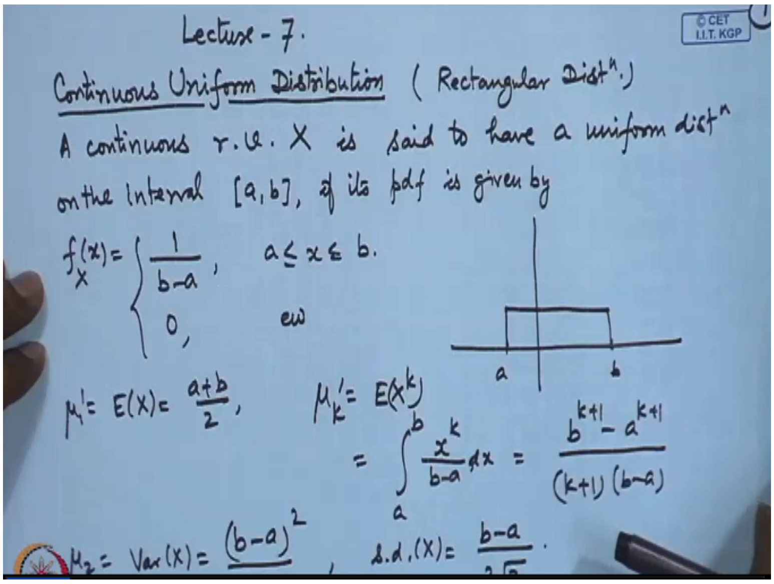
Statistical Methods for Scientist and Engineers
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology – Kharagpur

Lecture - 07
Special Distributions Contd.

In the last lecture, I had started the continuous distributions especially we did exponential distribution and Erlang or Gamma distribution and both of the distribution I showed that they arise as distribution for the waiting time of the incidences in the Poisson process. However today I will introduce some more continuous distribution which may not be related to the Poisson process.

The first one is one of the simplest distribution we call it uniform distribution. Now we have seen the uniform distribution in the case of discrete variable also where we allocate equal probability for each outcomes. Now in the case of continuous distribution if you have constant density over the region over a finite interval then it is called a continuous uniform distribution.

(Refer Slide Time: 01:15)



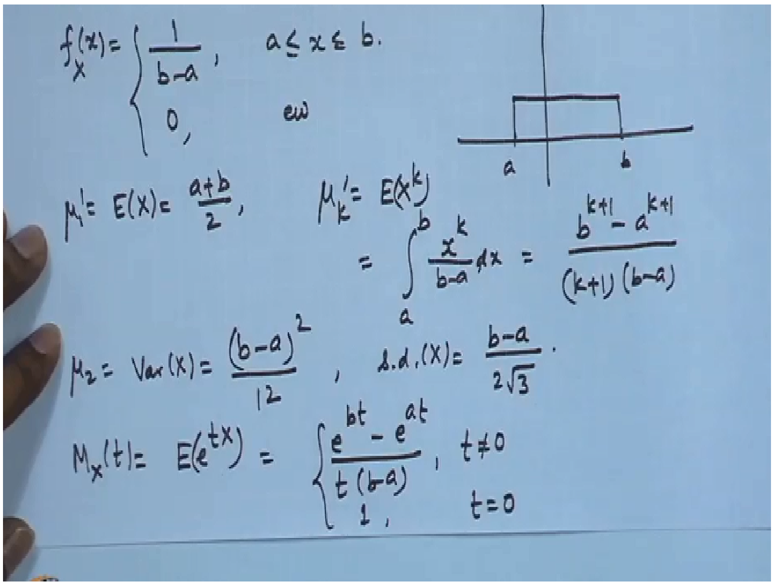
So we can define like this a continuous random variable X is said to have a uniform distribution on the interval say a to b. Now here one may take open interval or closed interval it will not make any difference if its probability density function is given by fx=1/b-a for a<or= equal to x <or = b and it is 0 elsewhere. So if you make a plot of this you can easily see how it will look like.

Suppose I consider a and b here so 1/b-a is this one. So that is why if you plot it looks like a rectangle so that is why there is another name to this distribution it is also called Rectangular Distribution. This type of distribution is applicable in various situations for example you may consider waiting at a traffic crossing when you go on a busy road. So signal waiting time may up to say 3 minutes so you may have to wait say 0 to 3 minutes.

For example, the time spent at telephone booth by the customer and many other applications of this nature can be considered as applications of uniform distribution. Now as you can see since the density is constant the mean or the first moment will be simply the mid value of the interval. One can easily calculate higher order moment also in fact moments of any order can be easily calculated.

I can consider mu k prime that is= expectation of x to the power k that is=integral x to the power k/b-A dx from a to b. So naturally this is= b to the power k+1-a to the power k+1/k+1*b-A. In particular, we can consider mu 2 prime and also the variance that is mu 2 in this case it will turn out to b-a whole square/12 that means the standard deviation will be= b-a/ 2 root 3.

(Refer Slide Time: 04:12)

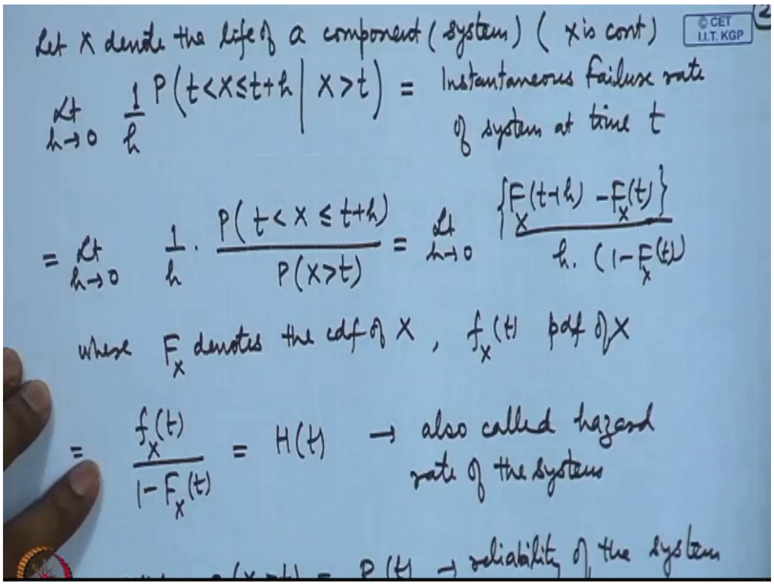


One can look at the moment generating function expectation of e to the power tx so that will be=e to the power bt-e to the power at/t*b-a of course this is for t not equal to 0. If t=0 then naturally it is=1. As I mentioned uniform distribution has uses when the density is assumed to be constant. Let us look at some other useful distribution. Let me introduce one concept here we have considered exponential distribution or gamma distribution.

So for example what is exponential distribution? Exponential distribution we have defined as the distribution of the waiting time for the first occurrence in Poisson process. So it can model various kind of phenomena for example life time of the components in a manufacturing process, life of an electronic system and so on. Now when we consider the life time interpretation of the exponential distribution then one of the important concept is that of a rate or you can say rate of occurrence.

Now in the context of lifetime that rate of occurrence of the incident can be called to be a for example if the system fails so you can call it failure so that means we are interested in failure rate. So we define formally what is a failure rate we connect it with the exponential distribution and in the light of exponential distribution then we like to see what further generalizations of this can be made.

(Refer Slide Time: 05:59)



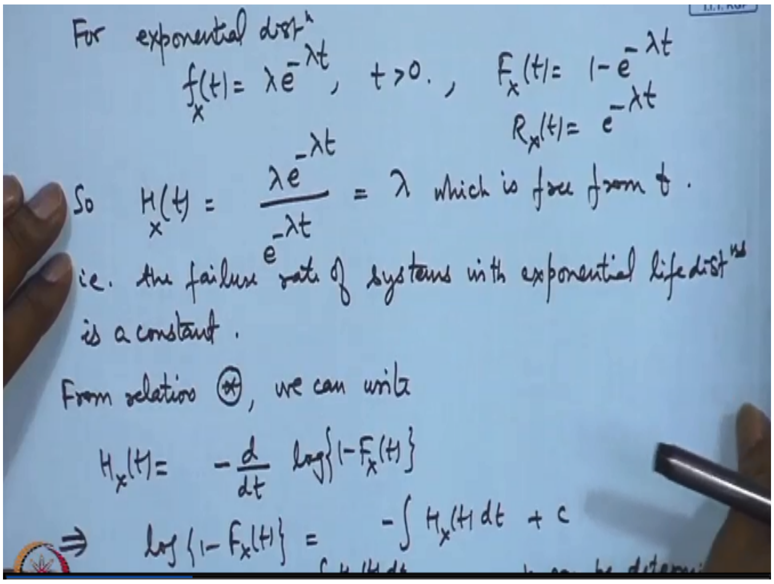
So we consider let us consider a quantity probability of say system is working till time t. So I am considering let X denote the life of a component or system etcetera. Now if the system is working till time t what is the probability that it fails in some time immediately after t that means from t to t+ h where h is a small quantity and if I want to look at the rate then I divide it by h and then I take limit h tends to 0.

Now you can easily give an interpretation to this. The system is working till time t and we assume that it fails immediately after time t and then we divide by the length of the interval in which it fails and then take the limit then this can be called instantaneous failure rate of system at time t. So this is a useful quantity let us evaluate it. So this is= h tending to 0 limit 1/h.

Now if you look at this one this is conditional probability, but in the numerator you have a event which is subset of the event in the conditioning so this will become probability of t<x < or= t +h/ probability of x>t. So this is= limit h tending to 0 and now this is nothing but f of t+ h- f of t/h and this quantity is 1-Fx of t. So x is continuous here okay then we assume capital F denotes the cumulative distribution function of x and let us take small Fx to be pdf of x then this limit as h extends to 0.

This will become the density function of x so this is=fxt/1-Fxt. So this we call failure rate we have some name like Ht this is also called hazard rate of the system. This 1-Fxt that is probability of x>t this is called reliability of the system at time t that means what is the probability of system surviving till time t. Now in the light of this let us firstly look at exponential distribution and then we will look at generalizations.

(Refer Slide Time: 09:48)



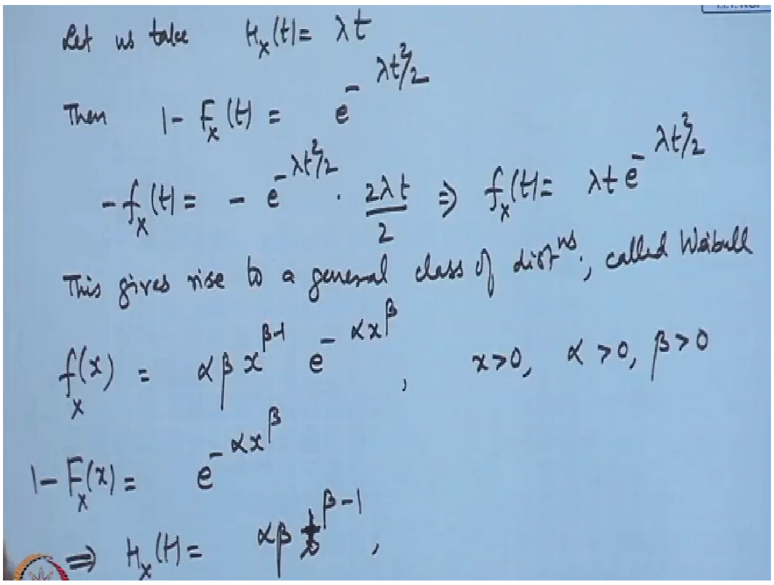
For exponential distribution let us consider one standard model we can consider say fxt=lambda e to the power –lambda t for t positive. So here capital Fxt is nothing but 1-e to the power-lambda t that is Rxt= e to the power-lambda t. So this Ht that will be= lambda e to the power –lambda t/ e to the power –lambda t that is=to lambda. Now you can easily note here that this is free from time t.

That means hazard rate for the system which follows exponential life time are constant. Now this is a very unique property which is free from t that is the failure rate of systems with exponential life distributions is a constant. One more thing that we can see here this particular relationship you can actually consider inverse relationship that means given the distribution we can evaluate the failure rate.

Given the failure rate also we can calculate the distribution. So let me call this relationship * from relation * we can write see Hxt that is= we can consider –d/dt log of 1-Fxt. So this implies I can consider log of 1-Fxt= integral of Hxt dt. So this means (()) (12:17) constant you will get Fxt that is= and you will get a –sign e to the power –integral Hxt dt*constant. For the exponential distribution you can see Hxt is lambda so you get e to the power-lambda t and this k can be determined from the initial condition.

For example, Fx0=0. So this can be calculated. Now you think of a situation where the failure rate is not constant that means it may depend upon t. Now the simplest situation you may think of for example it could be lambda t that means it is linear failure rate or you may think of parabolic failure rate like lambda t square etcetera. Now given any such thing you can evaluate the function.

(Refer Slide Time: 13:30)



Let us take 1 example here let us say Hxt=lambda t. In that case what will be 1-Fxt that is= e to the power-lambda t square/2. And if you consider say differentiation so the density function will give you- and this – will come on this side. So you will e to the power –lambda t square/2 twice lambda t/2 that is= that means the density function is= lambda t e to the power –lambda t square/2.

Now this gives you a more for example if I had taken lambda to the power K here then I would be getting lambda t to the power K/k e to the power k+1/k+1 and here I would have got again the derivative of t to the power k+1 that is lambda t to the power k here. So that gives a more general class of density called Weibull distributions. This give rise to general class of distribution called Weibull distributions.

That means we are considering fx=something like alpha beta x to the beta-1 e to the power –alpha x to the power beta for x>0 and of course alpha and beta positive. So if we have this density you can see that capital F or 1-Fx will be=e to the power –alpha x to the power beta. That means the hazard rate will become=alpha beta t to the power beta-1 which is nothing, but a power of t.

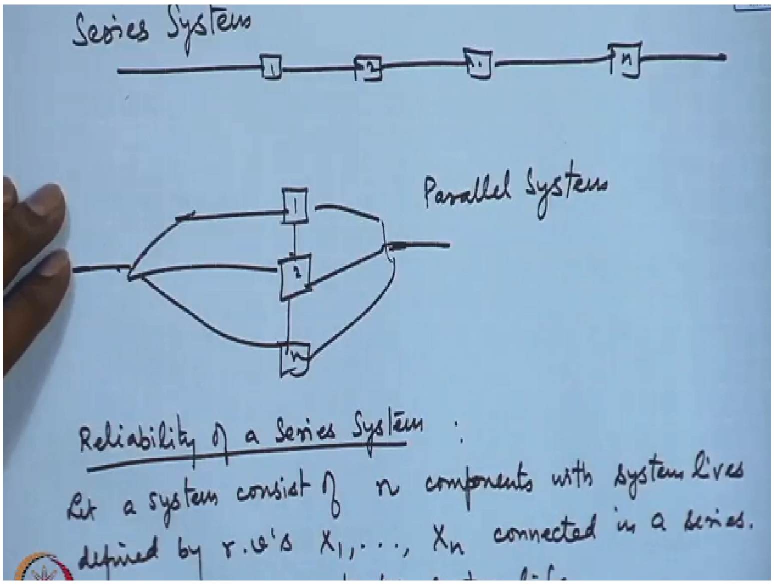
So when you have power of t you get a general Weibull distribution. Now here you can have 2 types of things. For example, if I consider beta>1 then what does this mean if beta >1 that means if time increases the failure rate increases. So that describes various for example systems which are used in industry. For example, any practical kind of system the failure rate will increase as the time increases be at television set or be any manufacturing system or be in any machine.

But we can also model here certain different kind of system for example I may consider beta <1 if I take beta<1 then this t will go in the denominator that means as the time increases the rate of failure decreases. Now this kind of things for example are applicable you consider any organic life forms. So in a organic life form for example a small kid after the birth the failure rate or you can say the death rate is very high mortality rate is very high.

But as the time progresses then his rate of death decreases. So for example from 0 to 5 years it is quite high then from 5 years to say 50 years it is much less. So that means you can model the systems where the initial failure rate maybe very high. That means as the time increases the failure rate decreases. So this gives you a flexibility. Now in the light of this reliability function.

And the failure rate we also consider systems in which multiple components are there.

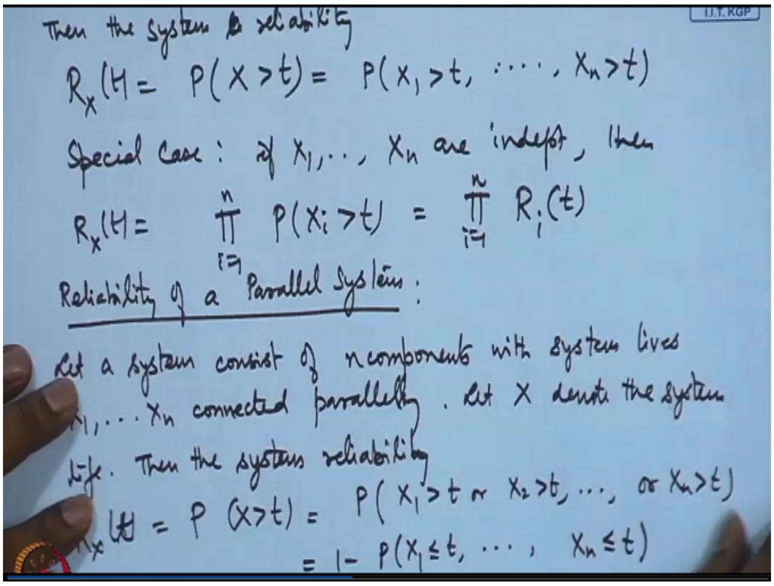
(Refer Slide Time: 17:45)



So we can consider say systems which we consider systems in a series so this is called a series system and you may have systems in parallel. So for example the system will be working if either of this system 1, 2, n is working. Here the system will work if all of the system 1 to n are working. So this is a series system this is a parallel system. So these are the system which are used quite often in engineering design.

Let us look at the reliability of such systems. Reliability of series systems. Let a system consists of n components with say system lives defined by random variable say x1, x2, Xn which are connected in a series. So let us consider say let us define x to be the system life.

(Refer Slide Time: 19:40)



Then if we define the system reliability that is let us call it Rxt that is probability of x>t. So if this entire system is functioning at time t that means each of the components x1, x2, Xn is functioning at time t. So this equivalent to saying probability of x1>t, x2>t, Xn>t. So this will define the system reliability. So if we know the joint distribution x1, x2, Xn this can be evaluated.

Now in special case we can consider that if x1, x2, Xn are independent then this Rxt can be written as product of this probabilities product of Xi>t I=1 to n. That means it is product of the reliabilities of the individual system. So if you have a system which is connected in a series which has components connected in a series and the components are independent then system reliability is nothing, but the product of individual component reliabilities.

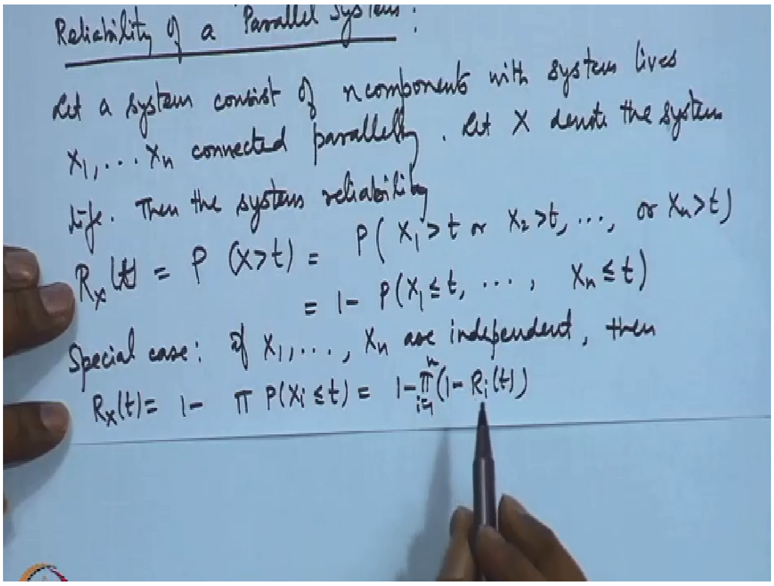
Now the physical interpretation you can think of for example let us take constant component reliability say p at a time t. Now suppose you have n components then the system reliability at time t for the entire system will become p to the power n. So naturally suppose I consider p=1/2 and I take n=3 then Pq that is a system reliability is 1/8 which is much less than ½.

That means if you add the systems in a series then the entire system reliability will continuously decrease. The reason is that if you have more systems then any of them can contribute to the failure of the system and therefore the entire system reliability will become weak or you can say it will become less. On the other hand, let us consider reliability of a parallel system.

So let a system consists of n components with system lives say x1, x2, Xn connected parallely that means for example if you have a circuit from A to B then the circuit will function as long as any of the systems 1 to 3 functions. So let us again x denotes the system life. Then the system reliability will be this is= probability x1>t or x2>t or Xn>t. Now we can consider it in a different way.

We can consider 1- if it is the at least one of them is working then if I consider the complimentary event. Complimentary event will be that all of them will be failing. So it will become probability of x1<or=t, x2 < or=t, Xn <or=t.

(Refer Slide Time: 24:14)



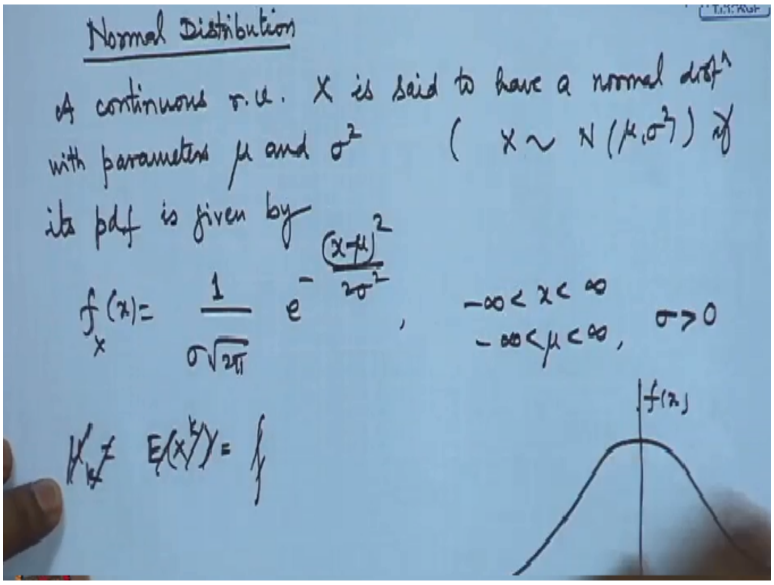
Now again if I consider the special case that is if x1, x2, Xn are independent then Rxt can be written as 1-product of probability Xi< or=t which is nothing but 1- now this will become Rit product I=1 to n. Once again let us look at through a practical example. Suppose I assume components to have identical reliabilities at time t suppose p=1/2. Suppose I have 2 components if I have 2 components then this will become 1-1/2 square that is=3/4 that means the system reliability increases if we connect the components in parallel.

The system reliability decreases for the series the components which are attached in a series, but if you put them in the parallel then it increases. So an important engineering design is that in which we create or you can say increase the system reliability by adding extra system as a backup in series. Generally, they call it like waiting time that as soon as something fails you put something more that means other system starts functioning.

So this concept of reliability failure rate, hazard rate has extremely useful applications in engineering studies because generally we are dealing with the system life. So these are the important quantities to be considered. Based on Weibull distribution there are some other distribution also which are used. As you can see here that we have introduced in the exponent a power here that is x to the power beta.

There are other distributions were in place of polynomial power we can consider exponential power also. Now naturally you can see that they will go to 0 very rapidly so they are also called extreme value distribution. So one can look at various extreme value distributions also. In this particular course we are just mentioning this point here.

(Refer Slide Time: 27:00)

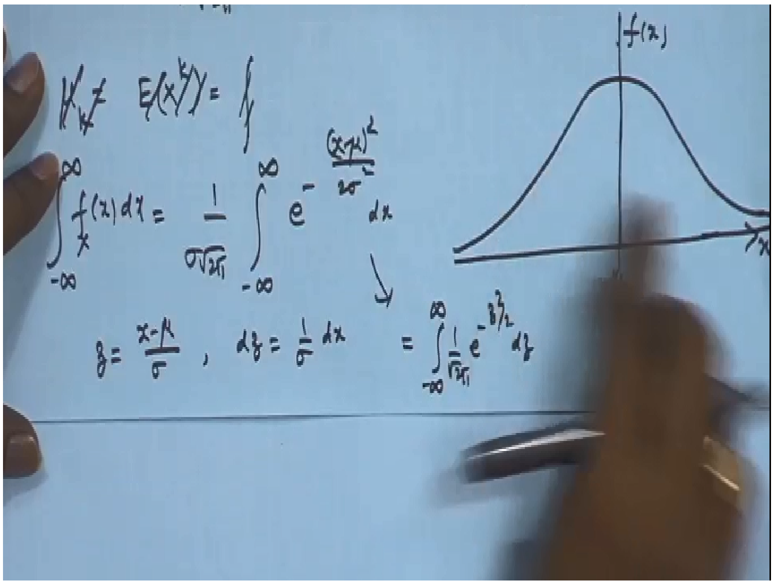


Now we move to one of the most widely used distributions in a statistical theory it is called Normal distribution. Firstly, let me introduce the distribution and then we look at its importance and then why it is actually considered to be most popular and why the name normal is coming. So a continuous random variable X is said to have a normal distribution with parameters mu and sigma square we write actually x follows normal mu sigma square.

If its pdf is given by 1/sigma root 2 pi e to the power – x-mu square/2 sigma square where the range of the variable is – infinity to infinity. The parameter mu is also from -infinity to infinity and the parameter sigma is positive. If we plot the curve it is something like this. As you can see this is symmetric about the value mu. Well let me demonstrate how to evaluate the integral related to this and then we can.

So let us consider the evaluations first. So for the evaluation I consider a general term for example mu k prime. So mu k prime is expectation of x to the power k that is= integral rather I will consider firstly.

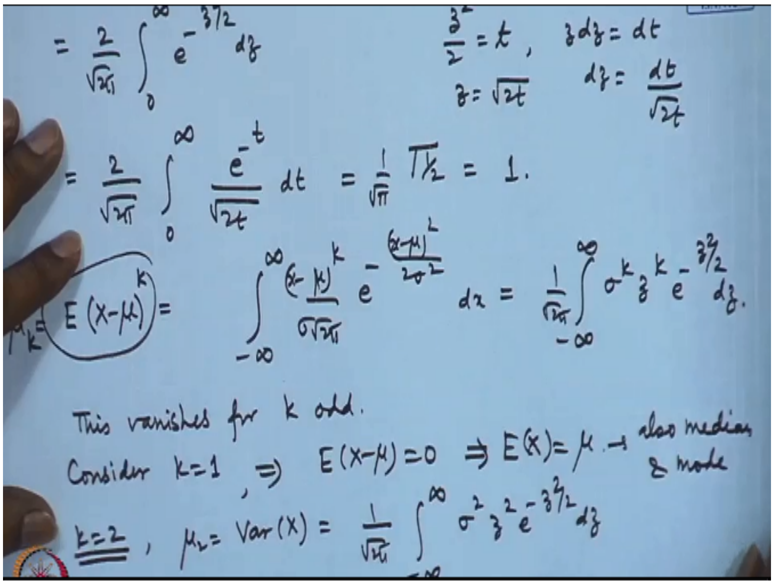
(Refer Slide Time: 29:20)



Let me just look the evaluation of the density form –infinity to infinity. Actually if it is a proper density then this should integrate to 1 so let us look at this one first 1/sigma root 2 pi. Now you consider a transformation z=x-mu/sigma. As you can see from –infinity to infinity this is one to one transformation. So this integral will then become =1/root 2 pi e to the power –z square/2 d z.

Now what we do we observe that this is a convergent integral that you can easily check because it will be rapidly diverging to 0 on both the sides.

(Refer Slide Time: 30:30)



So let us put then this will become=twice 0 to infinity e to the power –z square/2. Those who have done the theory of gamma function you already know its value, but I will just demonstrate the evaluation here. So z is actually=root 2 t. So we can also write it as this is to demonstrate to transform it into a gamma function. So this is becoming e to the power –t by root 2 t dt that is=simply gamma. So 1/root pi gamma 1/2 which is root pi so it is=1.

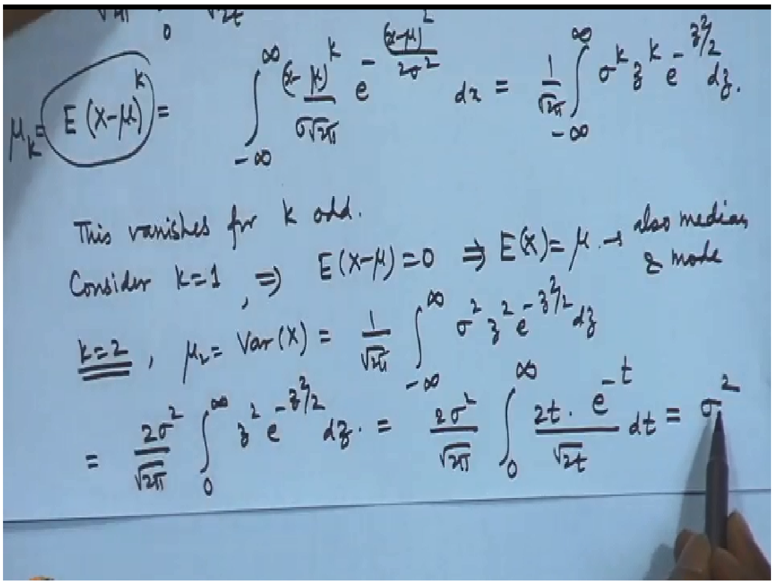
Now if we use this we can evaluate for example what is expectation of x-mu if I consider expectation of x-mu to the power k for example. As you can see this integral evaluation is much better compared to expectation of x to the power k because I have made the transformation x-mu/sigma=z. So that will simplify this term whereas in this one it will not get simplified.

So if I consider this term here 1/sigma root 2 pi e to the power-x-mu square/2 sigma square dx and here you have x-mu to the power k. So if I make this transformation x-mu/sigma=z then this becomes = –infinity to infinity 1/root 2 pi sigma to the power k z to the power k e to the power –z square/2 dz. So obviously this vanishes for k odd. Now let us consider k=1. This means expectation of x-mu=0 which means expectation of x=mu.

So the term mu here denotes the mean of normal distribution which is also you can look at it here from the shape of the distribution this is also the median and it is also the mode of this distribution also it is median and mode. Now let us consider say for example k=2. If I take k=2. So now this quantity because mu is the mean then this becomes actually the central moment here.

This expectation of x-mu to the k this is actually kth central moment because I approved here mu is the mean. So for k=2 mu 2 will denote the variance of x Now let us look at the value here it is-=1/root 2 pi-infinity to infinity sigma square z square e to the power –z square/2 dz. Now this one is an even function.

(Refer Slide Time: 34:18)

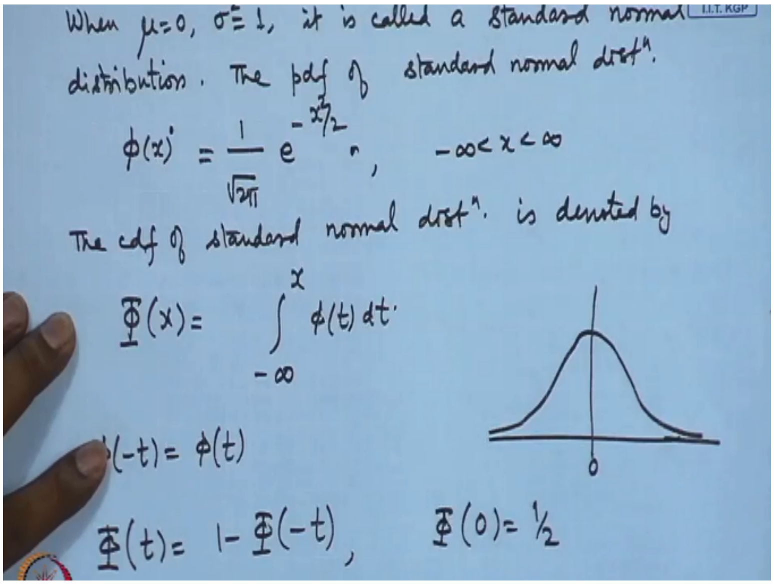


And therefore this is=2/root 2 pi 0 to infinity sigma square z square e to the power-z square/2 dz. We again look at the transformation that we considered here that is z square/2=t. So if we consider this transformation then this value is simply=2 sigma square/root 2 pi 0 to infinity 2 t e to the power –t/root 2 t dt. So easily you can see this term is giving you t to the power 1/2 that is gamma 3/2.

So gamma 3/2 is 1/2 gamma 1/2 that is 1/2 root pi and this 2 and this 2 and this 2 cancels so 1/2 and 2 cancels so you get simply sigma square. So we have shown that this parameter sigma square of the normal distribution is actually denoting the variance. So now when I write that x follows normal mu sigma square means the mean of the normal distribution is mu and the variance is sigma square.

Now as I have mentioned here the mu can be any real number and sigma can be any positive number. One of the important special case will be when mu=0 and sigma =1 that is called a standard normal distribution.

(Refer Slide Time: 36:01)

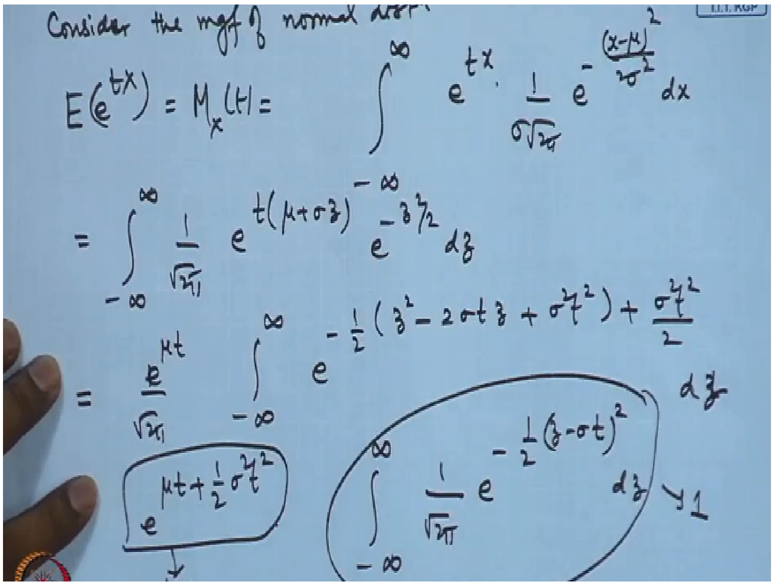


When mu=0 sigma square=1 it is called a standard normal distribution. The probability density function of standard normal distribution that is 1/root 2 pi. Since sigma is 1 e to the power –x square/2 for x between –infinity to infinity. In the statistical text special notation small phi is used for this. The cumulative distribution function of standard normal distribution is denoted by capital phi that is –infinity to x small phi t dt.

Now this function has some special property also. As you can see since the normal distribution is symmetric around its mean when I have a standard normal this will be symmetric around 0. That means small phi of-t=small phi of t. Now if you utilize this here you will get capital phi of t-=1-capital phi of-t and in particular capital phi of 0 will be=half that means 0 is the median which is true here.

Now there is another important point because of which standard normal distribution is considered. Given any normal distribution you can always shift it to a standard normal distribution for that I will prove one result here.

(Refer Slide Time: 38:26)



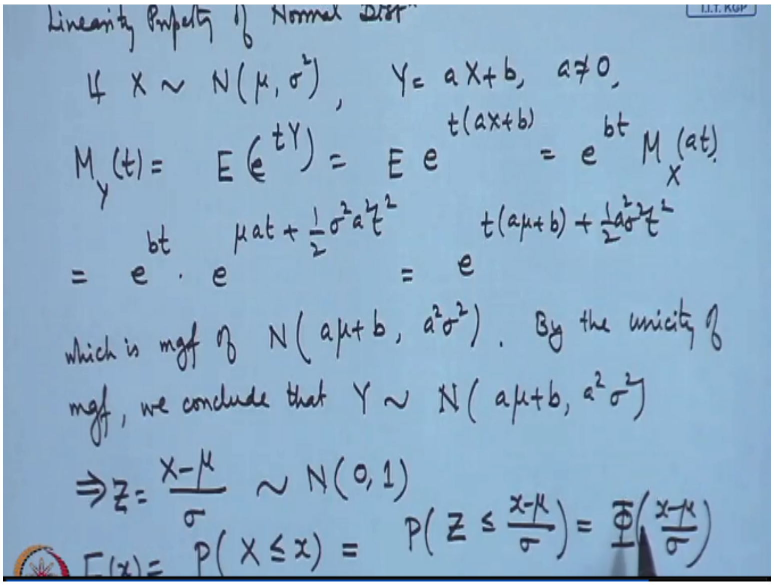
So firstly let us look at say consider the moment generating function of normal distribution. So expectation of e to the power tx that is Mxt. So that is=integral from – infinity to infinity e to the power tx 1/sigma root 2 pi e to the power –x-mu square/2 sigma square dx. Now as before we consider the transformation that is x-mu/sigma=z. Now a consequence of this is that x=mu +sigma z.

So if we consider this then I get here-infinity to infinity 1/root 2 pi e to the power t mu+ sigma z e to the power –z square/2 dz. So this I write as 1/root 2 pi and this e to the power mu t can be written outside and e to the power-1/2 z square-twice sigma tz. Now I add and subtract sigma square t square here. So if I subtract here and I take out it outside it will become.

So this I can express as e to the power mu t+1/2 sigma square t square-infinity to infinity 1/root 2 pi e to the power-1/2 z-sigma t whole square dz. If we look at the integrant this is nothing, but the probability density function of a normal random variable with mean sigma t and variance 1 that means in place of mu if I put sigma t and in place of sigma I put 1 then I get this.

Therefore, the integral of this should be simply =1. So this quantity becomes the mgf of a normal distribution with mean mu and variance sigma square.

(Refer Slide Time: 41:18)



Now using this we can prove the linearity property of a normal distribution. That means if x has a normal mu sigma square distribution. Let us define y=say a X+ b where a is not 0 then let us calculate the moment generating function of y that is expectation of e to the power ty that is= expectation of e to the power t ax+b so that is= e to the power bt moment generating function of x at the point at.

Now moment generating function of x is calculated as e to the power mu t+1/2 sigma square t square. So we substitute it here this gives us e to the power bt e to the power mu at+1/2 sigma square a square t square that is= e to the power t a mu+b+1/2 sigma square t square. So this is moment generating function of a normal distribution with mean a mu +b and variance this is a square sigma square so a square sigma square.

So what we have proved by the uniqueness property of mgf we conclude that Y follows normal a mu+ b and a square sigma square. So this is if you look at any random variable x with certain mean and variance then we know that it is mean is linear and variance will be square, but if we consider the distribution, the distribution itself may change. However, for the normal distribution any linear function is also having the normal distribution.

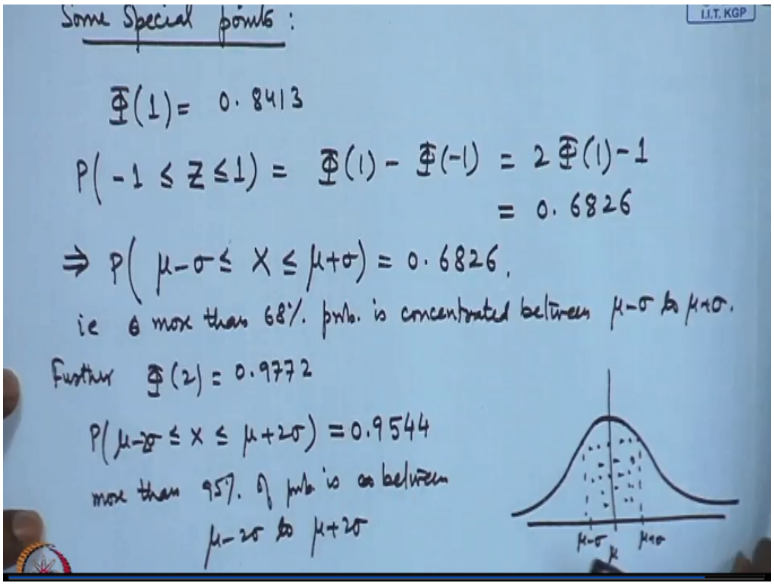
Now if I take the special case if I take x-mu/sigma then that will follow normal 0, 1. So from any normal random variable I can transfer it to normal 0, 1 this is called standardized variable or normalized value that means you subtract the mean and divide by the standard deviation then this will have normal 0,1 distributions. That means from any normal I can always shift to standard normal.

Now this property is useful to evaluate the probabilities related to a general normal distribution. So if I consider say cumulative distribution function of x that is probability of x<or=x. As you can see this is nothing, but the integral from –infinity 2 x of the density function here 1/sigma root 2 pi e to the power –x-mu square/2 sigma square which will lead to an incomplete gamma function.

So for every different mu and sigma it will be difficult to evaluate. However, if we use this linearity property we can consider this is=probability of Z < or= x-mu/sigma that is nothing, but phi of x-mu/sigma. The tables of capital phi are available in almost all the statistical text these tables are there. I will just show you the cdf of the standard normal distribution. So that means you have the normal curve and standard normal curve.

So what is the probability up to a point z that is cdf. So for different values of z this values are the probabilities are tabulated for example what is the probability up to 0 so this is=1/2. What is the probability up to say +1? it is 0.8413. Now if we look at this we also come across some very interesting phenomena about normal distribution which I will show you now.

(Refer Slide Time: 46:18)

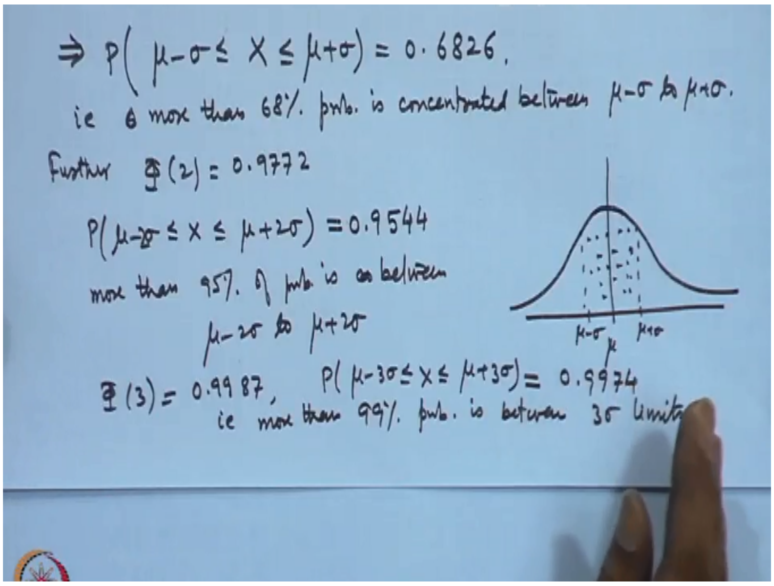


Let us consider some special points. We look at phi of 1 that is=0.8413. I also given you the relationship that phi of x=1-phi of –x. So if I consider what is the probability of -1 < or=z < or =1 that is= phi of 1-phi of-1. So if I use this property I get twice phi of 1-1. So that will give us 0.6826 that is I have just taken twice of this that is=1.6826 so then I subtract 1 so I get 0.6826.

Now if you write here z=x-mu/sigma then it is giving you mu-sigma < or=x < or=mu+ sigma is= 0.6826. Now this is interesting let me again draw the normal curve this is mu. So let us consider mu-sigma to mu+ sigma. So what we are saying that more than 68% of the probability is concentrated in the zone mu-sigma 2 mu+ sigma that is more than 68% probability is concentrated between mu-sigma to mu+ sigma.

Likewise let us consider say phi of 2 from the tables of standard and normal distribution you can see phi of 0.2 is 0.9772 that is= 0.9772. Once again if I consider probability of mu-2 sigma to mu +2 sigma then that will give me that means I have just multiplied this by twice and subtracted 1 like here. So what we are concluding that more than 95% of probability is between mu-2 sigma to mu+2 sigma. Finally let us write say phi 3.

(Refer Slide Time: 49:40)

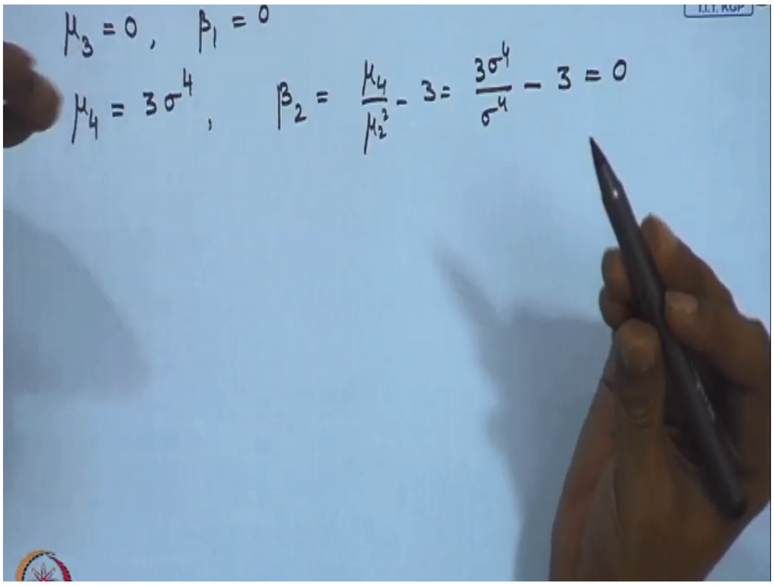


If you see from the tables of the normal distribution it is=0. 9987. So if I consider probability of mu-3 sigma < or=x < or=mu+3 sigma. Then I get point that is more than 99% probability is between 3 sigma limits that is mu-3 sigma to mu+3 sigma. So in the industrial quality that means when we are looking at the quality of the items then we try to see that our most of the material is between mu-3 sigma to mu+ 3 sigma.

So for very long time in the industry this 3 sigma limits are found to be very useful. So they say that process is under control if it is within the 3 sigma limit and of course nowadays there is further generalization we are considering 6 sigma in place of 3 sigma because if you consider 6 sigma then the probability of inclusion becomes it will be actually 0.999998. So that means probability having outside will be 1 in a million kind of thing.

So this 3 sigma limits are used in the industry. To conclude about this normal distribution actually I will be telling more in the following lecture.

(Refer Slide Time: 51:32)



Let us just look at here the measures of skewness and kurtosis for the normal distribution. You can easily check from the calculation because we have already calculated the general mu k. So mu 3 is 0. So measure of skewness will be 0. If I look at the measure of kurtosis for that I need mu 4 it can be checked it is=3 sigma to the power 4. So the measure of kurtosis that is mu 4/mu2 square -3 that is= 3 sigma to the power 4/sigma to the power 4-3 that is=0.

So when I was mentioning the peakedness or the kurtosis of the distribution I mentioned that there is a normal peak and higher peak that we called leptokurtic and flat peak we called it Platykurtic. So here we can see that the normal peak is actually the peak of the normal curve or the normal distribution. Now more about that why we are actually calling it a normal distribution I will be covering in the next lecture.

What we have observed here is that the normal distribution satisfies a linearity property it is also having symmetry and the probability is connected to any normal distribution can be calculated using standard normal probabilities which are actually available in the form of tables. We will in the following lecture show that normal distribution arises naturally as the limiting distribution of various distributions.

So the results are known as a central limit theorem. So in the next lecture I will be covering that.