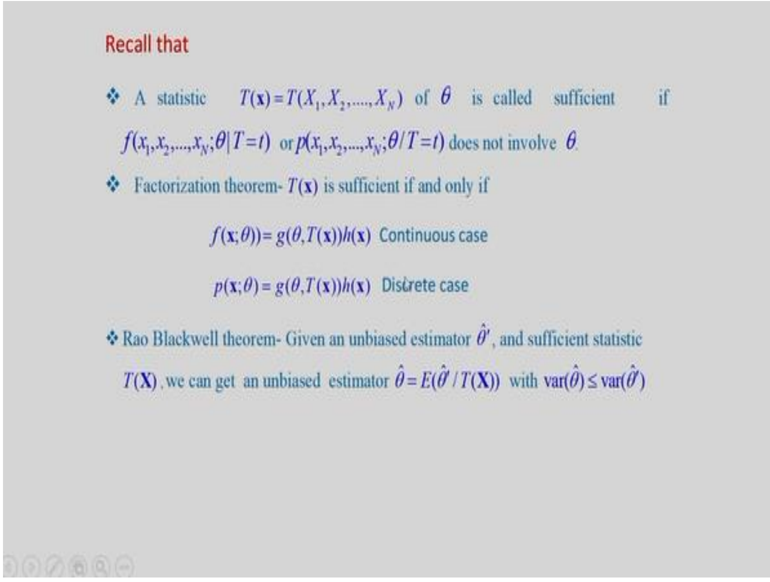
Statistical Signal Processing
Prof. Prabin Kumar Bora
Department of Electronics & Electrical Engineering
Indian Institute of Technology, Guwahati

Lecture No 12
MVUE through Sufficient Statistics II

Hello Students, welcome to lecture 12 MVUE through sufficient statistics II.

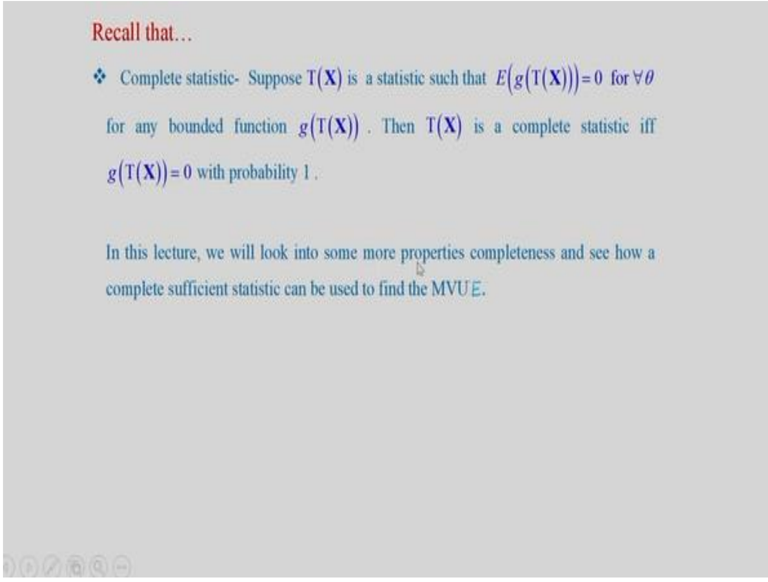
(Refer Slide Time: 00:42)



I recall that, a statistics Tx that is a function of X1, X2 up to XN is called sufficient, if the conditional period f or conditional period PMF does not involve theta. That is the definition of sufficient statistics and we discussed also Factorization theorem Tx is sufficient if and only if. The joint PDF is a function of theta can be factorized into two vectors. One vector g is function of theta and Tx, other vector is h that is a non-zero function of X1, X2 up to XN only.

So the same can be written in the case of discrete case also in that case probability mass function as a function of theta is product of these two vectors. Rao Blackwell theorem- given an unbiased estimator theta this head and a sufficient statistic TX we can get an unbiased estimator theta head that is given by E of theta head given TX which is also unbiased and which has variance, less than the variance of theta this head. So that way we can get a better estimate or using the Rao Blackwell theorem.

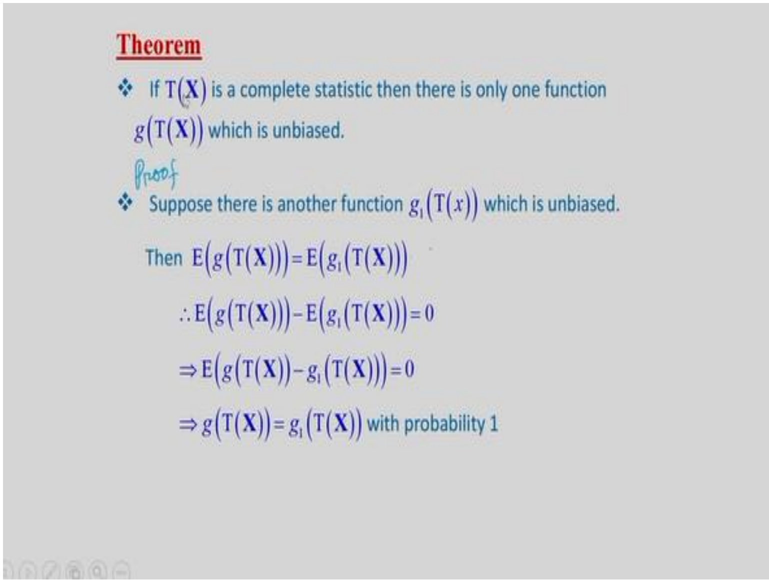
(Refer Slide Time: 02:15)



We also discussed complete statistic suppose T X is a statistic such that E of g TX is equal to 0, for all theta and for any bounded function g TX. Then TX is complete if and only if g TX is equal to 0 with probability 1. That means if this is equal to 0 expectation is equal to 0 then g TX must be equal to 0 with probability 1, so that way we define complete statistics.

In this lecture we will look into some more properties of completeness and see how a complete sufficient statistic can be used to find the MVUE, minimum variance unbiased estimator.

(Refer Slide Time: 03:08)

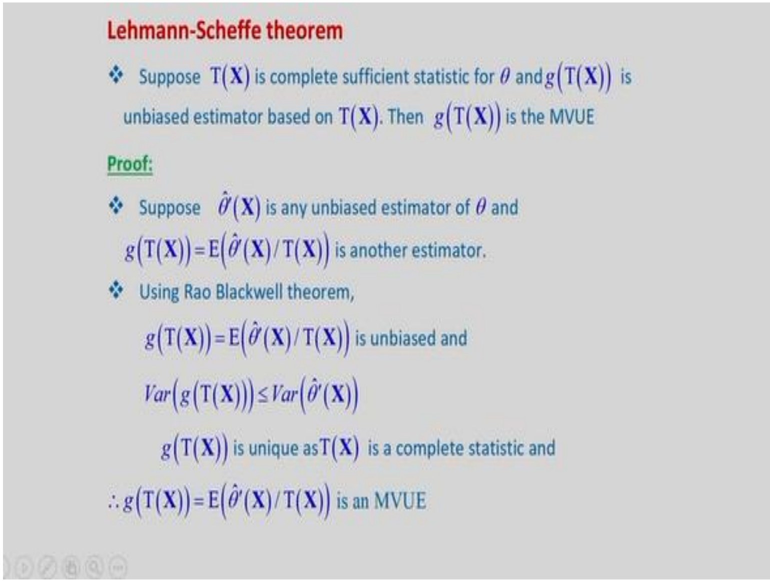


We will start with a theorem, if TX is a complete sufficient statistic then there is only one function g TX which is unbiased. So given TX there is only one function g TX which is unbiased, so that is important. Now, if TX is complete and unbiased, then we can find only one function g TX which is unbiased. We will proof this theorem, suppose there is another function g1 TX which is unbiased.

So already g TX is unbiased, suppose there is another function g1 TX which is unbiased. Then E of g TX must be equal to E of g1 TX, therefore if I take the E of g1 TX to the left hand side I will get E of g TX minus g of g1 TX, that must be equal to 0 and we can take E outside E of g TX minus g1 TX must be equal to 0. Now TX is a complete statistic therefore E of g of TX minus g1 TX is equal to 0. Then this argument must be equal to 0 with probably 1.

So this will replace that g TX is equal to 0 and TX with probability 1, therefore what do we conclude that g TX is a unique unbiased estimator. If TX is a complete sufficient statistic there exists only one unbiased estimate that is TX.

(Refer Slide Time: 05:07)

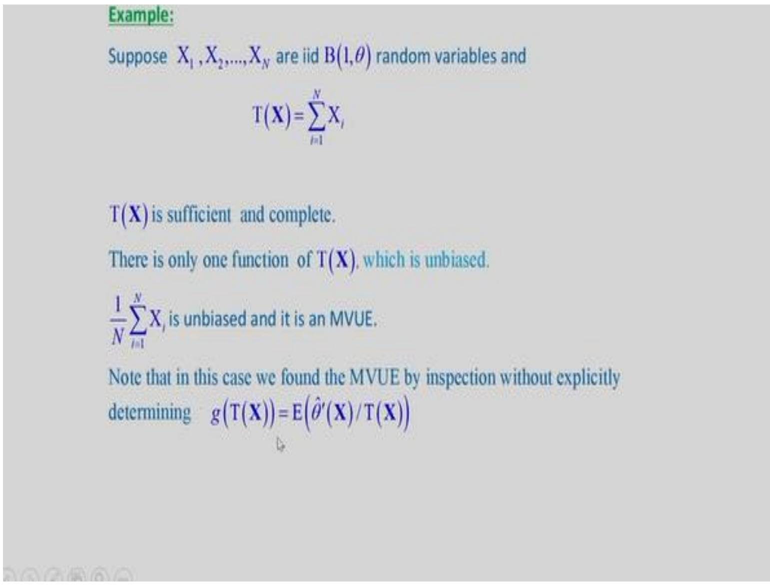


Now we will prove one classical theorem, which will help us to find out the MVUE. Suppose TX is complete and sufficient statistic for theta and g TX is unbiased estimator based on TX. So we have an unbiased estimator based on TX, then g TX is the MVUE will proof this; suppose theta dash head X is any unbiased estimator of theta. Then g of TX, E of theta dash head X given TX is another estimator.

And now using Rao Blackwell theorem, this estimator is unbiased and variance of g TX is less than equal to variance of theta dash head X but we know that TX is a complete statistics therefore there is only one function g TX which is unbiased. Therefore g TX is unique and its variance is less than the variance of any other estimator theta dash head X.

So therefore g TX equal to E of theta dash head X given TX is an MVUE, because it is unbiased and it is unique and its variance is less than equal to any other unbiased estimator therefore this will be the MVUE. So we saw that using the Lehmann Scheffe theorem, we can find out an unbiased estimator which has the minimum variance.

(Refer Slide Time: 06:57)



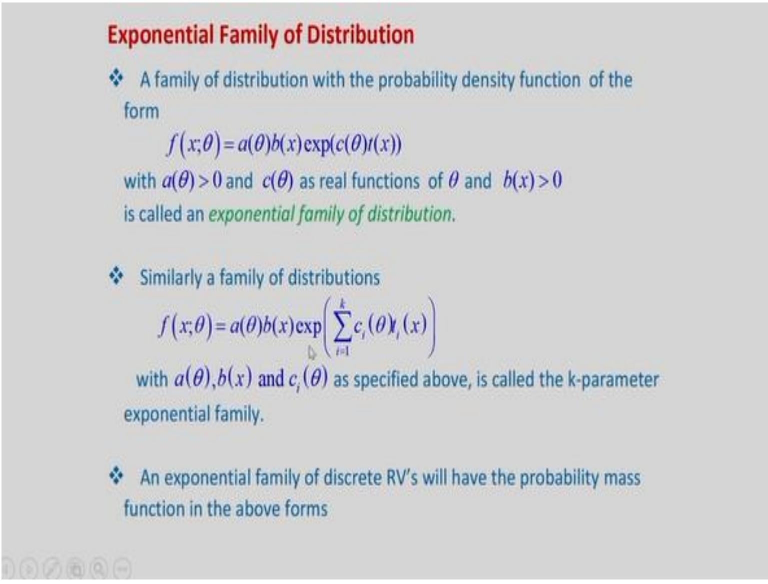
We shall consider one example; suppose X1, X2 up to XN are iid independent and identically distributed Bernoulli theta random variables, this symbol represents the Bernoulli distribution with parameter theta, theta the probability of success. So TX summation Xi, i going from 1 to N, this is the statistic we are considering we have already proved that this statistic is sufficient and it is complete also.

So there is only one function of TX which is unbiased, that is according to Lehmann Scheffe theorem. Now by inspection we get that if we divide this sum by N that is 1 by N summation Xi are going to N this will be unbiased, because you can take the expectation here and there will be N times expectation and then N, N will get cancelled. Therefore this is unbiased and I know that according to Lehmann Scheffe theorem there is only one unbiased estimator which is a function of the complete statistic.

Therefore this estimator must be an MVUE, so that we see how the completeness has helped us to find out the MVUE. Note that in this case we found the MUEVE by inspection without explicitly determining g TX. So this is the formulation I want without this formulation we could find out this quantity but in all situation it may not be possible to find a complete sufficient statistic which is unbiased by inspection.

So we have to compute this statistic or this MVUE through this operation E of theta head dash given TX, so we have to perform this operation to find out the TX.

(Refer Slide Time: 09:17)

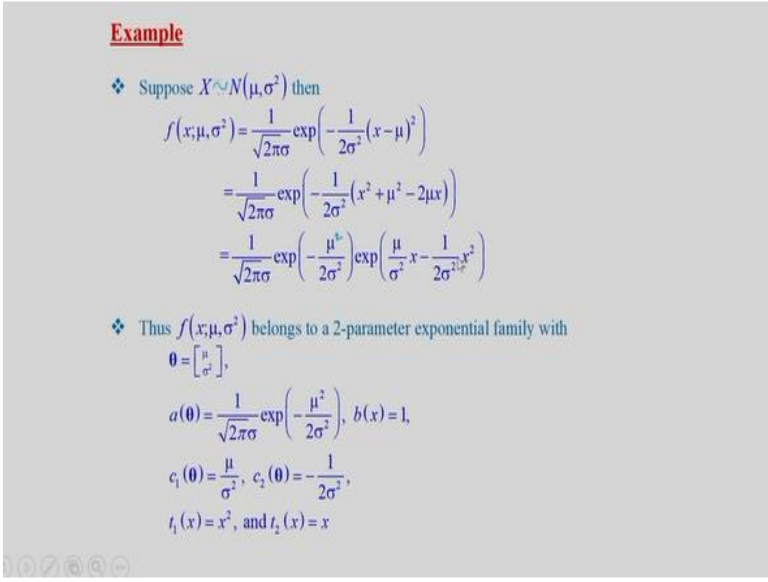


We will discuss one important family of distribution for whose finding out a complete sufficient statistic is easier such a family is the exponential family of distribution, a family of distribution with the probability density function of the form f x theta that is a function of theta and it can be written as a theta1 term is a theta another term is bx and third term is exponential of c theta into tx.

So if we can write in this form with a theta greater than 0 because it is a probability this quantity we get a theta greater than 0 and c theta is a real function of theta and bx greater than 0. Then this family will be called the exponential family of distribution with one parameter and now we can have a Muelti parameter exponential family of distribution, so suppose f of x theta is equal to a theta into bx into exponential of summation ci theta into ti x, i going from 1 to k.

Suppose in this form where as in here, a theta is greater than zero bx is also greater than zero, non-negative function and ci of theta is also earlier function like this and then this family is known as the k parameter exponential family of distribution. We can find out an exponential family of discrete random variables also in the similar manner if the probability mass function can be written in this form then that family will be called the exponential family of distribution.

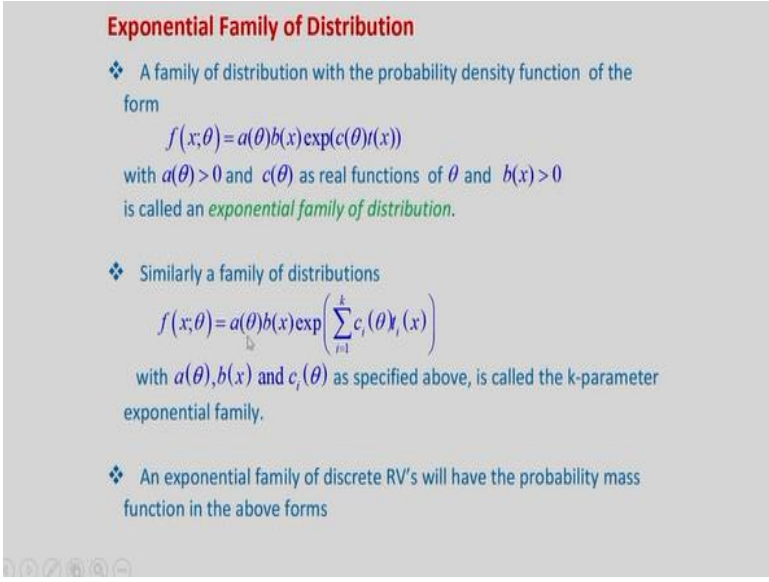
(Refer Slide Time: 11:15)



We will see one example; suppose X is the normal distribution with mean Mue and variance Sigma square and Mue and Sigma square are unknown, then the PDF we can write like this f of x, Mue, Sigma square it is a function of Mue and Sigma square that is equal to 1 by root over 2 Pi Sigma into exponential minus one by 2 Sigma square into x minus Mue whole square.

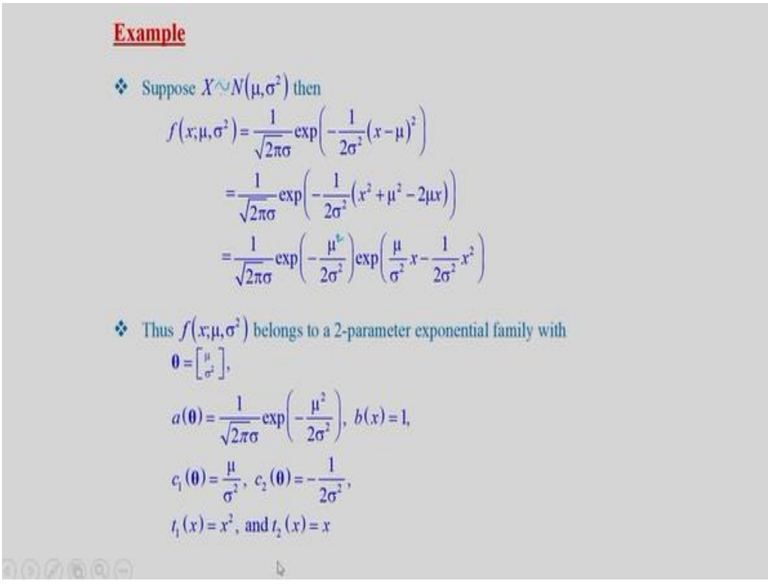
So this we can write like that is 1 by root over 2 Pi Sigma into exponential minus 1 by 2 Sigma square into S square plus Mue square minus 2 Mue X. Now this Mue square term we can take outside, so that way it will be 1 by root over 2 Pi Sigma into exponential minus Mue square by 2 Sigma square into exponential Mue by Sigma square X this part, minus because of this minus sign 1 by 2 Sigma square into X square.

(Refer Slide Time: 12:38)



So now we can compare this expression with this distribution and that is there should be a theta term bx term and then exponential term in this summation form.

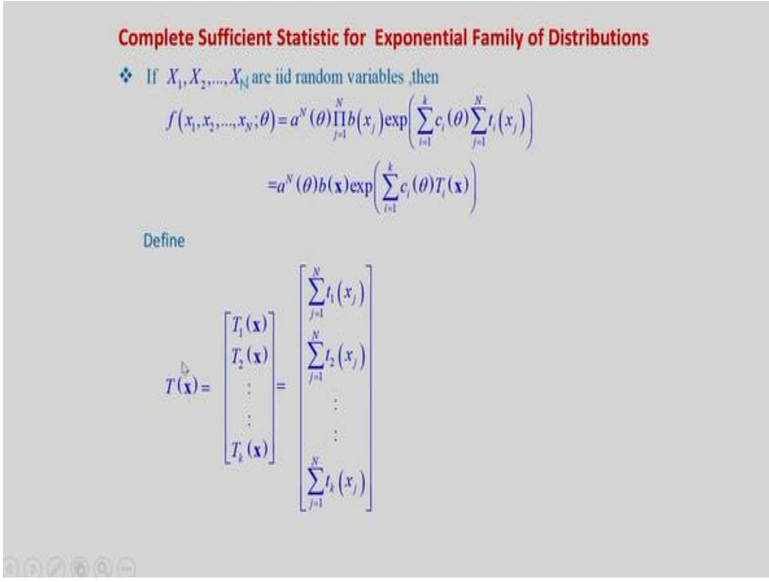
(Refer Slide Time: 12:47)



So there we see that they are sum of two quantities here and this is the a theta term that is it is a function of Mue and Sigma Square and bx will be equal to 1 and then exponential term is this summation of these two quantities. So therefore f x, Mue, Sigma square belongs to a two parameter exponential family with theta that is parameter vector is equal to Mue Sigma square it comprises of Mue and Sigma square.

Then, a theta will be equal to 1 by root over 2 Pi Sigma into exponential minus Mue square by 2 Sigma square. So it is a function of Mue and Sigma Square and bx is equal to 1, c1 theta will be equal to this is Mue by Sigma square, c2 theta will be minus 1 by 2 Sigma Square. T1 x is equal to x and t2 x is equal to x square. So that way normal distribution is a member of the exponential family of distribution.

(Refer Slide Time: 13:58)

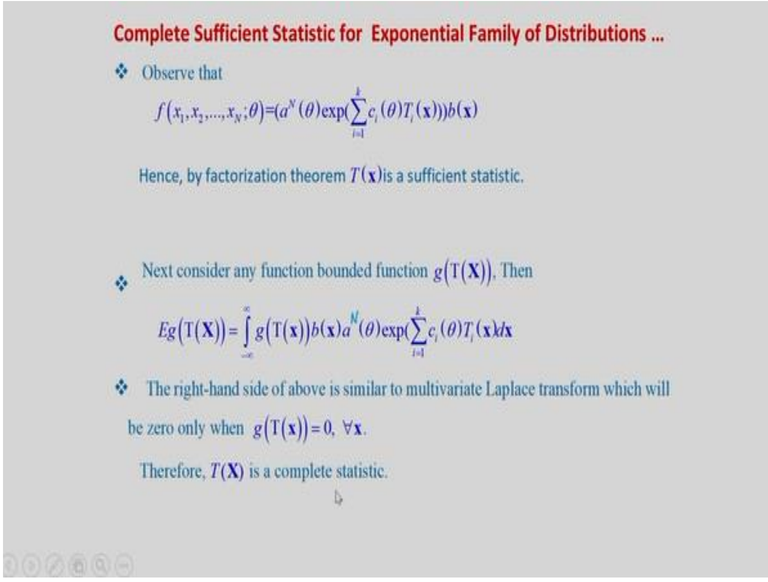


Now we will see that, complete sufficient statistics for exponential family of distribution, can be found from the expression for the joint PDF itself. Suppose, if X1, X2 up to XN are iid random variables of the exponential family, then we can write f X1, X2 up to XN as a function of theta, that will be because there are N terms now this term will be a to the power N theta into the product b of x j, j going from 1 to N into exponential i going from 1 to k, ci theta into summation j going from 1 to N, ti xj.

So this way and this is the concern of xa we are considering because there are n terms will get their expression like this; now, this we can write as a to the power N theta into b of x vector, because there is a product of xa terms, so that we can write as b of x vector. Similarly here also, this summation also we can write as a statistic Ti x. Now my complete statistic will be Tx that is equal to T1 x, T2 x up to Tk x, where T1 x will be the partial summation for j is equal to 1 to N, T1 xi and similarly T2 x will be summation T2 xj, j going from 1 to N.

So in that way we can find out N statistic from this and they can be written in a vector from T of x vector.

(Refer Slide Time: 15:52)

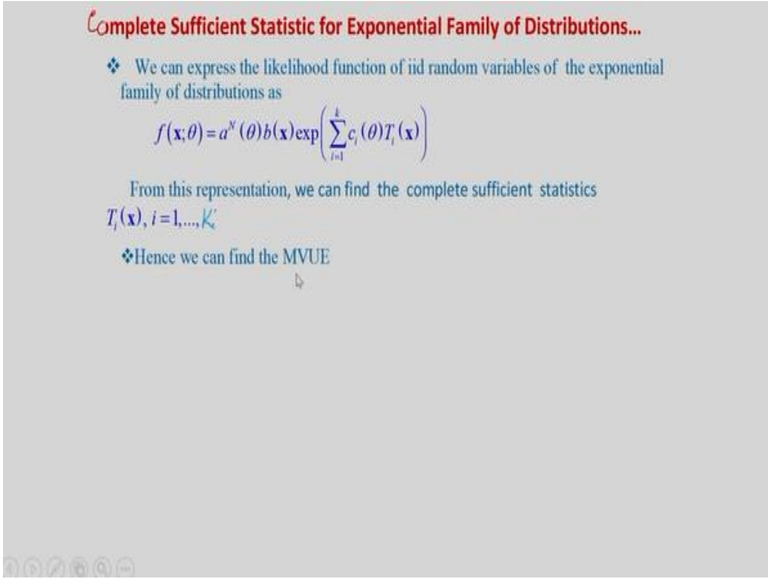


Now we observe that, this joint PDF can be written like this, a to the power N theta into exponential summation ci theta Ti x, i going from 1 to k, this is one vector. We see the function of Ti x that is sufficient statistic N theta and into bx, there is another term bx which is simply a function of x. So we can express joint PDF as a product of two vectors one vector is independent of any parameters, it is a simply a function of x and other factor is a function of the parameter and the estimator.

So therefore by factorization theorem Tx is a sufficient statistic, we have to prove that it is a complete statistic. Next consider any bounded function of the form g Tx, then we have to find out expected value of g Tx to find out the completeness therefore e of g Tx that is integration from minus infinity to infinity of g Tx, that is the function into PDF. Now PDF is bx into a to the power n theta into exponential summation ci theta Tx into bx.

So this integration is the expected value of g Tx. Now this right hand side because it is an exponential of Multiple terms so this can be interpreted as Multivariate laplace transform and which will be 0 only when g Tx is equal to 0 if this function is 0 only then this Multivariate Laplace transform can be equal to 0, therefore g Tx must be 0 with probability 1, therefore Tx is a complete statistic. So we saw that Tx is not only sufficient but it is also a complete statistic.

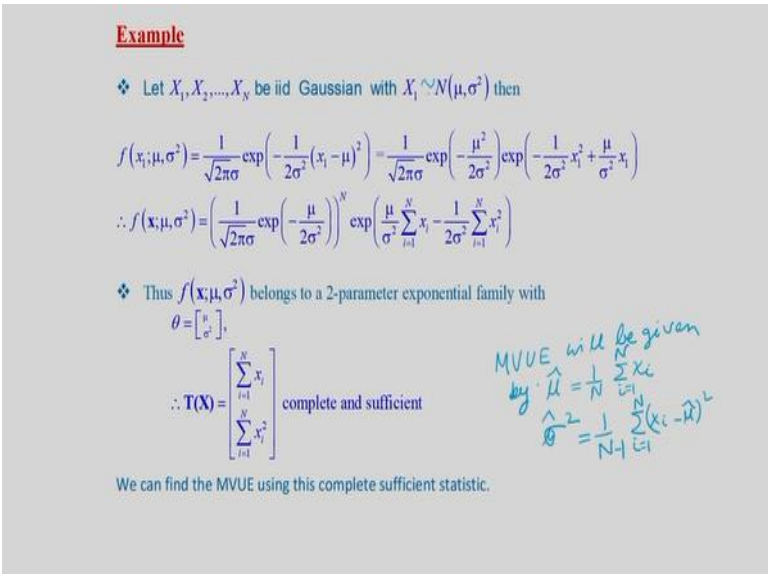
(Refer Slide Time: 18:08)



Now given a complete sufficient statistics for exponential family of distribution, we can find out the sufficient that is in this term itself we can express the likelihood function of iid random variables of the exponential family of distribution like this f of x theta is equal to a to the power N theta into bx into exponential summation ci theta Ti x, i going from i to k. From this representation we can find out the complete sufficient statistic Ti x, for i is equal to 1 to k.

Hence we can find the MVUE as a function of this Tx which is unbiased and we know that only one unbiased estimator exists and which has variance less than or equal to variance of any other unbiased estimator. Therefore that estimator will be MVUE.

(Refer Slide Time: 19:12)



So we will go back to example like again; let X1, X2 up to XN be iid Gaussian with X1 are distributed as normal distribution with mean Mue and variance Sigma square and therefore if I consider only one element, that is x1 only then it is PDF is given by this 1 by root over 2 pi Sigma into exponential minus half of 2 Sigma square into x1 minus Mue whole square.

So this we can express in the exponential family of distribution like this 1 by root over 2 Pi Sigma into exponential of minus Mue square by 2 Sigma square into exponential of minus 1 by 2 Sigma square x1 square plus Mue by Sigma square into x1. Now for N random variables the joint PDF as a function of Mue Sigma square which is the likelihood function.

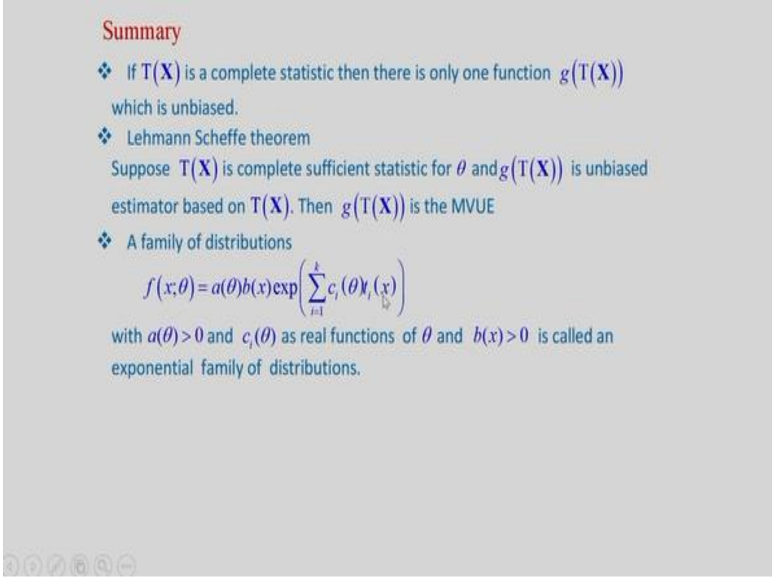
We can write this as this vector to the power N and then exponential because of the product now all summation will come exponential Mue by Sigma square into summation xi, i going from 1 to N minus 1 by 2 Sigma square into summation xi square, i going from 1 to N. So according to our earlier result this is an exponential family of distribution and here these are the statistics, these are the complete sufficient statistics.

Thus f of x, Mue, Sigma square is a function of Mue sigma square belongs to a two parameter exponential family with parameter vector is Mue by Sigma Square and TX is given by first component is summation xi, i going from 1 to N, second component is summation xi square, i going from 1 to N and these two statistics are complete and sufficient. So therefore you can find out MVUE using this complete sufficient statistics.

So MVUE will be given by, MVUE will be given by parties we have to find out MVUE be the Mue power is Mue head. This is simple we have just divided by N, so that will be 1 by N into summation xi, i going from 1 to N and the other one for this we can so that this function will be given by Sigma head square estimate that will be equal to 1 by N into summation xi minus Mue head whole square, i going from 1 to N, this is not N this is N minus 1.

Then this estimator will be unbiased and this is the MVUE for Sigma head square. So that is we see that for exponential family of distribution we can find out the MVUE from the expression for the joint PDF itself.

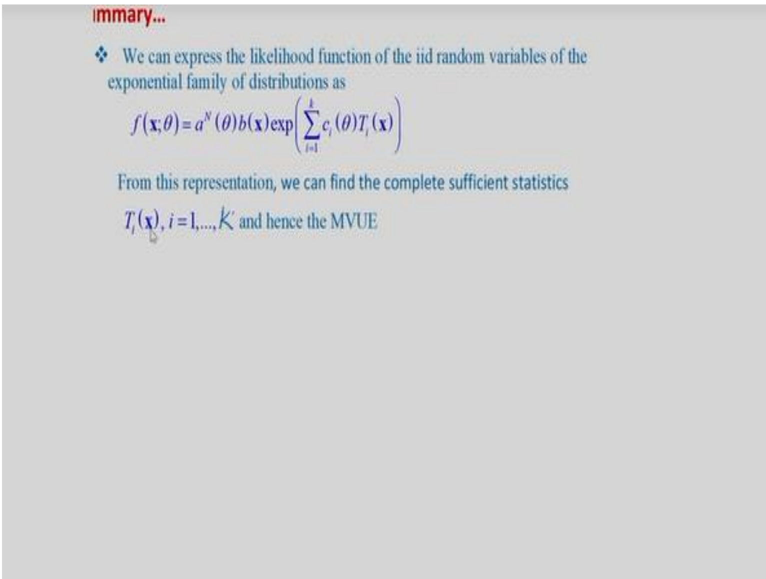
(Refer Slide Time: 22:43)



Let us summarize first point we note that if TX is a complete statistic then there is only one function g T x which is unbiased that we have established. Then Lehmann Scheffe theorems suppose, TX is a complete sufficient statistic for theta and g TX is unbiased, then g TX must be an MVUE. So if we have a function of complete sufficient statistic which is unbiased then that must be an MVUE.

Then we discussed about the family of distribution, which is known as the exponential family of distribution. So it is given by f of x theta is equal to a theta into bx into exponential of k sums that is i going from 1 to k of ci theta into Ti x. So this is the exponential family of distribution with a theta greater than 0, ci theta is a real function and bx is also greater than 0 because this is a probability therefore this must be greater than 0 and this is the exponential family of distribution.

(Refer Slide Time: 24:11)



We can express the likelihood function of the iid random variables of the exponential family of distribution in this form, that is the f of x vector theta as a function of theta this is the likelihood function or joint PDF this can be written as a to the power N theta into b of x vector exponential summation ci theta Ti x, i going from 1 to k.

We can write in this form and from this representation we can find out the complete sufficient statistics as Ti x, i going from 1 to k, we have found out their complete sufficient statistic from the exponential family of distribution and from that complete sufficient statistic we can find out the MVUE because we have to find out a function of Ti x which are unbiased that will give you the MVUE.

So that way we saw that we can find out the MVUE by employing two processes first one is if we can find out the expression for Cramer Rao Lower Bound, then that will give us to the MVUE. We next we saw that MVUE can be found out through complete sufficient statistics. But both this a process involve complicated analysis, we have to apply some simpler techniques to find out a good estimator for parameters, that we will be discussing in the next lecture, thank you.