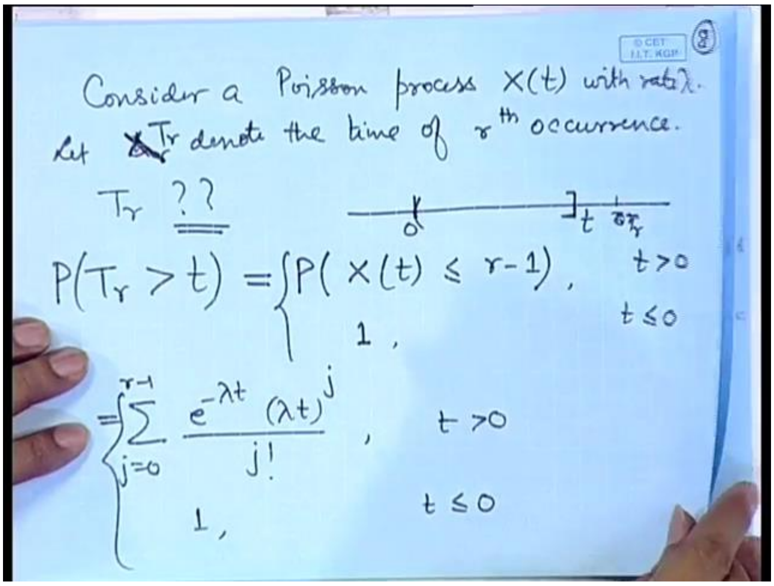
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 26
Special Continuous Distributions – III

Now, consider that we are not interested in a single occurrence or a single failure or a single happening in a poisson process.

(Refer Slide Time: 00:34)



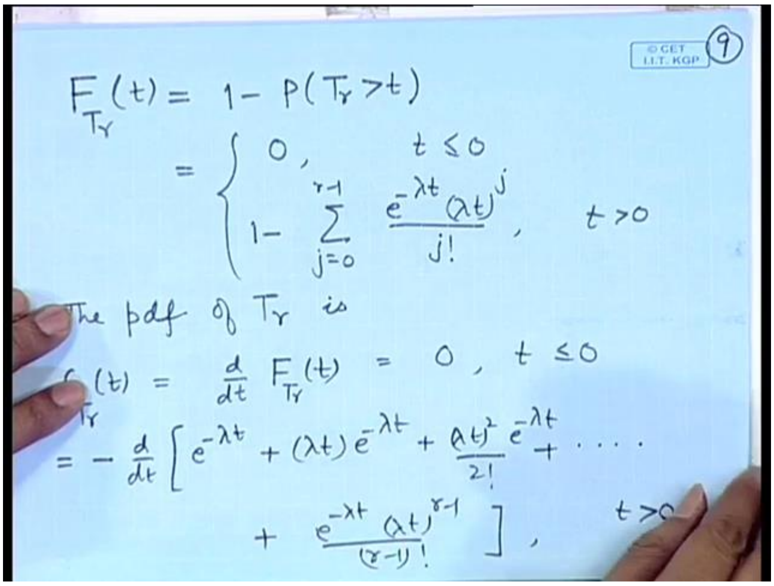
In place of that we are considering; so, consider a Poisson process, let us denote x t with rate lambda. And let X r denote the time well in place of X, we will use notation say T r because X is being used here, let T r denote the time of rth occurrence.

In place of the first occurrence, we are looking at a certain number of occurrences. As we are discussed yesterday also like in the negative binomial distribution that a certain major event may occur as a consequence of certain smaller events. For example, if thousand smaller intensity earthquakes occur then it may make the earth to crumble and a major earthquake may occur a sequence of certain mishaps may closed on the plant itself, a sequence of, so any smaller kind of events may lead to a major disaster. So, we may be interested in the vetting time for that of course, considering the events to occur in a poisson process. For example, a certain number of occurrences occurred, say certain number of people purchases certain tickets then we may how to close on the window because the seats are full.

What is the distribution of T r? Once again we can consider probability of T r greater than t. So, a starting from a time 0, consider time small t, if we say that rth occurrence has not taken place till this time, suppose this is the r th occurrence; that means, in the interval 0 to t less than or equal to r minus 1 occurrence will be there. So, this event is equivalent to probability of x t less than or equal to r minus 1, of course, here t is positive, it is equal to 1 for t less than or equal to 0.

Now, x t is having a Poisson distribution with parameter lambda t. So, this is equal to e to the power minus lambda t, lambda t to the power j by j factorial summation j is equal to 0 to r minus 1. This is for t greater than 0 and it is one for t less than or equal to 0.

(Refer Slide Time: 03:41)

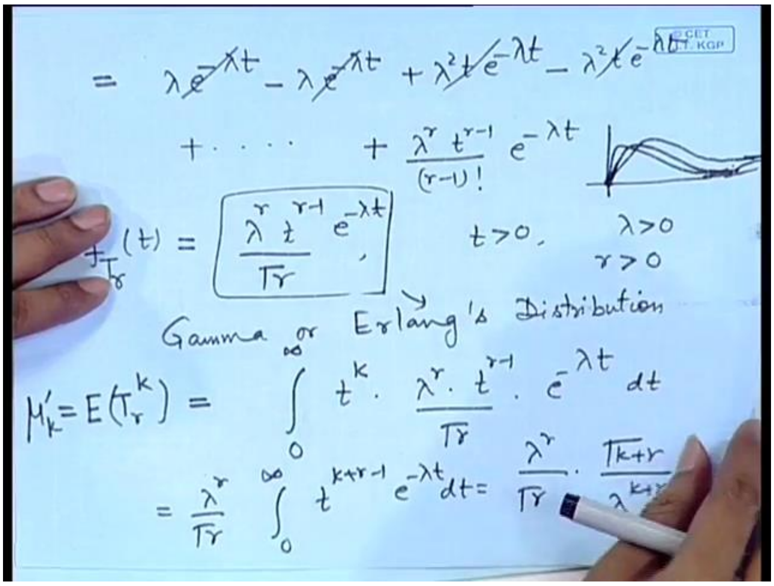


Consequently, we can write down the cumulative distribution function of T r as F of T r at the point t that is 1 minus probability of T r greater than t and this is equal to 0 for t less than or equal to 0, it is 1 minus summation j is equal to 0 to r minus 1, e to the power minus lambda t; lambda t to the power j by j factorial for t greater than 0.

You can observe here, it is a time variable and this is an absolutely continuous function. So, the probability density function can be obtained by differentiation of this. Now for t less than or equal to 0, it is 0 and in this particular portion we are having some of a series. So, we are do term by term differentiation. So, this is equal to 0 for t less than or equal to 0 and in this portion now, this is d by d t of e to the power minus lambda t plus lambda t into e to the power minus lambda t plus lambda t square e to the power minus lambda t by 2 factorial and so on plus e to the power minus lambda t lambda t to the power r minus 1 by r minus 1 factorial for T greater than 0.

Here you observe that when we differentiate here 1 term is there, but there after each term is a product of 2 terms involve in t. So, when we differentiate we have do by applying the formula of derivative of a product.

(Refer Slide Time: 05:44)



If we expand the terms the derivative of the first term will give us minus lambda e to the power minus lambda t and there is a minus sign here, consequently we will get it as lambda e to the power minus lambda t, if we look at the second term and we look at the derivative of this we will get lambda multiplied by e to the power minus lambda t and there is a minus sign here. So, we will get here minus lambda e to the power minus lambda t and observe that this can cancel out.

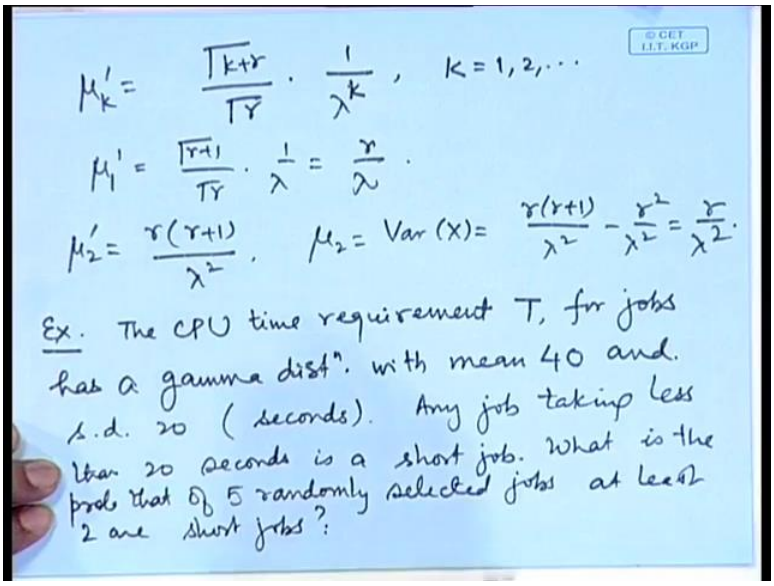
The next term will be the derivative of this. So, that will give minus lambda and minus is there. So, it becomes plus lambda square t e to the power minus lambda t. Now look at the third term and derivative of this. So, if we look at the derivative of this one here t square is coming. So, 2 t and this 2 will cancel out and there will be a minus sign. So, we will get with a, because there are as minus sign outside, it becomes minus lambda square t to the power minus lambda t. So, you can easily observe that this is a telescopic sum, the final term will be plus which will be contributed by this term that will become lambda to the power r, t to the power r minus 1 by r minus 1 factorial e to the power minus lambda t.

This will be the left out term. So, we are getting the probability density function of T r as lambda to the power r, t to the power minus 1, e to the power minus lambda t by gamma r for t greater than 0, here lambda is a positive parameter, in the derivation we have considered r to be a integer because its occurrence, but after writing down this you can see that it is valid for r greater than 0 any positive real number. So, this is known as a gamma distribution r; Erlang's distribution. So, this distribution arises as the distribution of the vetting time for the rth occurrence in a Poisson process in place of the first occurrence if you put r is equal to 1 you will get the exponential distribution because this term will vanish you will get lambda e to the power minus lambda t.

This is a generalization of the exponential distribution, but it has much more applicability because we are looking at a higher order of occurrences in a Poisson process apart from the regular application one can see the characteristic of this distribution if we consider a kth order non central moment that is mu k prime . So, mu k prime is equal to integral from 0 to infinity t to the power k lambda to the power r t to the power r minus 1 by gamma r e to the power minus lambda t d t.

Obviously, this is a gamma function you can write it as lambda to the power r by gamma r; t to the power k plus r minus, 1 e to the power minus lambda t d t. So, it is lambda to the power r by gamma r gamma k plus r divided by lambda to the power k plus r.

(Refer Slide Time: 09:51)



This is equal to that is mu k prime is equal to gamma k plus r of course, this is valid for any k greater than 0, but we will be more concerned with the positive integral moments.

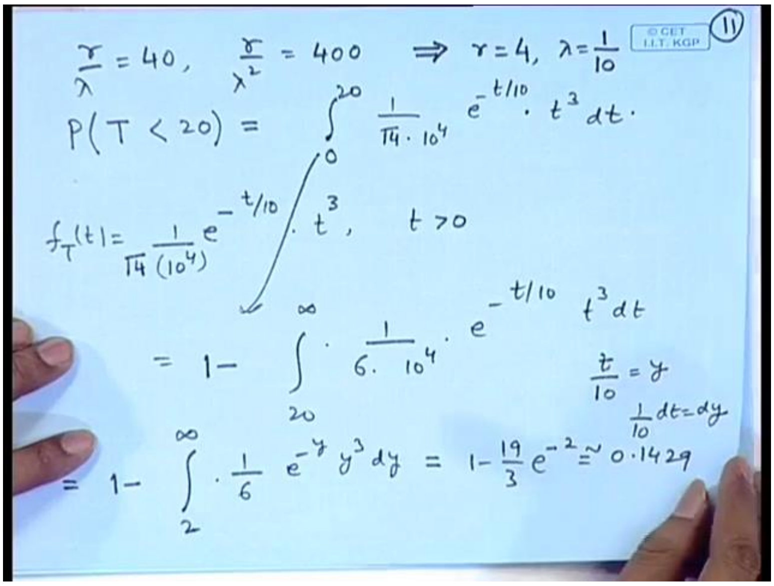
Mu 1 prime is equal to gamma r plus 1 by gamma r 1 by lambda that is r by lambda. So, if the vetting time for the first occurrence was average vetting time for the first occurrence was one by lambda then for the rth occurrence it will be r by lambda. So, this is again because of the memory less property of the exponential distribution because after first occurrence we can again consider it as the starting of the process observing from time 0. So, again for the second occurrence vetting time will be 1 by lambda and so on.

If we consider mu 2 prime then this becomes r into r plus 1 by lambda square and therefore, mu 2 that is the variance of this distribution is equal to r into r plus 1 by lambda square minus r square by lambda square that is r by lambda square which is again you can see, in the exponential distribution the variance was 1 by lambda square and if you are looking at the variance of the rth occurrence then it becomes r by lambda square.

Let us look at one application of this the CPU time requirement T for jobs has a gamma distribution with mean 40 and a standard deviation 20, this measure is in seconds, any job taking less than 20 seconds is a short job. What is the probability that of 5 randomly selected jobs at least 2 are short jobs? So, in order to answer this question let us consider the setup, here mean is given to be 40. So, that is r by lambda is 40 and standard deviation.

Here variance is r by lambda square. So, r by lambda square is 400. So, this is a 2 parameter distribution and we have now 2 equations.

(Refer Slide Time: 13:32)



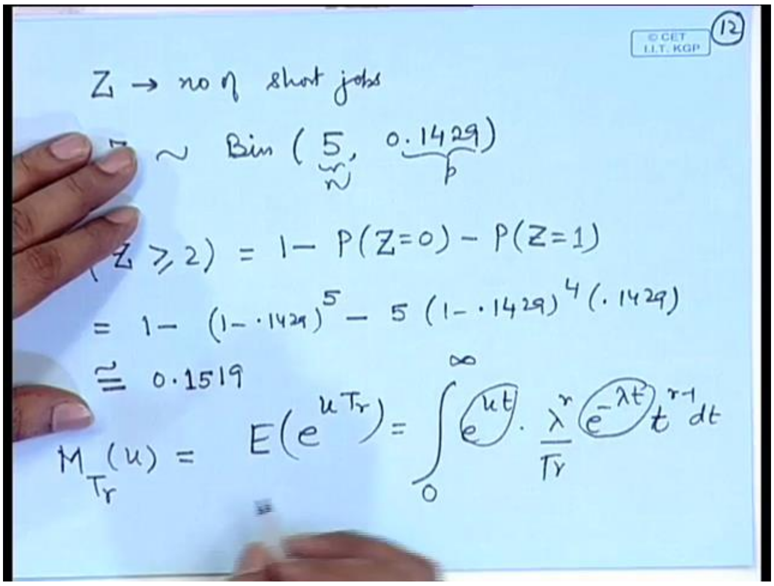
We can have r by lambda is equal to 40; r by lambda square is equal to 400. So, if we solve this, we get r is equal to 4 and so, lambda is equal to 1 by 10. So, that is right, what is a probability of a short job; that means T is less than 20.

We need to consider the density function of T; here f T is equal to e to the power minus lambda t so that is t by 100 t to the power r minus 1 that is 3 then lambda to the power r that is 1 by 10 to the power 4 and gamma r, this is for t greater than 0. So, we need to calculate 0 to 20, the integral of this density 1 by gamma 4, 10 to the power 4, e to the power minus t by 10 t cube d t, this is the probability of a randomly selected job to be a short job. So, if we want to evaluate this, we can consider since exponential function is involved the integral will be at 1 and something value 1 and at another point some value will be there.

It is convenient if we consider this as 1 minus 20 to infinite 1 by 6 into 10 to the power 4 e to the power minus t by 10 t cube d t. So, from the integrals, integral we can observe that it is convenient if we put t by 10 is equal to y that is 1 by 10 d t is equal to d y. So, this becomes equal to 1 minus, this will be the 2 to infinity, 1 by 6 e to the power minus y; y cube d y. So, this is not a complete gamma function, this is an incomplete gamma function because integral is from 2 to infinite. So, we can do integration by parts and after doing certain simplification, this one will turn out to be 1 minus 19 by 3 e to the power minus 2 which is approximately 0.1429.

Probability of a randomly selected job being a short job is only 0.1429; that means, in the gamma distribution if the mean is 40 and the standard deviation is 20 the probability of job being over in mu minus sigma is actually much smaller.

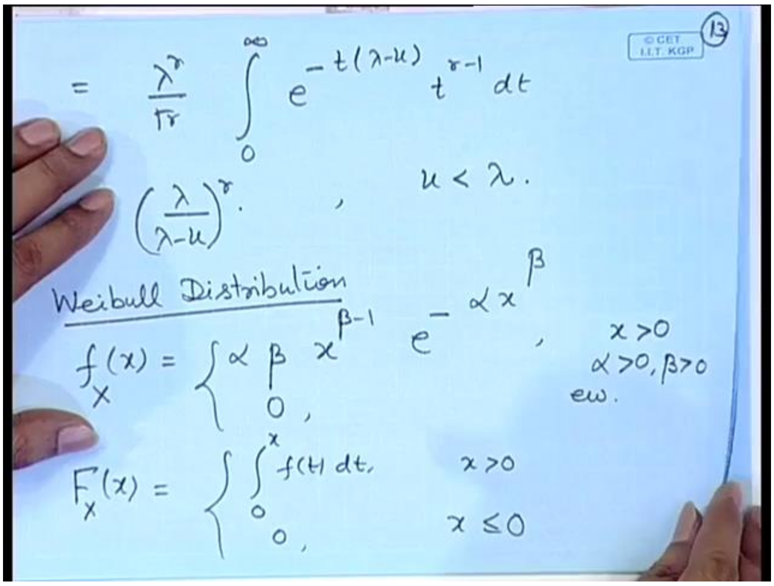
(Refer Slide Time: 17:15)



Now we can consider the random variable say Z is the number of short jobs out of 5 then Z is following binomial distribution with n is equal to 5 and p is equal to 0.1429, this is p this is n, we are saying what is a probability that at least 2 jobs are short jobs; that means, probability of Z greater than or equal to 2, this we can consider as probability Z is equal to 0 and Z is equal to 1 you subtract from 1 that is a complimentary event.

This is equal to 0.1 minus 0.1429 to the power 5 minus 5 c 1; 1 minus 0.1429 to the power 4 into 0.1429. So, this can be evaluated and it is approximately 0.1519. So, the probability that at least 2 out of the 5 jobs are short jobs is quite is small this is because the probability of a single job itself being short job is much smaller. So, we may look at the moment generating function etcetera of this distribution. So, moment generating function of the gamma distribution expectation of e to the power u T r that is equal to integral e to the power u t lambda to the power r by gamma r e to the power minus lambda t; t to the power r minus 1 d t; 0 to infinite. So, here if you see this term can be combined with this. So, it becomes directly a gamma function.

(Refer Slide Time: 19:32)



This is equal to lambda to the power r by gamma r, e to the power minus t lambda minus u, t to the power r minus 1, 0 to infinity. So, this is just the gamma function lambda by lambda minus u to the power r and gamma r will be cancelling out. So, this is valid for u less than lambda. You can see the similarity from the; a moment generating function of the exponential distribution where r was 1. So, it was lambda by lambda minus u.

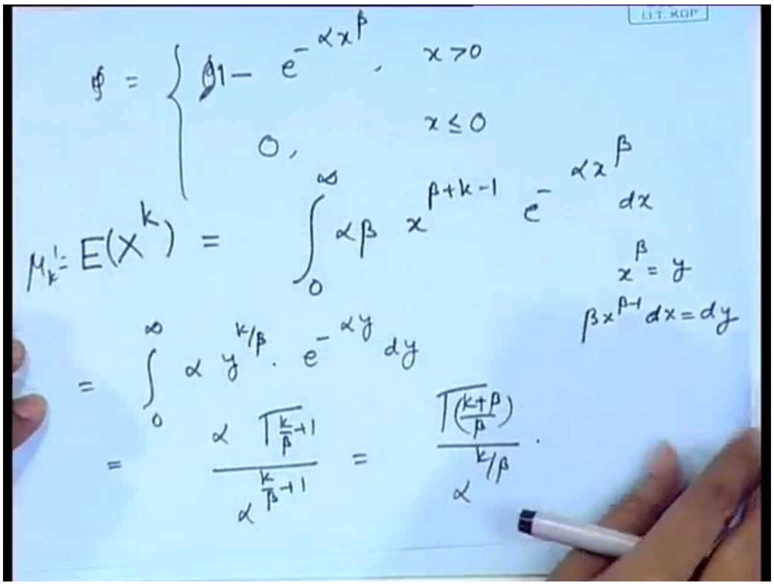
We will show some relationship between the gamma distribution and the exponential distribution later on. The moments of the gamma distribution as we have seen can all be calculated from the general expression for mu k prime? So, mu 3 prime mu 4 prime can be calculated and then mu 3 and mu 4 can also be calculated you can see that it will be actually positive. So, all the gamma distributions will be in fact, positively skewed the reason is that if you look at the density function it is lambda to the power t to the power r minus 1, e to the power minus lambda t divided by of course, gamma r.

In the beginning if you consider at t is equal to 0 apart from r is equal to 1, this value is going to be 0 and if you are considering then t becoming larger than in the beginning since it is t to the power r minus 1, it may increase little bit, but there after this term will dominate and therefore, the densities will always be positively skewed for various gamma distribution, of course, it will depend upon what is the value of r and what is the value of lambda for different shapes, but whatever be the shapes; they will be positively skewed, the peak will of course, depend upon the value of the r and lambda.

If you are considering the exponential distribution or the gamma distribution as the life of certain components then another interesting distribution in this same direction is the so called Weibull distribution, the general form of a probability density function of a Weibull distribution is given by alpha beta x to the power beta minus 1 e to the power minus alpha x to the power beta, where x is always positive and alpha and beta positive parameters. So, quite naturally one can see the form of the CDF here because if you integrate from since x is positive value random variable the integral will be from 0 to x f t d t for x positive and it is 0 of course, for x less than or equal to 0.

If you consider this integrant, it looks like a derivative of e to the power minus alpha x to the power beta.

(Refer Slide Time: 23:28)

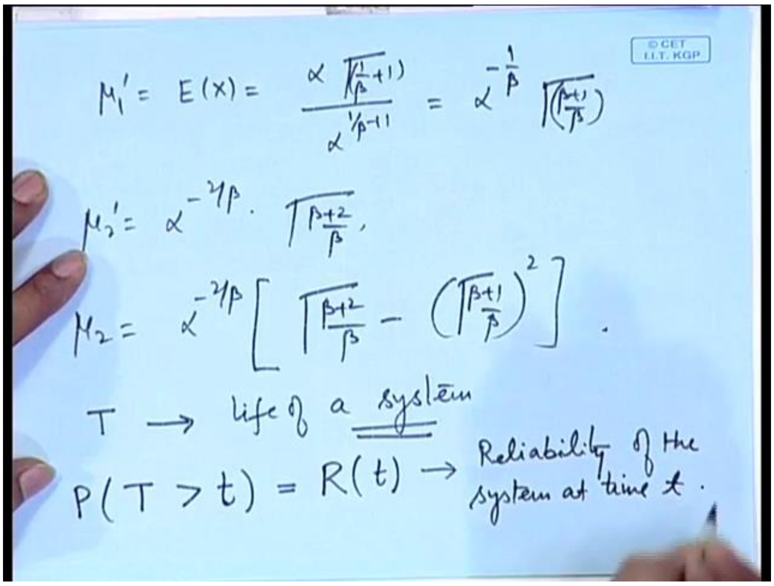


And therefore, when we calculate the CDF, this will be simply equal to 1 minus e to the power minus alpha x to the power beta for x positive and it is 0 for x less than or equal to 0. If we put beta is equal to 1 then this term vanishes, this term becomes e to the power minus alpha x. So, you get alpha e to the power minus alpha x so that becomes exactly the density of an exponential distribution.

This Weibull distribution is actually a generalization or extension of the exponential distribution. So, we will see that what is the significance of making power x to the power beta here because in the exponential we had alpha e to the power minus alpha x so, x has been replaced by x to the power beta. So, first is that we can look at its moment structure so; obviously, you can make use of the gamma functions, if I consider expectation of X to the power k, it is alpha beta x to the power beta plus k minus 1. So, quite; obviously, you can understand here that if you put x to the power beta is equal to y that is beta x to the power beta minus 1 is equal to d x is equal to d y then this becomes integral from 0 to infinity, alpha and x to the power beta minus 1 term is combined here. So, you have x to the power k that is y to the power k by beta e to the power minus alpha y d y.

This is simply a gamma function and the expression for this turns out to be alpha gamma k by beta plus 1 divided by alpha to the power k by beta plus 1; that means gamma k by or you can write it as k plus beta by beta divided by alpha to the power k by beta.

(Refer Slide Time: 25:54)

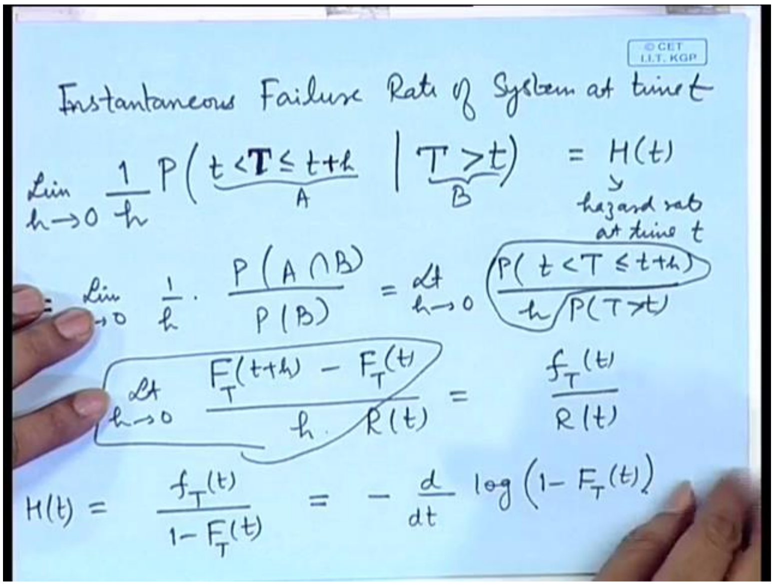


In particular, if I am looking at the mean. So, mu 1 prime is equal to expectation of X that is equal to alpha 1 by beta plus 1, gamma of this divided by alpha to the power 1 by beta plus 1 that is equal to alpha to the power minus 1 by beta, gamma of beta plus 1 by beta and mu 2 prime will become alpha to the power minus 2 by beta gamma of beta plus 2 by beta and therefore, the variance of the Weibull distribution will be alpha to the power minus 2 by beta gamma beta plus 2 by beta minus gamma beta plus 1 by beta whole square which looks slightly complicated, but nevertheless the functions of the gamma functions can be easily calculated using tables of the gamma distribution or the from calculators or computers now a days it can be easily.

However this distribution as more importance in the reliability analysis of certain mechanical are electronic systems. So, let us define, what do you mean by the reliability of the system? So, if T denotes the survival time or the life of a system so, we can consider probability of T greater than t, this means the system has been functioning till time t or it as not fail till time t or the system is working at time t, the probability of system working at time t is called the reliability of the system at time t; reliability of the system at time t.

Using this, we can define another quantity and of course, you can easily see that this is equal to 1 minus the CDF at the point T.

(Refer Slide Time: 28:24)



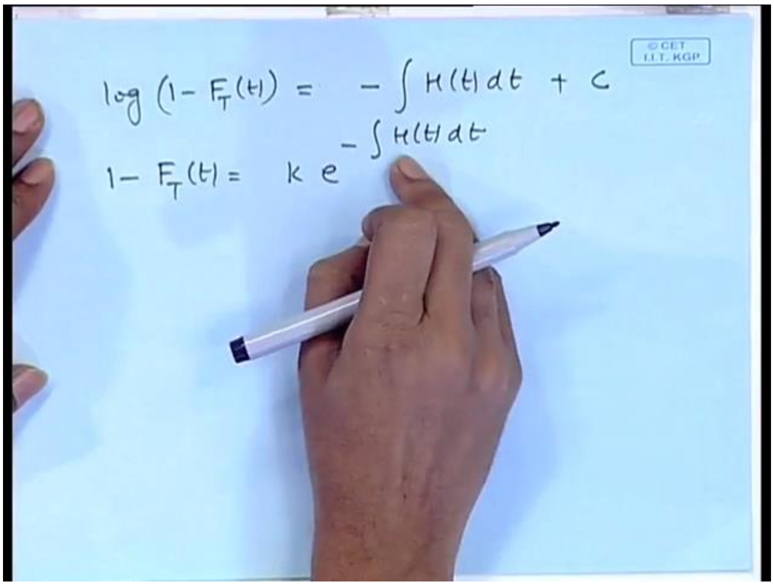
We also define what is known as instantaneous failure rate of system at time t. So, let us consider the interpretation of this. The system is functioning at time t and immediately after the time t, it fails; that means, in an interval from T to t plus H. So, if we are considering the rate, we consider it divided by h and take limit as h tend to 0, let us give some notation say H of t, you also call it hazard rate; hazard rate at time t.

This we define for any random variable which is denoting the life of a system. So, we call failure rate at time t or hazard rate at time t; that means, giving that the system is functioning at time t what is the probability that it will failure immediately after that and therefore, if you want to calculate the rate then we divide by the length of the interval and take the limit as it is. So, let us evaluate this. Now if you consider this as event A and this as event B then this is probability of A intersection B divided by probability of B.

Now, once again you consider A is a subset of B. So, this becomes probability of t less than capital T less than or equal to t plus h divided by probability T greater than t, h limit as h tends to 0. So, this term if you see it is nothing but the density function of the variable t because you are taking limit as h tends to 0 see you can expand it like this the numerator becomes F of t plus h minus F of t by h and this is R t. So, this term is nothing, but the F T by R t, R you can consider it as f T divided by 1 minus F T. So, the hazard rate of a life time distribution can be denoted by the density divided by the reliability or the density divided by 1 minus the cumulative distribution function of the random variable.

We are calling it as H t, we can notice another relationship here this is minus d by d t log of 1 minus F T. So, this looks like a first order differential equation, we can integrate it out.

(Refer Slide Time: 31:46)



And we will get log of 1 minus F t as equal to and therefore, plus of course, a constant of integration. So, 1 minus F T becomes some k times e to the power minus integral H t d t, this shows that given the distribution one can calculate the hazard rate, given the hazard rate function one can determine the CDF and hence the probability density function of the random variable.

We will look at this quantity that is the reliability, the hazard rate, in context of the Weibull distribution, the exponential distribution and try to see what does it signify? So, in the next lecture, we will be covering these issues.

Thank you.