Statistical Methods for Scientists and Engineers

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

Lecture – 04
Moments and Special Distributions

We have noticed that we can classify the random variables as discrete random variable as a continuous random variable or a mixture of the 2. Accordingly, if we have a discrete random variable the probability distribution of discrete random variable is described by a probability mass function. The probability distribution of a continuous random variable is described by a probability density function.

And for mixed random variable for the portion where we have the masses we have a probability mass function and for the portion where the distribution is located over an interval we describe it by a density function. We also saw that there is a more general function which is helpful in describing the various kinds of random variables that is called the cumulative distribution function of a random variable.

And we also saw its relations with the probability mass function and the probability density function in case the random variables are totally discrete or totally continuous. If we have the description of the probability mass function or the probability density function, then we can say that we know or we have all the information about random variable. However, there are certain characteristics, which are also helpful in describing a random variable or its properties.

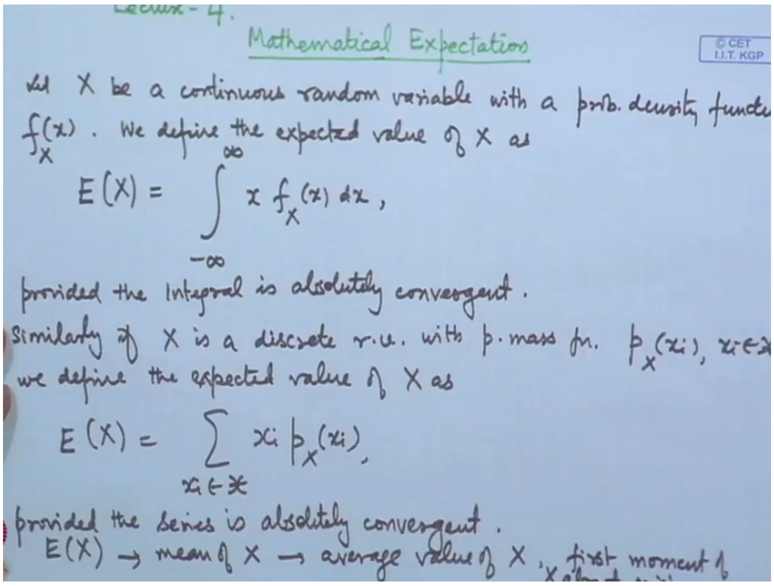
For example, all the time we may not need to describe the full probability distribution of a random variable, but we may be satisfied with certain characteristic. In common sense language we talk about say the mean of a distribution or the average value of a random variable. For example, if we are discussing the heights of individuals in a particular ethnic group, then we may say okay what is the average height?

If we are discussing the longevity or the average age at death, then a common question would be that what is the average age at death for example it may be of interest to the people who are dealing with the insurance companies and they want to fix up a premium for certain type of policies. Whether scientist talks about average rainfall, the average temperatures increasing in the day of global warming we talk about the average yield per hectare for certain crop in a particular year.

So therefore this quantity is like average also we talk about variability that we may say okay average rainfall was the same in the 2 years, but there was lot of variation for example there were lot of rain happened at one time and at another time it was like a drought kind of situation whereas in another place it was a continuous rain in mild amounts so then that means there is more variation.

So in day today language we talk about these terminologies such as average variability or dispersion etc. So we now formally define these characteristics which are helpful in understanding the nature of a random variable and its probability distribution. So we start with the concept of mathematical expectation. Let me define this now.

(Refer Slide Time: 03:50)



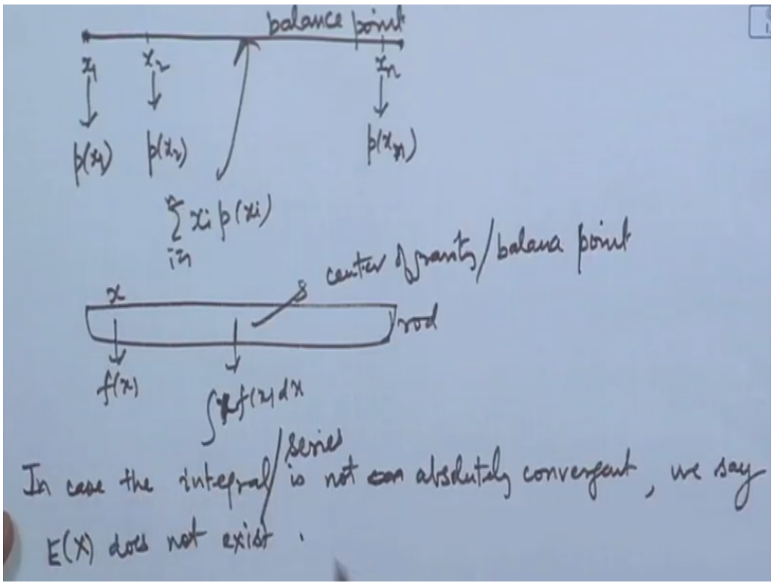
Let x be a continuous random variable with a probability density function say fx. As usual we have been putting a subscript here to denote the random variable capital x and the value taken by that is written by small x. So we define the expected value of x as expectation of x = integral x fx dx from - infinity to infinity, but there is a condition here since it is expressed in terms of an integral we say provided the integral is absolutely convergent.

Similarly, if x is a discrete random variable say with probability mass function let me call it px(xi) where xi belongs to x we define the expected value as expectation of x = sigma xi*px(xi) where xi belongs to x. Once again we impose the condition because this could be in finite series also so we put a condition here provided the series is absolutely convergent. So this expectation x which we have defined this is also called mean of x average value of x etc.

We have several names for this thing. This is also called first moment of x about origin. Now the question is that how do we interpret as scientist or engineers who will use this quantity how to they interpret this why we do call it average. You can see here. Let us look at the discrete case. In the discrete case, we are considering that random variable x takes value xi with probability p(xi).

So this gives that the value is multiplied by the probability and then we are considering the summation of this. Now if you interpret in this way that suppose we have a weightless bar.

(Refer Slide Time: 07:28)



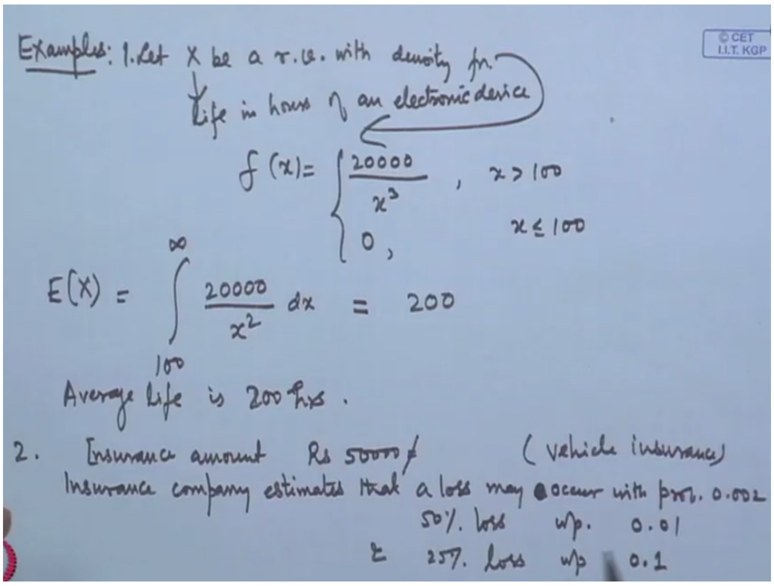
And we have the points here say x1, x2, and so on xn say and at x1 we place a mass px1, at x2 we place a mass px2 and pxn at xn. Then the sigma xipxi I = 1 to n this will be relating to a balance point of this weightless bar. Suppose this is a string which is hung from 2 ends here and then you place the masses. So this is like in mechanics we call it moment so you can have a nice interpretation of this here.

Similarly, in the continuous case we may consider it as a rod and if you consider the rod at x you have the density say fx then again this expectation this xfxdx this will denote the centre of gravity or balance point of this rod. So this has a nice physical interpretation here. In case, the integral is not absolutely convergent or series then we say that expectation does not exist.

Now you will see that this elementary concept is extremely helpful for looking at various day to day problems and interpretations of those things.

So let us consider say one example let me give.

(Refer Slide Time: 09:40)



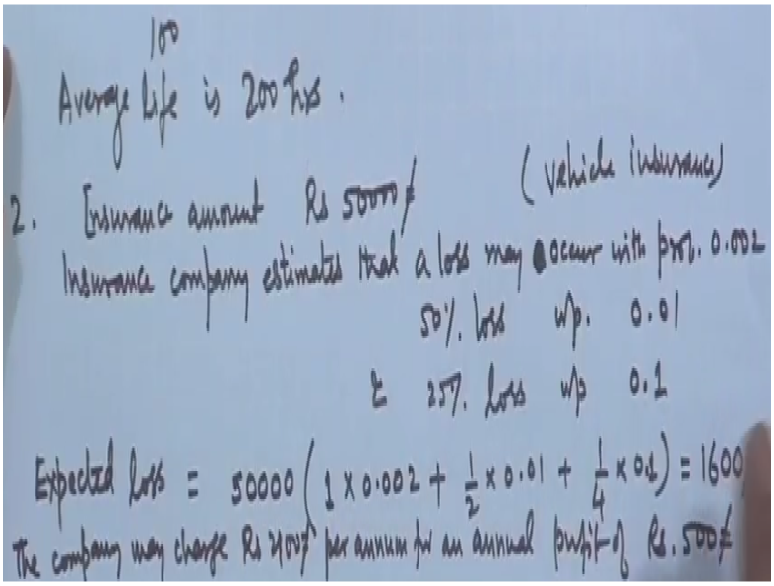
Let x be a random variable with density function so this is a continuous random variable. Let us consider say x is actually denoting life in hours of an electronic device. Suppose the density function is given by say fx = 20000/x cube for x > 100 it is 0 for x < or = 100. I want to find out what is the expected value of this. Now in this case, the range of the distribution is from 100 to infinity because for x < or = 100 the density function is 0 so it is 20000/x square dx.

You can easily see that this is nothing but 200. So that means we conclude that average life is 200 hours. So rather than bothering about the entire distribution how it is varying this is also useful information that we can say that the average life of the device is 200 hours because when we study the physical phenomena it is fine to have a problematic model and we can have full distribution, we can find various probabilities, but a certain person may not be completely interested in the entire thing.

He may be satisfied with the average value. So now you can see here we have average life and we may like to make certain decision based on that. Let me take one example for discrete case also. So let us consider say suppose we want to ensure so insurance amount say rupees 50,000 for certain thing. Now insurance company estimates that a loss may occur with probability say 0.002 complete loss a partial loss say 50% loss with probability say 0.01 and say 25% loss with probability 0.1.

This could be like some vehicle insurance. Suppose it is vehicle insurance so that means there is a 0.002 probability that there will be a complete damage to the vehicle that means it will not be useable, but the probability of that is small. 0.01 probability is that there may be a 50% loss or damage to the vehicle and there is a 0.1% chance that 25% loss will be there. That is the probability of 25% loss is 0.1.

(Refer Slide Time: 13:20)



Now the question is that how much premium the insurance company should charge. So for example that means company should calculate say the expected loss. So what is the expected loss here. So expected loss here will be that is equal to 50,000 then full loss means 1 we will put 0.002 + 50% loss means we will put half 0.01 and 25% loss so 0.1. So this is equal to 1600. So expected loss is actually 1600.

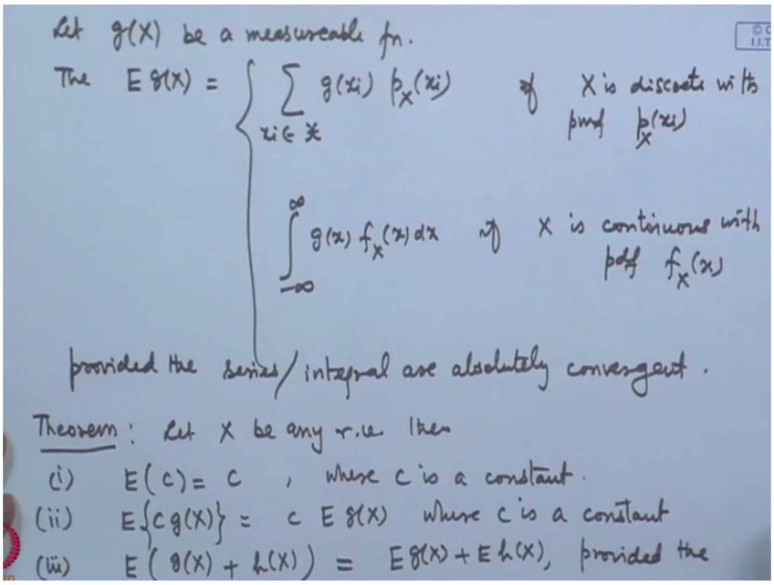
Now suppose the company wants annual profit of 500 rupees on this policy then how much the company should charge the premium so naturally you put 1600+500 that means 2100 should be the charge so if we say the company insurance company may charge rupees 2100 per annum for an annual profit of Rs. 500. So you can see that there is a practical usage of this mathematical expectation here.

So for example I have told here that an insurance company can calculate how much premium it should charge for a policy of a certain amount. So for example if the insurance amount is

50000 and if the losses have been estimated that how much the loss distribution is known then we can see that expected loss is 1600 rupees. Now if the company wants to survive with a profit of 500 rupees per annum then the per annum premium should be 2100.

Now we can extend this concept of expectation. Here we have seen that expectation of x that means whatever random variable originally we are considering for that we can calculate the average value, but many times we may not been interested in that value itself. We may be interested in some function of that. So then we can extend this concept.

(Refer Slide Time: 16:19)

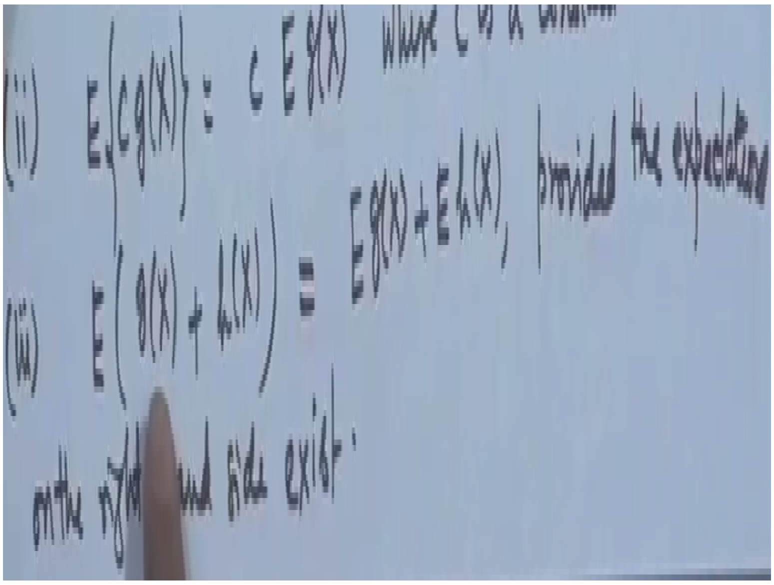


Left gx be a measurable function so in that case we will say the expected value of gx now suppose it is discrete then it will be gx(xi)*px(xi) for xi belonging to x if x is discrete with pmf px(xi) and it is equal to integral gxfxdx if x is continuous with pdf fx. So you have to say provided the series or the integral are absolutely convergent. Now there are certain elementary property is that this expectation will satisfy so I stated in the form of a theorem say which are quite easy to verify.

Let x be any random variable then if I am considering expectation of a constant then it is same constant. You can see the proof of this is extremely simple. If I put c here in place of gx(xi) then it is becoming c times the sum of the probabilities which is equal to 1 or here if I put c then I can out c and I will be integrating the density over the entire range which should give me c again therefore this is = c.

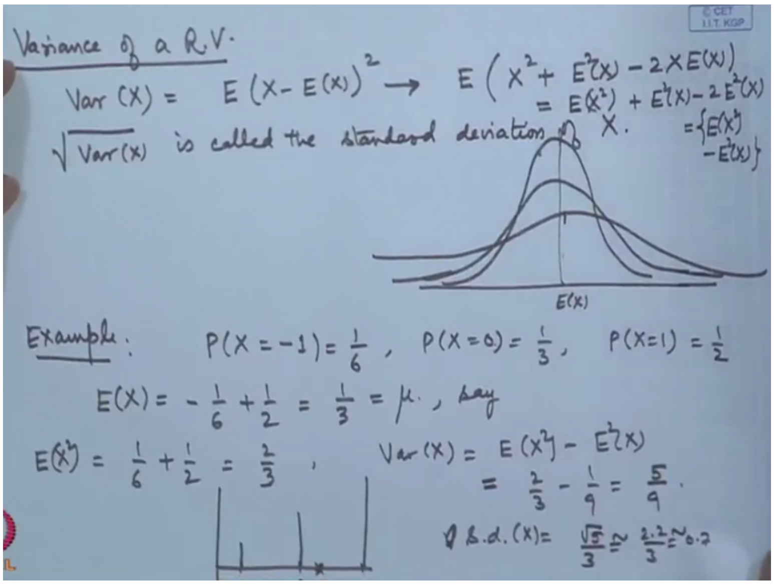
Expectation of c*gx = c* expectation of gx where again c is a constant. Again you can see the proof is extremely simple. Thirdly expectation of gx + some hx where g and h are measurable functions = expectation of gx + expectation of hx once again provided the expectations on the right hand side exist.

(Refer Slide Time: 19:05)



Once again the proof of this will be quite simple. If we substitute here, then we get gx+hx here and if the series are convergent then I can separate out and similarly in the integral part. Now using this we talk about some further characteristics. We define variance of.

(Refer Slide Time: 19:40)



Variance of a random variable. So we define it to be variance of x a popular notation is also sigma square which I will mention later on that is equal to expectation of x - expectation of x whole square and the square root of variance x this is called the standard deviation of x. Now what is the use of this or what is the justification for defining this quantity. Now you can see here see I may have 2 random variables whose balance points may be same, but you can see here this random variable has values with less variability.

There are more values concentrated in the middle whereas this random variable is having move variability. You may even have something like this. So all of them may have the same mean, but the variability may be different. So now what we do suppose this is the expectation value here so we consider the deviation from the mean and then we square it. See one may say that why not take the expectation of this, but then this is equal to 0 because expectation x - expectation x will become 0.

So another option is to take absolute deviation so one may take that that is called the mean deviation about the mean and here we are considering the squared so this is called the variance and a square root of this is called the standard deviation of x. Let me consider one example here for which we already calculated the expectations so let us consider this say we have of course this may have long decimal places so let me take a simpler one.

Suppose I consider x to be a random variable probability x = say -1 = say 1/6 probability x = 0 say 1/3 and probability of x = 1 = 1/2. So for example what is the expectation of this, expectation of this will become - 1/6+0*1/3 + 1/2 = 1/3. Let me use a notation say mu here. Now we can also see this simplification this quantity I can simply expectation of x square + expectation square x - twice x expectation x. Now this I can expand.

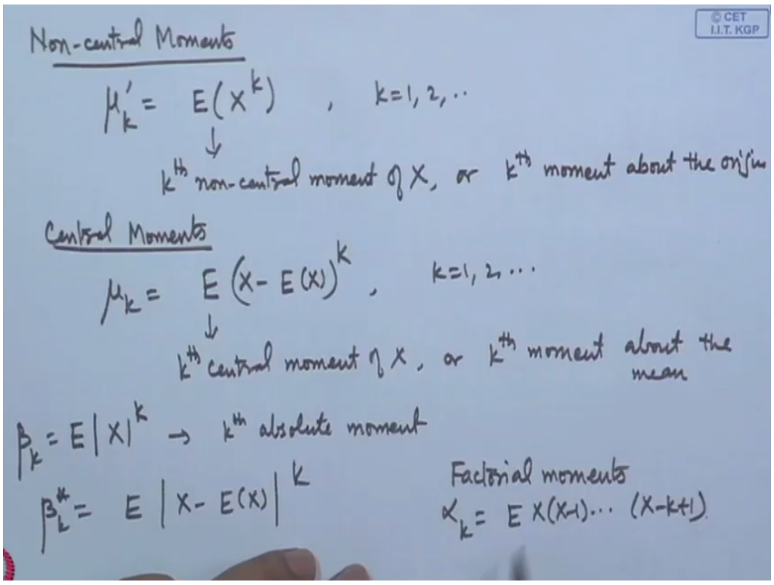
This becomes expectation x square + expectation square x - twice. Now here if I take expectation it becomes expectation x * expectation x that is again so this quantity actually becomes expectation of x square - expectation square x. So this is an alternative computational formula for the variance here. So let me take this. Expectation x square here it is equal to 1/6 + 1/2 = 2/3.

So if I look at the variance of x that is equal to the expectation of x square - expectation square x then = 2/3 - 1/9 so that = 5/9. Now look at this. See this is a discrete distribution and all the values are I have labelled here. so you look at this if I consider the bar chart here for the distribution let me consider this is 0, this is - 1, this is 1 so you are putting the let me draw through a bar diagram so 1/6 is say here, then 1/3 is double of that and 1/2 is triple of that so it is something like this.

So this is the shape of the distribution here and you look at the mean. Mean is 1/3 that is coming here that is the mean because of higher way to the right side the mean has shifted and if you look at the variability that is 5/9 is a variance and if I consider the standard deviation of x = root 5/3 which is < 1 basically because root 5 is 2.2 something. So it is approximately 2.23 that is roughly 0.7 kind of thing that is the variability here.

So by this we should not think that the variability is 1 here while differing by - 1, 0, and 1 because the weights are different and therefore the variability is actually 0.7 here. Now one may easily think that since I calculated expectation x square so that means we can consider expectation of any power of the random variable that gives the concept of the moments so let us consider then non-central moments.

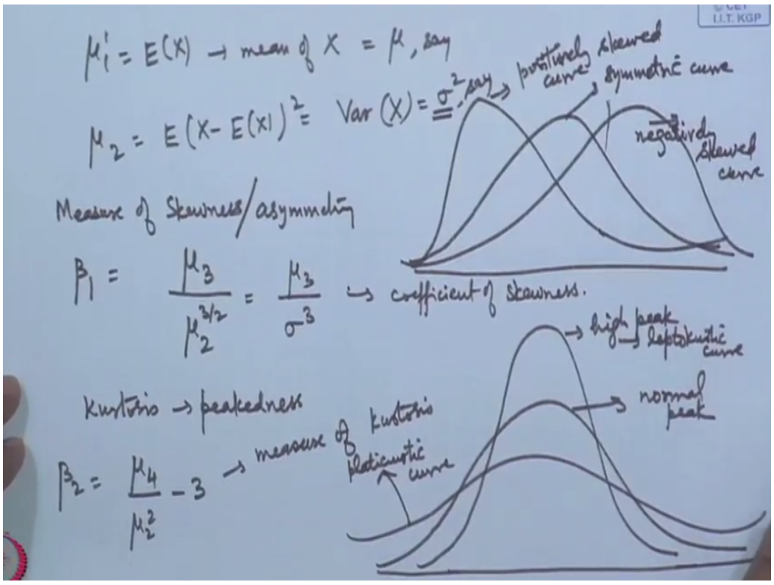
(Refer Slide Time: 25:42)



So we define mu k prime = expectation of x to the power k for any positive integer power of x this is called kth non-central moment of x or it is also called kth moment about the origin and similarly we can define central moments that is defined by muk that is expectation of x - expectation of x to the power k for k = 1 to 1 so on. This is called kth central moment of x or kth moment about the mean. Likewise, we can actually refine absolute moments also like expectation of modulus x to the power k.

Kth absolute moment and again you can define say beta k star and again you can define say beta K star say expectation of x - expectation of x to the power k so these are all various kind of moments one can think of. We can also think of something called factorial moments of course we have to see that what is the use of defining these kind of definition so factorial moments for example we define expectation of x * x - 1 up to x - k + 1. So I will tell you here. See if we consider the mu 1 prime. So what is mu1 prime.

(Refer Slide Time: 27:59)



If I put mu1 prime then this is nothing but expectation x that is the mean, mean of x similarly if I consider mu2 that is expectation of x - expectation of x square that is nothing but the variance of x. So you see that is concept of non-centre and central moment is simply a generalization of the concept of mean and variance and then the natural question that why should we consider higher order.

So as we have seen that we have some use of the characteristics such as mean for example it is telling you about the average similarly the variance or it is say square root tells about the variability in the values. Similarly, you can also talk about the shape of the curve for example you can consider the shapes like this. You may consider a shape like this and you may consider a shape of distribution like this. Now the purpose of drawing this curve is to look at the distinction.

If you look at this one a Layman's view will be that this looks like a symmetric curve. If it is a symmetric curve, the mean is centred and you can say that the values are the probabilities are equally distributed along the on both the sides of the mean. In this one if you see you have lot of weight age to the left side, but on the right side you have a long tail.

For example, if I consider the number of students appearing in a competitive examination for example an engineering entrance examination like JEE examination so the number of students who appear is very large. For example, 1,50,000 students appeared this year in the joint entrance examinations of IITs, but out of that only 10,000 or 15,000 students are declared qualified.

So a large shunk of a student is at the bottom and if you look at the average value that will be somewhere here and then so the average value will be pushed down because there is lot of weight age to the left side for low marks. There are large number of students with very low marks and there will be very few students with high marks which are actually declared as qualified.

So if there is a long tail to the right side then it is called positively skewed curve. This will be called a symmetric curve and on the other hand you have a similar thing here. For example, if you are looking at the lives of say human beings for example so a large number of people they live longer and there will be a certain number of people who die at an early age for example infant deaths or deaths of children below 5 years.

And so on, but there will be a majority of people who live longer say who complete the age of 60, 70, or 80 kind of thing. So this is negatively skewed. So here the average age is pushed up for example in developing countries you see that the average life is 62 years, 63 years. In developed countries the average life is 75 years or 76 years' kind of thing. So average life is pushed, but although you have people in all the age groups for example the people die very early infant deaths, death at births and so on.

So here the mean is pushed to the right, but there is a long tail to the left. This is called a negatively skewed curve. An empirical measure for checking the skewness of the curve is based on the third central moment. So we define a measure of skewness or we can say asymmetry. Let me give some name to it say beta1 that is defined to be based on mu3. Now, one thing that we notice for example if x is the say life in hours.

So expectation will be in the life in the hours. If I take mu2 variance, then it will be squared units and that is why we take square root of that to make it standard deviation so that will become again in the hours. Similarly suppose we are talking about average income, then it may be in rupees, and then the here it will be squared rupees, and if I take square root then it will be again in the rupees.

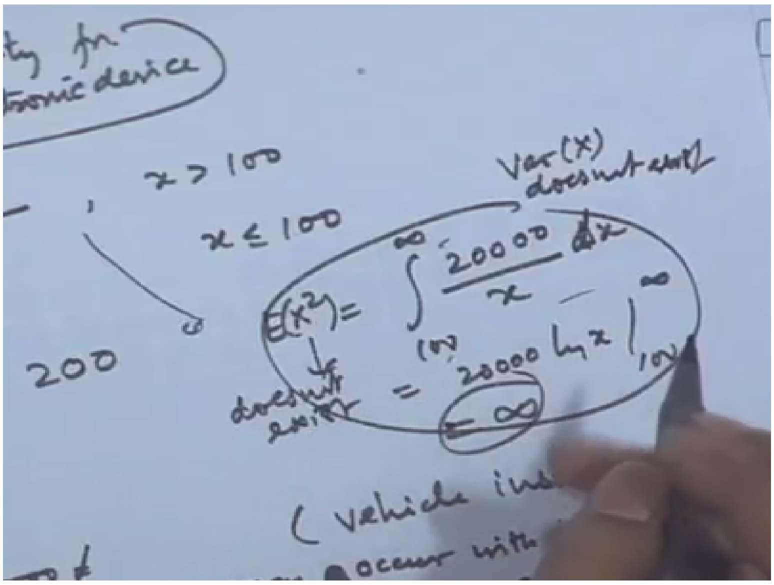
Similarly, if I consider mu3 so in mu3 you will have power 3 therefore it will be in the cubic therefore to make it free from the units of measurement we consider division by mu2 to the power 3/2. A standard notation for expectation x is mu and another standard notation for variance is sigma square. So these are some of the popular notations which are found in the books let me also use them.

So we can also write it as mu3/sigma cube here. So this is called coefficient of or measure of skewness. Let me also talk about another type of variation that may occur in the curve in the shape of the distribution. Once again it looks like some normal type of distribution, normal curve. You may have a curve like this and you may have a curve like this. Apart from the difference in the variability another striking thing could be the height of the curves.

See this curve takes more height. This curve has a flat kind of shape here. Now this is also a property. This is called property of kurtosis that is peakedness. So we say this is a normal peak, this is a high peak and in statistical terminology we call it a leptokurtic curve and this one we call us platykurtic curve. A measure an empirical measure for measuring the kurtosis is based on mu4 so beta 2 we define to be mu4/mu2 square - 3.

So this is a measure of kurtosis. When we do the spatial distribution I will show you that how what they inform about the shapes of the distributions. Apart from the characteristics fo the distribution which are based on the moments we can also base it on the distribution of the probability itself. What is the meaning of that? because sometimes the moments may not exit because we are putting a condition of the absolute convergence we can actually have an example let us say.

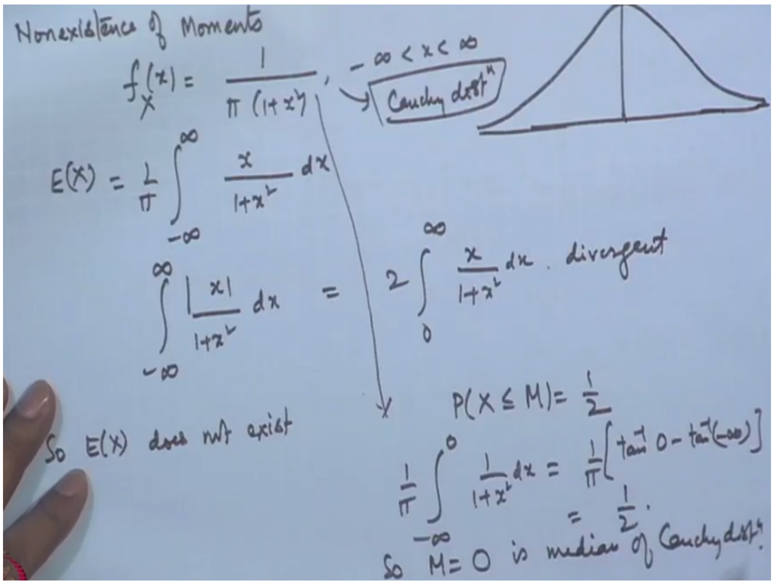
(Refer Slide Time: 36:20)



Non existence of moments. Let us look at this example. Here if I consider say expectation of x square for calculation of variance which I will need then it becomes 20,000/x from 100 to infinity. Now this is equal to 20,000 log x from 100 to infinity. Naturally you can see that this is infinite, this does not exist. So here in this case variance of x does not exist because expectation x square does not exist so expectation.

So variance will also not exist that means this does not exist. That is only there are certain measures which are defined on the basis of the distribution of the probability itself.

(Refer Slide Time: 37:35)

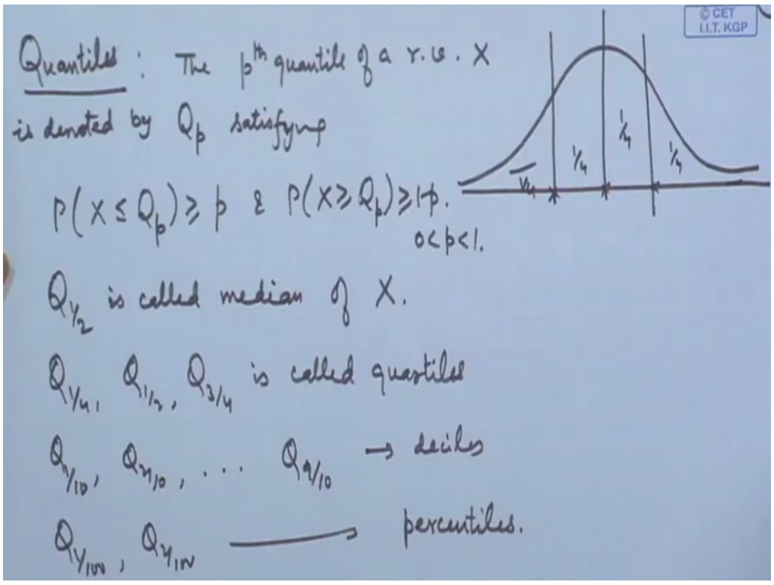


We talk about let us look at see nonexistence of moments another example. Suppose I consider this distribution fx = 1/pi 1+x square from - infinity to infinity. If I plot it, it looks like this at x = 0 it is having value 1/pi and as x tends to + infinity and - infinity it goes to 0.

If I look at say expectation of x here then this is equal to integral x/1 + x square dx, but this integral does not exist because if I consider modulus of this then this is equal to twice o to infinity x/1 + x square which is divergent.

So in this case, also expectation x does not exist. In fact, this is one of the well known distribution this is called a Cauchy distribution we will discuss later about this distribution that how it arises etc and therefore we need certain characteristics which may be useful when the moment structure is not known or it does not exist so they are called quantiles.

(Refer Slide Time: 39:00)



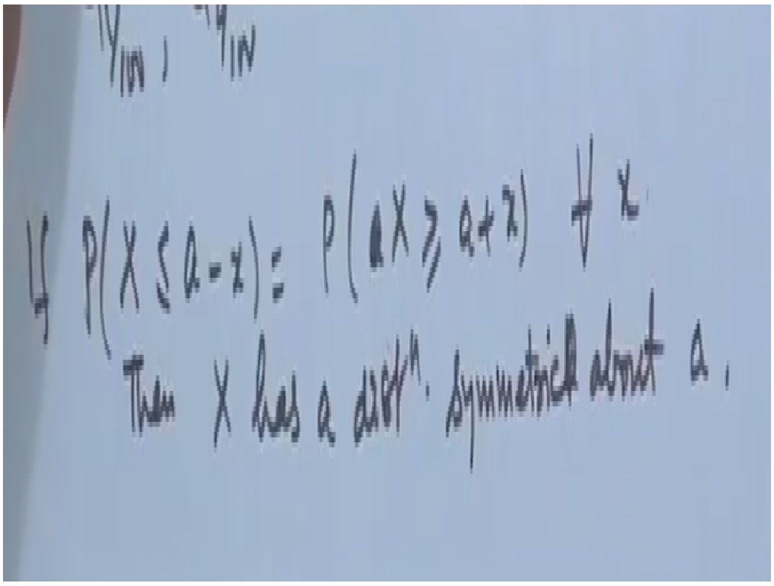
So by quantile we mean that we find out the points on the curve below which you have certain probability for example you may have half probability here, half probability here. You may have a half probability here, half probability here. You may have a probability say 1/4 here, 1/4 here and say 1/4 and 1/4 here. So that means the points which split the area under the curve into certain proportions they are called quantiles.

So we define in a rough way we can simply say that if I am finding out probability half then probability of x < or = that point should be equal to half, but at the same time you have to keep on the track on the right hand side also and to take care of the discrete distributions we define it in a slightly generalized way. The pth quantile of a random variable x is denoted by say Qp and it should satisfy probability x < or = Qp is > or = p and probability of x > or = Qp > or = 1 - p where of course p is a number between 0 and 1.

Q half that is called median. So you can see in this example mean did not exit, but if I look at median then what is the point such that you will have probability equal to half. You can easily see that this is symmetric around 0 so if I integrate from - infinity to 0 1/pi, 1/1 + x square dx then = 1/pi tan inverse 0 - tan inverse - infinity so = 1/2. So here 0 is the median. So m = 0 that is the median of Cauchy distribution.

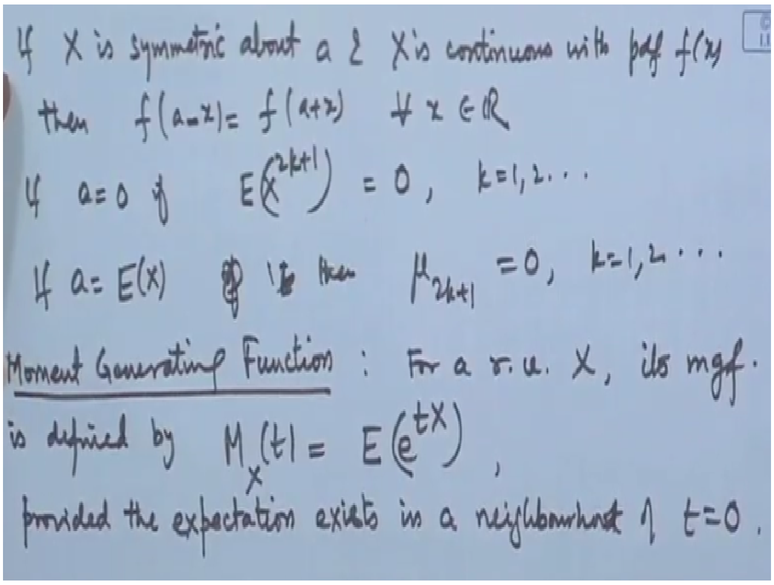
In a similar way, we can consider the points which are dividing into 1/4 etc that means Q 1/4, Q 1/2, Q 3/4 these are called quartiles Q 1/10, Q 2/10, so Q 9/10 etc. These are called deciles and so on. Q 1/100, Q 2/100 these are called percentiles because this will divide the entire probability distribution into 100 parts. We can have a formal definition of symmetry based on this. We will say that a distribution is symmetric about a if the probabilities are equally distributed on both the sides of or on either side of the point a.

(Refer Slide Time: 42:53)



So for example we may have probability of say x < = say a - x = probability of x > or = a + x for all x. Then we say x has a distribution which is symmetric about a. now consequence of the symmetry is that.

(Refer Slide Time: 43:35)

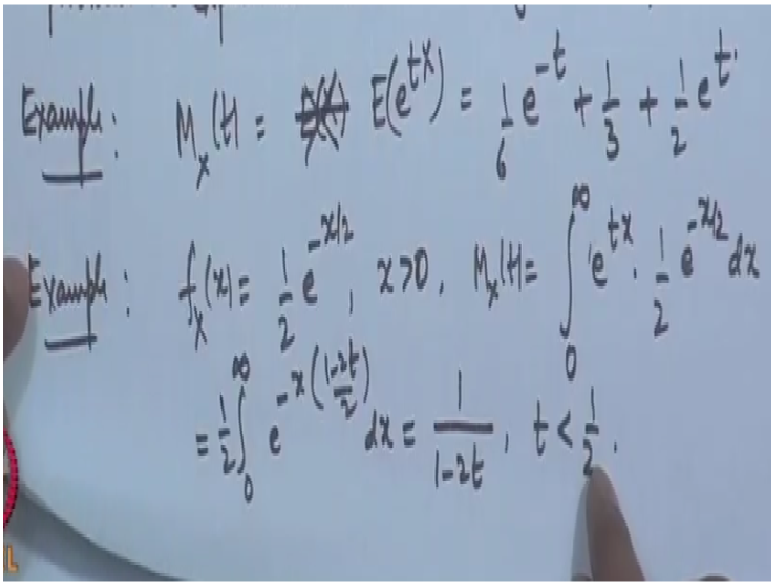


If x is symmetric about a and x is continuous with say pdf say fx then you will have f of a - x = f of a + x for all x. similarly if a = 0 then the odd moments will be 0. If you have a = expectation of x then mu 2k + 1 that is the central moments will be 0. Odd central moments will be 0. I will be discussing some problem later on. Right now let me develop the theory or all the properties of this moments, quantiles etc.

There is a useful function which can generate the moments. For example, here you need to calculate all the moments one by one by looking at the absolute convergence and so on. Now natural question arises if there is a function which will give me the values of all then it may be much better. Fortunately, there exist a function of this nature let me call it moment generating function.

So for a random variable x its moment generating function mgf is defined by mxt = expectation of e to the power tx provided the expectation exits in a neighbourhood of t = 0. You can see actually that t = 0 this will always exist because it will become 1. So we will say that mgf exist only for some non0 value of t it will exist and since it is a continuous function so it exists for some non0 value then there will be an interval in which it will exist. Let us take a very simple example.

(Refer Slide Time: 46:58)

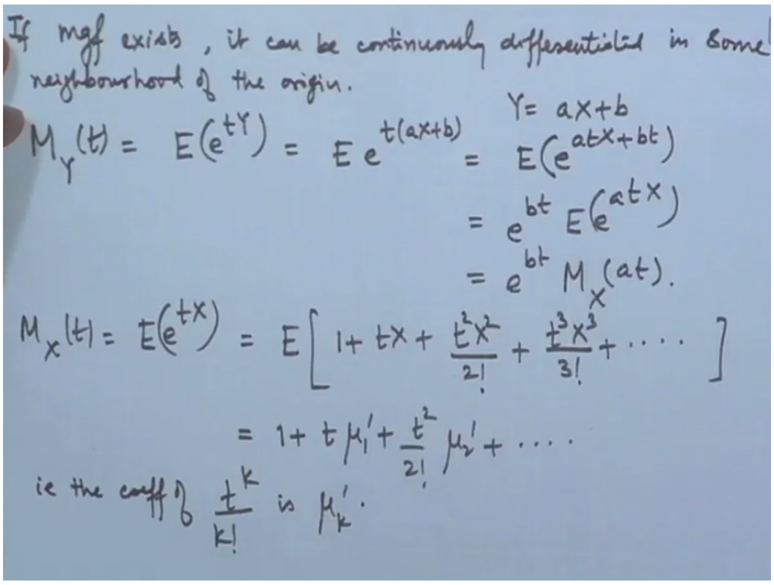


Let us consider this example what is moment generating function so that is equal to expectation of x sorry expectation of e to the power tx. So that is equal to if I put x = - 1 it becomes e to the power - t 1/6. When I put 0 this will become 1 so it is 1/3 + 1/2 e to the power t. so now you can see that this is the moment generating function and certainly the question arises that what are the uses of this.

Let me take another example also. So let us take say fx as a density function say half e to the power - x/2 for x positive. So what is the moment generating function here it will become expectation of e to the power tx that is the e to the power tx multiplied by the density function integrated over the range that is equal to 0 to infinity half e to the power - x 1 - 2 * t/2 dx so = 1/1 - 2 * t this is valid for t < half.

So easily you can see that for a non0 value of t it is there and therefore in an interval it is existing. Now a simple property of this will be that if mgf is existing it will be one can differentiate it any number of times.

(Refer Slide Time: 48:49)

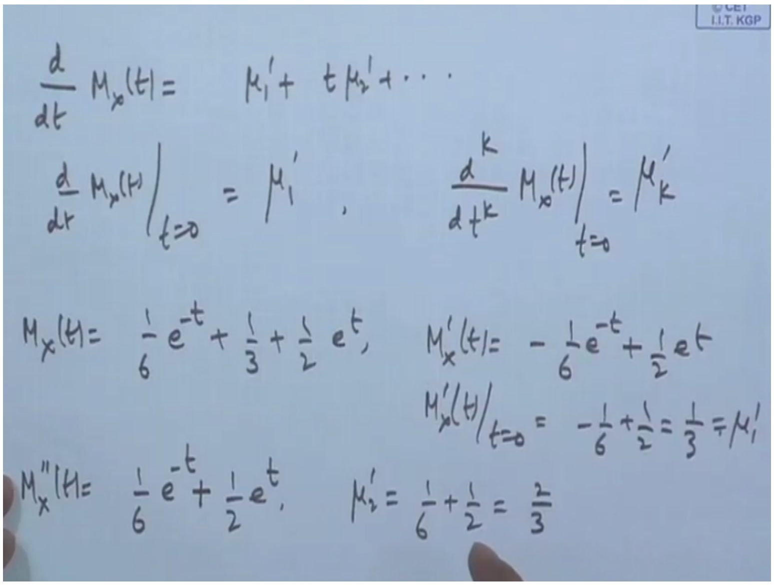


If mgf exists, it can be continuously differentiated in some neighbourhood of the origin in fact in the region of existence basically it will exist. You can also see some simple property suppose I say myt and I define y to be say ax + b. This is equal to expectation of e to the power ty that = expectation of e to the power t ax + b = expectation of atx + bt that we can write us e to the power bt expectation of e to the power atx = e to the power bt * the moment genetic function of x at the point at.

Now the question is that why we have introduced this function what is the use of this. So you can see here we can consider mxt as expectation of e to the power tx and let us consider the expansion of this so it becomes tx+t square x square by 2 factorial + t cube * x cube/3 factorial and so on = now if I assume that the mgf is existing that means this quantity is existing.

If this quantity is existing that means all the term by term expectations will exist that means I will get it as 1 + t expectation x that is mu1 prime + t square by 2 factorial mu prime and so on. That is the coefficient of t to the power k/k factorial is muk prime. Another way of looking at it is if I consider say.

(Refer Slide Time: 51:18)



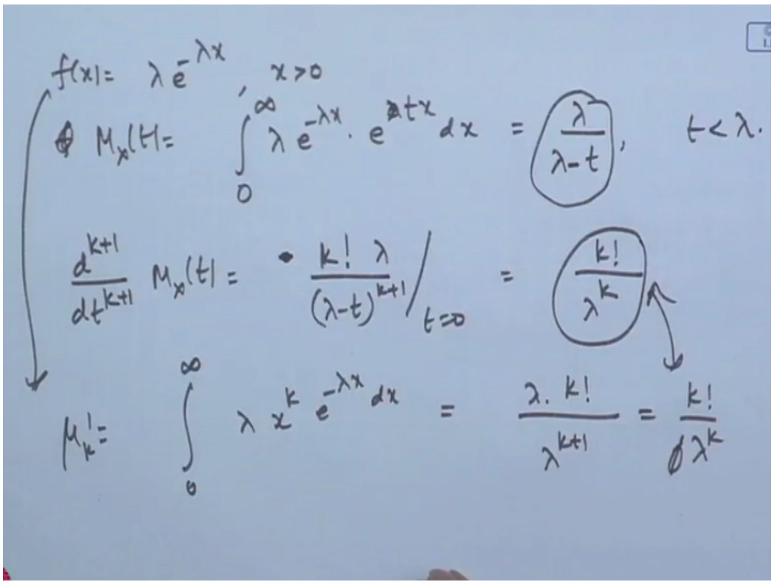
Derivative of this then I will get mu1 prime + t mu2 prime and so on. So if I substitute d/dt mxt at t = 0 then I will get mu1 prime. Similarly, if I consider second derivative then here I will get mu2 prime and this term will contain t so in general I can consider kth order derivative and if I put t = 0 I will get the kth non-central moment. So this is extremely useful result because if mgf is given to us we can find all the moments.

So that is why the name moment genetic function is justified for this. Let us take this example here. So let us look at this one. So if I consider the derivative of this it is 1/6 e to the power – t + 1/3 + 1/2 e to the power t. If I consider mx prime t = - 1/6 e to the power - t, + half e to the power t and if I put t = 0 here I get - 1/6 + half that is equal to 1/3 which was actually the mean of this distribution.

If you look at this, we had the mean here. So this is the mean. Similarly, if I consider say second derivative mx double prime t then I will get 1/6 e to the power - t + half e to the power t. So mu2 prime = the value of this at t = 0 that is 1/6 + half = 2/3 which was the value calculated here.

So you can see that having the moment generating function it is extremely helpful to calculate moments of various orders and since if the function is nice we can actually write it in such a form that the successive differentiation formula can be applied and then we can write the moments of any order. Let us look at say this example.

(Refer Slide Time: 54:01)

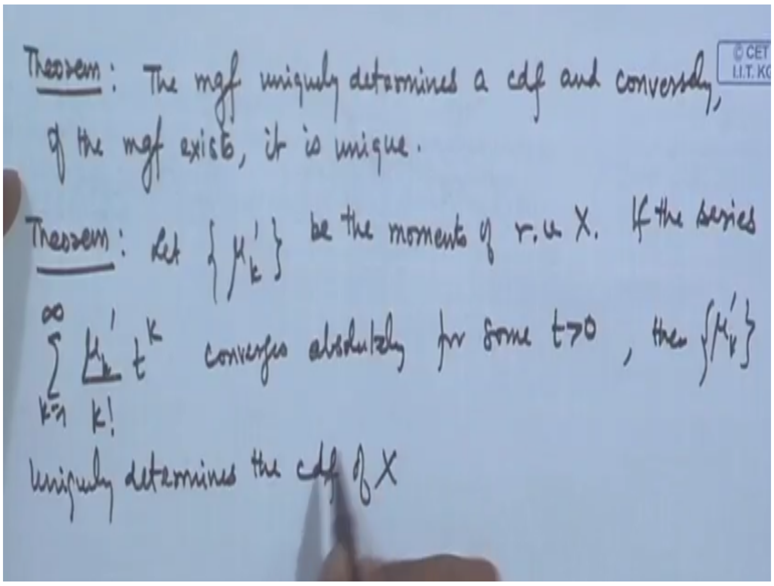


Say fx = say lambda e to the power - lambda x. so what is mxt here. If you look at mxt it is becoming integral lambda e to the power - lambda x, e to the power lambda t, e to the power tx dx from 0 to infinity. So you can easily see that it is convergent for t < lambda. Now you look at this. This is having very nice form. If I consider the k + 1 th order derivative = k factorial because every time if I differentiate there is a minus here so that will become plus.

So it becomes k factorial and then you will have lambda - t to the power k + 1 lambda. So if I put t = 0 then this is becoming k factorial/lambda to the power k which you can verify here if I look at say muk prime directly, then I will get lambda x to the power k e to the power -lambda x dx from 0 to infinity. So this is becoming lambda k factorial/lambda to the power k + 1 using gamma function so it is again = k factorial by lambda to the power k so you can see that these 2 things are same.

So this moment generating function is extremely useful function for determining the moments and therefore all other characteristics such as mean, variance, measures of skewness, and kurtosis etc. One or 2 other important properties of the mgf are there which I will list without proof here that the mgf.

(Refer Slide Time: 56:08)



The moment generating function uniquely determines a distribution so I am writing cumulative distribution function and conversely if the mgf exists it is unique that means no 2 distinct distributions may have the same mgf and therefore there is a one to one correspondence and that is extremely useful because many times what happens that we are considering distribution of certain functions of random variables.

And it may be possible to derive the mgf of that nor if we can assign that mgf to a certain distribution, then we know that that particular function of random variable has that distribution. So in identification of the distribution it is extremely useful. There is another one which is actually the moments convergence theorem but let me also mention this thing. Let mk prime be the moments of random variable x.

If the series sigma muk prime by k factorial t to the power k converges smoothly for some t then muk prime sequence uniquely determines the cdf of x. now this is actually a consequence of this uniqueness moment theorem because actually this is nothing but the if you expand expectation of e to the power tx we have see that it will consists of the towns of the moment sequent s so expecting the first time all other terms are written.

So if it is absolutely convergent that means the mgf will exist and therefore unique determination of the distribution will be there. I mentioned that factorial moment generating function etc, but I will not spend too much time here. We will over the special distribution and in the next lecture I will give the motivation for each distribution which are more commonly used and for those distribution we will look at the moment structure that what it is the mean variance and measures of skewness, kurtosis etc. so in the following lecture I will be taking up this.