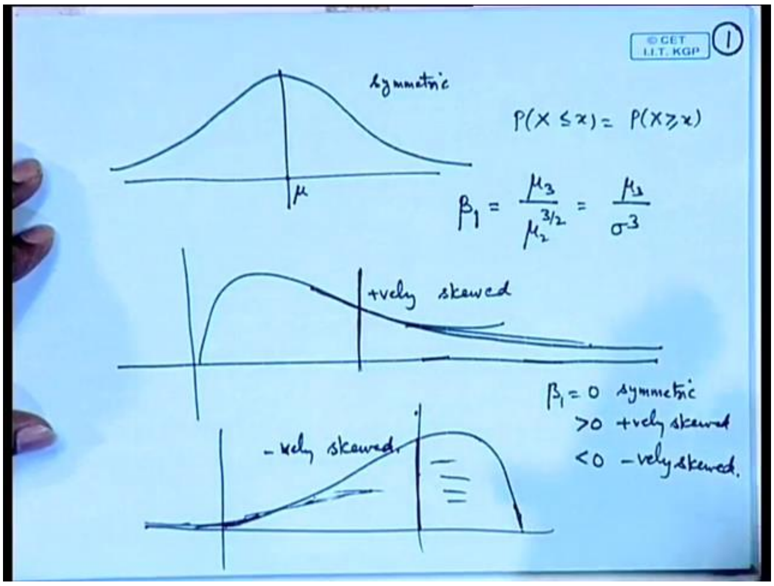
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 17
Characteristics of Distributions – I

In the last lecture I have introduced certain characteristics of the probability distributions, such as its expected value that is the mean, variance and also some higher order moments. So, the mean of the variance distribution denotes the measure of central tendency are the measure of the location for a distribution, the variance are the standard deviation denote they tell about the variability of the values of the distribution.

We may also be interested in some further characteristics of the probability distribution such as a say is skewness; let us define what is known as skewness.

(Refer Slide Time: 01:01)

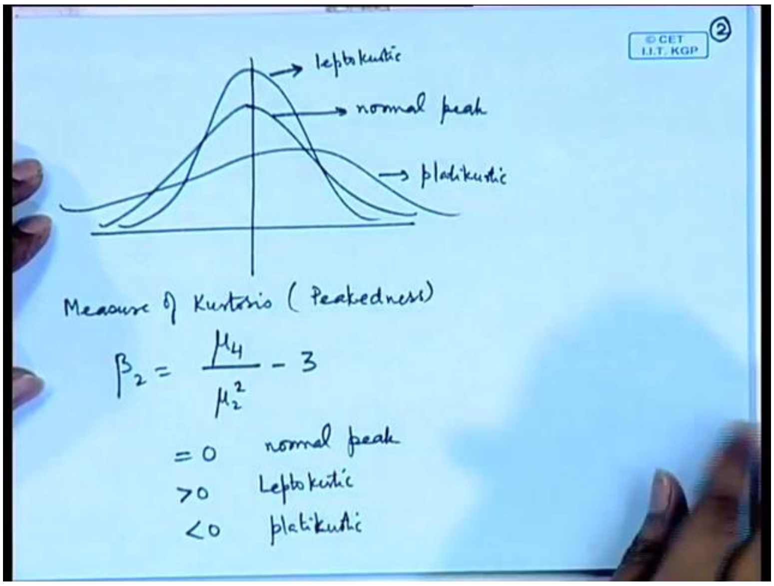


So, consider a distribution of this type, let us consider another distribution and let us consider. Now if we compare the shape of the curves, the first reaction after looking at the first curve that is symmetric about a certain axis say mu. If you look at the second curve then there is lot of concentration of probability on the left hand side and there is a long tail on the right side. That means there is a long tail to the right of the mean of the distribution. Whereas, if we look at the third curve here then there is a long tail to the left; that means, there are more concentration of values on the right side and there is a large variation towards the left of the mean.

So, we will call this as a symmetric curve; we considered the definition of symmetric distribution earlier that is probability that X is less than or equal to x is equal to probability X greater than or equal to x for a certain; suppose if it is a symmetric about 0 then we should have this kind of thing. So, if it is not symmetric we will call it is skewed. So, this one will be called positively skewed and this one we will call as negatively a skewed distribution. A measure for this can be defined in terms of say let me call it beta 1 that is equal to mu 3 divided by mu 2 to the power 3 by 2; we are consider this division by mu 2 to the power 3 by 2 that is sigma cube, where sigma denotes the standard deviation of a distribution, this is to make it free from the units of measurement.

So, if beta 1 is 0 we have symmetric; if it is greater than 0 it is positively skewed, if it is less than 0 it is negatively is skewed.

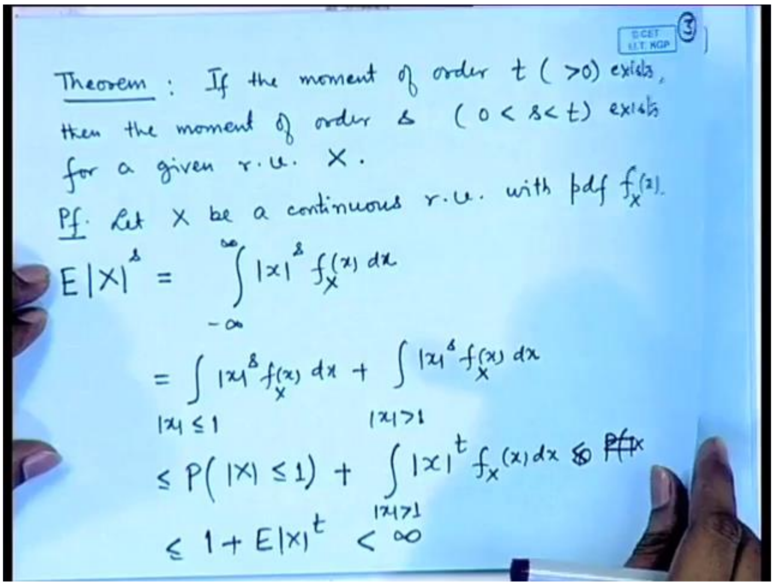
(Refer Slide Time: 03:38)



We also defined another characteristic called peakedness of a distribution, compare the curves. So, if we look at it this as a high peak, this is somewhat in the middle or average or normal and this is more flat curve; we consider this property as the kurtosis. So, this we call as a normal peak, this is called leptokurtic that is high peak and this is called platikurtic that is the flat peak; a measure of kurtosis are peakedness is defined to be beta 2 is equal to mu 4 by mu 2 square minus 3.

So, if it is 0 we have a normal peak, if it is greater than 0 it is leptokurtic and if it is less than 0 we call it platikurtic. The peak of a normal distribution which will be defined later on that will have the coefficient beta 2 is equal to 0. So, peak of any distribution is actually compared with the peak of a normal distribution. Now we have already seen that sometimes moments of the distributions may not exists; are a lower order moment may exists, but higher order moment may not exists.

(Refer Slide Time: 05:27)



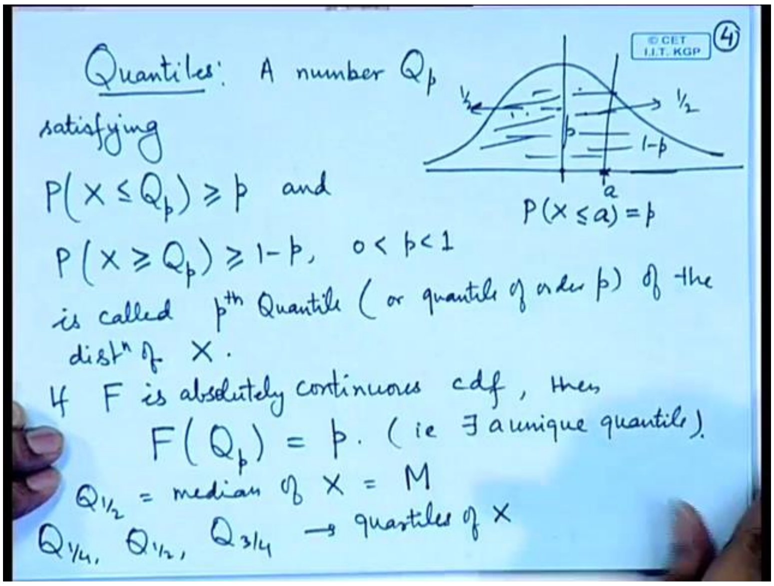
We have a general result in this direction, if the moment of order say t; where t is greater than 0 exists then the moment of order s where 0 is less than s is less than t exists for a given random variable X.

So, if a positive order moment exists, then all lower order positive moments will exists for the given random variable; a let us look at the proof of this one, for convenience let me take X to be a continuous random variable; let X be a continues random variable with say probability density function f x, let us write down expectation of modulus X to the power s, this is equal to integral minus infinity to infinity modulus x to the power s f x dx this one we split into two regions modulus x less than or equal to 1 and modulus x greater than 1, in this region where modulus x is less than or equal to 1, I can replace this by 1. So, this is less than or equal to integral of modulus x is than or equal to 1 f x dx, which is nothing, but the probability that modulus X is less than or equal to 1; when modulus X is greater than 1 if I replace this power s by power t, I will get a bigger quantity. So, this becomes modulus x to the power t f x dx.

Now, this is less than the expectation of modulus. So, the first terms itself this is less than or equal to 1 and this is less than expectation of modulus X to the power t, since we are assuming that the moment of order t exists this is finite and therefore, expectation of modulus X to the power s is finite; that means, the moment of order s exists, this is the condition for existence of the moment of order s. Now sometimes when the moments do not exists, then it may be difficult to find out the measures of the central tendency or measurer of location or measure of variability or say measure of symmetry or kurtosis etcetera.

So, we may look at the points on the distribution itself, which divide the curve into certain reasons with certain proportions these are called Quantiles of the distribution.

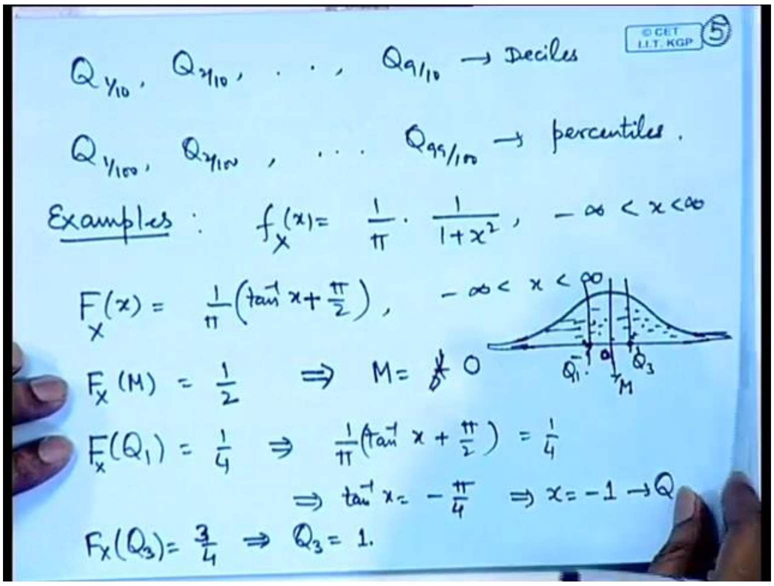
(Refer Slide Time: 08:52)



To explain the concept let us consider some distribution with a particular shape, suppose I have a point here let us call it is a and if I say probability that X is less than or equal to a is equal to p; that means, this weight age P and this weight age is 1 minus P, then this a is called the pth quantile, it is easy to explain the concept of median for example, that it divides the distribution into 2 parts the probability is half in this portion the probability is half in this portion.

So, roughly speaking a pth quantile is the point up to which the probability of random variable taking a value is P and the probability beyond that is 1 minus p; however, to take care of the discreet distributions, we give the formal definition of a quantile as follows. A number let me call it Q p satisfying, probability X less than or equal to Q p greater than or equal to P and probability X greater than or equal to Q p, greater than or equal to 1 minus P for 0 less than P less than 1 is called Pth Quantile or quantile of order P of the distribution of X so; obviously, if F is absolutely continues distribution function then you will have f of Q p is equal to P; that is there will be a unique quantile. So, Q half is called median of X we use a notation M, also Q 1 by 4 Q half and Q 3 by 4 these are called quartiles of X.

(Refer Slide Time: 11:53)



We also defined say Q 1 by 10, Q 2 by 10 and so on Q 9 by 10 etcetera these are called deciles, you can write Q 1 by hundred, Q 2 by 100 and so on these are called percentiles; that means, if you want to divide the distribution into 10 parts, the distribution into 4 parts, the distribution into 100 part etcetera so they are having different notations. In various problems we are interested in different kind of qunatiles for example, various studies we may be interested in percentage of the people leaving below poverty line etcetera.

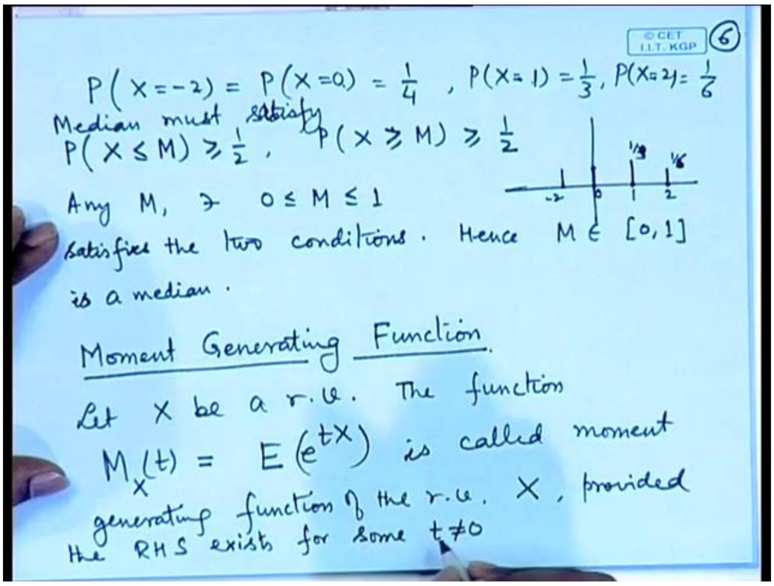
So, this is some particular percentile, but suppose we say 25 percent people lie below a certain thing are 75 percent of the items are above something then it becomes quartiles.

Lets us explain through certain examples; let us considered f x is equal to 1 by pi, 1 by 1 plus x square minus infinity less than x less than infinity. We have seen that for this distribution the mean does not exists therefore, there is no question of higher order moments also existing. However, if we look at f x then it is equal to 1 by pi tan inverse x minus pi by 2 and if we calculate this is plus pi by 2.

So, if we calculate say the point F x M is equal to half, then this is correspondent to simply X or M is equal to half M is equal to 0, which is clear if we plot this distribution so it is symmetric about the point 0. So, if the distribution is symmetric about a given point, then that point will be actually the median of the distribution and we can also calculate the quartiles here, suppose we look at F x Q 1 is equal to 1 by 4, then this gives 1 by pi tan inverse x plus pi by 2 is equal to 1 by 4 this means tan inverse x is equal to minus pi by 4, that is x is equal to minus 1; similarly if I calculate. So, this is Q 1 that is the first quartile in this distribution is minus 1, second quartile that is median is 0, and similar way if I look at f x Q 3 is equal to 1 by 4 3 by 4 then this will be give me Q 3 is equal to plus 1 .

So, we are able to determine the measures on the curve. So, it roughly tells that 25 percent of the observations lie below minus 1, 25 percent of the observations lie between minus 1 and 0, 25 percent of the observations lie between 0 and 1, and 25 percent of the observations lie beyond 1. So, it has a very long tail because in fact, 50 percent of the probability is between minus 1 to 1and rest, 50 percent is disbursed over minus infinity to minus 1 and 1 to infinity.

(Refer Slide Time: 16:01)



Let us considers a probability X is equal to minus 2, and probability X is equal 0 is say 1 by 4, probability say X is equal to 1 minus is equal to say 1 by 3, and probability X is equal to 2 is suppose 1 by 6. So, this is the discrete distribution concentrated on 4 points minus 2, 0, 1 and 2.

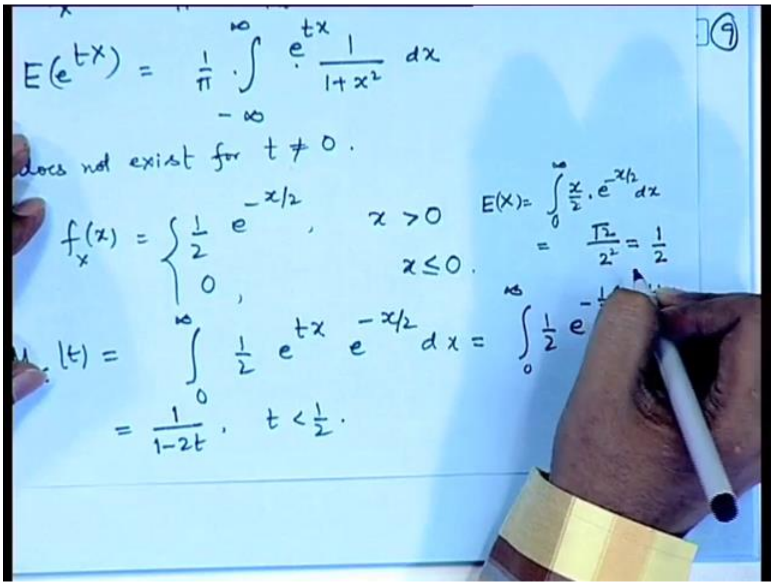
So, if we apply the definition of the median probability X less than or equal to M is greater than or equal to half and probability X greater than or equal to M is also greater than or equal to half, median must satisfy this condition. So, you look at which points satisfy this condition; now here the probability of X being less than or equal to 0 is half, because probability of X equal to minus 2 and probability of X equal to 0 both are 1 by 4. So, as soon as we approach 0, if we look at the up to at minus 2 you have 1 by 4 at 0 you have 1 by 4, at 0 you have 1 by 4, at 1 you have 1 by 3, and at 2 you have 1 by 6. So, any point after 0 this will have the condition probability X less than or equal to M greater than or equal to half satisfied.

If we look at the second condition; here the probability is 1 by 3, here the probably is 1 by 6 and 1 by 3, if you add this becomes half. So, if I consider M to be any point before 1, then probability that X greater than or equal to M is greater than or equal to half is satisfied; in fact if I considered probability X greater than or equal to 1, then it is equal to probability X plus X is equal to 1 plus probability X is equal to 2 which is equal to half.

So, any point which is less than or equal to 1 will satisfy the second condition, any point which is greater than or equal to 0 will satisfy the first condition.

So, any m such that 0 less than or equal to m less than or equal to 1 satisfies the 2 conditions, hence m belong into 0 to 1 is a median. So, this is a case where the median is not unique. So, particular in discrete distributions we may not have a unique quantile, in continues random variable case, there will be a unique quantile, there is another function called moment generating function which tells something about the distribution. So, let us consider that let X be a random variable the function M x at t is find to be expectation of E to the power t x is called moment generating function of the random variable X provided it exists, provided the right hand side exists for some t not equal to 0. As you see here at t is equal to 0 this will always exists. So, for t not equal to 0 it should exists; that means, in a neighborhood of the origin if it exist, then we say that the moment generating function is well defined we may have a case that moment generating function may not exist let us consider.

(Refer Slide Time: 21:06)

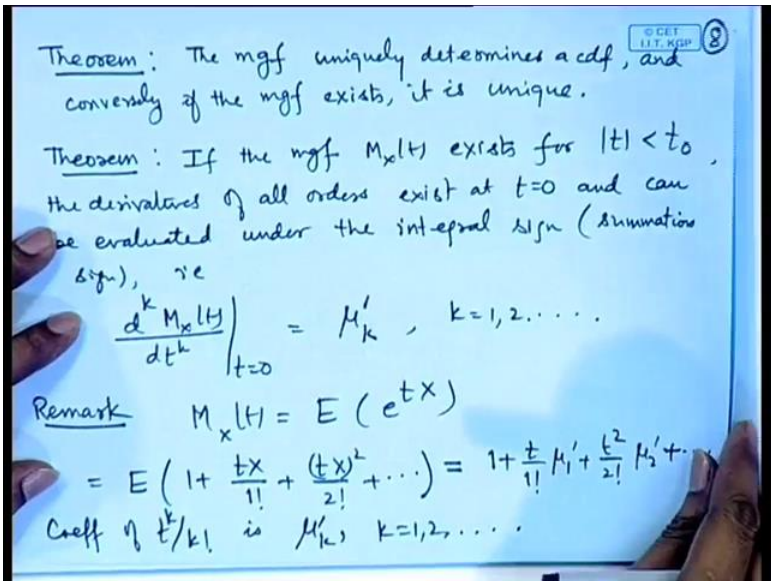


Suppose you take f x is equal to 1 by pi, 1 by 1 plus x square. So, if I look at expectation of E to the power t x, then it is equal to minus infinity to infinity 1 by pi, 1 by 1 plus x square, E to the power tx dx.

So, if you look at this one then this does not always exists, because in the numerator you have a exponential term and in the denominator you have only polynomial. In fact, we have seen that the mean itself does not exist; that means, if I put here x in place of e to the power tx that itself does not, exist for t not equal to 0. Let us take another example say f x is equal to half, e to the power x by 2 for a x greater than 0, 0 for x less than or equal to 0.

Let us consider M x t. So, it is equal to integral 0 to infinity half, e to the power tx, e to power minus x by 2 dx. Now this event combine, so it becomes 0 to infinity, half E to the power minus half 1 minus 2 t, x, dx that is equal to 1 by 1 minus 2 t for t less than half. So, here the moment generating function exists in a neighborhood of 0. The points that why we are interested in a function called moment generating function is that, it gives first of all it uniquely determines a distribution also it gives lot of information about the moments that is why the name moment generating function is there let us look at that thing.

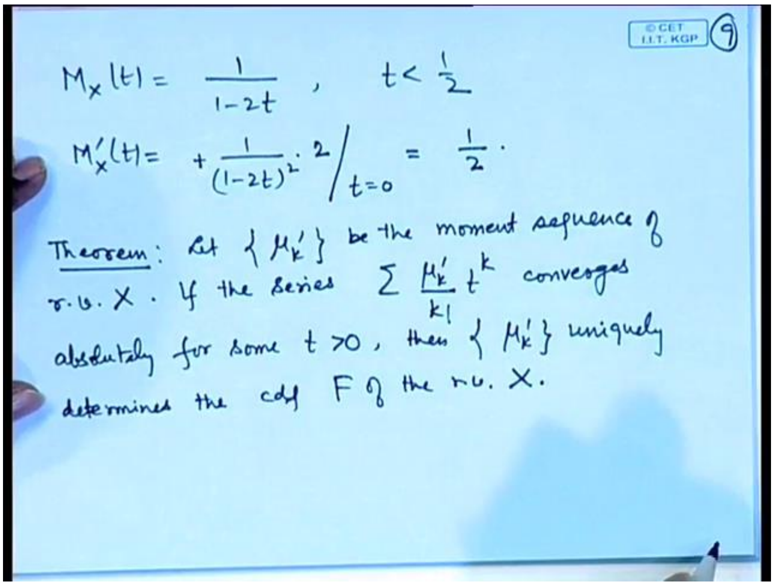
(Refer Slide Time: 23:36)



So, we have the following results, the moment generating function uniquely determines a cdf, and conversely if the moment generating function exists it is unique. If the moment generating function M x t exists for modulus t less than t naught; that means, in a neighborhood of 0, the derivatives of all orders exist at t equal to 0 and can be evaluated under the integral sign integral sign or you can say summation sign depending upon whether the discrete or continuous distribution is there.

So, derivative of the moment generating function of order k at t equal to 0 gives you the kth non central moment, that is why it is known as a moment generating function, you can see this fact if I say that the moment generating function exists then I can consider the expansion of e to the power tx in a maclaurin series as 1 plus tx by 1 factorial tx square by 2 factorial and so on; this is equal to 1 plus t by 1 factorial mu 1 prime, t square by 2 factorial mu 2 prime etcetera; that means, coefficient of t to the power k by k factorial is the kth order non central moment for k equal to 1 2 and so on.

(Refer Slide Time: 26:27)



Let us consider this example your M xt is equal to 1 min 1 divided by 1 minus 2 t for t less than half; let us consider derivative of this so that is equal to minus 1 by 1 minus 2 t square then minus so it becomes plus, and then you are multiplying 2, you put t equal to 0 here, that is equal to 1 by 4 and 2 it is half; let us check here directly from the distribution, if I calculate the expectation of x, this is equal to integral x by 2 e to the power minus x by 2 dx 0 to infinity. So, if we integrate this by parts are if we use the gamma function, then it is gamma 2 divided by multi this is 2 here, so you will get 2 that is equal to 1 by 2 square and then this will be 2 here 1 by 1 minus 2 t whole square and then you have 2 here.

Let mu k prime be the moment sequence of random variable X, if the series sigma mu k prime by k factorial t to the power k converges absolutely for some t greater than 0 then mu k prime uniquely determines the cdf F of the random variable X that is all in today's lecture.