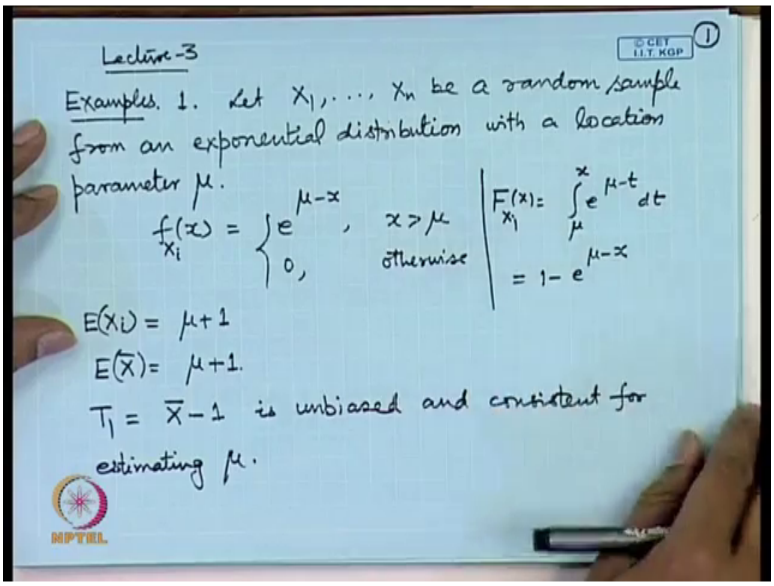
Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 05
Basic Concepts of Point Estimations-III

In the last lecture I introduced two Basic Concepts of Point Estimation namely unbiased estimation and consistency. One of the properties that is unbiasness it is related to the estimator being equal to the true value on the average; that means, if we have many samples then the average of that will be equal to the true value. Whereas, consistency is a large sample property; that means, if we take a sample to be large enough then the probability that it is close enough to the true value of the parameter is almost equal to one and. So, the two properties have somewhat different applications and as well as implications and many times we try to combine various properties of the estimators.

So, I had shown in the last lecture some sort of invariance of the consistent estimators for example, if T is a consistent estimator for theta then g of t and where g is a continuous function will be consistent for g theta. Similarly, if I have T n to be consistent for theta and I have sequences of numbers a n and b n such that a n converges to 1 and b n converges to 0 then a n T n plus b n also is a consistent estimator for theta. So, let me give a few examples.

(Refer Slide Time: 02:04)

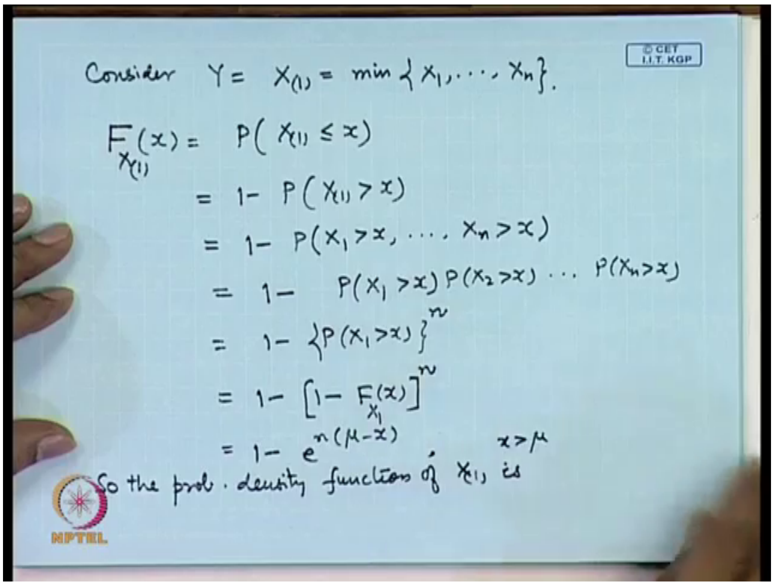


Let us consider say let X 1 X 2 X n be a random sample from an exponential distribution with a location parameter mu; that means, I am considering the density function of say X i to be e to the power mu minus x where x is greater than mu and 0 otherwise.

So, this is actually the well known shifted exponential distribution here mu denotes the minimum guarantee time of the component or the life. So, here if you see expectation of X i is equal to mu plus 1. So, mu plus 1 is the first moment and therefore, if I consider expectation of X bar that is also going to be mu plus 1.

So, by weak law of large numbers we get X bar as a consistent estimator for mu plus 1 and if I take 1 to the left hand side it then we get let me call it T 1. So, T 1 is equal to X bar minus 1 is unbiased and consistent for estimating mu, that is the minimum guarantee time Now, in this problem let me introduce another estimator.

(Refer Slide Time: 04:18)

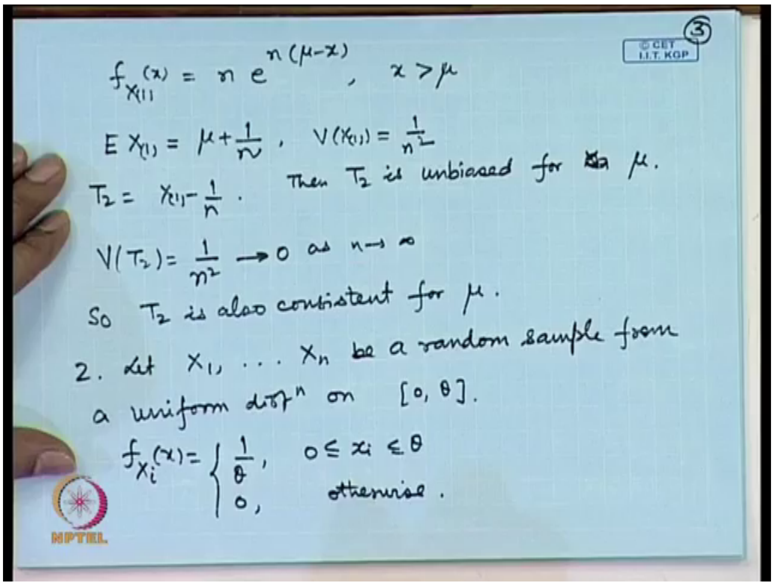


Let us consider say Y is equal to X 1 here this X 1 denotes the minimum of the observations X 1, X 2, X n one can derive the distribution of X 1. In fact, in general if I want to find out the distribution of this I can find it in the following way, I consider say c d f of this that is probability of X 1 less than or equal to x. This can be written as 1 minus probability of X 1 greater than x that I can write as 1 minus probability that now, if the minimum is greater than X this is equivalent to saying each of the X n is are greater than x.

Now, here X 1 X 2 X n are is a random sample therefore, X 1 X 2 X n are independently and identically distributed random variables. So, this can be actually written as 1 minus probability of X 1 greater than x into probability of X 2 greater than x and so on. Probability of X n greater than x that is 1 minus probability of X 1 greater than x to the power n that is equal to 1 minus, now this is again 1 minus c d f of X 1 itself. Now if I have this as the probability density function I can write down the corresponding c d f here that is integral from mu to x of e to the power mu minus t d t that is equal to 1 minus e to the power mu minus x.

So, if I substitute 1 minus e to the power mu minus x here I will get 1 minus e to the power n times mu minus x. So, the probability density function of X 1 is, now this can be obtained by considering derivative of this of course, this value I have written for x greater than mu, if x is less than mu then this is going to be 0.

(Refer Slide Time: 07:21)



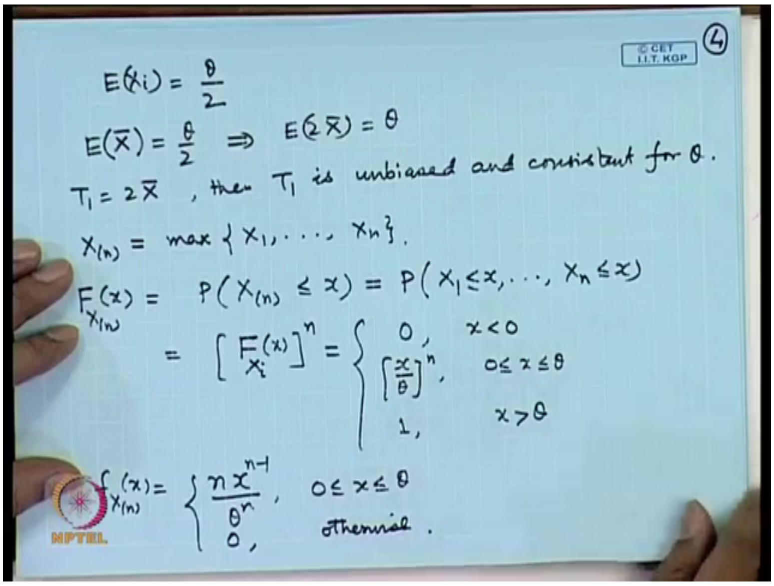
So, if we consider derivative of this I will get the density function of X 1 as n times e to the power n mu minus x where x is greater than mu. So, this is the probability density function of the minimum of the observations if I have considered a random sample from an exponential distribution with a location parameter mu.

Now, this is the usual two parameter exponential distribution, here the scale parameter is 1 by n and location parameter is mu. So, if I consider the expectation of X 1 that is equal to mu plus 1 by n. So, if I take T 2 as X 1 minus 1 by n then T 2 is also unbiased for X 1 for mu. So, I have got another unbiased estimator, now in this one I can consider variance also, what is variance of X 1 for example, why variance of X 1 here is 1 by n square. So, variance of T 2 is also 1 by n square because it is variance of X 1 itself, the variance of a function does not change if I make a change of origin.

Now, consider the result that if expectation is equal to the parameter and the variance goes to 0 T 2 is unbiased and its variance goes to 0 as n tends to infinity. So, T 2 is also consistent for mu. So, in this problem we have considered two estimators, one is based on the sample mean this is unbiased and consistent and at the same time we have considered T 2 which is based on the minimum of the observations this is also unbiased and consistent.

So, that brings us to the question that if I have more than one estimator satisfying certain given desirable properties then which one we should use. So, in this direction I will give you one more example, let us consider say X 1, X 2 X n be a random sample from a uniform distribution on the interval 0 to theta. That means, I am considering the density of x i as 1 by theta 0 less than or equal to x i less than or equal to theta it is equal to 0 otherwise.

(Refer Slide Time: 11:01)



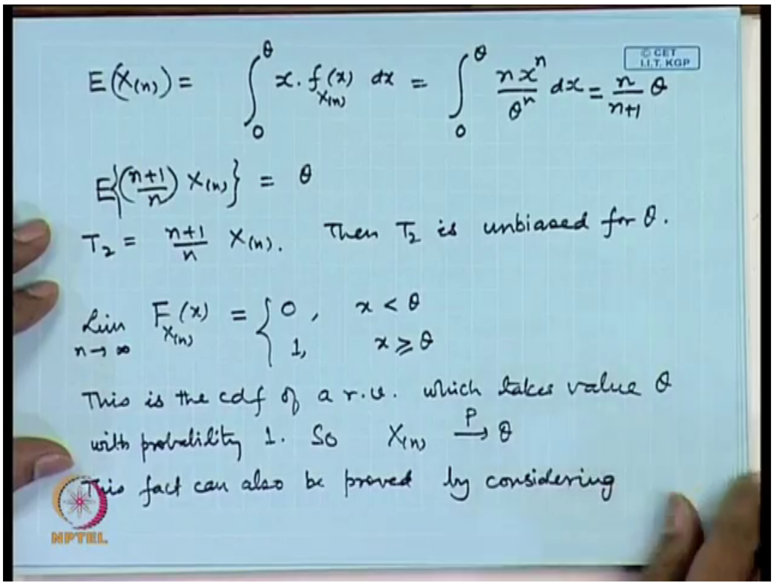
Now in the uniform distribution we know expectation of X i is the middle point of the interval that is theta by 2.

So, immediately we conclude that expectation of X bar is theta by 2 this implies that expectation of 2 X bar is equal to theta. So, if I call T 1 is equal to 2 X bar then T 1 is unbiased and consistent for theta. Now in this problem let me consider another one, let us consider say X n now X n I am calling to be the maximum of the observations. As in the previous case we can derive the distribution of X n let us consider the c d f of this. So, this is equal to probability of X n less than or equal to x.

Now, this is statement that the maximum is less than or equal to x is equivalent to that each of the observations is less than or equal to X and once again using the fact that X i's are independently and identically distributed this is equivalent to saying each of the X i's c d f at x. So, this is simply this to the power n, now for the uniform distribution the c d f is it is equal to 0 if x i is less than 0, it is equal to x by theta if 0 is less than or equal to x is less than or equal to theta it is equal to 1 if x is greater than theta. So, if we use this c d f here this becomes 0 if x is less than 0 it is equal to x by theta to the power n if 0 less than or equal to x is less than or equal to theta it is equal to 1 if x is greater than theta. One may find out the probability density function from here, by considering the derivative because this is a continuous distribution.

So, you get the density function as n x to the power n minus 1 by theta to the power n if x lies between 0 to theta and it is 0 otherwise.

(Refer Slide Time: 13:53)

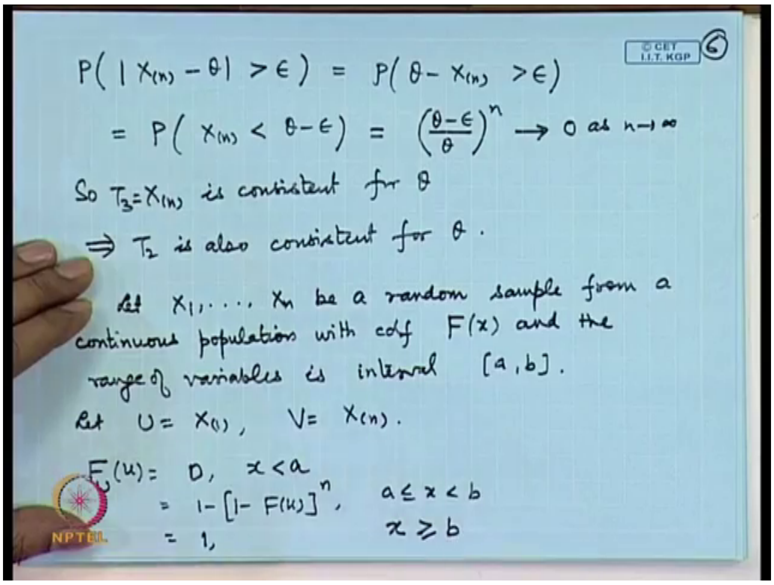


Let us consider say expectation of X n now. So, expectation of X n is equal to integral x into the density function of X n from 0 to theta that is equal to integral 0 to theta n x to the power n by theta to the power n d x. So, as we can see easily the integral of x to the power n will be x to the power n plus 1 by n plus 1 and if we substitute the limits from 0 to theta I will get theta to the power n plus 1. And in the denominator I have theta to the power n so that will cancel and therefore, this value will be equal to n by n plus 1 theta.

If I adjust this n by n plus 1 on the left hand side I get this is equal to theta. So, if I use the notation T 2 as n plus 1 by n x n then T 2 is unbiased we had obtained estimator T 1 as 2 X bar which is unbiased and now I have obtained T 2. Let us check say probability of x n minus modulus theta greater than epsilon whether it tends to 1 or not. That we can check from here also if we take the limit of this cumulative distribution function now here x is less than or equal to theta. So, this value will tend to 0 if x is less than theta and whenever x is greater than or equal to theta it is becoming 1.

So, if I take the limit of if I consider say limit of F X n as n tends to infinity then this is equal to 0 for x less than theta and it is equal to 1 for x greater than or equal to theta. Now this denotes the distribution of a random variable which takes value only theta, this is the c d f of a random variable which takes value theta with probability 1. So, basically we have proved that x n converges to theta n distribution, but theta is a constant therefore, it is equivalent to saying X n converges to theta in probability. In fact, this fact can also be proved in a different way, I will consider directly the definition of convergence in probability.

(Refer Slide Time: 17:25)



Let me take say probability of modulus X n minus theta say greater than epsilon.

Now, the distribution of X n is in the interval 0 to theta; that means, X n is always below theta. So, if we consider this modulus of X n minus theta this is same as theta minus X n. So, this is equivalent to probability of theta minus X n greater than epsilon which I can write as probability of X n less than theta minus epsilon. Now this is nothing, but the distribution function of X n at the point theta minus epsilon since X n is having a continuous distribution whether we put less than or less than or equal to it does not make a difference.

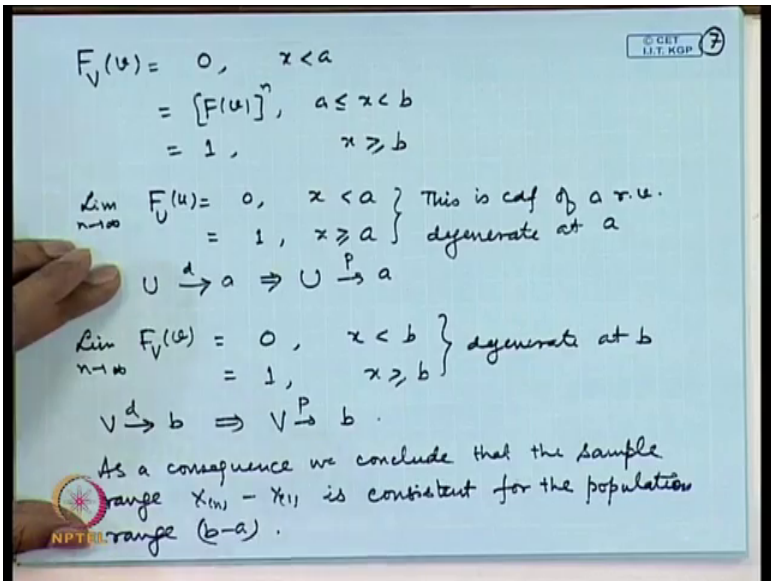
Therefore, this value is equal to theta minus epsilon by theta to the power n as we have derived just in the previous sheet here. So, now you can see epsilon is a positive number. So, theta minus epsilon is less than theta therefore, if I take the limit as n tends to infinity this will go to 0. So, X n is consistent for theta let me call it T 3; if we look at the coefficient n plus 1 by n this goes to 1 as n tends to infinity.

So, we have the result that if T n is consistent for theta and a n goes to 1 then a n T n is also consistent therefore, if we use this fact T 2 is also consistent for theta. So, T 2 is unbiased and consistent T 1 is also unbiased and consistent. So, I have given you two different distributions where for estimating of one parameter I am getting 2 different unbiased and consistent estimators. So, that shows that actually we need additional criteria to distinguish between or to choose between various competing estimators. From the previous two exercises we can also find something more important. If we look at the form of the distribution function of the maximum and the minimum there is some a specific structure here for example, when I took the limit here I got only 0 1 here.

Similarly, in the distribution of the minimum we had that X 1 is converging to mu. So, minimum was converging to the lower limit and here the maximum is converging to the upper limit theta. In fact, if we have any continuous distributions then this is a general fact. So, I will a state it in the following results. So, let me give it as exercise let X 1, X 2 X n be a random sample from a continuous population with c d f say capital F x and the range of variables is interval a to b. Of course, this interval may be open or closed that does not make any difference if we are handling a continuous distribution. Let us define say U is equal to the minimum and V is equal to X n then the claim is that U is a consistent estimator for a and V is a consistent estimator for b.

So, the proof will use the steps which we have derived just now, that is the c d f of u that is 1 minus. So, actually it is equal to 0 for x less than a it is equal to 1 minus 1 minus F u to the power n for a less than or equal to x less than b it is equal to 1 for x greater than or equal to b.

(Refer Slide Time: 22:42)



Similarly if I consider say F V then it is equal to x 0 for x less than a it is equal to F v to the power n for a less than or equal to x less than b it is equal to 1 for x greater than or equal to b notice here that this equality or inequality does not make any difference here because it is continuous distribution.

So, if I take the limits here now F is a number between 0 to 1. So, if I take the limits here this number will go to 0; so, this is going to 1. So, if I take the limit as n tends to infinity of F U I am getting 0 for x less than a and it is equal to 1 for x greater than or equal to a.

So, this is c d f of a random variable which is simply degenerate at a. So, we can conclude that u converges to a in distribution and therefore, u converges to a in probability because convergence in distribution and probability are equivalent if the right hand side is a constant. Similarly if I consider limit of F V as n tends to infinity then this is also 0 for x less than b and it is equal to 1 for x greater than or equal to b.

So, once again V is converging to b this random variable is degenerate at b. So V tends to b in distribution or V tends to b in probability. So, if a continuous distribution is having a range a to b then the smallest order statistics converges to the lowest value or the lowest value in the range and the largest order statistic converges to.

So, these can be treated as the consistent estimators of these respective parameters. So, this actually gives some easy applications basically for example, we want to find out a consistent estimator for the range. For example, here range may be b minus a then easily you can say that V minus u that is the maximum minus the minimum, sample range is the consistent estimator for the population range.

As a consequence we conclude that the sample range that is X n minus X 1 is consistent for the population range b minus a. We have some further special cases here for example, lower limit could be minus infinity or the upper limit could be plus infinity. In that case for example, if the lower limit is minus infinity then X 1 does not converge in probability. Similarly if the highest value is unbounded; that means, b is infinity then x n does not converge in probability.