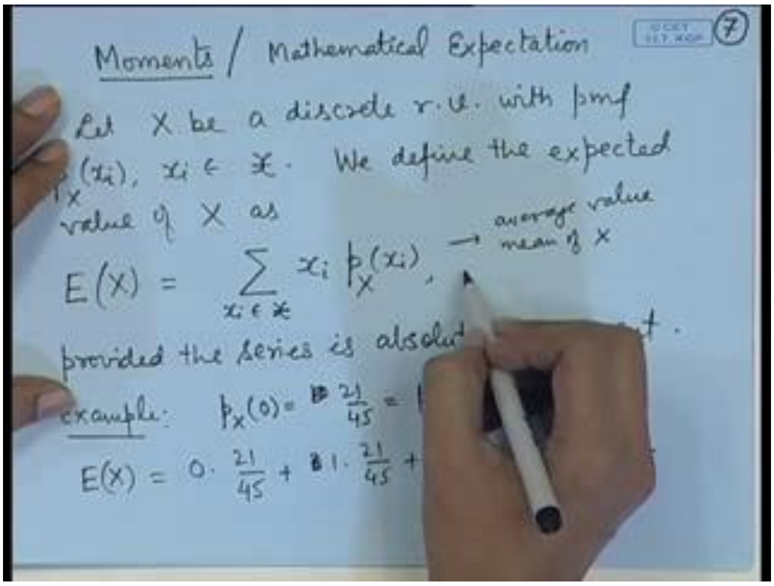
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 16
Moments

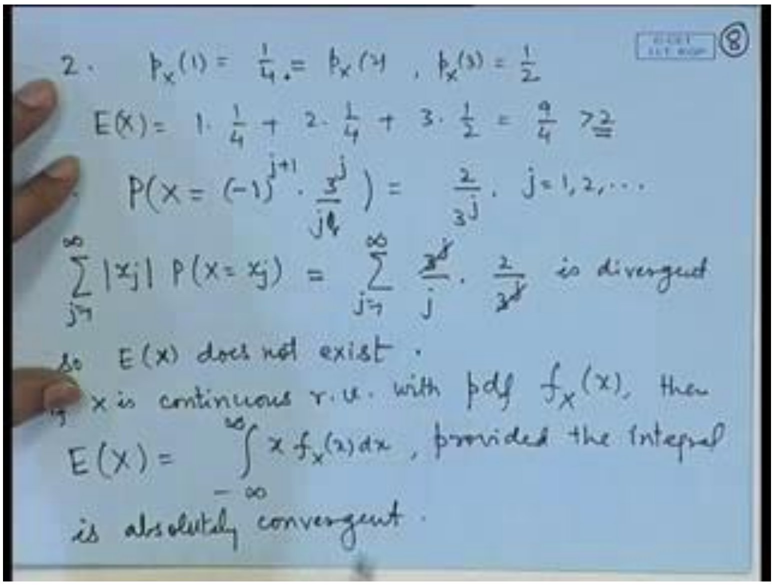
(Refer Slide Time: 00:25)



Concept of moments or mathematical expectation; let X be a discrete random variable with probability mass function say P X x i for x i belonging to some set effects. So, we define the expected value of X as E of X, this is defined as sigma x i into P X x i where x i belongs to x provided the series is absolutely convergent.

Let us consider some examples, now let us take the example of defective computer purchases; that means, a store had certain number of computers and a person purchase 2 computers. So, in that example we had probability of x equal to 0 was 10 by 21 by 45 and that was same as probability of x equal to 1 and probability x equal to 2 was 3 by 45. Let us look at expectation X. So, that will be equal to 0 and into 21 by 45 plus 1 into 21 by 45 plus 2 into 3 by 45. So, that is equal to 27 by 45 are 3 by 5; that means, on the average is purchased may include less than 1 defective computer.

(Refer Slide Time: 03:12)

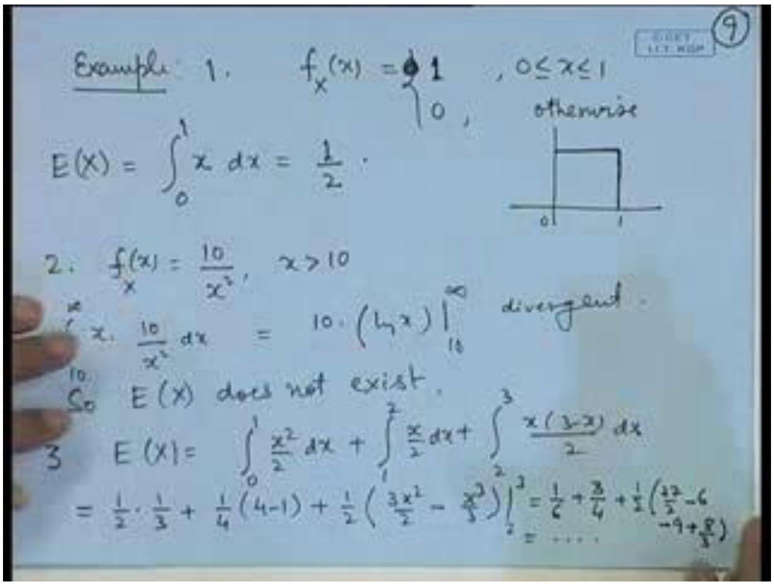


Let us look at the next problem of that is p x 1 was 1 by 4, p x 2 was 1 by 4 and p x 3 was half that is a number of trials needed for getting the defective able. Here expectation of x is equal to 1 into 1 by 4 plus 2 into 1 by 4 plus 3 into 1 by 2, which is equal to 9 by 4 which is actually greater than 2; that means, on the average you will need more than 2 testings for getting the defective able.

You notice here that we have not checked the condition of absolute convergence here, because that number of terms is only finite. The condition is needed actually in order that the expectation is well defined, consider say probability that x is equal to minus 1 to the power say j plus 1, 3 to the power j by j factorial say j is equal to 2 by 3 to the power j, j is equal to 1 2 and so on. Let us look at sigma modulus of x j probability X is equal to x j, j is equal to 1 to infinity, then this is equal to sigma 3 to the power j by j factorial by j 2 divided by 3 to the power j; now this is divergent.

So, expectation X does not exist. Although one may write here sigma x j probability X equal to x j and here you will get minus 1 to the power j plus 1 2 by j, which is having value log 2, but it is not absolutely convergent therefore, expectation X does not exist here. So, this expectation X is also called average value, it is also called mean of X are arithmetic mean of X etcetera. So, several names are there it is also called the first moment about the origin. If X is a continuous random variable with pdf f x, then expectation X is defined as integral minus infinity to infinity x f x d x provided the integral is absolutely.

(Refer Slide Time: 06:48)



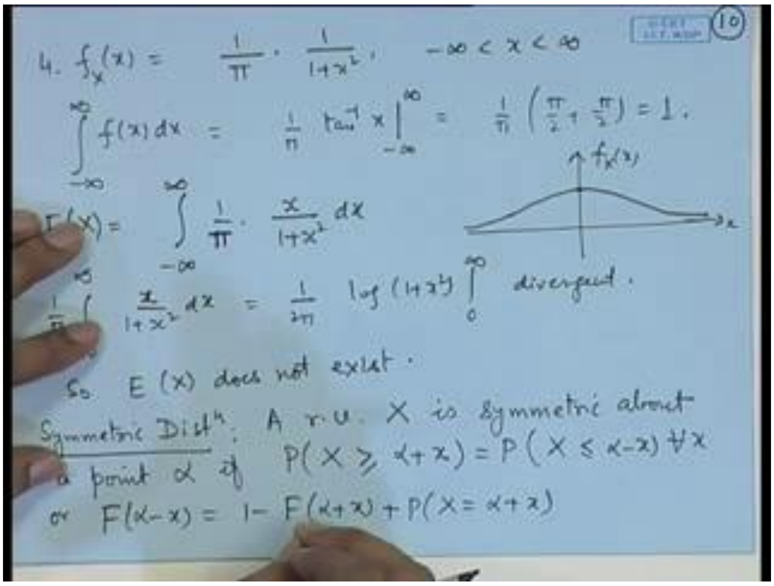
Let us look at some of the examples; the first example here we had in the continuous case f x was 0 for it was 1 for 0 less than or equal to x less than or equal to 1 and it is 0 otherwise. So, expectation X will become equal to integral x into f x, that is equal to 1 in this interval and therefore you will have basically only this term which is equal to half, which looks reasonable also because if you plot this distribution from the 0 to 1 interval, it is actually constant value and therefore the mean value must be middle value that is half. Let us look at the next example.

F x is equal to 10 by x square for x greater than 10 and I skipping the other part here 0. So, integral x 10 by x square on d x for 10 to infinity. So, that is equal to 10 and integral of 1 by x is log x for 10 to infinity this is divergent. So, here expectation X does not exist. Let us consider another example, f x is equal to x by 2, for 0 to 1 it is half for 1 to 2 and 3 minus x by 2 for 2 to 3. For this example expectation X will become equal to integral x square by 2 d x plus integral x by 2 d x for 1 to 2 plus integral 2 to 3 x into 3 minus x by 2, in the remaining portion f x was 0 so there will be no term here.

So, this is equal to half 1 by 3 plus here the integral will become x square by 4. So, 1 by 4 four minus 1 plus 1 by 2, integral of 3 x is 3 x square by 2 minus x square will give you x cube by 3 for 2 to 3. So, it is equal to 1 by 6 plus 3 by 4 plus half the value were this is

27 by 2 minus 6 minus 9 plus 8 by 3, which can be actually simplify. So, there may be some other case also where it looks that the expectation will exist as we have seen here in the second example, the integral itself is divergent and therefore, we are directly concluding that the expectation does not exist. However, there may be a case where it looks that the integral will exist, but actually it does not exist.

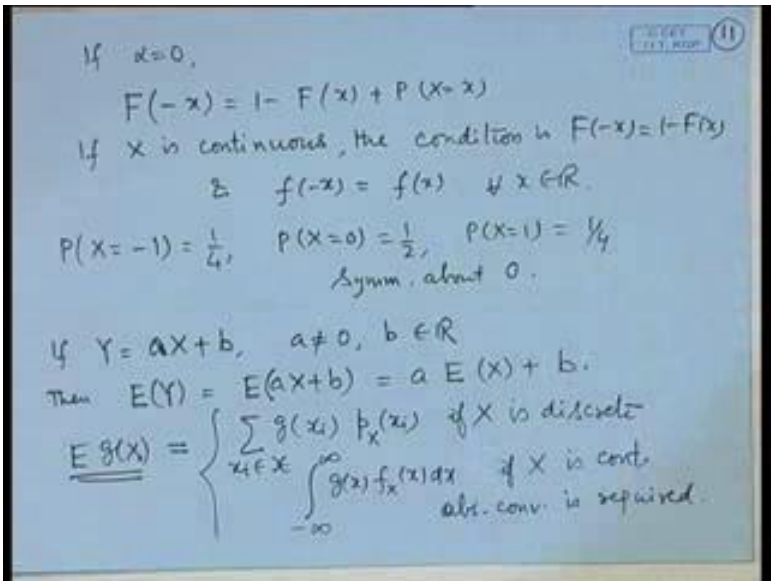
(Refer Slide Time: 10:45)



Let us consider the density 1 by pi into 1 by 1 plus x square. So, it is a valid probability density as can be seen easily the integral will be tan inverse x that is equal to 1 by pi, pi by 2 plus pi by 2 that is equal to 1. So, it is a valid probability density; however, if I look at expectation of X, a common mistake here is that here one can think that it is an odd function and it is over a symmetric region, so it should be 0; however, the property of the even function or odd function is applicable when the integral is convergent; here if we look at integral 0 to infinity 1 by 1 plus x by 1 plus x square then it is equal to log of 1 plus x square, which is actually divergent.

Therefore expectation X does not exist. If we look at the shape of this curve actually at x equal to 0 it is 1 by pi and if we look at the shape of this curve then here it is at x equal to 10 it is 1, and there after it is reducing. From various curves here for examples we found the expectation to be the middle value, here the expectation does not exist here it is somewhat different. So, we are tempted to consider something like a concept of symmetry, we can define symmetric distribution as a random variable X is symmetric about a point alpha if probability of X greater than or equal to alpha plus x is equal to probability of X less than or equal to alpha minus x for all x.

In other words we can say F of alpha minus x is equal to 1 minus F of alpha plus x plus probability of x is equal to alpha plus x. If the random variable is continuous this term will vanish and it will be simple relationship between in the CDF at minus and plus point (Refer Slide Time: 14:22)



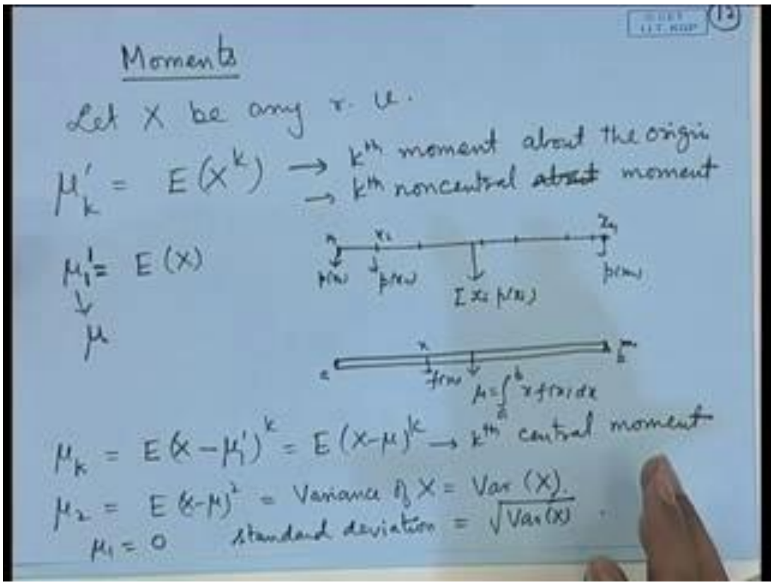
Suppose I take alpha is equal to 0 in this definition, if we take alpha is equal to 0 then this condition is reduced into simple F of minus x is equal to 1 minus F of x plus probability x equal to x. If X is continuous the condition is F of minus x is equal to 1 minus F of x and if it is density, so we will have. So, if you see this distribution it is symmetric about 0, here the distribution is symmetric about half etcetera.

We may also define a discrete symmetric distribution; probability x equal to say minus 1 is 1 by 4, probability x equal to 0 is half, probability x equal to 1 is 1 by 4. So, this is symmetric about 0; we notice here some properties about the expectation function, for example, if I define Y is equal to a x plus b, where a is any non 0 real number and b is any real number, then if we consider expectation of Y then it is expectation of a x plus b it will become equal to expectation a times expectation of a x plus b.

In general we can define expectation of a function of random variable also for example, if i have expectation of g x suppose x is a discrete random variable then we can consider it as if x is discrete; we can define it as integral g x, f x, d x, if x is continuous. The definition is subject to the condition that these summations are this integral must be absolutely convergent. So, absolute convergence is required in order that this expectation of g x is well defined. Another question arises at this point that how do we define the expectation of a mixed random variable. So, for a mixed random variable, the expectation would be simply the value multiplied by the probability plus the integral of the density multiplied by the value so; that means, in the discrete and the continuous case we separately evaluated.

Let us look at example of this mixed random variable. So, here expectation X will become equal to 0 into 1 by 4 plus x into 3 by 4 d x 0 to 1. So, this is equal to 3 by 8 which is actually less than half, this would have been the mean if the random variable was completely defined as 1 for 0 to 1, the probability density function; however, here the 1 by 4 probability is taken over by the point x equal to 0 therefore, the average value as average value waiting time as reduced and it is now 3 by 8 it is not half.

(Refer Slide Time: 19:00)



Next we discuss the concept of moments of distributions. So, let X be any random variable, we define mu k prime is equal to expectation of X to the power k, this is called kth moment about the origin or kth non central moment about kth non central moment. Observe here that if I put say k is equal to 1, then mu 1 prime is nothing, but expectation of X that is the expected value. So, here we are getting the if I am considering x equal to k equal to 2, k equal to 3 etcetera. So, basically we are able to define higher order non central moments. So, if you look at expectation of x, in what way it is a moment.

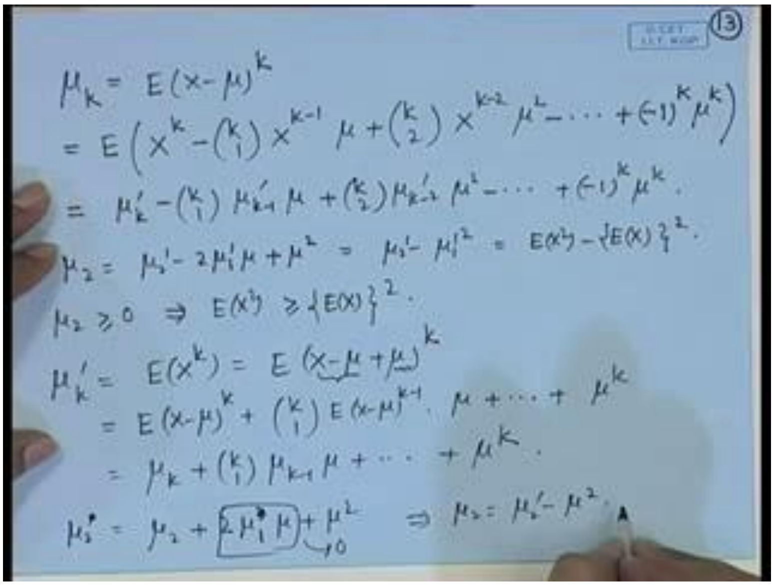
So, if I consider say a weightless bar with say weights attached here at the point say x 1 we attached where p x 1, at x 2 point we attached the weight p x 2, at the point say x n we attached the weight p x n. Suppose it is a attached to 2 in (Refer Time: 20:48) with 0 fiction and if we consider the balance point at the point of equilibrium or the center of gravity that will be sigma x i p x i that is the moment of the first moment of this. In a similar way if we consider a metallic bar with the density say f x at the point x.

Suppose this is point a this is point b, then if we consider a to b x, f x, d x then this is denoting the balance point of this or the center of gravity of this bar. So, that is why this mu 1 prime is actually the first moment about the origin of the random variable; we also define mu k that is equal to expectation of X minus mu 1 prime to the power k are expectation of x minus mu to the power k. The first one we can usually denote by mu the arithmetic mean, this is called kth central moment; in particular mu 2 is called variance of the distribution.

Let us look at the significance of this mu 1 prime and mu 2 in particular. So, mu 1 prime as I mentioned it is denoting the measure of central tendency or the center of gravity or the point of equilibrium of the distribution, we may also be interested in knowing how the values of the random variable are varying with respect to it is mean. To gets it is measure 1 may look at the values of x i minus mu. Now if you take the average value of x i minus mu or x minus mu then expectation of x minus mu is expectation X minus mu which is actually 0. So, this does not give you any information, this is basically because of the fact that the plus deviations and the minus deviations from the mean they cancel out each other.

So, mu 1 is actually 0; however, to get a better measure of variable t 1 may look at the squared differences. So, we look at x i minus mu square and then we take the average that is known as the variance of the distribution. We also define a quantity called a standard deviation that is equal to a square root of variance of x. So, this gives a measure of the variability of the distribution of the random variable. It is obvious that if we are considering k to be positive integer integral values, then there will be a relation between mu k and mu k prime, which is expressed as a follows.

(Refer Slide Time: 24:22)

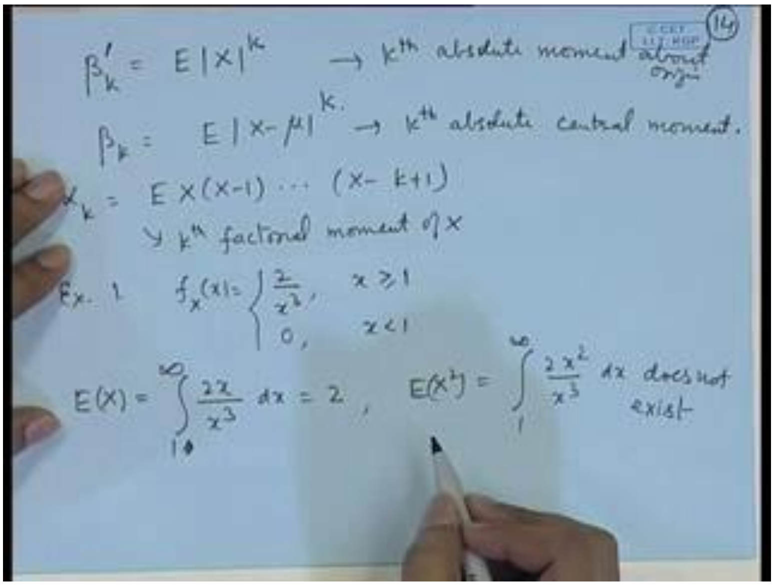


So, if we consider say mu k is equal to expectation of x minus mu to the power k. So, using the binomial expansion this becomes X to the power k minus k c 1, X to the power k minus 1 mu plus k c 2, X to the power k minus 2 mu square minus and so on plus minus 1 to the power, so the first term is k c 1. So, k c k and you will have minus 1 to the power k, mu to the power k. So, if you look at this it becomes mu k prime minus k c 1, mu k minus 1 prime into mu plus k c 2 mu k minus 2 prime mu square. In particular we can write say mu 2; mu 2 is equal to mu 2 prime minus 2 mu 1 prime mu plus mu square which is equal to mu 2 prime minus mu 1 prime square or expectation x square minus expectation x whole square.

We may also observe one thing here that, since mu 2 is greater than or equal to 0 this implies that expectation of x square is always greater than or equal to expectation x whole square; we also have an relationship between non central moments and central moments in the reverse direction; that means, we may interpret mu k prime as, expectation of x minus mu plus mu to the power k; here we consider this as one term and this as another term so this becomes expansion of plus k c 1, expectation of x minus mu to the power k minus 1 mu and so on that is equal to mu k plus k c 1, mu k minus 1, mu plus mu to the power k.

In particular mu 2 prime is equal to mu 2 plus 2 mu square mu 1. So, 2 mu 1 and mu plus mu square, now this term is actually 0. So, this means that mu 2 is equal to mu 2 prime minus mu square.

(Refer Slide Time: 28:32)



We also define absolute moment; so kth absolute moment of is defined as expectation of modulus X to the power k, we also consider say beta k is equal to expectation of modulus X minus mu to the power k, we also define what is known as some factorial moments expectation of X into X minus 1 up to X minus k plus 1 this is called kth factorial moment of X; this is kth absolute moment about origin and this is kth absolute central moment.

In all the definitions of the moments the basic thing is that these expectations must exist for example, expectation of modulus x to the power k must exist. Expectation of modulus x minus mu to the power k must exist; that means, the corresponding integrals are the summations must be absolutely convergent; in some cases a lower order moment may exist a higher order moment may not exist. Let us take 1 example if I consider f x is equal to say 2 by x cube, for x greater than or equal to 1 and 0 for x less than 1; if we consider expectation of x then it is equal to integral 1 to infinity d x, that is equal to 2; if we consider expectation of x square then that is equal to 2 x square by x cube 1 to infinity d x clearly this is divergent.

So, a lower order moment may exist, but a higher order moment may not exist. In the next classes we will define some further characteristics of distributions, basically this moments or other characteristics, actually like mean or variance they explain the nature of the random variables values, how the random variable is taking values over the range of the values with what probabilities. So, we will look into this thing in the next lecture.

Thank you.