal distribution. So, let us take say X 1, X 2 X n following normal mu sigma square. Now, as before we will consider different cases say; sigma square is equal to sigma naught square that is a known ok. In that case we want estimate of say UMVUE of mu. So, if we write down the distribution here 1 by sigma root 2 pi. So, here it will become sigma naught e to the power minus 1 by 2 sigma naught square X minus mu whole square.

So, log of f is equal to minus log of sigma naught minus half log 2 pi minus x minus mu square by 2 sigma naught square. So, if we consider derivative of this with respect to mu, we get simply x minus mu by sigma naught square. So, expectation of del log f by del mu whole square that is equal to expectation X minus mu square by sigma naught to the power 4. Once again in the normal distribution this is reducing to the variance term that is expectation of X minus mu square is variance. That is sigma naught square by sigma naught to the power 4 that is; 1 by sigma naught square.

(Refer Slide Time: 02:14)

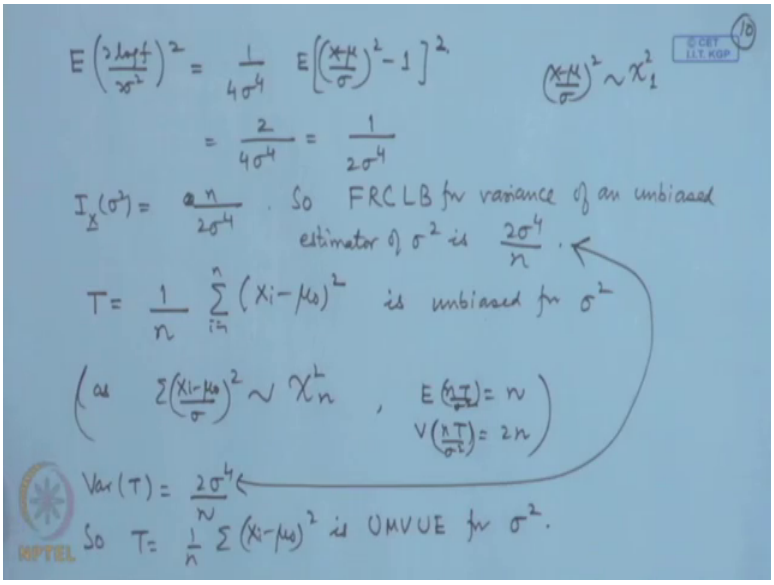


So, information content in this is information contained in this will be n by sigma naught square in the sample. So, the Fisher, Rao, Cramer lower bound for variance of an unbiased estimator of mu is sigma naught square by n. Now if you consider say X bar, then expectation of X bar is mu. And what is variance of X bar? That is sigma naught square by n that is equal to this value.

So, X bar is UMVUE of mu. Let us take another case when say mu is known and we want to estimate, say mu is equal to mu naught is known and we want sigma squares estimator. So, here the density function will be written as a function of sigma square 1 by sigma root 2 pi e to the power minus x minus mu square by 2 sigma square. So, log of f becomes minus half log sigma square minus half log 2 pi minus x minus mu square by 2 sigma square.

So, differentiation of this with respect to sigma square gives minus 1 by 2 sigma square plus x minus mu square by 2 sigma to the power 4 which I can write as, x minus mu square by sigma square minus 1; 1 by 4 sigma 1 by 2 sigma to the power 1 by 2 sigma square. So, if we consider expectation of.

(Refer Slide Time: 05:00)



So, if we consider expectation of del log f by del sigma square whole square. That is equal to 1 by 4 sigma to the power 4 expectation of X minus mu by sigma whole square minus 1 whole square. Once again you look at this X minus mu by sigma is a standard normal variable X minus mu by sigma square it will follow chi square 1. So, expectation of this is equal to 1 and therefore, this term reduces to the variance.

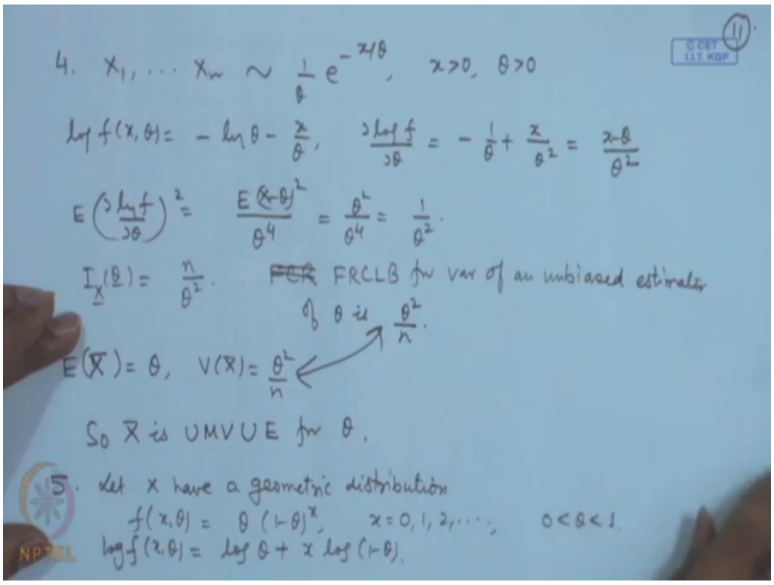
So, variance is twice the degrees of freedom; that is equal to 2 by 4 sigma to the power 4. So, you get 1 by 2 sigma to the power 4. So, the Fishers information in this problem will be a n by 2 sigma to the power 4. So, the Fisher, Rao, Cramer lower bound for variance of an unbiased estimator of sigma square is 2 sigma to the power 4 by n. Now, in this case let us consider see the maximum likelihood estimator for example or the method of moments estimator.

So, that would be for example, 1 by n sigma X i minus mu naught square. So, this is now you can see here X i minus mu naught by sigma that will follow a standard normal. So, sum of squares will be chi square n. So, expectation of that is n so, this divided by n will have expectation 1. So, if we multiply by sigma square we will get sigma square. So, this is unbiased for sigma square because, we can see here as sigma X i minus mu naught by sigma whole square that follows chi square on n.

So, expectation of n T is equal to n and variance of n T. Sorry this divided by sigma square and this divided by sigma square that will be equal to 2 n. So, we will get variance of T as equal to twice sigma to the power 4 by n because, this will go here and n square will come to know below. So, we will get 2 sigma to the power 4 by n which is same as this value once again here.

So, T is equal to 1 by n sigma X i minus mu naught square. It is the minimum variance unbiased estimator for; obviously, you can see here that if mu was not known then you could not have used this estimator. So, this solution is specific to this problem that is when we are dealing with one parameter case mu naught is known to us.

(Refer Slide Time: 08:33)



Let us consider say a random sample from exponential distribution with mean say theta. Now, in this case the density function is this. So, log of the density function is minus log of theta minus x by theta. So, the derivative with respect to theta will be minus 1 by theta plus x by theta square that is x minus theta by theta square. So, expectation of del log f by del theta whole square that is expectation of X minus theta square by theta to the power 4. In the exponential distribution with mean theta variance is equal to theta square.

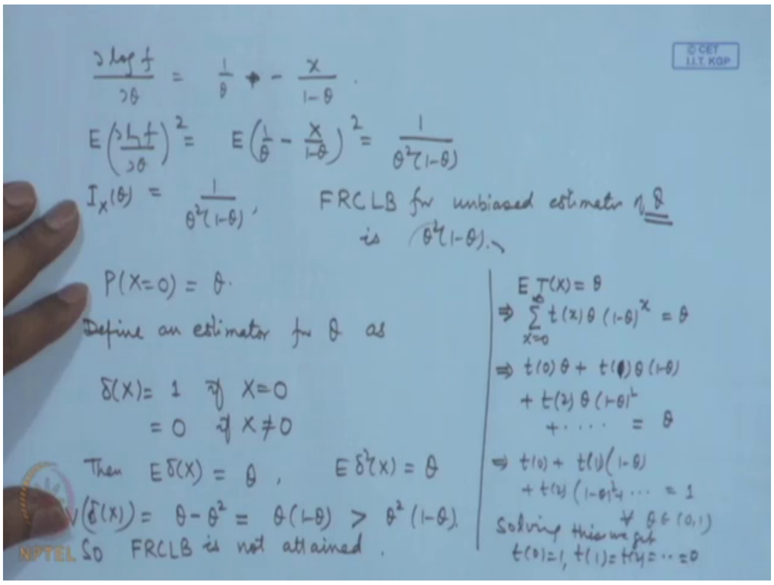
So, this term becomes theta square by theta to the power 4 that is equal to 1 by theta square. So, the information in the sample about theta is n by theta square and the lower bound Fisher, Rao, Cramer lower bound for variance of an unbiased estimator of theta is theta square by n. If we consider X bar then expectation of X bar is equal to theta and variance of X bar is equal to theta square by n. So, this will prove that X bar is minimum variance unbiased estimator for theta in the case of negative exponential distribution.

It is not necessary that the lower bound is always attained. In fact, if you see carefully in each of these problems we have calculated the derivative here. So, S function you can see here for example, S would have become here n by minus n by theta plus sigma x i by theta that is n x bar. So, this is linearly related with x bar and therefore, x bar must attain the variance lower bound for its expectation. If you see the previous problem, for the estimation of sigma square here del log f by del sigma square is this function. So, if we look at S function S X i sigma square that would have become minus n by 2 sigma square sigma X i minus mu square by something which is a linear function of sigma X i minus mu whole square.

And therefore, it is natural that sigma X i minus mu naught whole square by n will attain the lower bound here. So, if you see the estimation of the Poisson distribution case here the derivative is equal to minus 1 plus x by lambda. So, if you look at S function it would have become minus n plus sigma x i by lambda, which is again linearly related with x bar. Therefore, X bar must attain X bar must attain the lower bound for the variance of its unbiased estimation. So, in all these problems it is naturally coming; let me take another example where it may not be natural and therefore, the lower bound may not be attained.

Let us consider say let X have a geometric distribution and we consider the following form theta into 1 minus theta to the power x; where theta is any number between 0 and 1. So, here the problem is of estimation of theta. So, let us look at log of f x theta that is equal to log of theta plus x times log of 1 minus theta. So, if we consider del log f by del theta we get 1 by theta plus with a minus sign x by 1 minus theta here.

(Refer Slide Time: 13:25)



So, if we look at the expressions here expectation of del log f we can use the moment structure of the geometric distribution. So, if we use that this is equal to expectation of 1 by theta minus X by 1 minus theta whole square. So, after simplification this turns out to be 1 by theta square into 1 minus theta. So, since I have taken only one observation here the information will remain the same. And the lower bound for unbiased estimator of theta is theta square into 1 minus theta.

Now, here theta is not the mean actually if you look at the mean of this distribution that will be 1 minus theta by theta. So, x will attain the lower bound for that for the variance of an unbiased estimator for 1 minus theta by theta. But suppose we are considering estimation of theta, if we are estimating theta here then it will not be attained. So, you can see here what is the interpretation of say theta here, theta is the probability of X is equal to 0. Because if in the probability mass function we put X equal to 0 here I get theta.

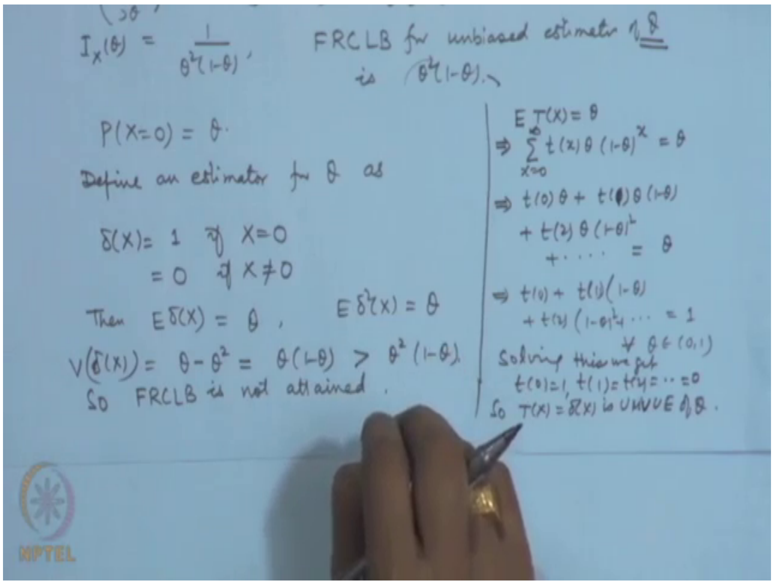
So, if I define an estimator for theta as say delta X is equal to 1. If X is equal to 0 it is equal to 0 if X is not equal to 0; that means, if X is equal to 1 2 and so on. Then expectation of delta X will be equal to 1 into probability X equal to 0 plus 0 into probability X not equal to 0. That will means it will be simply equal to theta and what is expectation of say delta square X that will also be theta. So, variance of delta X that will be equal to theta minus theta square that is equal to theta into 1 minus theta.

Now here if you compare with this lower bound here, lower bound is theta square into 1 minus theta and theta is any number between 0 and 1. So, this one will be naturally bigger than this. So, the lower bound is not attained. So, we do not know whether delta is minimum variance unbiased estimator here. We may try another approach here. Let us consider expectation of T X is equal to theta. If we consider this then we will get sigma t x into theta into 1 minus theta to the power x is equal to theta as x varies from 0 to infinity.

That will give me t 0 in to theta plus t 0 into theta into 1 minus theta plus sorry plus t 1 plus t 2 into theta into 1 minus theta square and so on is equal to theta. Now, you look at this; what we are getting that coefficient of theta here if you see so, this you can cancel out actually t 0 plus t 1 into 1 minus theta plus t 2 into 1 minus theta square and so on is equal to 1. This is true for all theta belonging to the interval 0 to 1. Now, if you see this carefully, what is the solution? See if you look at the coefficient of say theta here. Theta will have coefficient t 1; see for example, if I look at the coefficient of the constant term, constant term is t 0 plus t 1 plus t 2 and so on that should be equal to 1.

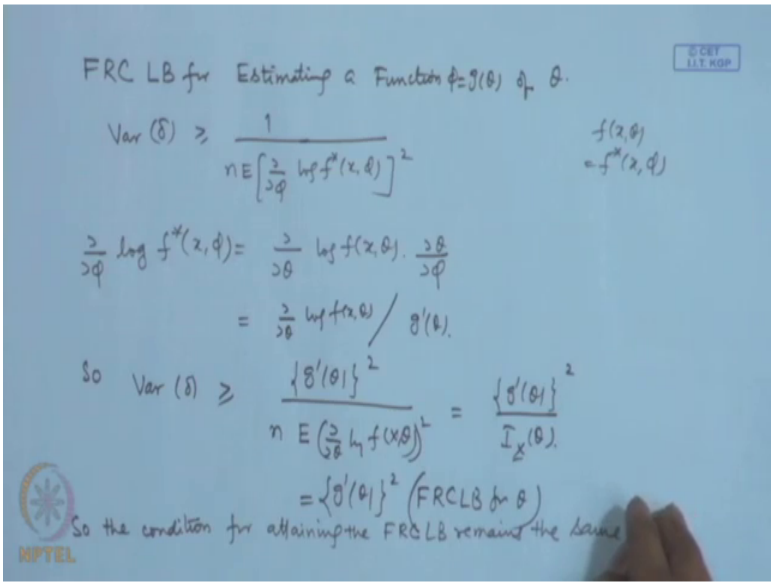
If you take coefficient of theta then you get minus t 1 minus 2 t 2. Then in the next one also minus 3 t 3 and so on that should be equal to 0. Then if you look at the coefficient of theta square you will get t 2 then here in the second one it will become 3 t 3 and so on. So, if you solve this solving this we get t 0 is equal to 1 t 1 t 2 and so on is equal to 0 which is nothing, but this t function that is becoming same as this. So, we have proved otherwise that T X that is equal to delta X is UMVUE because; this is the only unbiased estimator which we obtain through solving the equation itself.

(Refer Slide Time: 19:11)



However, using the method of lower bounds we are not able to prove this result here. Now many times we may not be interested directly in the theta itself, we will be interested in some function say g theta of theta; in that case what we can do is we can modify this lower bound formula like.

(Refer Slide Time: 19:56)

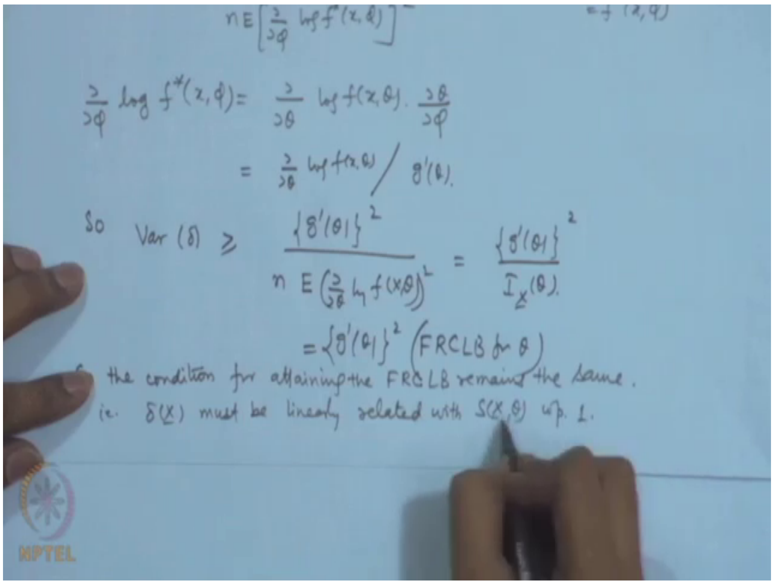


So, FRC lower bound for estimating a function g theta of theta, so let me call it phi. So, we will write variance of delta greater than or equal to 1 by n time’s expectation del by del phi log of f star x phi. Because f x theta density, now I am writing as f star x phi because we have substituted theta by g inverse phi in whatever form we are able to do that. So, if we look at this derivative here del by del phi log of f star x phi; you can apply the chain rule we can write it as del by del theta log of f x theta into del theta by del phi.

This you can write it as del by del theta log of f x theta divided by g prime theta. So, if you substitute this function here, we get variance of delta greater than or equal to g prime theta square divided by n time's expectation of del by del theta log of f X theta whole square. That is equal to g prime theta whole square by the information in the sample about theta. That means, if we have the lower bound for the variance of an unbiased estimator of theta. Then from there we can derive for any other function what we have to do, we have to multiply by the lower bound by g prime theta square.

So, this we can say it is equal to g prime theta square into the Fisher, Rao, Cramer lower bound for theta. So, this new formula can be obtained. Moreover the condition for obtaining the lower bound for attaining the lower bound, that will remain the same because the condition is coming only from the Cauchy Schwarz inequality which was dependent upon the estimator being linearly related with S X theta. Now the g theta function does not affect that thing. So, the condition for, the condition for attaining the F remains the same.

(Refer Slide Time: 27:08)



That is your delta X that is delta X must be linearly related with S X theta with probability 1. Tomorrow's class we will be considering further properties and further ramifications of this lower bound as well as we will see some extensions. There can be two types of extensions. One is the extension to the higher dimensions; that means, if in place of one dimensional parameter I have several dimensional parameter; then what will be the form of the Rao, Cramer inequality.

Similarly, here we have used first order derivative in the lower bound. Now, if we consider second and higher order derivatives then the level of the inequality can be changed. So, they are generalization into another direction. Another thing is that whenever we are considering differentiation, in some sense we are taking the limits. Suppose we do not take the limits in place of that we write the difference. For example, we are saying derivative so we are writing down the value of the function at 2 points theta and theta plus delta say.

So, we consider the difference there and then look at the inequality that inequality will be called the equality without the regularity condition. Because when we are having regularity conditions then we are considering the derivative and other things. But if that is not satisfied then what? So, we will have another extension in that direction. So, in the next lecture we will be considering extensions to these things and then further applications of this.

That is all for today's lecture.