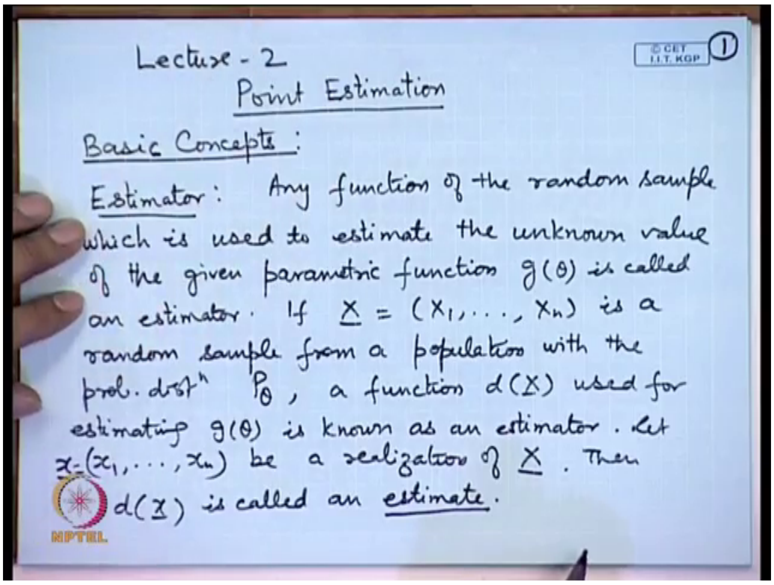
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 03
Basic Concepts of Point Estimation –I

In the last lecture we had discussed what is the problem of statistical inference, what is the motivation for a studying statistical inference problems. We had seen that the problem of statistical inference can be broadly categorized into two parts: one is the problem of estimation and the problem of testing of hypothesis.

So, we will start with the problem of estimation. Now, in the problem of estimation we had seen two parts are there: one is the problem of point estimation and that is and there is a problem of interval estimation.

(Refer Slide Time: 00:55)

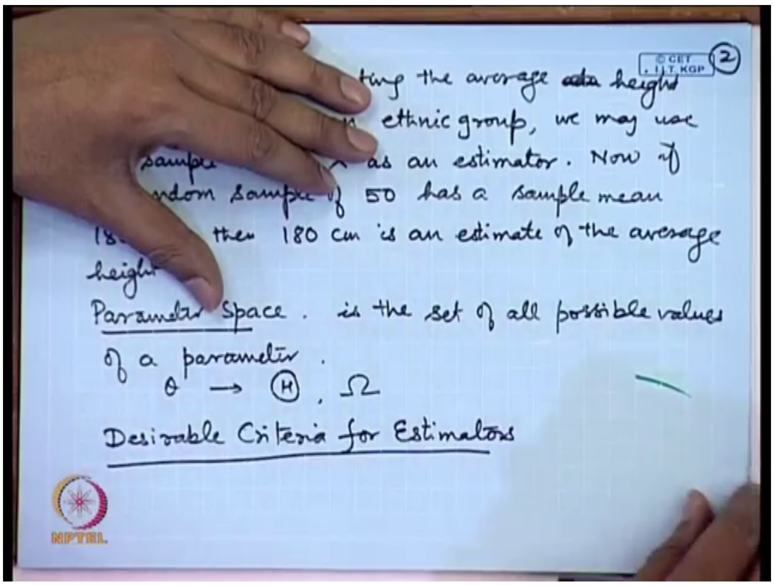


So, we start with the problem of point estimation and let us look at what are the basic concepts that are needed. In a general problem of statistical inference we have seen that the concept of population, the concept of a sample, the idea of a parameter and that of a statistic. Now, when we talk about point estimate in estimation then the first thing is that we have to identify an estimator to estimate the unknown parameter of the population. So, for that we define what is a point estimator. So, any function of the random sample which is used to estimate the unknown value of the given parametric function is called.

So, suppose we say parameter theta and we may consider a parametric function g theta then this is called an estimator. So, in practice we will have a random sample. So, if X say X 1 X 2 X n is a random sample from a population. Now, a probability distribution is identified with the given population. So, a population with the probability distribution say; P theta then a function say d X which is used for estimating g theta, then this is known as an estimator. Now, in reality what will happen that this X 1 X 2 X n will take some values because, when you go to the field and collect the data this X 1 X 2 X n will correspond to some numerical observations.

So, let x 1 x 2 x n be a realization of say X. So, let me call it a small x then the corresponding value of the estimator which is evaluated at this realization this is called an estimate. So, we have two important concepts here: one is estimator and a realized value of the estimator that is called an estimate. So, let me explain it through some example.

(Refer Slide Time: 04:41)



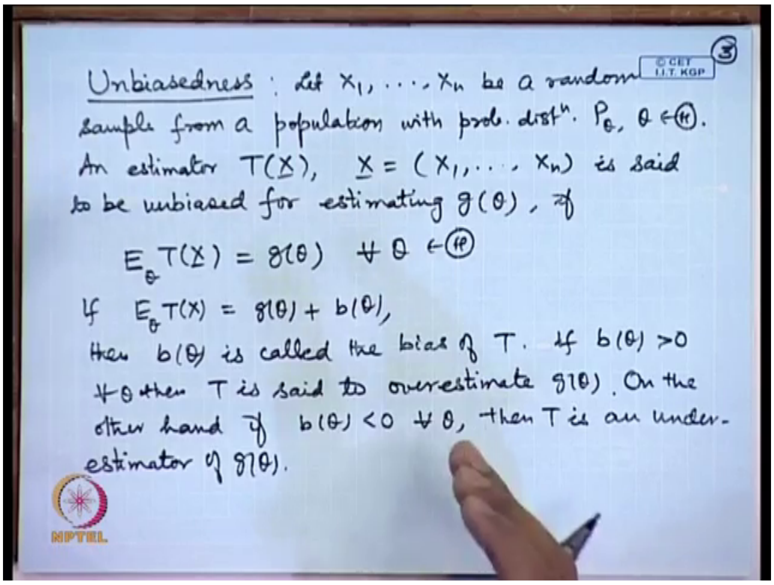
Let us consider we want to estimate say average height of adult males in a ethnic group, in estimating the average height of say adult males in an ethnic group. So, we may use the sample mean say X bar as an estimator. Now, if a random sample of say 50 has a sample mean say 180 centimeter then 180 centimeter is an estimate of the average height. So, in a given statistical problem of point estimation we will be proposing some estimators which are obtained through some concepts, through some rational reasoning or through some rational decision making procedure. And, the realized values of that function which we call now estimator will be used as estimates of the given parameters.

So, in short this is what we do it in a point estimation of certain parametric functions. So, we are talking about parameter repeatedly. Now, this parameter of a population for example, when we say average height of adult males and now this parameter theta is the parameter of the corresponding population of the heights. Now, that population is described by certain distribution it could be a gamma distribution, it could be a normal distribution. So, this parameter lies in a certain range that range is called a parameter a space. So, in a given problem we have to be careful that our estimator should take values in the given parameter a space.

So, a parameter space is the set of all possible values of a parameter. So, if I use the parameter theta then a space we can denote by say capital theta or omega etcetera. So, these are the usual notations. Now, the question is that one can obtain estimators, as I just now mentioned to estimate average height we may use sample mean as an estimator. If we are estimating say average a speed of a vehicle we may use a harmonic mean, we may use median. We if we are estimating variability of a population then we may consider range of the sample, we may consider mean deviation about the mean, we may use the standard deviation about the mean etcetera.

So, we may be able to propose various estimators for a given parametric function of interest. The question is that which one should be used. Therefore, the first point that comes to the mind is that we should be identifying certain criteria which will tell; that means, you can say certain desirable criteria that should be there present in the given estimators. That means, if the estimator satisfies one or more of those criteria it is supposed to be a good estimator. So, we may in the beginning I will mention certain desirable criteria for estimators.

(Refer Slide Time: 09:51)



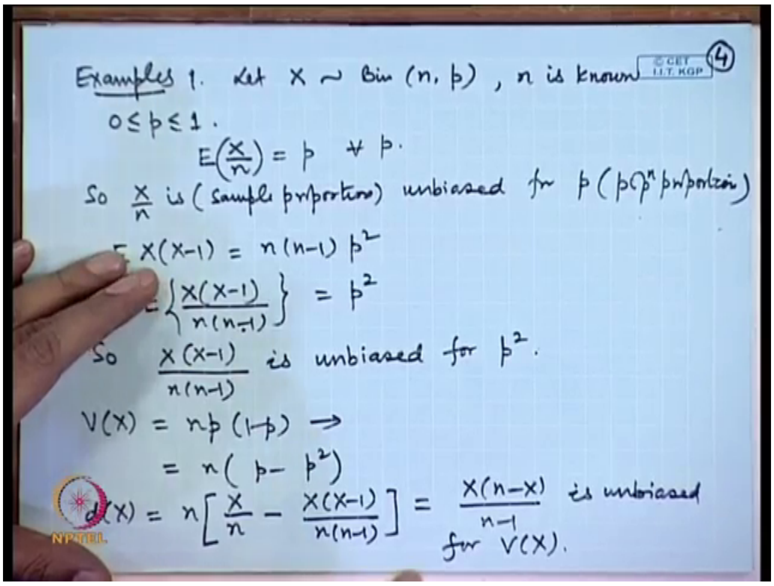
The first such criteria is that of unbiasedness. So, as the name unbiasness suggests; that means, we are looking at the estimators which do not show any bias towards anything. That means, a rational thinking person should be able to use it by saying that he is not biased by any other criteria other than the data itself. So, but in a statistical terminology the unbiasness means that on the average the estimated value is equal to the value of the parameter.

So, let us go back to our model that let X 1 X 2 X n be a random sample from a population with probability distribution P theta, where theta belongs to the parameter space a script theta. An estimator T X where X denotes X 1 X 2 X n is said to be unbiased for estimating the parametric function g theta, if expectation of T X is equal to g theta for all theta. That means, on the average the estimator equals the parameter; that means, if sufficiently large number of samples are considered then the average value of the estimator calculated from those many samples will be actually equal to the true parameter value.

Now, if expectation is not equal to g theta, but we can write it as say g theta plus b theta then b theta is called the bias of T. If b theta is always positive then T is said to be is said to overestimate g theta. On the other hand if b theta is always less than 0 then T is an under-estimator. So, in different estimation problems it may be desirable to have unbiased estimator or sometimes the situation may demand that we may overestimate or sometimes we may underestimate. And, also the consequences of over estimation or under estimation may be disasters in different ways.

For example, if you are considering building of a bridge then, if we say overestimate the strength of the concrete that is being used to build the bridge then it may be very disastrous. Because, it may break down when the vehicles are plying on the bridge and it may lead to serious accidents. Similarly, in certain other cases underestimation may be more serious. So, one has to be careful that how to control the bias of a given estimator. So, let us take some examples.

(Refer Slide Time: 14:13)



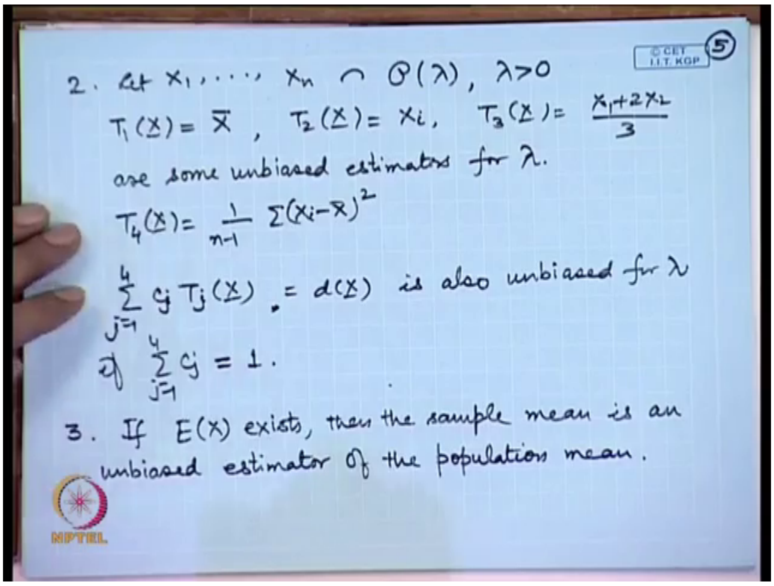
Now, let us consider say a binomial random variable n p so; that means, there is an experiment where the outcomes are Bernoullian trials and the number of successes X has been recorded. So, the outcomes of the random sample is recorded in the form of the total number of successes. Here we may consider n is known; that means the parameter of interest is the probability of success or the proportion of successes. The parameter space is the interval 0 to 1 and we may be interested in estimating the proportion p.

So, then you may consider the properties of the binomial distribution expectation of X is equal to n p. So, expectation of X by n is equal to p. So, here we conclude that X by n that is the sample proportion sample proportion is unbiased for p which is the population proportion. Now, there may be a problem where we may be interested in estimating the squared proportion. So, we may be interested for estimating p square then we further notice the expectations here expectation of X into X minus 1 is equal to n into n minus 1 p square.

So, this implies expectation of X into X minus 1 by n into n minus 1 is equal to p square. So, unbiased estimate of p square is X into X minus 1 by n into n minus 1. And, yet another application we may be interested to estimate the variability of this binomial distribution; that means, variance of X that is n p into 1 minus p. Suppose we are interested to estimate the variance of the binomial distribution then we can make use of the estimators of p and p square and substitute here, because this is equal to n times p minus p square. Now, for p and p square we have already obtained the unbiased estimators.

So, if we write d X is equal to n times for p we write X by n and for p square we write X into X minus 1 divided by n into n minus 1. If we simplify this, it turns out to be X into n minus X divided by n minus 1. So, then this is unbiased for variance of X. So, this is actually one of the common approaches to obtain the unbiased estimators; that means, we consider the moments of the given distribution. For example, in the binomial case we have considered the first two moments which are helpful in obtaining the unbiased estimators of the population proportion a square or the variance of this. Let us take up some other examples.

(Refer Slide Time: 17:55)



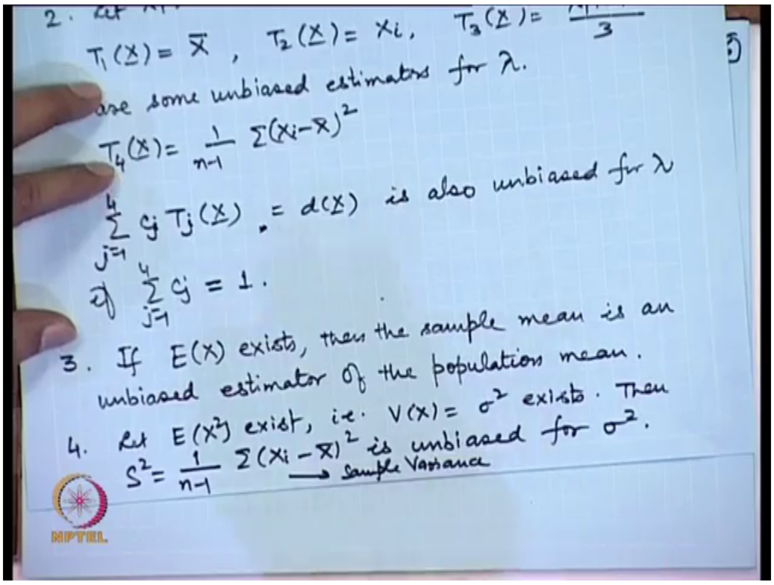
Suppose we are having a random sample X 1 X 2 X n from a Poisson distribution with parameter lambda; now naturally we may be interested to estimate lambda itself. So, we may consider say T 1 X as X bar; we know that the first moment of the Poisson distribution is X is lambda. So, expectation of X 1 is lambda and therefore, expectation of X bar is also lambda so, this is unbiased. However, we may even consider any of the X i's also we may consider say X 1 plus 2 X 2 by 3. Now, this is also unbiased for lambda because expectation of X 1 is lambda expectation of X 2 is lambda.

So, it becomes lambda plus 2 lambda that is 3 lambda by 3 that is equal to lambda. So, these are some unbiased estimators for lambda. Now, this brings us to a point that for the same parameter we may obtain several unbiased estimators. And therefore, we may look for further criteria to restrict the class of unbiased estimators also. So, we will consider that in a short while. Here we may also consider that lambda is also the variance of this distribution. If we see this as the variance then another unbiased estimator can be written as say 1 by n minus 1 sigma X i minus X bar whole square.

Now, if there are several unbiased estimators we may consider say sigma C i T i, i is equal to 1 to 4 and let me put j here because I have already used here. So, let me call it say d X then this is also unbiased because each of this is unbiased, then if sigma C j is equal to 1. If sigma C j, j is equal to 1 to 4 is equal to; that means, if we are having more than one unbiased estimator then we can construct a large number or you can say an infinite number of unbiased estimators also. Therefore, we need to put some further we need to qualify with certain other criteria so, that we can restrict attention to few of them only ok.

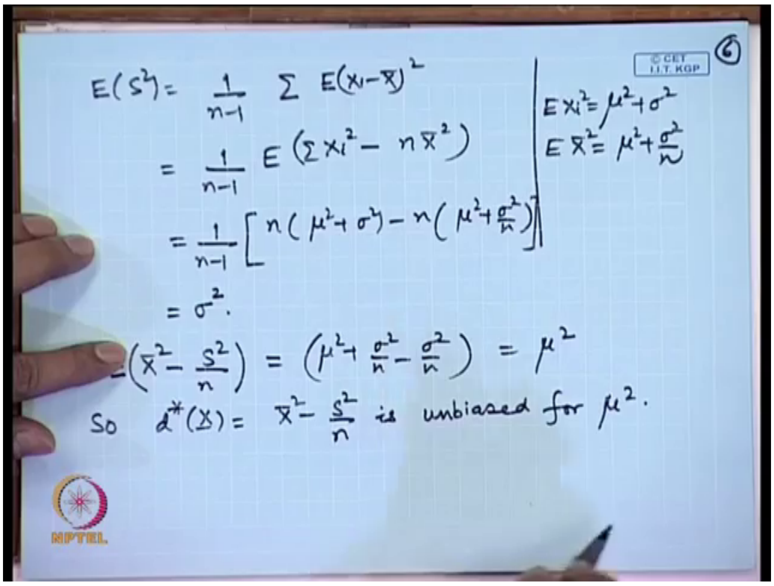
Now, if we notice this first 2 examples it is very clear that the sample mean will be unbiased estimator for the population mean. As in the previous case we have seen here lambda is the mean of this Poisson distribution and the sample mean X bar is unbiased for this. So, is it true in general? The answer is yes, if the first moment exists then always the sample mean will be unbiased estimator for the population mean. So, you can write that, if expectation of X exists then in any given estimation problem the sample mean is an unbiased estimator of the population mean. Similarly, we may consider the population variance.

(Refer Slide Time: 22:01)



So, let us specify let expectation X square exist ok; that means, variance of X let us say sigma square exists. Then the sample variance which we will denote by 1 by n minus 1 sigma X i minus X bar whole square is unbiased.

(Refer Slide Time: 22:37)



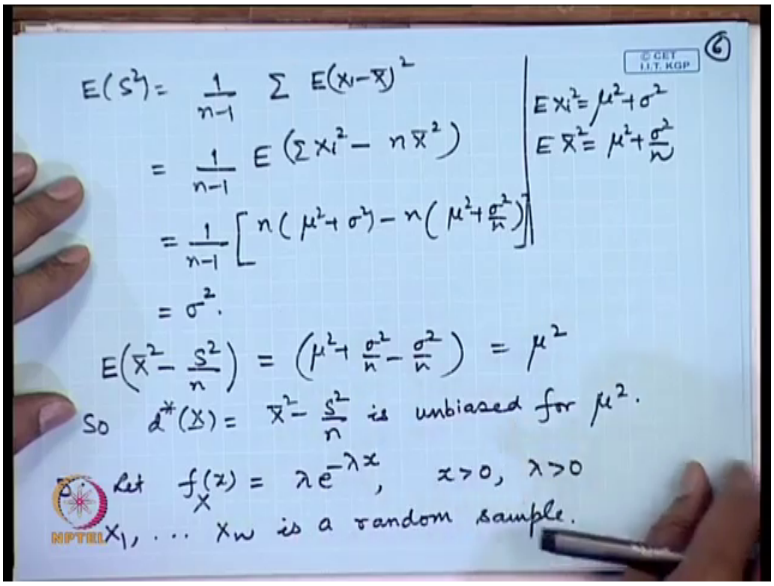
So, let me look at a proof of this, expectation of S square that is equal to 1 by n minus 1 sigma expectation of X i minus X bar whole square. This we may write as 1 by n minus 1 sigma expectation of sigma X i square minus n X bar square. Now, here we may use the property that expectation of X i square will be equal to mu square plus sigma square.

And, a similarly expectation of X bar square will be mu square plus sigma square by n because variance of X by is sigma square by n. So, these things if we substitute here it becomes 1 by n minus 1 n times mu square plus sigma square minus n times mu square plus sigma square by n.

So, after simplification you can see here this n mu square cancels out and n minus 1 sigma square by n minus 1 that is equal to sigma square. So, this quantity that we have defined here that is 1 by n minus 1 sigma X i minus X bar whole square this is termed as sample variance, because this is an unbiased estimator for the population variance. We may also notice here suppose, I want to estimate mu square. So, I have already obtained unbiased estimators for sigma square, unbiased estimator for mu is available to us. So, we consider expectation of X bar square by minus S square by n this is equal to mu square plus sigma square by n minus sigma square by n so, this is equal to mu square.

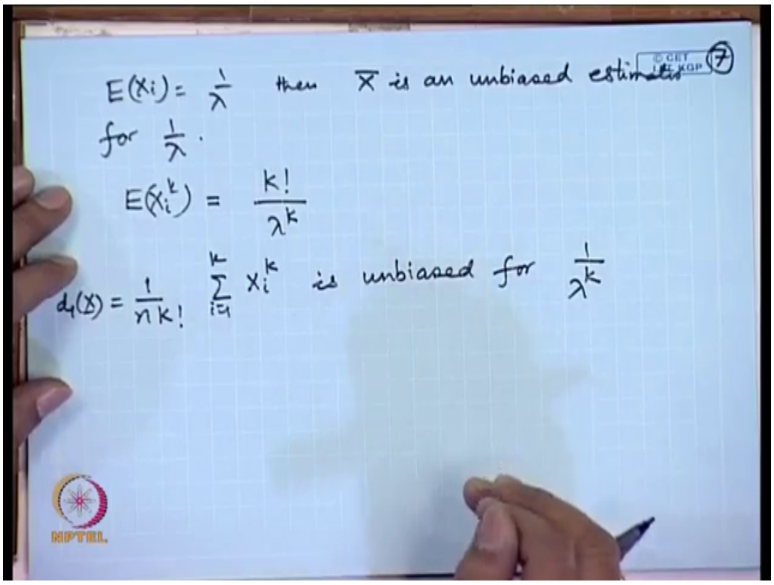
So, that shows that how we can estimate some related parameters in a given estimation problem using the concept of unbiasness. Let me give you a name here say d star X that is equal to X bar square minus S square by n is unbiased for mu square.

(Refer Slide Time: 25:13)



Let us consider say X having a density of exponential distribution lambda e to the power minus lambda x. Let X 1 X 2 X n be a random sample from this population. Now, here we know that the mean of the exponential distribution is reciprocal of the rate that is 1 by lambda.

(Refer Slide Time: 25:55)



So, if we use this here expectation of X i is equal to 1 by lambda then this gives us that X bar is an unbiased estimator for 1 by lambda. Not only that we may consider estimation of the higher order moments also. For example, we may look at say expectation of X i to the power k, in the exponential distribution this is equal to k factorial divided by lambda to the power k.

So, we may write 1 by n k factorial sigma X i to the power k, i is equal to 1 to k. Let us call it say d 1 X then expectation of this if you consider it will become 1 by n k factorial and this will become k factorial by lambda to the power k and n will come. So, then this becomes unbiased for 1 by lambda to the power k. So, that shows that moment of any order can be evaluated in the case of exponential distribution. So, we can obtain unbiased estimators for each of them.

So, these and other topics we will be covering in the forthcoming lectures.