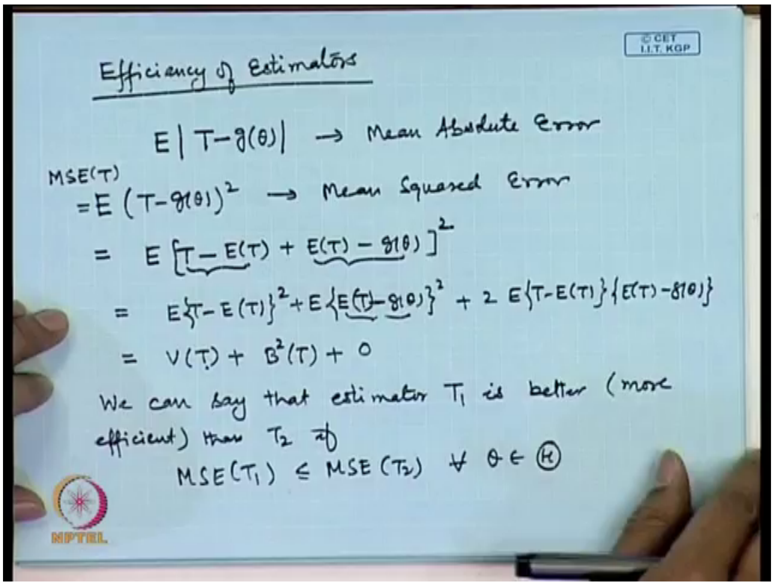
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 06
Basic Concepts of Point Estimations –IV

(Refer Slide Time: 00:43)



Next we introduce the concept of efficiency. As we have seen that there can be situations where we have more than one consistent estimator, we may have more than one estimator which is a unbiased as well as consistent. So, in that case, we introduce the concept of efficiency of estimators. The for judging the efficiencies of the estimators we consider something called as expected error. We have seen unbiasness; so, in unbiasness we had expectation of T is equal to the given parametric function say g theta.

So, if it is not unbiased expectation of T minus g theta is a bias or you can say expected error. But in the there is a danger in using one bias as a simple in a criteria for a goodness of an estimator, because sometimes the negative bias and the negative errors and the positive errors may cancel out each other. So, on the average the estimator may become unbiased, but actually it is not a good estimator.

We have seen the examples for example, in the estimation of e to the power minus 3 lambda we had an estimator minus 2 to the power x in Poisson distribution which was taking values always away from the range. But the errors were positive and negative both very large errors and they were cancelling out each other. So, simply using expectation of x minus g theta that is bias as a measure is a dangerous thing.

So, one may look at other measures for example, why not consider absolute error and then take expectation. So, one may consider expectation of say T minus g theta absolute value. So, this is called the mean absolute error or one may consider expectation of T minus g theta whole square which is called the mean squared error.

So, I will pay some attention to this in the definitions. In the first case, we are simply looking at the amount or you can say magnitude of the error that we have committed in estimating g theta by T and then we take the average of that. In the second one, we are considering the a squares. So, if you think as a layman, then probably we feel that the first one is an appropriate measure for the error or you can say average error. However, in practice the evaluation of expectation of modulus T minus g theta is quite complex.

The second point is that if you look at mathematically, this function is not easy to handle. The main problem is that modulus function is not a smooth function, because it is having a corner that is at T is equal to g theta it is not a smooth. Whereas, if you look at the mean squared error, it is easy to evaluate and it has a simple interpretation which is quite. So, what I do, I add and subtract expectation T here.

So, let us consider this as one term and this has one term. So, this becomes expectation of T minus expectation of T square plus expectation of T minus g theta expectation of that square plus twice expectation T minus expectation T into expectation T minus g theta. So, let us look at these terms. The first term is simply the variance of T. The second term is fixed terms so, expectation will be the same value because we have already taken expectation here. This term is nothing but the bias of the estimator T.

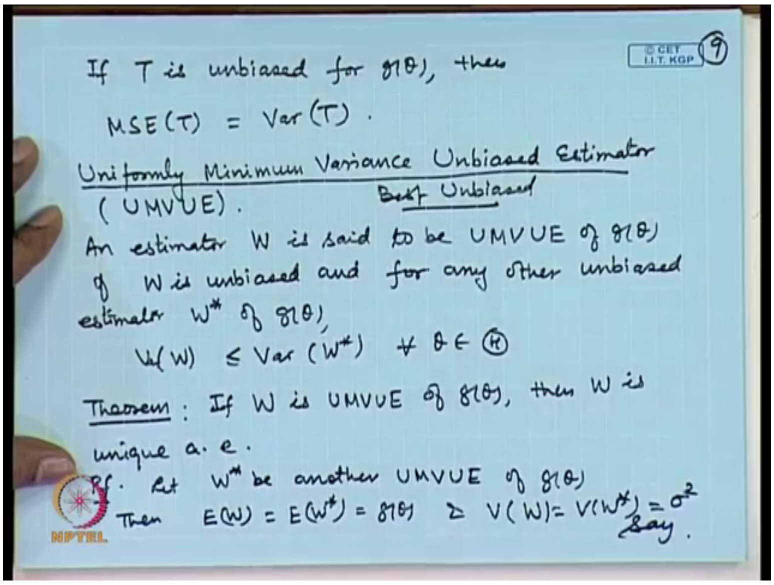
And if you look at the cross product term here, then this term is a constant. So, expectation applies to this, and this becomes 0. So, we have that mean squared error let me call it MSE of T that is equal to variance plus the bias. Now, this is quite significant interpretation. If I have to estimate a set T 1 and T 2 and we only say that variance of T 1 is less than variance of T 2, then we are controlling only one quantity

However, it may turn out that there is another estimator say T 3, which is which may be actually biased, but it is variance is much less, so that the overall mean squared error is a smaller. So, the average a squared error will be less. So, one can use mean squared error is as a good criteria for judging the goodness of an estimator.

So, we will say that we can say that estimator say T 1 is better which is actually a terminology for more efficient than T 2, if mean squared error of T 1 is less than or equal to mean squared error of T 2 for all theta. So, if the two mean squared errors are equal, then they will be same.

Now, in the context of unbiased estimation this concept of mean squared error being smaller is equivalent to variance being smaller. For example, if the estimator T is unbiased then bias will be 0 and this mean squared error will be equal to the variance.

(Refer Slide Time: 07:22)

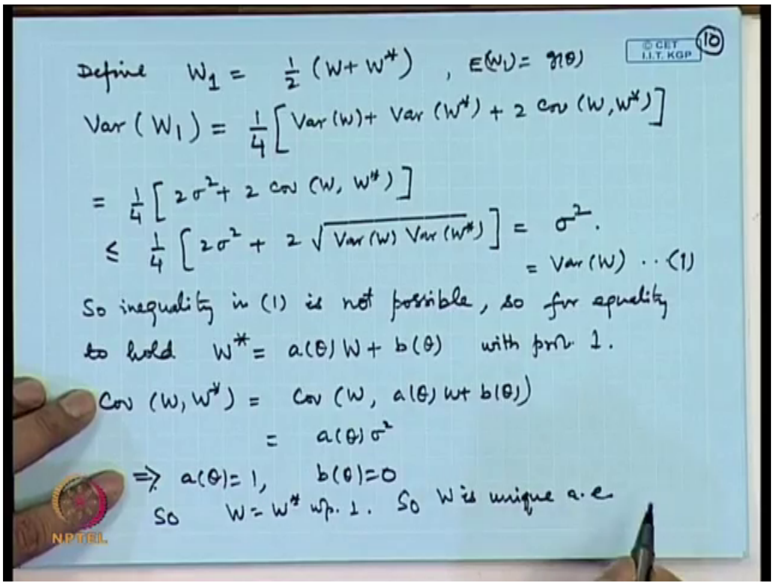


If T is unbiased for g theta, then mean squared error of T is called to be variance of T. Now, we define Uniformly Minimum Variance Unbiased Estimators that is UMVUE. So, an estimator W is said to be UMVUE of say g theta if W is unbiased and for any other unbiased estimator say W is star of g theta variance of W will be less than or equal to variance of W star that means, it will have the minimum variance throughout the parameter a space. The first result in this direction is about the uniqueness of the UMVUE.

If so we also use the terminology best unbiased estimator etcetera. So, if W is UMVUE of say g theta, then W is unique almost everywhere. So, let W is star be another

UMVUE. Then by definition expectation of W expectation of W star both are same as g theta and variance of W and variance of W star are also same let us call it say sigma square fine.

(Refer Slide Time: 10:14)



Now, let me define say W 1 as half W plus W star. Then what is the variance of W 1, we can apply the formula for a linear combination of variables. So, variance of a constant times that is that constant square times variance of W plus W star which is becoming variance of W plus variance of W star plus twice covariance between W and W star.

Now, we are assuming variance of W and variance of W is star to be sigma square, so it becomes 1 by four 2 sigma square plus twice covariance W, W star. Now, covariance square is less than or equal to the product of the variances the well known Cauchy Schwarz inequality. So, this becomes 1 by 4 2 sigma square plus twice a square root of variance W into variance of W star, but these are both sigma square. So, this is simply becoming sigma square, so 2 sigma square plus 2 sigma square, so it becomes sigma square which is the variance of W or W star.

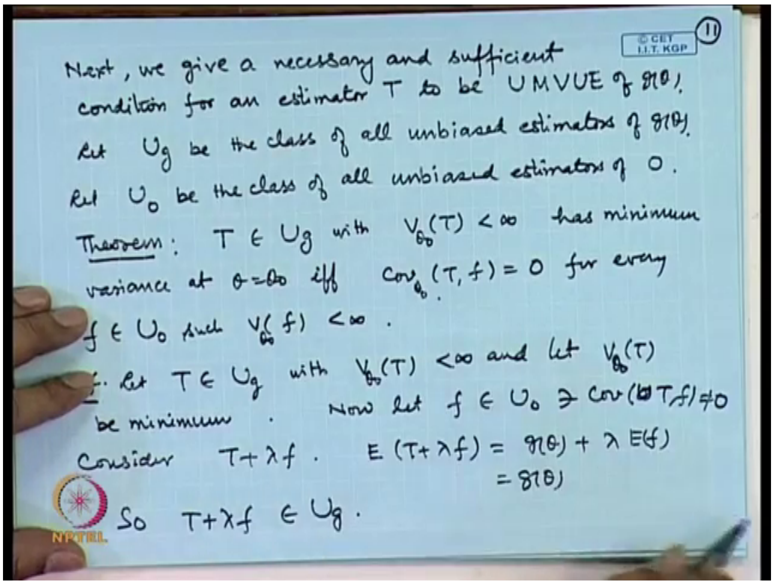
So, what we are proving? If W is UMVUE W star is another UMVUE, then I am able to get another estimator W 1 which is also unbiased because, if I take expectation of W 1 here that is again g theta as both W and W star are unbiased and it is variance is less than or equal to the variance of W. So, let me call this equation number 1. Inequality in 1 is not possible because our original claim is that W and W star are UMVUE. So, another unbiased estimator cannot have variance less than them. So, at the most it can have equal, so that means, we should have equality.

Now, how this inequality came inequality came from this condition of the correlation between the W and W star being less than 1, so that means, correlation must be one that is covariance is equal to the square root of the variances, that means, W and W star are linearly related with probability 1. So, for equality to hold W star must be linearly related to W with probability 1.

Now, once again you have unbiasness, so if you are saying unbiasness, then what should be the condition here. And also if I look at say covariance here between W and W star, then that is equal to covariance between W and a theta W plus b theta, so that is equal to a theta into sigma square. So, that means, because this covariance between W W star is equal to variance W so, a theta is 1 and b theta will be 0, because unbiasness is there because expectation W star must be a theta, so that is simply becoming g theta plus b theta so, b theta must be 0.

So, what we are concluding here that W is equal to W star with probability 1, that means, W is unique almost everywhere. So, you cannot have two different unbiased MVUE's in if they are two different then they are equal almost everywhere.

(Refer Slide Time: 15:06)

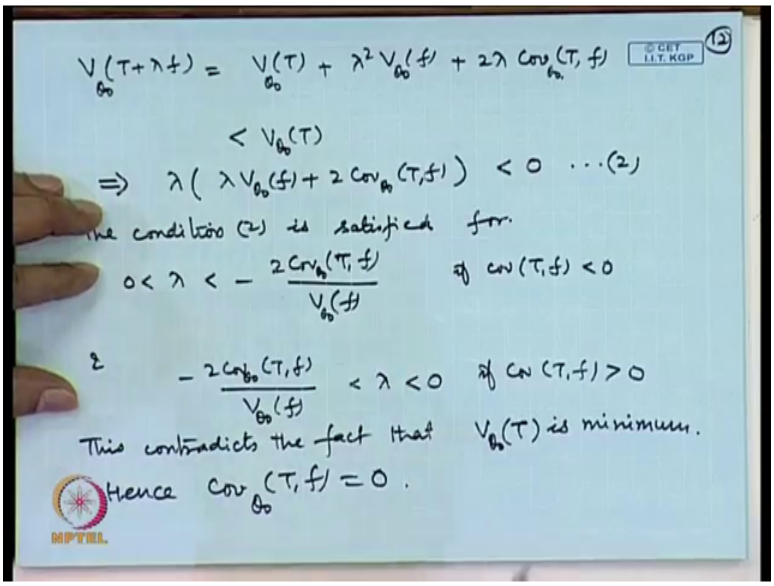


Now, next I give a necessary and sufficient condition for an estimator to be UMVUE. So, let us consider let U g be the class of all unbiased estimators of g theta. Let U 0 be the class of all unbiased estimators of 0. So, we have the following necessary and sufficient condition. So, T belongs to U g with variance of T to be finite. So, this has minimum variance at theta is equal to theta naught. If and only if covariance of T with say f is 0 for every f belonging to U 0 for which variance of f is finite. That means, if an estimator is having covariance 0; that means, it is uncorrelated with every unbiased estimator of 0, then this will be UMVUE of a function g.

Let me prove this here. So, let T the unbiased estimator of g and it is variance be finite and let variance theta naught T be minimum. Now, let us consider f belonging to U 0, such that covariance between T and f is not 0. So, I am assuming contrary to what we have to prove. So, we will arrive at a contradiction.

So, let us consider say T plus lambda f now if I take expectation of T plus lambda f then it is equal to expectation T that is g theta plus lambda times expectation f that is 0. So, it is equal to g theta so, this new function which I have created T plus lambda f is also unbiased.

(Refer Slide Time: 18:42)

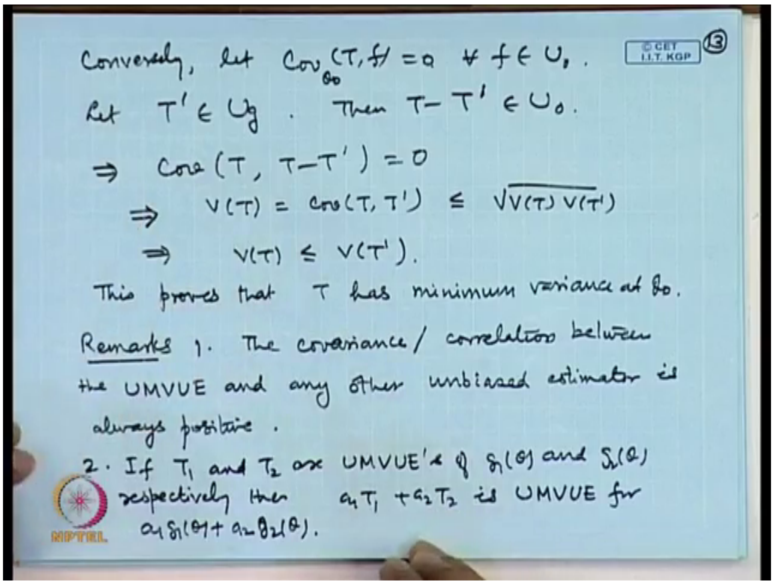


Now, let us take variance of T plus lambda f, so that is equal to variance of T plus lambda square times variance of f plus twice lambda covariance between T and f. Now, if I put a condition here that this is less than variance of theta naught T, then this thing cancels out and it is reducing to a quadratic being less than 0, that means, this condition is equivalent to lambda into lambda V theta naught f plus twice covariance theta naught T f less than 0. So, this condition, obviously, can be satisfied.

The condition two is satisfied for 0 less than lambda less than minus twice covariance T f by variance f, of course, all these evaluations are at the point theta naught if covariance of T and f is negative. And for minus twice covariance theta naught T f by variance theta naught f less than lambda less than 0 if this is positive. That means whatever be the value of covariance between T and f whether it is positive or negative, I am able to obtain a range of lambda values such that the variance of T plus lambda f is less than variance of T. This is a contradiction to the fact that I assumed that variance of T is minimum at theta naught.

So, where is the mistake? The mistake is that I am assuming that covariance between T and f is not 0. So, this is wrong. So, this contradicts the fact that variance theta naught T is minimum, hence we must have covariance between T and f equal to 0.

(Refer Slide Time: 21:42)

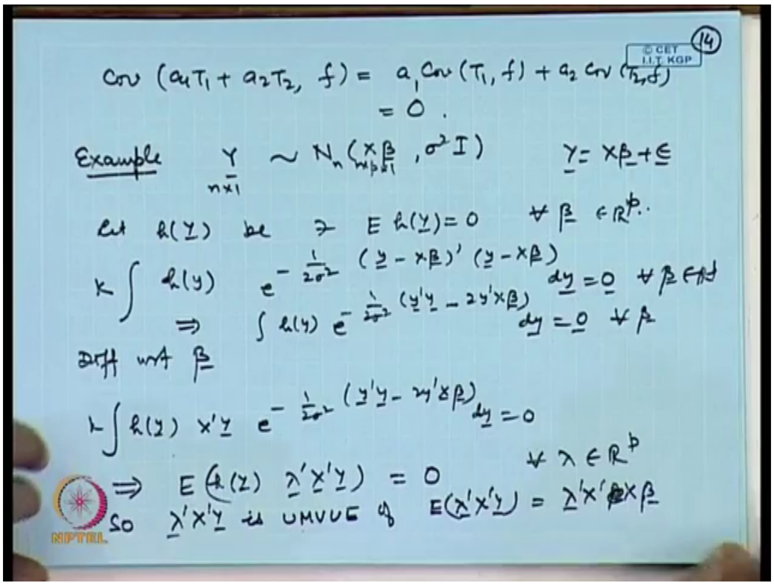


Now, let us take the converse of this. Conversely let between T and in two variance of T prime. So, obviously, this is equivalent to saying that variance of T is less than or equal to variance of T prime. So, if I am taking covariance of T to be 0 with every unbiased estimator of 0 and I am taking another unbiased estimator T prime of g theta then I am getting that the unbiased the variance of T is less than or equal to variance of T prime. This proves that T has minimum variance at theta.

Another thing which you can conclude from here I have proved that if T is UMVUE, then covariance between T and T prime that is equal to variance of T; that means, this is always positive. So, we are also concluding from here that the covariance or you can say correlation between the UMVUE and any other unbiased estimator is always positive.

And other interesting property about the UMVUE is that if T 1 and T 2 are UMVUE's of g 1 theta and g 2 theta respectively, then a 1 T 1 plus a 2 T 2 is UMVUE for a 1 g 1 theta plus a 2 g 2 theta that means some sort of linearity property is also true for the UMVUE. Although it is true for the unbiased estimation, but it is not clear that it will be true for UMVUE's, but that is true here.

(Refer Slide Time: 24:42)



In fact, one can look at a very simple proof of this. If I consider say covariance of a 1 T 1 plus a 2 T 2 with an unbiased estimator of 0 then it is equal to a times covariance between T 1 and f plus a 2 times covariance between T 2 and f. Now, if T 1 and T 2 are UMVUE's, these are 0 so, this is simply 0. So, by the previous theorem this result follows.

As an application of this theorem let us consider linear model and try to obtain UMVUE. So, let us consider the Gauss-Markov linear model. So, y is an N by 1 vector with mean x beta and variance covariance matrix as sigma square y. So, actually it is the part of the Gauss-Markov linear model where we write it as plus epsilon and epsilon follows normal 0 sigma square I. So, let us consider say h y be a real valued function such that expectation of h y is say 0 for all beta. This may be say N by p this may be p by 1 etcetera.

So, if you write this a statement expectation h y 0 it is equivalent to h y into the density function of y this is a multivariate normal distribution. So, it is e to the power minus 1 by 2 sigma square y minus x beta prime y minus x beta. And some coefficient will come which I am writing as a constant. This is equal to 0 for all beta belonging to the R p. This is a multivariate integral here.

Now, you differentiate both the sides with respect to beta, then I will get h y, then derivative of this will give this term into the derivative of this with respect to y that gives me x prime y e to the power minus 1 by 2 sigma square. In fact, here I can simplify beforehand I can write the term which is not involving y I can separate out and take to the other side. So, this is reducing to, so if you differentiate this, you will get x prime y here, and the same term here.

So, this is equivalent to saying expectation of h y into some coefficient lambda prime x prime y is equal to 0, for all lambda belonging to R p. So, what is this one this is a linear function. So, by the previous theorem, what we are saying is that lambda prime x prime y is UMVUE of expectation of lambda prime x prime y that is lambda prime x prime beta.

In the Gauss-Markov theory of linear models, we had proved that lambda prime x prime y is the best linear unbiased estimator of lambda prime x prime beta. Here we are proving that it is not only best linear unbiased it is actually best unbiased that is it is the UMVUE for this. Although I have made a small mistake here, it is lambda prime x prime x beta. So, for this it is becoming best unbiased estimator.

In the forthcoming classes we will consider methods for finding out estimators. Just now in the previous two classes we have considered the properties of the estimator some desirable criteria. However, there must be some methods by which we can derive these estimators. So, we will do some well-known methods.