Statistical Inference

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

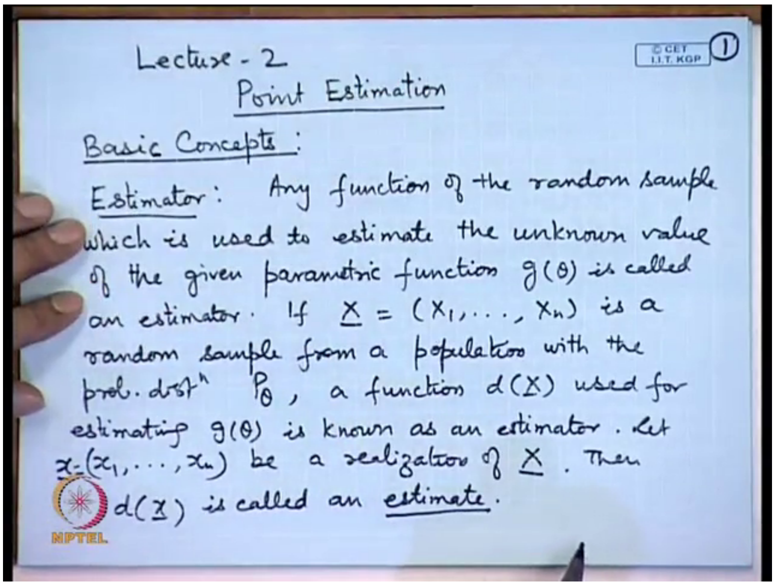
Lecture No. # 02

Basic Concepts of Point Estimations – I

In the last lecture, we had discussed the; what is the problem of statistical inference? What is the motivation for 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. 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.

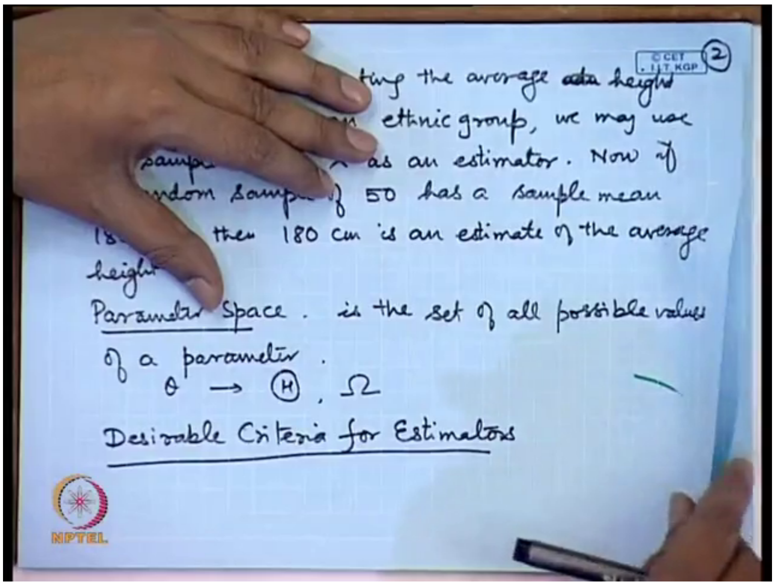
(Refer Slide Time: 01:25)



So, for that we define, what is a point estimator? So, any function of the random sample, (No Audio From: 01:40 to 01:46) 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 (No Audio From: 02:40 to 02:48) 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 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.

(Refer Slide Time: 04:37)

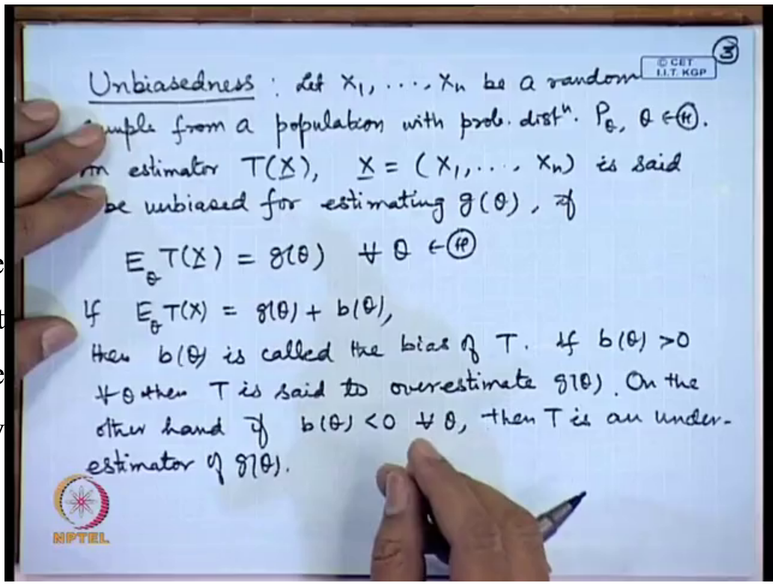


So, let me explain it through some example, let us consider, we want to estimate say average height of adult males in a ethnic group. So, we may use, the sample mean say X bar as an estimator. Now, if a random sample of say 50 has a sample mean say 180 centimeter, then 180 centimeter is an estimate of 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 value is 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 space. So, in a given problem we have to be careful that our estimator should take values in the given parameter space. So, a parameter space is the set of all possible values of a parameter. So, if I use a parameter theta then the space we can denote by say capital theta r 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, you may use sample mean as an estimator. If we are estimating say average speed of a vehicle, we may use say harmonic mean, we may use median. 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 suppose to be a good estimator. So, you may, in the beginning I will mention certain desirable criteria (No Audio From: 09:28 to 09:35) for estimators.

(Refer Slide Time: 09:49)

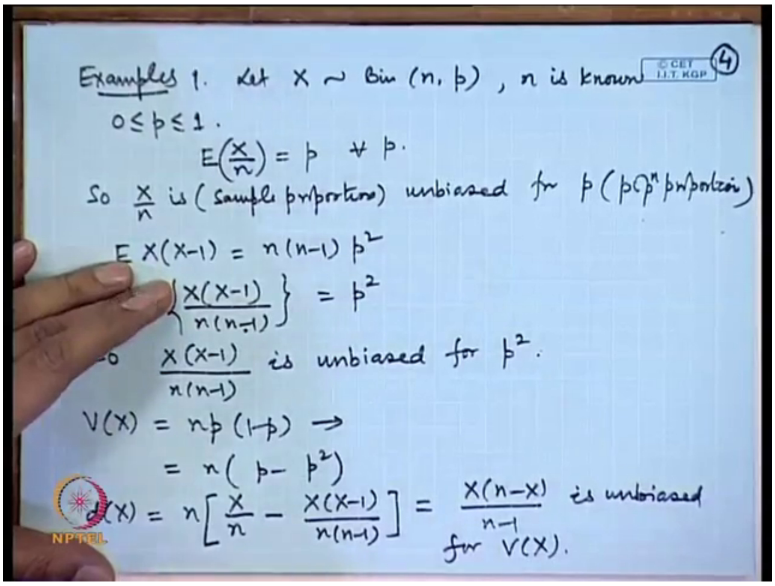


The first such criteria is that of unbiasedness. So, as the name unbiasedness suggests; that means, we are looking at the estimators which do not show any bias towards anything; that means, a rational thinking person should be able to use it by saying that he is not biased by any other criteria, other than the data itself. So, but in statistical terminology the unbiasedness 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 (No Audio From: 10:49 to 10:56) with probability distribution p theta, where theta belongs to the parameter space 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 overestimate g theta. On the other hand, if b theta is always less than 0, then T is an under estimator.

(No Audio From: 12:55 to 13:03) 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 overestimation or underestimation may be disastrous 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.

(Refer Slide Time: 14:11)

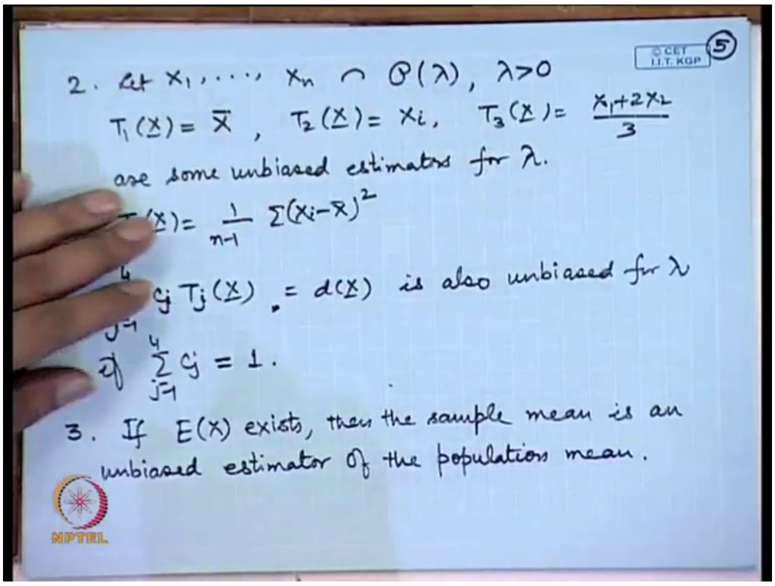


So, let us take some examples now, let us consider say a binomial random variable n p so; that means, there is an experiment where the outcomes are Bernoullian trials and the number of successes X has been recorded. So, the outcomes of the random sample is recorded in the form of the total number of 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 is unbiased for p which is the population proportion. Now, there may be a problem we are 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. (No Audio From: 16:08 to 16:18) 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 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.

(Refer Slide Time: 17:52)



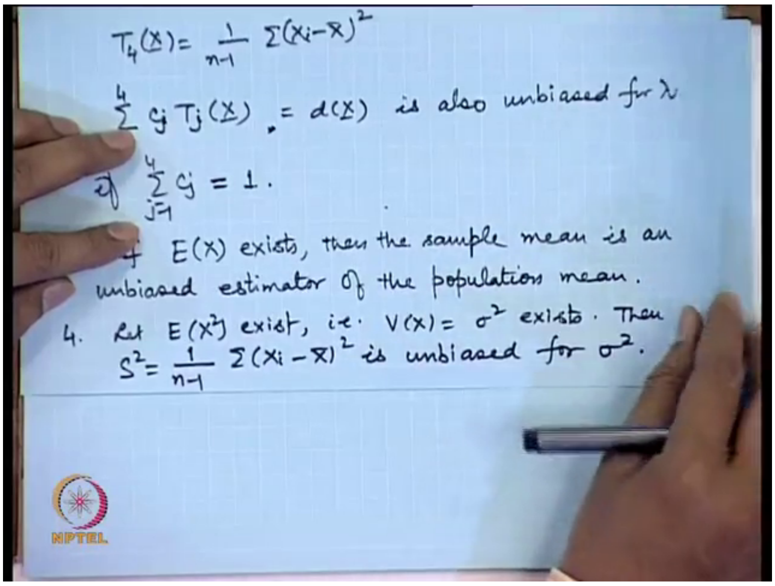
Let us take up some other examples, suppose we are having a random sample X 1 X 2 X n from a Poisson Distribution with parameter lamda. Now, naturally we may be interested to estimate lamda itself. So, we may consider say T 1 x as X bar, we know that the first moment of the poisson distribution is; X is lambda. So, expectation of X 1 is lamda and therefore, expectation of X bar is also lambda. So, this is unbiased, however, we may even consider any of the X I is also we may consider say X 1 plus 2 X 2 by 3. Now, this is also unbiased for lamda, because expectation of X 1 is lamda expectation of X 2 is lambda. So, it becomes lamda plus 2 lamda that is 3 lamda by 3 that is equal to lambda.

So, these are some unbiased estimators for lamda. 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 lamda is also the variance of this distribution. If we say 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 one, 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 a 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. Now, if you notice this first two 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 lamda is the mean of this poisson distribution and the sample mean X bar is unbiased for this.

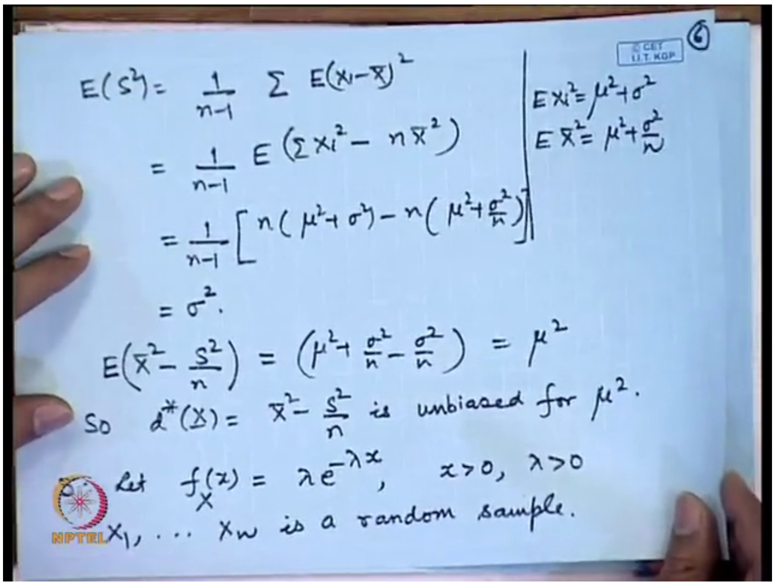
So, is it true in general the answer is yes, if the first moment exists then always the sample mean will be unbiased estimator for the population mean. So, you can write that, if expectation of X exists then in any given estimation problem, the sample mean is an unbiased estimator of the population mean.

(Refer Slide Time: 17:52)



Similarly we may consider the population variance. So, let specify, let expectation X square exist; 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: 22:32)



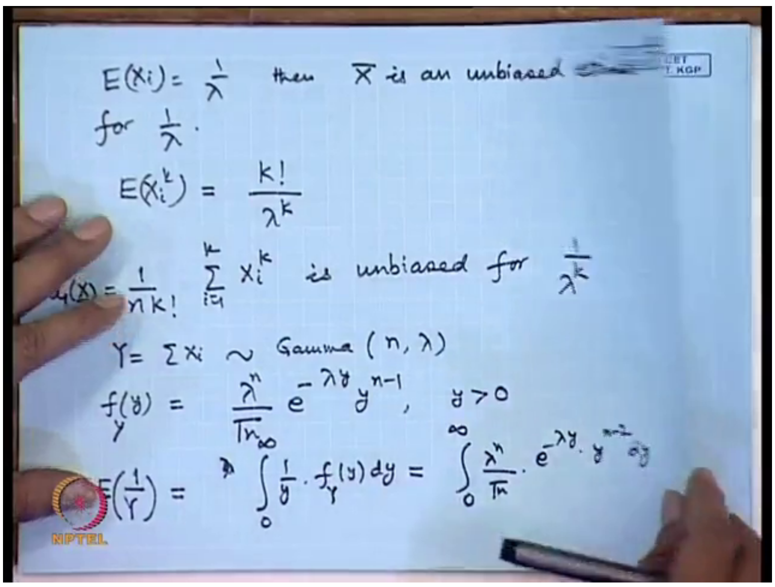
(No Audio From: 22:25 to 22:31) So, let me look at a proof of this expectation of s square that is equal to 1 by n minus 1, sigma expectation of X i minus X bar whole square. This we may write as 1 by n minus 1 sigma expectation of sigma X i square minus n X bar square.

Now, here we may use a property that expectation of X i square will be equal to mu square plus sigma square and similarly expectation of X bar square will be mu square plus sigma square by n, because variance of X bar is sigma square by n. So, these things if you 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 unbiasedness, let me give a name here. So, d star x that is equal to X bar square minus s square by n is unbiased for mu square. Let us consider say x having a density of exponential distribution lamda e to the power minus lamda 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 one by lambda.

(Refer Slide Time: 25:50)

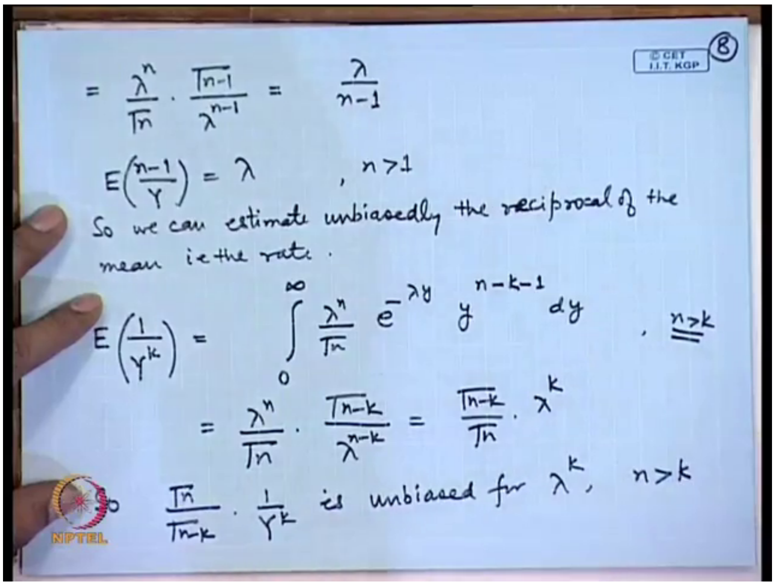


So, if we use this here, expectation of X i is equal to 1 by lamda then this gives us that X bar is an unbiased estimator for 1 by lambda. Not only that we may consider estimation of the higher order moments also. For example, we may look at say expectation of X i to the power k, in the exponential distribution this is equal to k factorial divided by lamda 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 1 k factorial and this will become k factorial by lamda to the power k and n will come. So, then this becomes unbiased for 1 by lamda 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. Sometimes we may be interested in lamda itself, which is the rate of this. Now, in that case we may have to do little bit of introspection here, because if we are considering direct moments I am getting the powers of one by lamda. If you want to estimate lamda itself then it suggests that we may have to consider reciprocal of X r X bar. Now, that is a curious thing here if I consider expectation of 1 by X in the exponential distribution here that is not existing. That means, I cannot obtain averaging in the way that I have done in these two cases.

On the other hand if we consider say the distribution of say sigma X i that is giving a gamma distribution with parameters n and lambda. That means, if I want to write down the density of y that is equal to lamda to the power n by gamma n e to the power minus lamda y y to the power n minus 1, where y is greater than 0. Now, let us consider expectation of say 1 by y now, that is equal to lamda to the power well 1 by y into this density from 0 to infinity. So, here let me substitute here 0 to infinity this form for this density so, I will get lamda to the power n by gamma n, e to the power minus lamda y, y to the power n minus 2 d y.

(Refer Slide Time: 29:15)

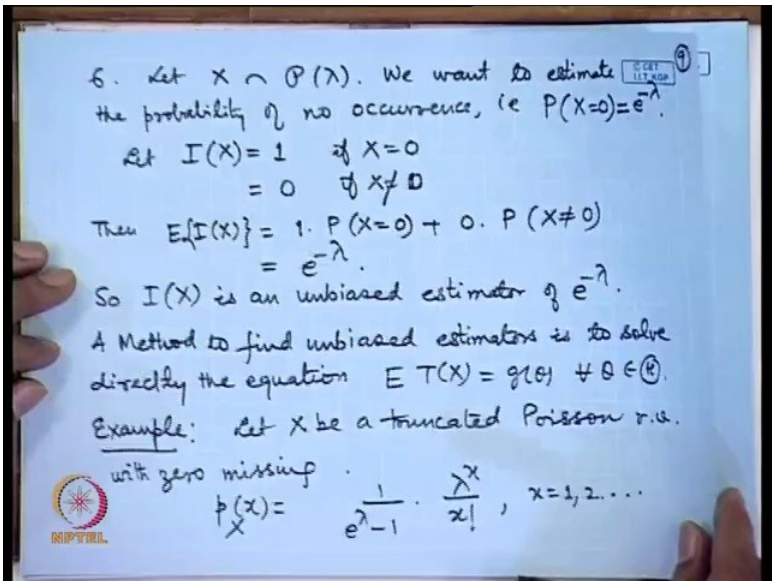


Now, that is equal to lamda to the power n by gamma n now, this can be evaluated using the gamma function formula. So, it becomes gamma n minus 1 divided by lamda to the power n minus 1 that gives us clearly lamda divided by n minus 1. So, what do we conclude expectation of n minus 1 divided by y is equal to lamda of course, here n has to be greater than 1 otherwise this value will not exist. So, if I have more than one observation from an exponential distribution, I can estimate even the reciprocal of the rate which was not possible. If I am considering only one observation because expectation of 1 by X does not exist.

But, here I am considering n observations where n is greater than 1 then expectation of n minus 1 by y that is sigma X i that is equal to lambda. So, we can estimate unbiasedly the rate, the reciprocal of the mean, not only that if we want to now consider some powers of lamda that also can be considered the corresponding powers of y in the denominator. For example if I consider say expectation of 1 by y to the power k now, that becomes lamda to the power n by gamma n e to the power minus lamda y, y to the power n minus k minus 1 d y. So, if n is greater than k then this is a gamma function and we can straight forwardly evaluate it as gamma n minus k divided by lamda to the power n minus k that is equal to gamma n minus k by gamma n lamda to the power k.

So, after adjustment of this coefficient we get gamma n by gamma n minus k 1 by y to the power k is unbiased for lamda to the power k, if n is greater than k. So, in most of the typical estimation problems, the structure of the moments keeps the unbiased estimators for the usual parametric functions. So, the usual parametric functions I mean either the moments or some linear functions of the moments, or some other type of functions. For example, in this case we have considered even the reciprocal of the moments and we are able to evaluate, but of course, we use the structure of the gamma distribution here. Sometimes the parameter to be estimated may not be in the form of a moment or it is a or any it may not have any relation with the moments. In that case, you may have to identify the kind of parameter that we are having

(Refer Slide Time: 32:56)



Let us take one example here, say X having poisson distribution and we want to estimate in many of the Poisson problems, because Poisson Distribution generally denotes the arrival rate in a service queue etcetera. Therefore, it is of interest to the organizers or the service providers to know when there will be no arrivals, accordingly they can provide the or you can say distribute the service personal in such a way that when there is a slack period that is you can estimate that this much period there will be no arrivals or no customers or no persons in the queue. Then the person which who suppose to do the duty there, he can be posted elsewhere or that slot may be kept free also.

So, we need to estimate say the occurrence of zero, or the probability of no occurrence that is probability say X is equal to 0 so, we want to estimate this. Now, in the Poisson Distribution this is equal to e to the power minus lamda, if I am considering poisson lamda distribution the probability of x equal to 0 is e to the power minus lamda. Now, you can easily notice that this is not a moment. In fact, if I expand it I will get it as a series 1 minus lamda plus lamda square by 2 factorial and so on. And if I substitute the corresponding estimates one by one then we are not sure of the convergence of the series. So, it is not a good idea to substitute directly in the expansion of e to the power minus lambda.

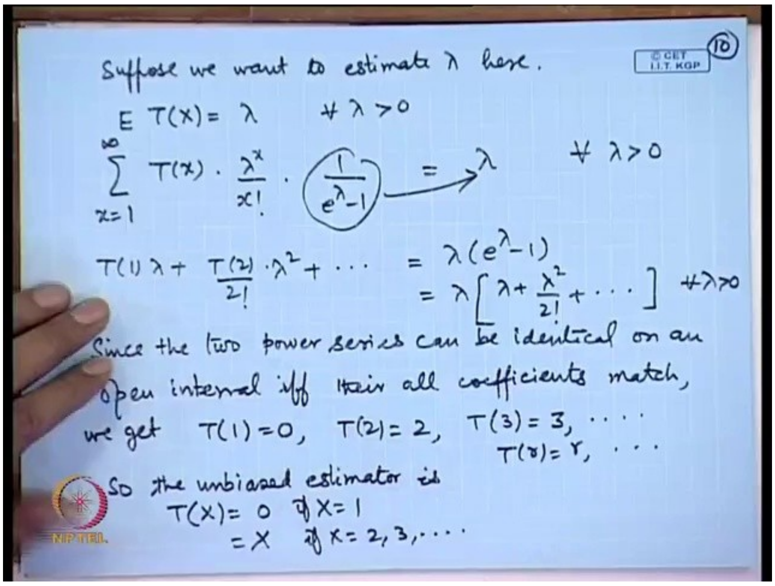
However, we can notice something like this, let us consider say I x the indicator function as 1, if X is equal to 0 and it is equal to 0 if X is equal to 1, if X is not equal to 0. Then, what is expectation of I X? That is equal to 1 into probability of X is equal to 0 plus 0 into probability X not equal to 0. So, this cancels out and we get only probability X equal to 0, which is our required parameter. So, here the indicator function of the set where X equal to zero that itself becomes an unbiased estimator for the parameter or you can say the probability of e to the power minus lamda. Of course, at this point you may raise the question that; this is not a proper estimator, or it may not be very informative in the sense that I just conduct the trial once and use the estimator as one if X is equal to zero otherwise I use it as a zero.

So, this is not a very good estimator for the function which is actually lying between zero and one. So, of course, that question remains and we may do further ramifications of this. So, right now we are able to obtain an unbiased estimator. (No Audio From: 36:24 to 36:36) So, we may look at the methods of finding out the unbiased estimators here. (No Audio From: 36:45 to 36:57) Right now, I have told two cases one is that we may use moments or reciprocal of the moments or some functions of the moments. Another thing could be to use the form of the parameter, which is coming in the form of the probability so, we took the indicate or function of that set.

So, let us consider some more examples here, we may directly write down the equation expectation of T x is equal to the g theta. Now, by using some methods of analysis we may be able to solve this equation; that means, we need to get the solution for the function T x here, which function T x will satisfy this equation so, that is another one. So, let me describe this thing, A method to find unbiased estimators is to solve directly the equation, expectation of T X is equal to g theta for all theta. Let us take one example here, let x be a truncated Poisson random variable with zero missing.

That means, the probability mass function is of this form say 1 by e to the power lamda minus 1 lamda to the power X by X factorial for X equal to 1 2 and. So, on the standard poisson distribution is X equal to 0 1 2 and so on. So, suppose zero is missing this type of situation may arise where we know from the given setup that the assumptions of the poisson process are satisfied for the arrival distribution. However, in the physical setup, it may turn out that when we are actually doing the sampling we cannot record the occurrence of zero. So, in that case the distribution which will be actually recorded will be of this form. Now, suppose we are considering again say estimation of lamda.

(Refer Slide Time: 39:41)

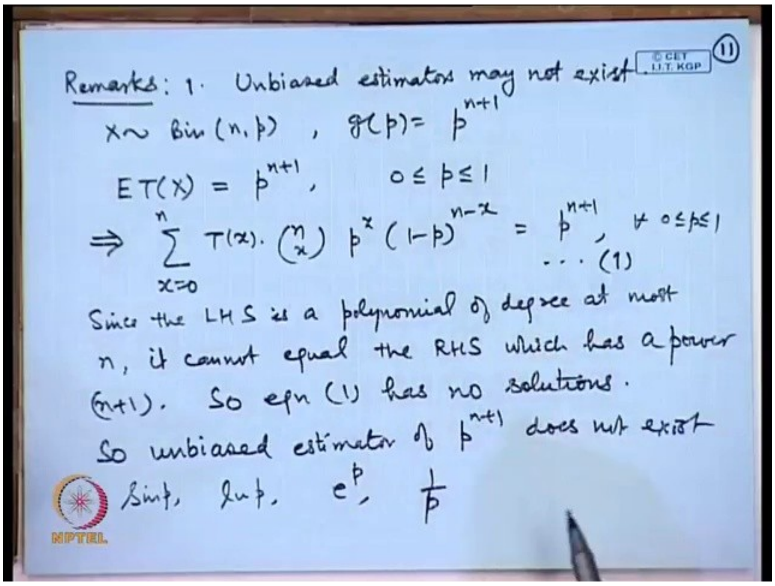


So, suppose we want to estimate here. (No Audio From: 39:44 to 39:52) Now, notice here that here the expectation of X is not equal to lamda so, we cannot directly use the moment here. So, we write down an equation expectation T X is equal to lamda for all lamda greater than 0. So, now substitute here T x, where X is equal to 1 to infinity lamda to the power X by X factorial 1 by e to the power lamda minus 1, this is equal to lamda for all lamda greater than 0. Let us elaborately write down this equation so, I take this term to this side this becomes lamda into e to the power lamda minus 1 and then we expand this. So, I get T 1 lamda plus T 2 by 2 factorial lamda square and so on. That is equal to lamda into e to the power lamda minus 1, which I can expand this becomes lamda plus lamda square by 2 factorial and so on.

Now, this statement I am writing for all lamda greater than 0, the left hand side is a power series in lamda the right hand side is a power series in lambda. So, the power series are equal on a open subset of a real line if and only if all the coefficients are equal. So, if we do that then we can compare the coefficient, since the two power series can be identical on an open interval if and only if their all coefficients match, we get. So, you compare the coefficient of lamda on the right hand side there is no coefficient of lambda. So, we get T 1 is equal to 0 now T 2 the coefficient of lamda square is one here the coefficient of lamda square is T 2 by 2 factorial.

So, T 2 becomes two factorial that is 2 then coefficient of lamda q 1 the left hand side will be T 3 by 3 factorial and here it is 1 by 2 factorial. So, T 3 by 3 factorial is equal to 2 factorial; that means, T 3 is equal to 3 factorial by two factorial that is equal to 3 and so on. That means, in general I will get t r is equal to r that is r factorial divided by r minus one factorial. So, the unbiased estimator is (No Audio From: 42:36 to 42:44) T X is equal to 0, if X is equal to 1 it is equal to X, if X is equal to 2 3 and so on. So, here we have seen that by solving an equation directly one can obtain unbiased estimators. However, this type of technique may not be very useful in continuous distributions etcetera.

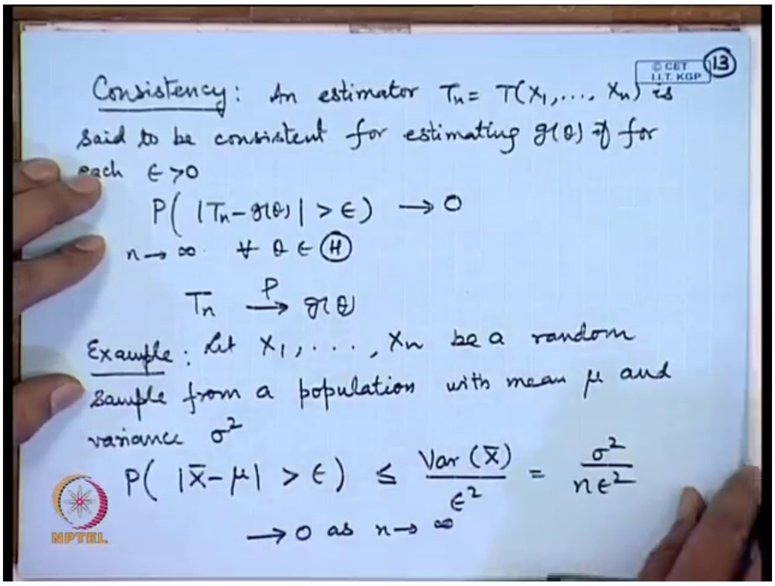
(Refer Slide Time: 43:12)



Let me give brief remark on the existence of unbiased estimators, it may turn out that in some situations there may not be any unbiased estimators, unbiased estimators may not exist. So, for example, consider say X following binomial n p, I want to estimate say p to the power n plus 1, let my g p be equal to p to the power n plus 1. So, if I consider say T be an unbiased estimator so, I will write expectation of T X is equal to p to the power n plus 1 for p in the interval 0 to 1, this means n c x p to the power X into 1 minus p to the power n minus X. (No Audio From: 44:15 to 44:22) Now, notice this equation here, on the left hand side you have a polynomial in p which is having degree at most n, because the maximum power that p can have is p to the power n or here 1 minus p to the power n.

Whereas in the right hand side you have p to the power n plus 1. So, now the two polynomials can agree on an interval if and only if all their coefficients agree, but here that does not seem to be possible. Therefore, this situation or you can say this equation has no solution. Since the left hand side is a polynomial of degree at most n, it cannot equal the right hand side which has a power n plus 1. So, this equation let me call it 1, equation 1 has no solutions. So, unbiased estimator of p to the power n plus 1 does not exist. See similarly we may consider a function say sin p, we may consider say l n p, suppose we consider e to the power p suppose we consider 1 by p in all these cases unbiased estimators will not exist.

(Refer Slide Time: 46:19)



There may be yet another type of situation that unbiased estimators are not reasonable. By reasonable, I mean that if I say my parameter lies between zero to one then, my estimator should also take values between zero and one. If I say my parameter is positive then, my estimator should also take positive values. If I have my parameter to lie in a given range say from minus m to m then my estimator should also be between minus m to m. So, these are some physical constraints that the estimator must satisfy. So, there may be some situations where unbiased estimators actually does not satisfy this. I gave you one example for estimation of the probability of zero occurrence.

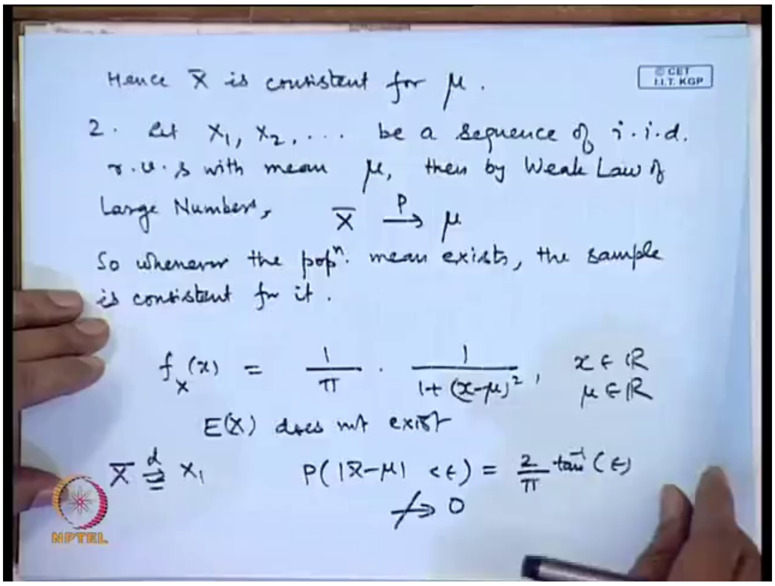
So, here you can see e to the power minus lamda this is lying between 0 to 1 whereas, the estimator is taking value either 0 or 1 so, this is not a very proper estimator here. We may consider another example, say I want to estimate say e to the power minus 3 lamda. Now, if I consider expectation of minus 2 to the power X in poisson distribution then it is equal to minus 2 to the power x e to the power minus lamda, lambda to the power x by x factorial. So, if you sum this you get e to the power minus 3 lambda. So, now, you see the values of this minus 2 to the power X, it will take values corresponding to X is equal to zero it is taking value 1, corresponding X equal to 1 it is taking value minus 2, corresponding X equal to 2 it is taking value 4, corresponding x equal to 3 it is taking value minus 8,16 and so on.

So, you can see here this is never taking values, which is prescribed for this this parameter lies between zero to one e to the power minus 3 lamda because lamda is a positive parameter here. But the estimator is taking absurdly different values starting from 1 then minus 2 4 minus 8 and so on, so this is not a reasonable estimator. So, in unbiased estimation, we may have to be careful that the estimator should be reasonable, more about unbiased estimation we will take up in the next class. Now, we take up another desirable criteria that is called consistency. So, this is a large sample property by a large sample property we mean that if n is large, what is the behavior of the estimator?

So, an estimator T n that is equal to T of X 1 X 2 X n is said to be; so, here we are showing dependence upon n here that n observations are used. So, this is said to be consistent for estimating g theta, if for each epsilon greater than 0 probability of modulus T n minus g theta greater than epsilon this goes to 0 as n tends to infinity for all theta. Actually in probability theory when we discuss the concept of convergence of random variables, this is equivalent to saying that T n converges to g theta in probability. So, this means here essentially that as n increases; that means, if I have a sufficiently large sample. Then, the probability that my estimator is quite close to the true value of the parameter, because I am saying that the probability of this being greater than epsilon is actually almost negligible.

So, there is a very high probability that in a large sample my estimator will be almost equal to the, or it will be very close to the true value of the parameter. So, this is actually a essentially a large sample property you can say it is a slightly relaxed property compare to the unbiasedness. So, let us take an example here, let X 1 X 2 X n be a random sample from a population (No Audio From: 51:08 to 51:21) with mean mu and say variance sigma square. Let us consider probability of modulus X bar minus mu greater than epsilon then by shapes in equality it is less than or equal to variance of X bar by epsilon square that is equal to sigma square by n epsilon square. Now, you can see that this quantity goes to 0 as n tends to infinity.

(Refer Slide Time: 52:09)

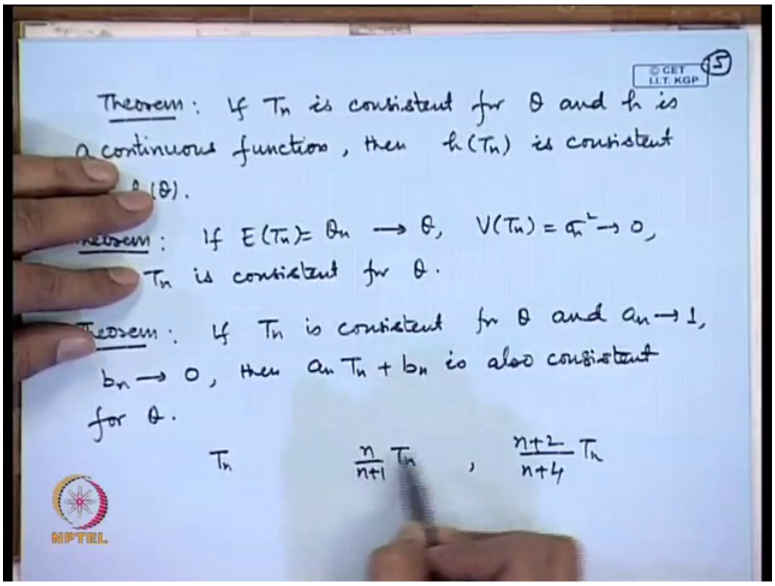


That means, if I am assuming the mean and variance then, the sample mean is a consistent estimator for the population mean. X bar is consistent for the population mean mu, another way of looking at it could be through the weak law of large numbers. Let us consider let X 1 X 2 and so on be a sequence of independent and identically distributed random variables, say with mean mu then by weak law of large numbers X bar converges to mu in probability. So, in fact, the existence of the second moment is not required here I used shapes in equality that is why I assumed sigma square here, but for actual weak law of large numbers that is not required.

So, in general whenever the population mean exists, the sample mean is a consistent estimator for the population means. So, like unbiasedness criteria the consistency for the sample mean is also a very nice property that is holding. So, we can say that whenever, the population mean exists the sample mean is consistent for it. You may say that this property may be a trivial, but actually it is not. So, there are distributions such as Cauchy Distribution, see if I consider a Cauchy Distribution (No Audio From: 53:57 to 54:06) then we know that here the mean does not exist, expectation x does not exist. In fact, if I consider the distribution of X bar that is same as X 1; that means, in distribution these two are same.

So, if I look at the probability of modulus X bar minus mu less than epsilon that is equal to basically 2 by pi tan inverse epsilon. So, this does not go to 0, because there is no role of n here. So, here the criteria of consistency fails to hold.

(Refer Slide Time: 54:50)



There may be also cases where for certain range of parameter the first moment will exist in some other range it will not exist. Therefore, consistency of the sample mean or the unbiasedness of the sample mean will be holding only in that region, another important property which is satisfied by the consistency is the invariance. So, we have the following result that if t n is consistent for say theta and h is a continuous function, then h T n is consistent for h theta. So, for example, here I am mentioning say T is say consistent for theta and I am looking at say estimation of theta square then T square will be consistent for theta square.

Now, you can notice here the difference from the unbiasedness here in unbiaedness this type of invariance was not there except for the linear invariance. If I am considering say one by T, then if T is unbiased for theta one by T is not necessarily unbiased for one by theta. In fact, in most of the cases it will not be whereas, inconsistency this will be true. Another important result in consistency is that if expectation of T n is equal to theta n, which actually converges to theta. And variance of Tn converges to 0, then t n is consistent for; that means, it need not be unbiased. But in limit it is unbiased and if the variance of the estimator is negligible, or it becomes negligible as sample size increases then T n becomes consistent for theta.

So, as I was mentioning this is slightly relax property and it is quite helpful in the large samples that many estimators. Which may not look very reasonable from the point of view of unbiasedness etcetera, but they become alright for the consistency property. Similarly, if T n is consistent and a n is a sequence of numbers which goes to 1, b n goes to 0, then a n T n plus b n is also consistent for theta. That means, if I say T n then I can consider say n plus; n by n plus 1 T n, I may consider say n plus 2 by n plus four. So, these are all consistent, because these are all going to suppose I put here plus 1 by n plus 1 plus 1 by n etcetera then they are also consistent estimators for the parameter theta.

So, today we have discussed two important criteria one is unbiasedness and another is consistency. However, it is important to know, what are the methods for determining the estimators? Or how to find out the desirable estimators? So, in the next lectures we will be covering important methods for finding out the estimators. And then, we will come back further to the topic of criteria; that means, we will look at the efficiency of the estimators which estimator should be used, suppose there are more than one estimator available for the same problem satisfying the same criteria. Then what extra criteria should be introduced to choose one over the other. So, these and other topics we will be covering in the forth coming lectures.