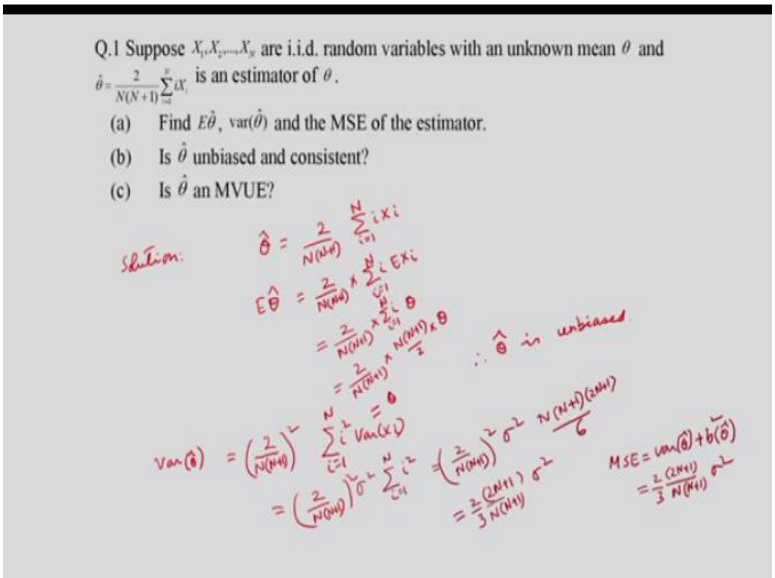
Statistical Signal Processing
Prof. Prabin Kumar Bora
Department of Electronics and Electrical Engineering
Indian Institute of Technology - Guwahati

Lecture – 28
Solution to Review Assignment 2

(Refer Slide Time: 00:48)



Hello students, welcome this session on solution to review assignment 2, question 1; suppose X1, X2 up to XN are iid random variables with an unknown mean theta and theta hat this is the estimator is given by 2 by N into N + 1 into summation i X i going from 1 to N, find E of theta hat variance of theta hat and MSE of theta hat, is theta hat unbiased and consistent, is theta hat an MVUE?

Solution; so we are given that theta hat is equal to 2 by N into N + 1 into summation i Xi i going from 1 to N, therefore E of theta hat will be equal to 2 by N into N + 1 into; now expectation of this sum is same as the sum of the expectation, so that way into summation i going from 1 to N, i is a constant into E of Xi, so that way now, unknown mean is theta, therefore this will be equal to 2 by N into N + 1 into theta summation i; i going from 1 to N into theta.

Theta is common everywhere, so this summation is now given by N into N + 1 by 2, so that way 2 by N into N + 1 into N into N + 1 by 2 into theta, so this will be equal to theta, therefore theta hat is unbiased. We can find out the variance of theta hat because it is an unbiased estimator, so variance simply we can find out, that is equal to 2 by N into N + 1 whole square.

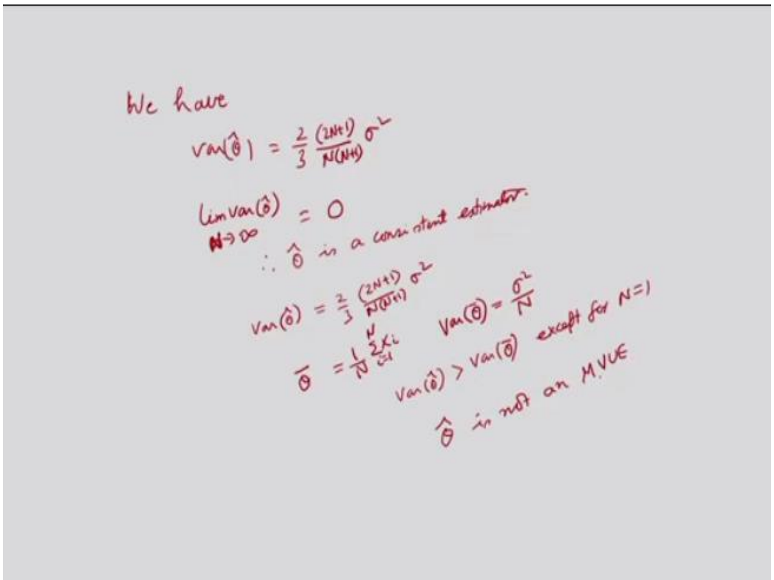
Because this is a constant, if we take the variance this constant will be square into summation i again a constant, so it will be i square into variance of Xi i going from 1 to N and all the cross terms will become 0 because Xi's are independent, so that way this will be equal to 2 by N into N + 1 whole square. Now, variance of Xi, this is iid, therefore variance of Xi is same for all Xi’s.

So, we can take it out, so this will be sigma square suppose, then this will be summation i going from 1 to N, summation of i square i going from 1 to N and this is equal to; so 2 by N into N + 1 whole square sigma square and this is N into N + 1 into twice N + 1 divided by 6. So, this if we simplify this expression, we will get this is equal to; so 2/3rd will be here, 2/3rd of then here N is there, N is there, 1N will get cancelled.

So, N into N + 1 here and this part will be there twice N + 1 here and then this sigma square, so this is the variance of theta hat. So, we have found out the E of theta, variance of theta, we have found out E of theta hat and variance of theta hat. Now, we will see MSE of theta hat, now MSE of theta hat, so we know MSE is equal to variance; var of theta hat + bias of theta hat square.

But since E of theta hat is equal to theta, bias is 0, so that way this will be simply this expression only, 2 by 3 twice N + 1 divided by N into N + 1 into sigma square, so that theta hat is biased. Is theta hat unbiased and consistent; so next question is whether theta hat is consistent, so since it is an unbiased estimator, now we have variances.

(Refer Slide Time: 07:37)



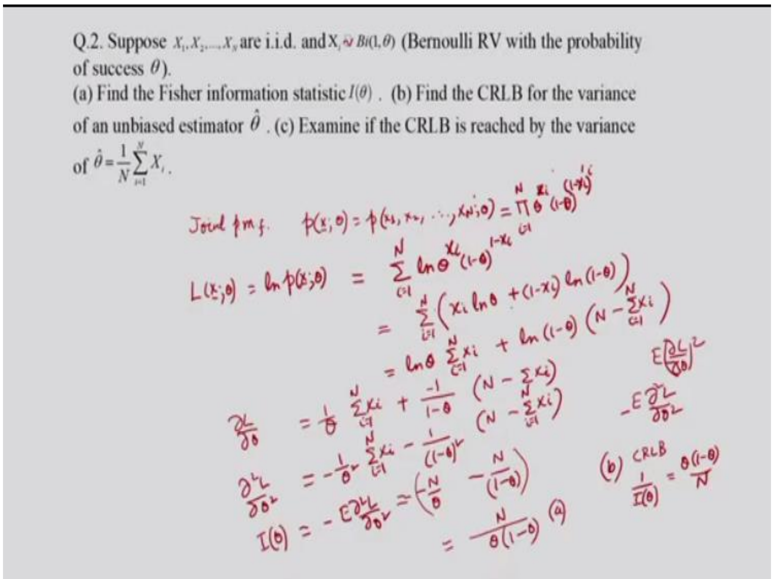
So, to prove consistency, we have variance of theta hat is equal to that is 2/3rd of twice N + 1 divided by N into N + 1 into sigma square, so power to take because it is an unbiased estimator to take whether it is consistent or not, we have to see limit variance of theta hat as N tends to infinity. So, we see that there here N is there, so this part as N tends to infinity, it will become 0.

So, we see that limit of variance of theta has N tends to infinity is equal to 0, therefore theta hat is a consistent estimator. Next question is; is theta hat an MVUE, so we have variance of theta hat is equal to; we know that there is another estimator suppose, simple mean that is theta bar is equal to 1 by N into summation Xi i going from 1 to N and in this case, variance of theta bar this will be equal to, if another estimator theta bar is there, this variance of theta bar will be equal to, we can show that this is sigma square by N.

Now, this quantity is less than this, so variance of theta bar is less than variance of theta hat, we can because here this is, this expression; this expression is always larger than 1 by N, of course for N is equal to 1 both will be same but otherwise, this variance of theta hat will be greater than variance of theta bar except for N is equal to 1; N is equal to 1 in that case this will be also sigma square, this will be also sigma square.

So, therefore they are exist an estimator theta bar with variance lower than the variance of this estimator, therefore this theta hat is not an MVUE, it is not an minimum variance unbiased estimator, it is unbiased estimator but it does not have the minimum variance.

(Refer Slide Time: 11:16)



Next question; suppose, X1, X2 up to XN are iid and Xi distributed as Bernoulli distributed, Bernoulli random variable, this is the symbol for Bernoulli random variable, this is binomial with parameter 1 and theta. Find the Fisher information statistic I theta, find the CRLB for the variance of an unbiased estimator theta hat, examine if the CRLB is reached by the variance of theta hat is equal to this summation Xi i going from 1 to N divided by N.

So, this is the estimator whether the variance of this estimator satisfy CRLB; Cramer Rao lower bound, now in this case there are N random variables are there, therefore the joint PMF; probability mass function that is also likelihood function that is equal to p of; if I write in vector notation, X theta, so that is equal to p of X1, X2 up to XN and as a function of theta and this is equal to now because it is iid independent and identically distributed.

So, it will be product i is equal to 1 to N and each random variable is Bernoulli distributed, therefore this will be theta to the power Xi and 1 - theta to the power 1 – Xi, so this is the probability; joint probability mass function, so we can take the logarithm; L of X theta this is a vector, X theta that is equal to log of p of X theta and if I take the logarithm, this will be summation log of theta to the power Xi 1 - theta to the power 1 – Xi summation i going from 1 to N.

And this I can write as whole summation, this is i going from 1 to N, this is Xi into log of theta and plus because it is logarithm, 1 – Xi into log of 1 - theta and this I can write as log of theta into summation i is equal to 1 to N Xi + log of 1 - theta into N - summation Xi i going from 1 to N. Now, we have to take the first partial derivative del L del theta, therefore del L del theta that will be equal to 1 by theta here.

Then summation Xi i is equal to 1 to N plus; now this is log of 1 by theta, its derivative will be 1 - theta and then because it is minus is there, so there will be minus 1 will be there, - 1 by 1 - theta into this term N - summation Xi. Now, we can find out information statistic by finding either E of del L del theta whole square or we can go to the second order partial derivative E of del 2 L del theta square.

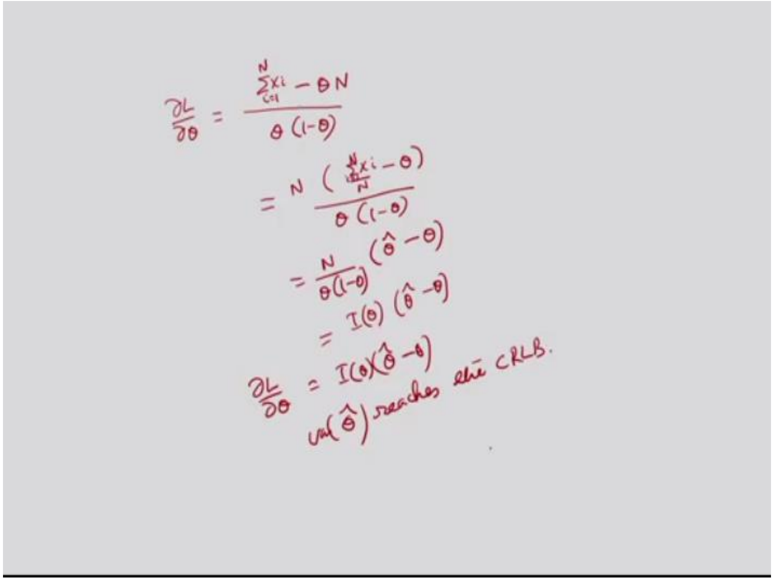
So, we can carry out the second order partial derivative, del 2 L delta theta square will be equal to, this will be - 1 by theta square into summation Xi i going from 1 to N and this one will be minus is there, then so minus, minus will be plus again, minus will come, so that way it will be - 1 by 1 - theta whole square into summation N - summation Xi, this is i going from 1 to N, so we have to take the expected value of this.

Because I information; I theta is equal to minus of expectation of del 2 L del theta square, so and therefore I theta is equal to that is equal to - of E of del 2 L del theta square, so now if I take the expectation here; E of Xi is equal to theta, therefore N theta will be there and 1 theta will get cancelled, so that way it will be; first term will be N by theta square and similarly, here this will be N theta; N - N theta, so that way 1 - theta term will get cancelled.

So, here we will simply have N divided by 1 – theta, this one is N by theta, so that way this is the expression and we have also a minus term, so minus of entire thing, so that will be equal to; if I carry out the simplification, theta into 1 - theta and this one will be, so 1 - theta + theta so that way it will be simply N. So, therefore I theta is equal to N divided by theta into 1 - theta, so this is I theta.

Second part; b is, this is part a, this is a, part b is we have to find out CRLB; CRLB for variance of theta hat, unbiased estimator theta hat will be equal to because 1 by I theta that is equal to theta into 1 - theta divided by N.

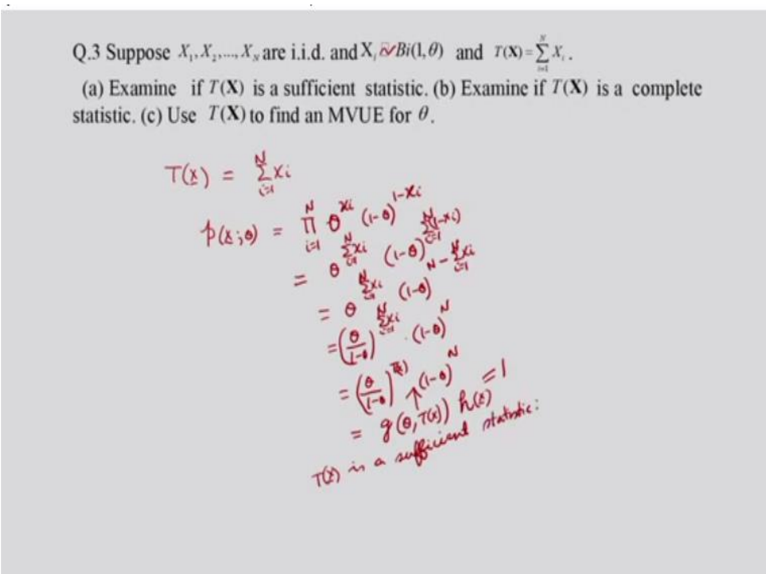
(Refer Slide Time: 19:19)



So, to see if CRLB is reached, so we will find out that del L del theta; del L del theta we have already found out let me write down, del L del theta is equal to; if we simplify this will be equal to summation X i going from 1 to N minus; this will be theta into N divided by theta into 1 - theta that way we will can simplify this and if I take N common, so N into summation Xi divided by N, this i is going from 1 to N - theta divided by theta into 1 – theta.

And this is equal to N divided by theta into 1 - theta and this quantity is our theta hat, minus theta, so therefore we see that equality condition is satisfied and this quantity is the I theta; I theta into theta hat – theta, so del L del theta, therefore del L del theta is equal to I theta into theta hat – theta, therefore variance of theta hat reaches the CRLB, so that way we proved this result.

(Refer Slide Time: 21:04)



Next question is suppose, X1, X2 up to XN are iid and same distribution Xi is distributed as a Bernoulli with parameter 1 and theta that is binomial with parameter 1 and theta that is same as Bernoulli distribution and TX is equal to summation Xi i going from 1 to N, TX is equal to summation Xi i going from 1 to N. Examine if TX is a sufficient statistics and if TX is a complete statistic.

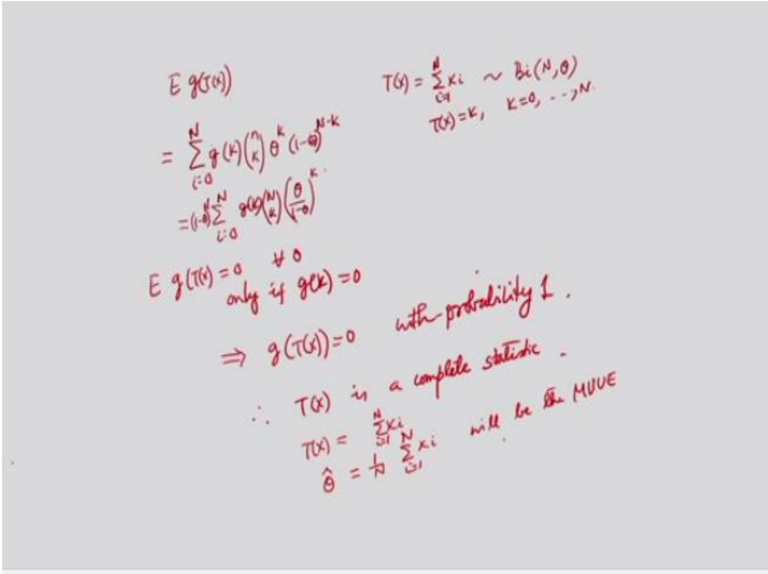
Use TX to find an MVUE for theta, so in this case we have TX is equal to summation Xi i going from 1 to N and the likelihood function we will see the joint PMF here, already we have find out p of X as a function of theta that is the joint PMF, so that is equal to product of i going from 1 to N because iid and individual distribution is theta to the power Xi into 1 - theta to the power 1 – Xi.

So, this I can write as theta to the power summation Xi and i going from 1 to N and similarly, 1 - theta to the power summation 1 – Xi i going from 1 to N, so that way this will be theta to the power summation Xi i going from 1 to N 1 - theta to the power, this first term will become N - summation Xi i going from 1 to N and this I can write, this 1 by theta to the power - summation Xi i going from 1 to N that I can write here.

So, that way theta by 1 by theta to the power summation Xi i going from 1 to N into 1 - theta to the power N, so this is TX, so this is equal to theta by 1 - theta to the power TX and then 1 - theta to the power N and this is same, we can see that this factorization theorem is satisfied. We can take these as the g of theta TX into hx, this hx is equal to 1, so hx is equal to 1 and g of theta TX is equal to this.

So that way we see that, that factorization theorem is satisfied because we know that in the case of factorization theorem, if the joint PMF or joint PDF can be expressed into 2 factors like this, one simply function of X, another is a function of the parameter and the statistic, so that way this PMF satisfy this factorisation theorem, therefore TX is a sufficient statistic.

(Refer Slide Time: 25:34)



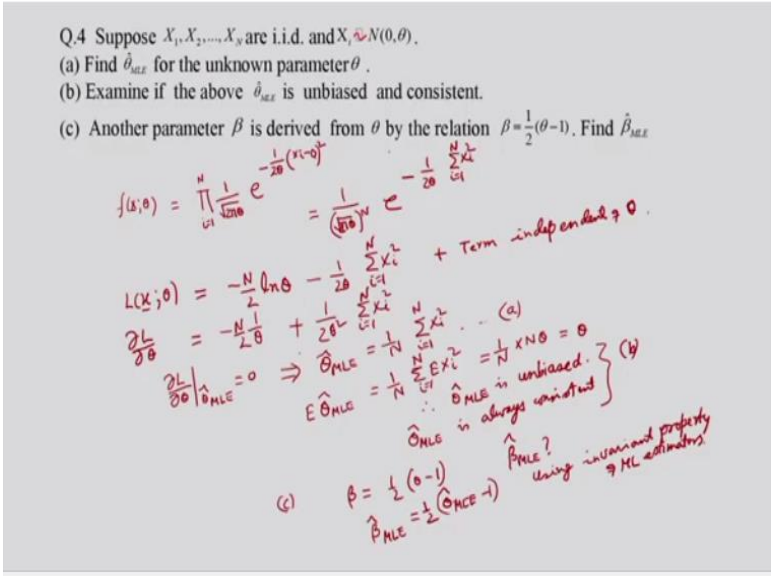
To see whether TX is complete or not, we have to consider this quantity E of g TX, so now TX is equal to summation Xi i going from 1 to N, since Xi's are Bernoulli, TX will be equal to K; K going from 0 to up to N, so that way this will be equal to summation and TX is distributed as binomial with parameter N and parameter theta, so summation i going from 0 to N, this will be g of K into; now this is binomial distribution, so that way NCK, NK that is the combination NK into theta to the power K into 1 - theta to the power N – K.

So, this we can write as summation i is equal to 0 to N, gK NCK theta by 1 - theta to the power K and this 1 - theta to the power N, I can write here 1 - theta to the power N, now we want E of gTX is equal to 0 for all theta and but this quantity is a polynomial in theta by 1 - theta to the power K and this can become 0, only if this gK is 0, why; because this quantity is a polynomial, so a polynomial can be 0, if only its coefficient is 0.

But this quantity is nonzero, so gK must be equal to 0 implies that g of TX is equal to 0 with probability 1, so therefore TX is a complete statistic, T is a complete statistic that we have proved. Use TX to find an MVUE for theta, so we have to use TX to find an MVUE for theta now, TX is equal to summation Xi, therefore any unbiased estimator involving TX will be the MVUE.

So, there is only one unbiased estimator involving this TX, therefore theta hat is equal to 1 by N summation Xi, i going from 1 to N will be d MVUE, so according to that Lehmann-Scheffe theorem we can have an unique unbiased estimator with minimum variance property and there is only one unbiased estimator here that is summation Xi, i going from 1 to N divided by N, therefore this is the MVUE.

(Refer Slide Time: 29:48)



We will go to question 4; suppose, X1, X2 up to XN are iid, identically; independent and identically distributed and Xi is Xi is distributed as normal with mean 0 and variance theta. Find theta hat MLE for the unknown parameter theta, theta is the unknown parameter we have to find out the ML estimator for theta. Examine if the above theta hat MLE is unbiased and consistent.

Another parameter beta is derived from theta by the relation, beta is equal to 1/2 of theta – 1, find beta hat MLE, so here the joint density function is given by f of X, this is the vector as a function of the unknown parameter theta is equal to because it is iid, it is the product of the individual PDF that will product i going from 1 to N and its density is 1 by root over 2 pi theta into e to the power - 1 by 2 theta into Xi - 0.

Because mean is 0 whole square okay, so this is equal to 1 by root over 2 pi theta to the power N into e to the power - 1 by 2 theta into summation Xi square; i going from 1 to N, now if I take the log of this expression that is L of X theta will be equal to; now there is a theta to the power N by 2 is there, so log of theta to the power N by 2 that will be N by 2 log theta.

So, that way, and 1 by is there, so that way N by 2 log of theta and this term will come as - 1 by 2 theta summation Xi square, i going from 1 to N plus some term independent of theta because we will be taking the derivative with respect to theta, independent of theta. Now, we will take the derivatives; del L del theta that will be equal to - N by 2, this is log theta 1 by theta that will be and this will be minus, minus will be plus, so 1 by 2 theta square summation Xi square; i going from 1 to N.

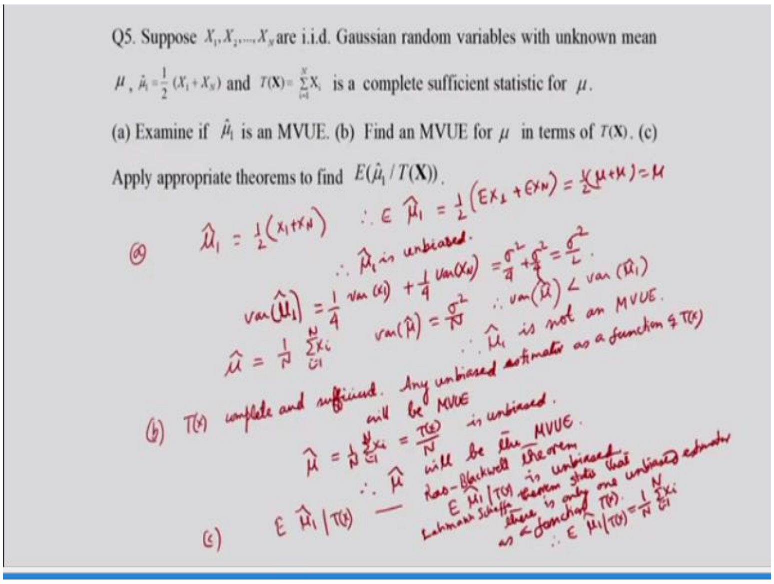
And this term derivative will become 0, now for maximum likelihood estimator del of del theta at theta hat MLE is equal to 0, so del L del theta at theta hat MLE is equal to 0, putting theta hat MLE here, so we will get that theta hat MLE is equal to 1 by N summation Xi square, i going from 1 to N, so this is the maximum likelihood estimator for theta. Next is examine if the above theta hat MLE is unbiased and consistent.

So, E of theta hat MLE, this is the solution to part a, now for part b, E of theta hat MLE is equal to 1 by N into summation i going from 1 to N E of Xi square, now this is a zero mean, therefore E of Xi square is variance and iid, so that way and there will be 1 by N into E of Xi square is theta, there will be N such term, so N theta and that will be equal to theta, therefore theta hat MLE is unbiased.

So, this is the solution b and there is also we know that ML estimator is always consistent theta hat MLE is always consistent, therefore this is true, theta hat MLE here also it will be consistent, so this is the solution of part b. Now, in part c, we have another parameter that is beta is; for part c beta is equal to 1/2 of theta – 1. Now, question is what is beta hat MLE? Now, we will use the invariant property of ML estimators that is beta is a function of theta.

Therefore, we know that if we know the theta hat MLE, then we can find out beta, beta hat MLE because you have to put just theta hat MLE here, therefore beta hat MLE will be is equal to 1/2 of theta hat MLE which is given by this theta hat MLE; theta hat MLE is given by this one, so theta hat MLE – 1, so how do I get it using invariant property of ML estimator.

(Refer Slide Time: 37:18)



So, therefore beta hat MLE will be given by this, we will go to the next question; suppose X1, X2 up to XN are iid, Gaussian random variables with unknown mean mu, mu 1 hat is equal to 1/2 of X1 + XN and TX is equal to summation Xi, i going from 1 to N is a complete sufficient statistic for mu. So, you are given that Xi's are iid Gaussian with unknown mean mu and mu 1 hat is an estimator given by 1/2 of X1 + XN.

And TX is a complete sufficient statistic for mu and it is given by this summation Xi, i going from 1 to N, examine if mu 1 hat is an MVUE, then part b is find an MVUE for mu in terms of TX and c is apply appropriate theorems to find E of mu 1 hat given TX. So, here mu 1 hat is equal to 1/2 of X1 + XN, therefore E of mu 1 hat will be equal to 1/2 of E of X1 + E of Xn, now Xi's are iid, so all will have the same mean mu.

Therefore, this will be 1/2 of mu + mu that is equal to mu, therefore mu 1 hat is unbiased, variance of mu 1 hat will be equal to again these are iid, therefore this will be 1 by 4 into variance of X1 + 1/4 into variance of XN because Xi's are iid independent, therefore we can write like this and this will be equal to variance of X1 is equal to suppose, sigma square, then it will be sigma square by 4 + sigma square by 4 is equal to sigma square by 2.

So, this is the variance of mu 1 hat, now we know that suppose mu hat is equal to 1 by N summation Xi, i is equal to 1 to N, then in this case we know that variance of mu hat is equal to sigma square by N, which is lower than sigma square by 2, therefore this mu hat is an estimator which has variance lower than the variance of mu 1 hat, therefore variance of mu hat which is given by this is lower than variance of mu 1 hat.

Therefore, mu 1 hat is not an MVUE, this is the part a, now we will solve part b, given TX is complete and sufficient, therefore any unbiased estimator as a function of TX will be MVUE and we know that, that is mu hat is equal to summation Xi, i going from 1 to N divided by N that is equal to TX by N, so it is a function of TX, therefore and it is unbiased, therefore it must be MVUE, therefore mu hat will be the MVUE.

So, we will go to part c now, apply an appropriate theorem to find E of mu 1 hat given TX, now according to Rao Blackwell theorem, in this quantity E of mu1 hat given TX is unbiased and Lehmann-Scheffe theorem states that there is only 1 unbiased estimator as a function of TX. Now, this quantity is E of Mu 1 hat given TX is unbiased and it is a function of TX, therefore this must be equal to 1 by N summation Xi, i going from 1 to N.

Therefore, E of mu 1 hat given TX must be equal to 1 by N into summation Xi, i going from 1 to N, so here we first apply the Rao Blackwell theorem, he says that this conditional expectation is unbiased, then we know that since TX is a complete statistic, therefore it has only 1 unbiased estimator as a function of TX and we know that 1 by N summation Xi i going to 1 to N is a function of TX and it is an unbiased estimator.

Therefore, this quantity E of mu 1 hat given TX that conditional expectation must be equal to this estimator itself, summation Xi i going from 1 to N divided by N, thank you.