Statistical Inference

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

Lecture No. # 09

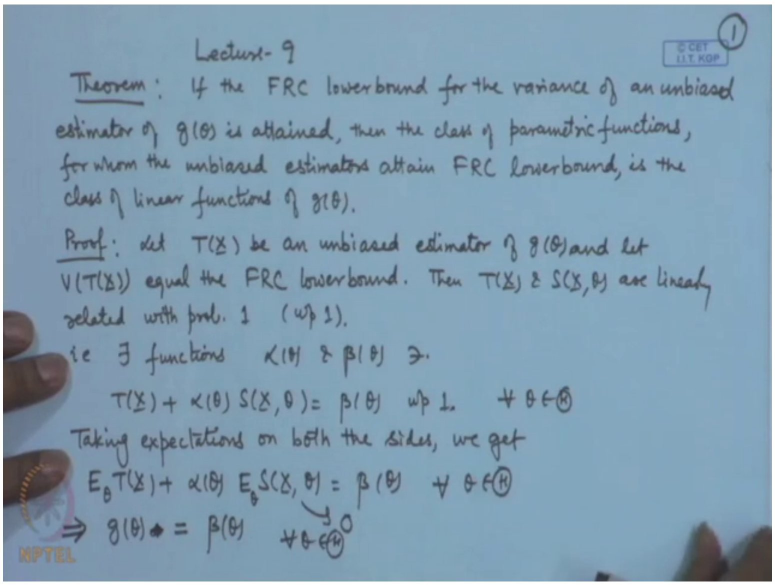
Lower Bounds of Variance - II

In the previous lecture, I explained the method of finding out a lower bound for the variance unbiased estimator for a given parametric version. As I mentioned it was derived independently by three statisticians Frechet-Rao and Cramer and therefore, we have named it as a Frechet-Rao-Cramer lower bound, that is FRC lower bound for the variance of an unbiased estimator.

We have seen that there are cases where we can find out an estimator, for which this lower bound is attained, there are also cases where it is not attained. We gave a condition, under which an unbiased estimator will attain this lower bound. The condition was in the terms that it should be linearly related with a function S X theta with probability 1, this method as I had explained this method of lower bounds is very very useful from two points of view, one is that given any estimators we can compare it is variance with the lower bound and therefore, we know that how far we are from the actual.

That means what could be the best possible way minimum variance and where are we; that means, where is our estimator standing in it is for relative position and second thing is that, if we are able to obtain an estimator, for which it is equal to the lower bound then certainly it is a minimum variance and biased estimator that is among the unbiased estimator it will certainly the best. So, from this point of view this method of lower bound say extremely useful. We have seen that FRC lower bound as I call it is dependent upon certain regularity conditions, that is when the density are the mass function under consideration satisfy conditions, then only this lower bound is valid. We also seen this what are the p arametric functions for which this lower bound is attained.

(Refer Slide Time: 02:26)



So, let me give it in the form of a theorem. So, we have a random sample X 1, X 2, X n and we know that the FRC lower bound, for the variance of an unbiased estimator of g theta is attained, then what are the parametric functions? Apart from g theta for which we have attained, then the answer is that they are actually the linear functions of g theta and then the class of parametric functions for whom the unbiased estimators attain this FRC lower bound, then this class is the class of linear functions of g theta.

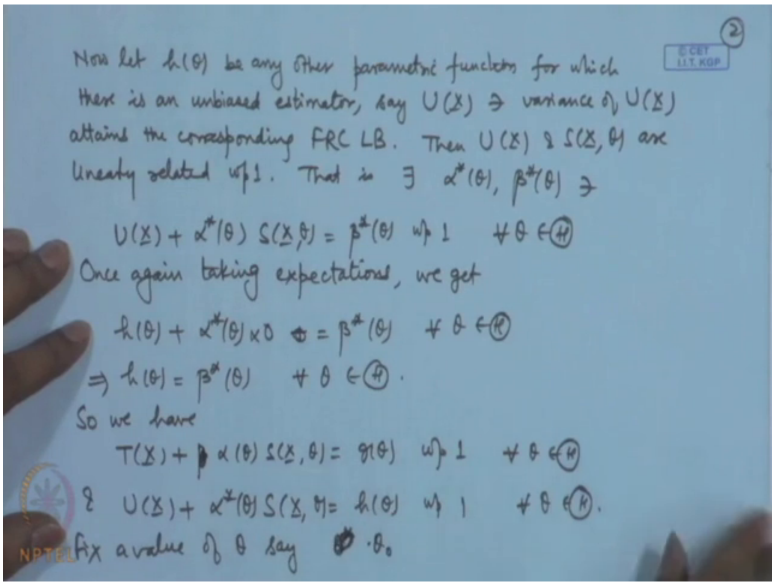
Like I said what is the unbiased estimator for which the lower bound will be attained that should be a linear function of S X theta with probability 1. Now, what are the parametric functions for utility which will be attain and then they should simply be the linear functions of g theta that is the statement of this theorem.

Let me prove this theorem here. So, let us consider T X, let T X be an unbiased estimator of g theta and let variance of T X equal the FRC lower bound, then certainly we know that T X and S X theta, they are linearly related with probability 1, we will use this with probability one as an abbreviation here.

So; that means, there exists functions say alpha theta and beta theta, such that say T X plus alpha theta S X theta is equal to say beta theta with probability 1, this should be true for all theta. Now, in this relation let us take expectations on both the sides. So, expectation of T X plus alpha theta expectation of S X theta is equal to beta theta for all theta since, this statement is true for all.

That means for random variable X here it is true with probability 1 therefore, it is possible to take the expectations basically, expectations means either we have taken summations or we have taken the integrals or a mixture of the two therefore, we will get expectation of these equal to beta theta. Now, T is unbiased estimator for g theta; that means g theta now, expectation of S X theta that is 0. Therefore this is simply giving you beta theta, because this is equal to 0. So, in this relationship beta theta has turned out to b g theta here.

(Refer Slide Time: 07:30)

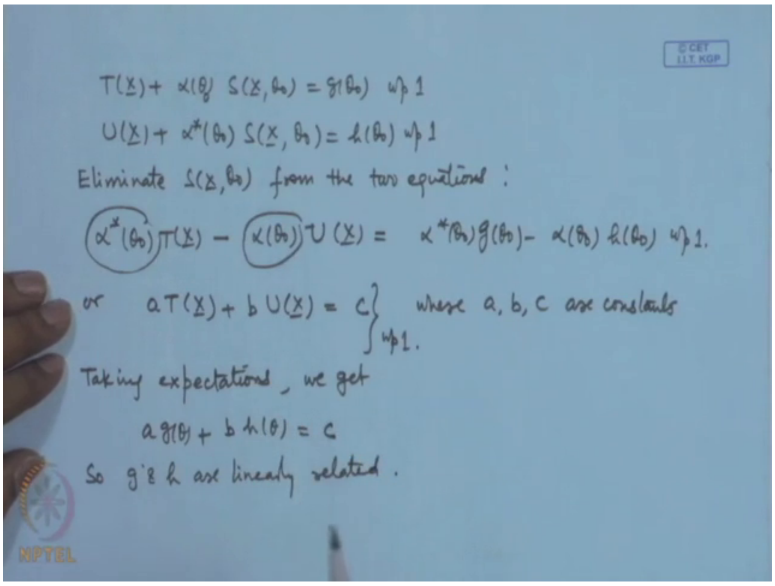


Now, let h theta b any other parametric function, for which there exist an unbiased estimator, for which this lower bound is attained. So, for which there is an unbiased estimator say U X, such that variance of U X attains the corresponding FRC lower bound. We have seen that even if we change the parametric function, the lower bound is changed, but the conditions for attaining the lower bound remains the same therefore, So, there will exists that U X and S X theta are again linearly related with probability 1; that means, we can say that there exist say functions alpha star theta and beta star theta, such that U X plus alpha star theta into S X theta is equals to beta star theta with the probability 1 for all theta.

Once again since this statement is true with probability 1, we can take expectations. So, if we take expectations, we get expectations of U X will be equal to h theta plus alpha star theta into expectation of S X theta is 0 is equal to beta star theta. So, we are getting h theta is equals to beta star theta. So, if we look at this two equations now, T X plus alpha theta S X theta that will be equal to g theta and U X plus alpha star theta S X theta equals to h theta.

So, we have T X plus alpha theta S X theta is equal to g theta with probability 1 for all theta and U X plus alpha star theta, S X theta is equals to h theta with probability 1, for all theta belonging to theta. If this relationship is true for all theta, we can fix a value of theta fix a value of theta say theta star or let me put theta not, because already stars are there.

(Refer Slide Time: 11:19)

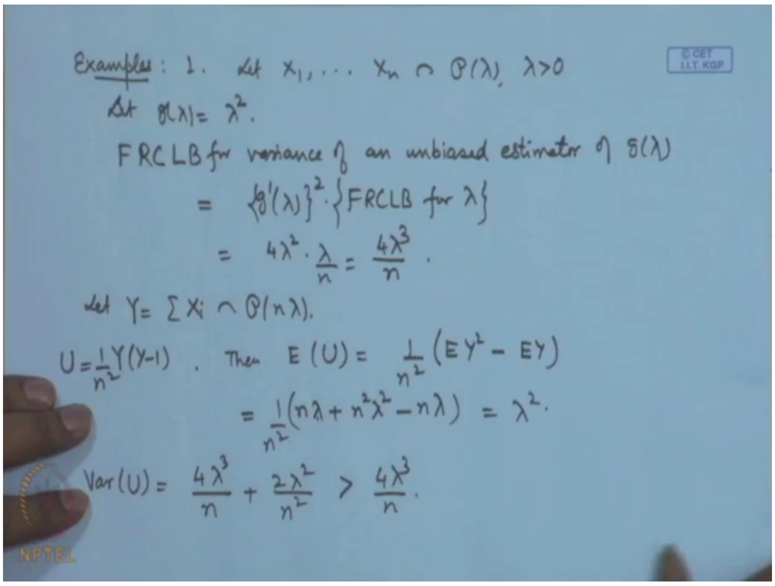


So, in that case we can write the relationship as T X plus alpha theta naught, S X theta naught is equals to g theta naught with the probability 1 and the U X plus alpha star theta naught, S X theta naught is equals to h theta naught with probability 1; that means, what I have done is that these two relations I have written for a fixed value of theta that is theta naught. Now, in both of these equations S X theta naught is appearing. So, I can eliminate that. So, eliminate S X theta naught from the two equations that, is in the first equation multiply by alpha star theta naught, in the second equation multiply by alpha theta naught and then subtract. So, we get alpha star theta naught, T X minus alpha theta naught, U X is equal to alpha star theta naught, g theta naught minus, alpha theta naught h theta naught with probability 1.

Now, once again you can take the expectation, because what is happening here is that this coefficient is a fixed number, this coefficient is a fixed number and right hand is also a fixed number. So, we can say that a times say T X plus say b times U X equals to c, where a, b, c are constants and this statement is true with probability 1. So, we can again take expectations, if we take expectations we get a time g theta that is expectations of T X plus b times h theta is equals to c.

Now, you look at the significance of this I started with a function g, for which the FRC lower bound was attain. I assumed h theta to be another parametric function for which the lower bound is attained and now, we are getting that such g and h will b related using linear relationship here. So, j and h are linearly related therefore, all functions for which the FRC lower bound will be attained, they will be linear functions of g. Now, in yesterdays lecture I have given examples, in some examples the lower bound was attained.

(Refer Slide Time: 14:50)



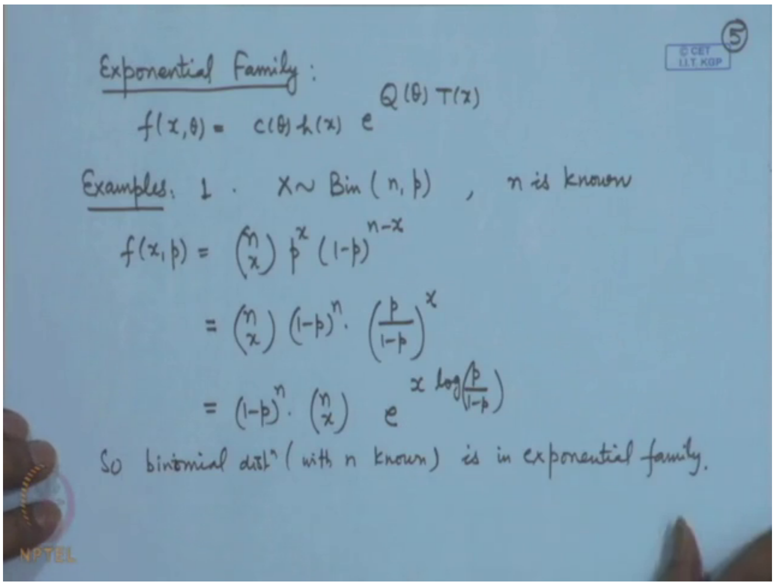
Let us take one such example, say Poisson distribution. So, we had X 1, X 2, X n, following Poisson lambda where lambda is positive, we have seen that X bar was unbiased for lambda and variance of X bar was lambda by n, which was also the FRC lower bound far unbiased estimator of lambda. So, if I consider say lambda square, let g lambda b equal to lambda square. In that case what we will get? The FRC lower bound for variance of an unbiased estimator of g lambda now, that will be equal to g prime lambda square into the FRCLB for lambda. So, this will become two lambdas square that is 4 lambdas square and this is lambda by n. So, it is equal to 4 lambda cube by n.

So, now let us consider say y is equal to sigma X i of course, this will follow Poisson n lambda and you can look at y into y minus 1 by n square let me call it to be say U, then expectation of U it is equal to 1 by n square expectation of y square minus expectation of y, that is equal to now, this will become equal to now, n lambda plus n square lambda square minus n lambda expectation of y square is n lambda plus n square lambda square, because we can see that Poisson distribution with parameter lambda. The second moment is lambda square plus lambda and expectation y is equal to n lambda. So, this divided by n square. So, that is equal to lambda square.

But if we consider say variance of U that will be equal to this can be calculated easily, that will turn out to be, because this will involve expectation of U square minus, expectation of U whole square. Now, expectation of U is lambda square and expectation of U whole square will be involve expectation of y to the power 4, expectation of y Q and expectation of y square which is available. All the expression there further Poisson distribution after simplification you get it as, 4 lambda cube by n plus twice lambda square by n square.

Now, you can easily see that this is bigger than 4 lambda cube by n, it is understood that this statement should be true, because lambda square is not linear function of lambda here. We have already shown that for lambda the variance of the unbiased estimator are tensed the lower bound. Therefore, all other function for which we will attain for the form a lambda plus b and this is lambda square. So, suddenly this cannot be attained, later on we will show that actually this is minimum variance and y estimator using another method.

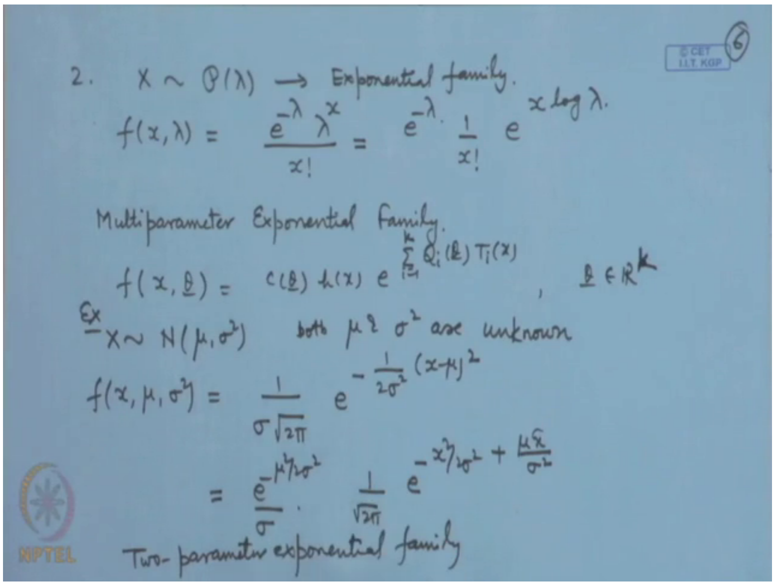
(Refer Slide Time: 18:34)



Let us consider general form of a distribution in the exponential family. So, let us consider a density in the exponential family. What is an exponential family? The densities of the form c theta, h x e to the power, Q theta T x. Now, if we have a distribution of these form, it is said to be distribution in the exponential family, we can see examples here say x follows binomial n p here n is known, then the form of the distribution is n c x, p to the power x, 1 minus p to the power n minus x.

This we can write as n c x, 1 minus p to the power n, p by 1 minus p to the power x, this we write as 1 minus p to the power n, n c x, e to the power X log p by 1 minus p. So, if you compare it with this pair here you have a function of the parameter that is c theta here theta is p, h x is n c x, here e to the power Q theta T x. So, here Q theta is a function here log p by 1 minus p and x is the term T x. So, this is a distribution so, binomial distribution. Binomial distribution with n known is in exponential family, let us take some more popular examples in the statistics.

(Refer Slide Time: 21:06)



Let us consider say x following Poisson lambda distribution. The form of the probability mass function is f x lambda, it is equal to e to the power minus lambda, lambda to the power x by x factorial, this we express as e to the power minus lambda, 1 by x factorial, e to the power x log lambda. Once again if we compare it with this particular form you can see hear, e to the power minus lambda is a function of lambda, 1 by x factorial is a function of x. So, you can call it a h x.

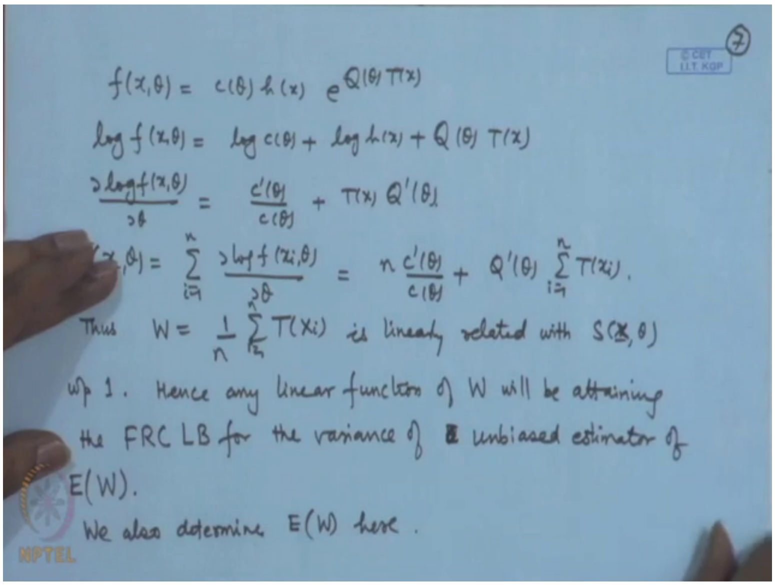
X can be written as T x and Q theta it is log lambda here. So, you can easily see that this also a distribution in exponential family. We can actually also consider as a 1 parameter exponential family, we may also consider multi parameter exponential family, here parameter could be multi parameter here. So, here we write c theta h x, e to the power sigma Q I theta T I x, i equals to 1 to k. So, theta could be say p dimensional and we may have this particular form here. So, this is actually called.

See if we have the same dimension here k, then this is called a k parameter exponential family. Let us consider say k x following normal mu sigma square, here both mu and sigma square are unknown, f x mu sigma square we can write as 1 by sigma root 2 pi, e to the power minus 1 by 2 sigma square x minus mu whole square, this we express in the following fraction.

If we expand this term you get a term mu square. So, U get minus mu square by 2 sigma square and there is 1 by sigma here 1 by root 2 by you have e to the power minus x square by 2 sigma square, plus mu x by sigma square now, this is a function of parameters here. So, this can be considered as a c theta function. This constant 1 by root 2 pi can be considered as a function of a x alone and then you have you can write here T 1 x equals to X square and Q 1 theta equals to minus 1 by 2 sigma square. Similarly, here T 2 x can be taken to be x and Q 2 theta can be considered to be mu by sigma square. So, this is a distribution in two parameter exponential family.

Most of the standard distributions in a statistics that we use for example, gamma distribution with R known and lambda are known that is a distribution and exponential family. If we consider a negative exponential distribution with a scale parameter that is also in the exponential family. So, there are various distributions which are actually in the exponential family. Now, exponential families have some important features and in particular with respect to the FRC lower bound.

(Refer Slide Time: 25:37)



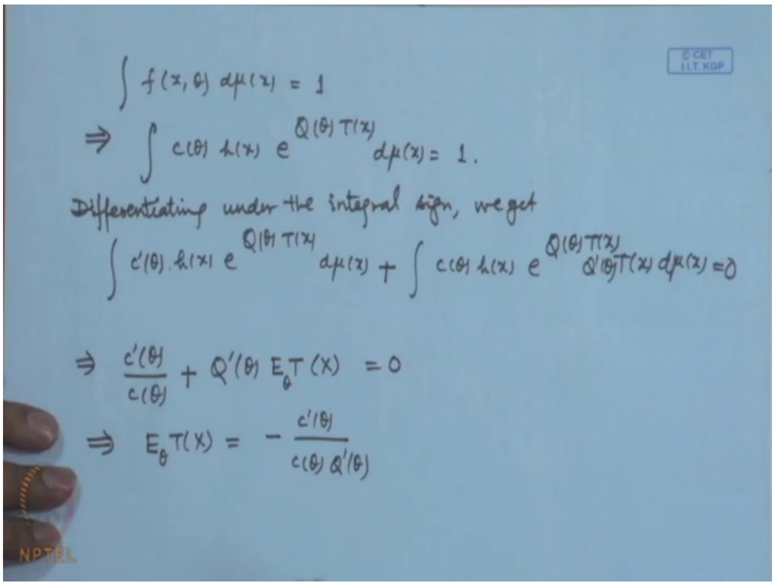
So, let us consider in the context of the lower bound. So, if we are writing 1 parameter exponential family. Let us take log of this that is equal to log of c theta, plus log of h X, plus Q theta T x, if we consider the derivatives of this with respect to theta, we get c prime theta by c theta, plus T x into Q prime theta. Now, if you remember your S X function it is nothing, but sigma dell log f x i theta by del theta for I equal to 1 to n. So, this becomes simply n times c prime theta by c theta plus Q prime theta sigma T X i.

Now, you see here this is constant as far as variable is concerned. So, this is actually a linear function of sigma T X i. So, S X theta is a linear function of sigma T X i. So, in the distributions which are in the exponential family, the variables or you can say the estimators which are linear functions of sigma T X I, the variances of them will be attaining the lower bound for the estimation of the expectations of these. So, what we are saying is, let us call it say w that is 1 by n sigma T X i. So, this is linearly related with S x theta with probability 1.

Hence, any linear function of w will be attaining the FRC lower bound for the variance of expectations, for the variance of unbiased estimators of expectations w. We can also see that what will be this expectation in general, see in this particular case see we discussed some examples like a Poisson distribution. Now, in this Poisson distribution if you see c theta is e to the power minus lambda, it is derivatives will also be equal to e to the power minus lambda.

So, you will get minus n here and Q is log lambda. So, Q prime will become 1 by lambda. So, you are getting minus n plus lambda and this will become sigma x i. So, when we say v, v is equal to x bar and this w is equal to X bar here. So, X bar is attaining the FRC lower bound for expectation of X bar that is lambda. So, we have already proved this statement, I am just once again just demonstrating, that if the distribution is in the exponential family then all the linear functions of 1 by n sigma T x i they will have variance equal to the FRC lower bound. So, this is a remarkable thing whenever we are having distribution in the exponential family there will be certain parameters for which the lower bound will certainly be attained. Now, let me now also obtain the expression for, what is the expectation of w. So, let us also we also determine expectation of w here.

(Refer Slide Time: 30:26)



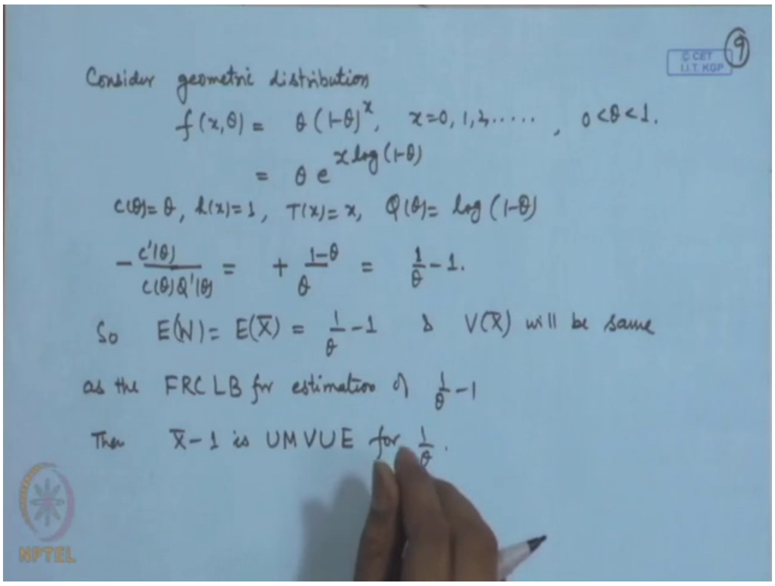
So, let us consider the integral or the summation of the density function or the mass function will be equal to 1. So, I general just integral meaning that it covers the discreet and continuous cases both. So, c theta h x, e to the power Q theta T x is equal to 1. Now, we may have certain assumptions here like differentiation to the integral signs will be assumed, because in the Rao-Cramer lower bound itself we make certain assumptions certain regularity assumptions. So, that assumptions should be true here also. So, if we assume that, then we can differentiate under the integral sign. So, we will get here there are two terms which involve theta. So, if we take the first one we get c prime theta h x, e to the power Q theta T x, d mu x.

And if you differentiate the second term, you will get c theta h X, e to the power E theta T X into T X, d mu X and of course Q prime theta will also come this is equal to 0 the right hand side is 1. So, the derivatives is going to be 0. Now, this term we can write as divided by c theta multiplied by c theta, then that will be integral of the density once again. So, that will become equal to 0, if you look at the second term this density is as such then you are forgetting this term as additional term. So, Q prime theta expectation of T X this equal to 0.

That means what we are saying, expectation of theta expectation of T x is actually equal to minus c prime theta by c theta into Q prime theta consider for example, the case of Poisson distribution, in the case of Poisson distribution c was e to the power minus lambda. So, c prime theta by c theta will become equal to minus 1 that is minus minus becomes plus Q prime theta that will become 1 by lambda.

So, if you put it in the down later you will get lambda here, then in the case of Poisson distribution will become lambda sigma of T x i by n was X bar. So, the statement is that, that x bar will attain FRC lower bound for the estimation of lambda. So, that statement we verify directly now, if we are having that estimation that exponential family this will be always true.

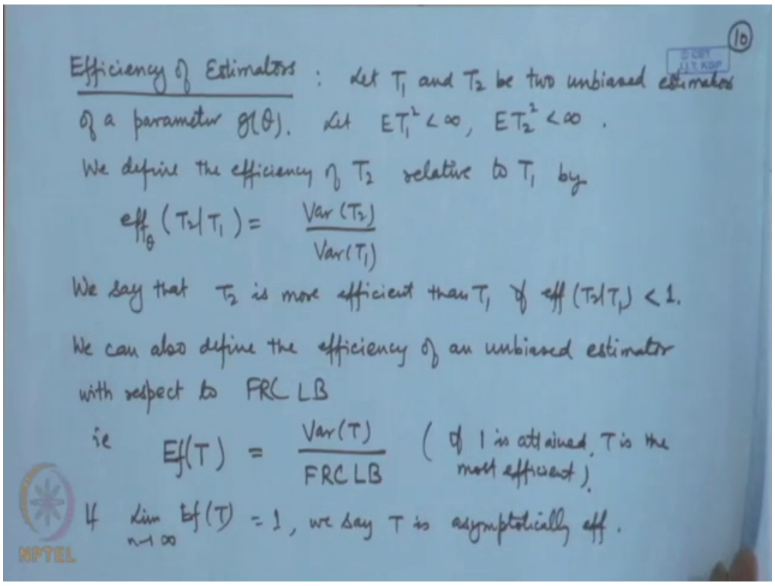
(Refer Slide Time: 33:52)



Let me take one more application here, consider say geometric distribution. Yesterday we have seen, here the form of the distribution taken theta into 1 minus theta to the power x for x is equal to 0, 1, 2 and so on. So, we can write this is equal to theta e to the power x, log 1 minus theta. So, here c theta is equal to theta, if you compare with the distribution here with the exponential family h X is 1, T x equals to x and Q theta is equal to log of 1 minus theta. So, naturally minus c prime theta by c theta Q prime theta, that is going to be equal to minus 1 c theta is theta Q prime theta will become equal to minus 1 by 1 minus theta. So, it is equal to 1 by theta minus 1; that means, and here x T x equals to x. So, w is equal to X bar.

So, expectation of w, that is equal to expectation of X bar is equal to 1 by theta minus 1 and variance of X bar will be attaining the Rao-Cramer lower bound will be same as the FRC lower bound for estimation of 1 by theta minus 1. So, now if you can see there a linear function of it is 1 by theta also we can consider. So, we can say that X bar minus 1 is minimum variance unbiased estimator for 1 by theta. So, this statement is also true now, we have discussed the concept of minimum variance; that means, among unbiased estimators, the estimators which has the minimum variance is considered to be the best. In general we can always compare to unbiased estimators by comparing the variances; that means, the one which has the smaller variance is considered to be more stable or better.

(Refer Slide Time: 36:43)



So, there is a classical concept of efficiency of estimators based on this, let me discuss that here efficiency of estimators. So, let T 1 and T 2 be two unbiased estimators of a parameter say g theta and let us assume that, they have the finite second moment this condition is required, because the variance is must exists. So, we define the efficiency of T 2 relative to T 1 by. So, we use a notation of E f T 2 given T 1, it is equal to variance of T 2 divided by variance of T 1. Naturally, if the variances are equal then the efficiency will be equal to 1.

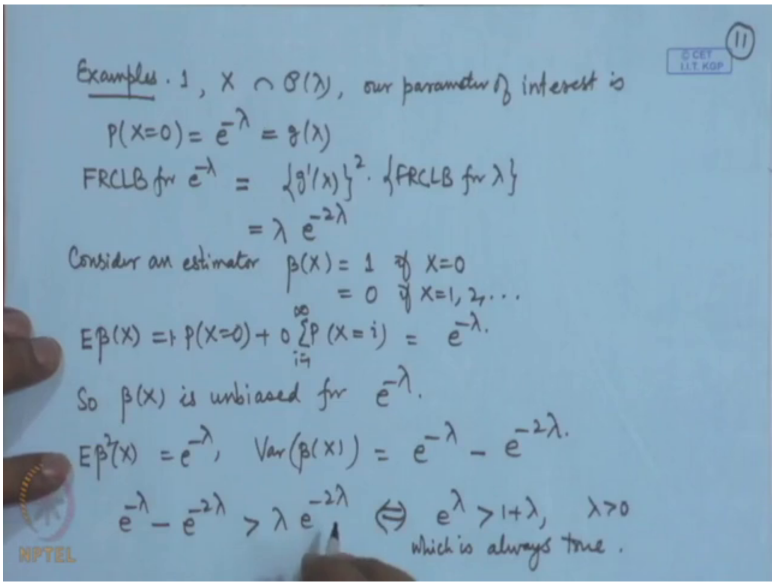
If the efficiency is less than 1; that means, the variance of T 2 is less than variance of T 1; that means, T 2 is more efficient than T 1 conversely, if the efficiency is more than 1 variance of T 2 will become bigger than variance of T 1; that means, T 1 is better than T 2. So, we say that we say that T 2 is more efficient than T 1, if efficiency functions is less than 1.

Now, this is regarding any two estimators now in general given anyone estimators we can consider its efficiency with respect to the Rao-Cramer lower bound. So, for example, we can consider estimators which attains the FRC lower bound if that is. So, then that is a benchmark or you can say the best thing. So, anything which is bigger than that it is efficiency will be considered with respect to that.

That means its efficiency will be bigger than one. So, we can also define the efficiency of an unbiased estimator with respect to FRC lower bound, that is we may say let me give another notation we may call it E notation. So, efficiency of an unbiased estimator E f efficiency of an estimator T, we define as variance of T divided by FRC lower bound for the variance of unbiased estimator for that parameter. Suddenly, we know that sometimes this may be attained and sometimes it may not be attained.

So, these definitions are not full proof another thing is that in certain cases we may not consider unbiased estimators, because if we consider only mean square as a criteria it may turn out that the means square error is less than the variance by combining certain terms. We can also consider that although this may not be attain asymptotically it attain. So, we can give a definition that, if limit of this is equal to 1 then we say that T is asymptotically efficient. So, here if 1 is attained T is the most efficient.

(Refer Slide Time: 41:22)



Let us look at some examples here, let us go back to the Poisson example and for convenience. Let me restrict attention to 1 observation, suppose X follow Poisson lambda and here our parameter of interest is say probability X is equal to 0 that is e or minus lambda of course, we may ask the question that, why we are considering this function. Now, usually a

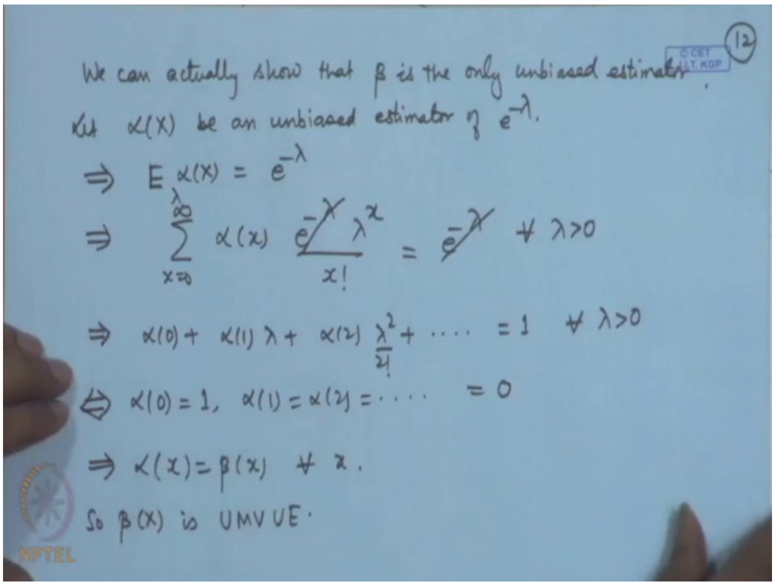
Poisson distribution is the distribution of the number of arrivals number of occurrences during a given time interval or during a given area or during a given space etcetera. Now, what happens for example, considering a Q, service Q then how many people are arriving that will denote the number X, then certainly it is of interest to know that if X is equal to 0; that means, there is a slab period.

Because there is in a service Q it may happen that we may have to employee a service personal that is the persons who will be giving the service for example, it is a railway ticket counter, it is a ticket counter at a cinema hall or it is a service counter at a popular say café. So, therefore, persons are required there are personal are required, in then when there are no person; that means, when X equals to 0 we need not deploy the people or we may deploy less number of people.

So, certainly in such cases it is of interest to know or estimate the probability of 0 occurrences. So, this gives us this parametric function, e to the minus lambda certainly, it is a non-linear function of lambda therefore, the variance of unbiased estimator of e to the power minus lambda can never attain the lower bound. Let us look at this what will be the lower bound FRC lower bound for e to the power minus lambda that is equal to g prime lambda square into the FRC for lambda, for lambda it is lambda by n and if n equal to 1 then it is simply lambda. The derivative of g lambda is e to the power minus lambda with a minus sign when we square it with e to the power minus 2 lambda.

So, this is lambda this is a lower bound. So, now let us consider an estimator say beta X is equal to 1, if X equals to 0 it is equal to 0, if X equal to 1 2 and. So, on then if you look at expectation of beta X that is equal to 1 into probability X equal to 0 plus 0 into probability is equal to X say I is equal to 1 into infinity. So, this becomes 0. So, this is e to the power minus lambda. So, beta X is unbiased for e to the power minus lambda; however, if you look at expectation beta square. Now, this will again be same and therefore, variance of beta that is also e to the power minus lambda minus e to the power minus 2 lambda. Now, if you compare this with the lower bound, e to the power minus lambda minus e to the power minus 2 lambda greater than lambda e to the power minus 2 lambda. Because this is equal into e to the power lambda greater than 1 plus lambda for lambda positive which is always true. So, you can see that this lower bound is not attained.

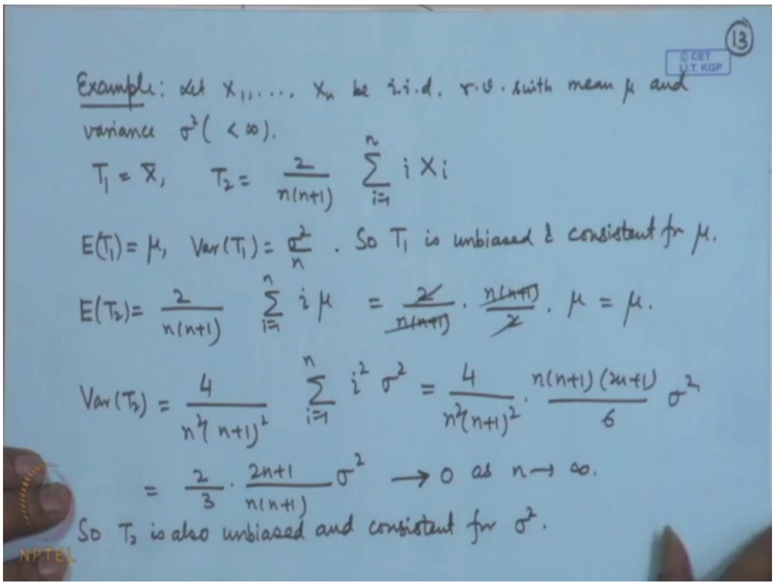
(Refer Slide Time: 45:53)



However, we can use another argument to actually prove that beta X is we can actually show, that beta is the only unbiased estimator. We can proceed by the basic principles let us consider say alpha X, let alpha x be a unbiased estimator of e to the power minus lambda, then we should have expectation of alpha x equal to e to the power minus lambda. Now, let us write down this relation alpha x, e to the power lambda, and lambda to the power x by x factorial is equal to e to the power minus lambda for all lambda. Now, this e to the power lambda you can remove from both the sides, because this is positive term. So, this reducing to then alpha 0 plus alpha 1 into lambda, plus alpha 2 into lambda square by 2 factorial and.

So, on is equal to 1. So, left hand side is a power series in lambda, hand right hand side is simply constant. So, this is true, if an only the coefficients match; that means, alpha 0 must be 1 and alpha 1 alpha 2 and so on. All of them must be 0 which is the same as the function beta, because beta 1 beta 0 was 1 and beta 1, beta 2 and. So, on when all of them were 0. So, this alpha function and beta functions are the same. So, beta must be UMVUE. So, although here the lower bound is not attained, but actually beta will be the most efficient estimator here.

(Refer Slide Time: 48:08)



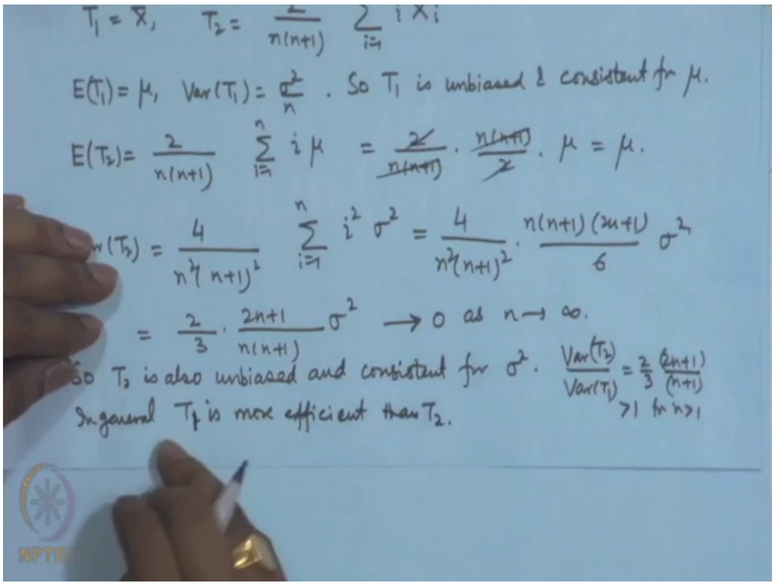
Let me give an example of comparing two unbiased estimators with respect to their variances. The estimators may both may be unbiased, both may be consistent etcetera. So, let us tale another example, I am not taking any distributional form let us consider say X 1, X 2, X n, d independent and identically distributed random variables with say mean mu and variance sigma square; obviously, we are assuming that variance is finite here now, you consider two unbiased estimators.

Let me take T 1 equal to X bar and T 2 is equals to 2 by n into n into plus 1, sigma I, x I, I equals to 1 2 n now; obviously, if you look at expectation of T 1 this we have seen that a sample mean is unbiased for the population mean, the variance of this is equal to sigma square by n. So, if estimators is unbiased it is variance converges to 0, then we also know that it will be consistent. So, what we are seeing is that T 1 is unbiased and consistent for estimating mu.

Now, if you look at T 2. So, that is equal to expectation of T 2 is 2 by n into n plus 1, sigma I is equal to I to 2 n I, expectation of x i that is again mu. So, sigma I is n into n plus by 2. So, you get 2 by n into n plus 1 into n into n plus 1 by 2 into mu. So, this terms cancel out you get only mu. So, T 2 is also unbiased let us look at variance of T 2 now, variance of T 2 if you take this is constants this will become square, 4 n square into n plus 1 square sigma I equals to n I square into variance X i.

Variance of x i is sigma square since, we have assume independence of the observations the correlation of co variance term will not come here, you will get this now sigma square we have the formula. So, you get 4 by n square into n plus 1 whole square n into n plus 1 into 2 n plus 1 by six sigma square. So, after simplification you get it as 2 by 3, 2 n plus 1 divided by n into n plus 1 sigma square. So, as n tense to infinity this goes to 0.

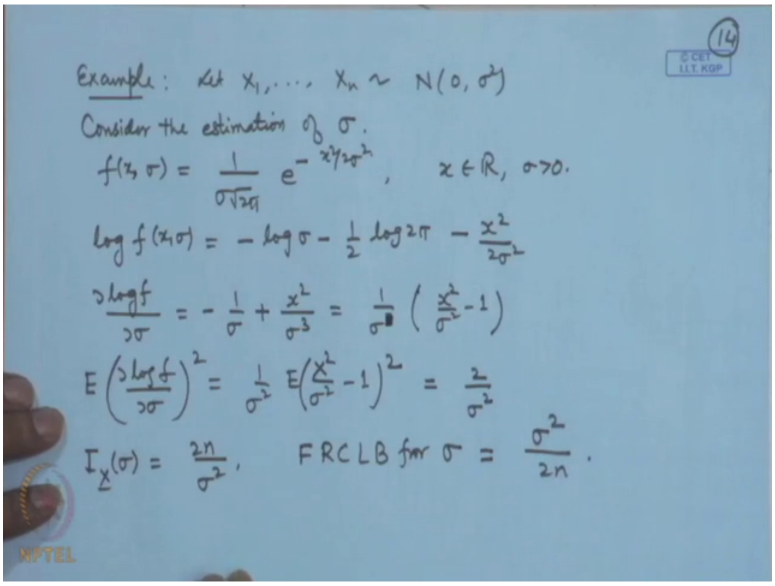
(Refer Slide Time: 48:08)



So, T 2 is also unbiased and consistent for sigma square; however, let us compare variances, what is variance of T 2 by variance of T 1, variance of T 2 divided by variance of T 1. So, sigma square is coming here sigma square is appearing here by n by n. So, that will cancel out.

So, you get the term as 2 by 3, 2 n plus 1 divided by n plus 1; obviously, this is always greater than 1 for n greater than 1, if n equals to 1 of course, this will be equal to 1 and if n equal to 1 actually T 1 and T 2 are both equal to X 1. So, that case is of not interest. So, in general T 2 T 1 is more efficient than T 2. So, here you are seen we have two estimators both of which are unbiased as well as consistent for the sample mean, but for 1 of them can be preferred over the other, if we are applying the criteria of smaller variance.

(Refer Slide Time: 52:56)



So, now let me also take another distribution say suppose, we consider a random sample from a normal distribution where mean in assume to be 0 and variance is sigma square. We have already discussed this example in the context of estimation of sigma square when mu was some fixed value mu not now. Whenever, mu is some fixed value mu not you can always shift the observations. So, the mean can be made to be 0.

Now, suppose my interest is not to consider estimation of sigma square, but the estimation of sigma. So, consider the estimation of sigma say now, let us look at the lower bound. The density function is of the form when by sigma root 2 pi by e to the power minus x square by 2 sigma square where x is of course, any real value.

Log equal to minus log sigma minus 1 by 2, log 2 pi minus x square by 2 sigma square. So, derivative of this with respect to sigma that is minus 1 by sigma minus now, derivatives of this will become 0 and derivative of 1 by sigma square is minus 2 by sigma cube. So, it will become x square by sigma cube that is equal to 1 by sigma cube, we can write it as 1 by sigma x square by sigma square minus 1.

So, expectation of del log f by del sigma is equal to 1 by sigma square expectation of x square by sigma square minus 1 whole square. Now, if x follows normal 0 sigma square then x by sigma follows normal 0 1, x square by sigma square will follow chi square on 1 degree of freedom. So, therefore, this will have expectation 1 and therefore, expectation of the variable minus it is mean square that is going to be the variance. Now, variance of a chi square is twice it is degrees of freedom. So, this term will become equal 2. So, this is simply equal to 2 by sigma square. So, if we consider the information that will be equal to 2 n by sigma square. So, the FRC lower bound for estimation of sigma that will be equal to sigma square by 2 n.

In the following class, I will consider two estimators for this see whether they any of them attain the lower bound and also compare them. So, that I will be doing in the following lecture.