

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

Module No.#01
Lecture No. #13
Special Distributions-IV

In the previous lecture, we have introduced negative exponential distribution and we gave its origin as the distribution of the waiting time for the first occurrence in a Poisson distribution.

(Refer Slide Time: 00:39)

Lecture-13

$$P(T > a) = e^{-\lambda a}$$
$$P(\underbrace{T > a+b}_A | \underbrace{T > b}_B) = \frac{P(A \cap B)}{P(B)}$$
$$= \frac{P(A)}{P(B)} = \frac{P(T > a+b)}{P(T > b)} = \frac{e^{-\lambda(a+b)}}{e^{-\lambda b}}$$
$$= e^{-\lambda a} = P(T > a)$$

Memoryless property of the exponential distⁿ.

So, we considered probability of T greater than T , if T is having an exponential distribution. So, if we consider probability of, say T greater than a number a , then it is equal to e to the power minus λa .

Now, if we consider, so this means that the first occurrence has not taken place till time a . That means, in the interval 0 to a , it has not taken place. So, if we consider probability of, say T greater than a plus b , given that T is greater than b . So, if we consider this as event A , this as event B , then it is equal to probability of A intersection B divided by probability of B . Now, A intersection B , so here, A is a subset of B . Therefore, this becomes probability of A divided by probability of B , that is probability of T greater than a plus b divided by probability of T greater than B . So, by the formula that we have developed,

this one will become equal to $e^{-\lambda a}$ plus b divided by $e^{-\lambda a}$ plus b divided by $e^{-\lambda a}$. So, after simplification, it becomes $e^{-\lambda a}$, which is nothing, but the probability of T greater than a .

So, this phenomena represents that, if the event has not taken place till time b , then it will not take place till an additional time a , the probability of that is same as, that the event will take place after time a starting from time 0. That means, if we are considering the waiting time for certain thing, then the probability of waiting for an additional time is same irrespective of the starting point.

So, like geometric distribution, this is called the memoryless property of the exponential distribution. You can see that exponential distribution is analogous to the geometric distribution of the discrete case. In the geometric distribution, we were considering Bernoullian trials and we were waiting for the first success or first occurrence. So, here also, it says Poisson process and we are waiting for the first occurrence. So, in a sense, this exponential distribution is a continuous analog of the geometric distribution.

So, if we consider, say occurrence to be failure of certain component in a mechanical system, and if we know, that the system has not failed till that time, then the probability of failure after a certain time is same irrespective of the starting point where we have taken. So, this type of thing is interesting because many times we buy, say second hand items, second hand transistor, second hand computer or calculator because if it is still working, then the probability of its failure after a certain time will remain the same. So, this is because of the memoryless property of the exponential distribution.

Let us also consider some modifications of this exponential distribution because here, we are starting from the time 0, but many times, for example, you consider certain item which we are purchasing from the market, then the market, when you are purchasing, then the shopkeeper gives you a guarantee for certain time like 1 year or 2 year etcetera. That means, the item is not supposed to fail before that because if it is failing, he is taking it back. So, it is again as if you are starting afresh.

(Refer Slide Time: 05:33)

Shifted Exponential Distribution (Generalization) ②

$$f(x) = \frac{1}{\sigma} e^{-\frac{(x-\mu)}{\sigma}}, \quad x > \mu, \quad \sigma > 0, \quad \mu \in \mathbb{R}$$

$$E(x-\mu)^k = \int_{\mu}^{\infty} (x-\mu)^k \cdot \frac{1}{\sigma} e^{-\frac{x-\mu}{\sigma}} dx$$

$$= \sigma^k \int_0^{\infty} y^k e^{-y} dy = k! \sigma^k, \quad \begin{matrix} \frac{x-\mu}{\sigma} = y \\ \frac{1}{\sigma} dx = dy \end{matrix}$$

$$E(x-\mu) = \sigma \Rightarrow \mu_1' = E(x) = \mu + \sigma$$

$$E(x-\mu)^2 = 2\sigma^2 \Rightarrow E(x^2) - 2\mu E(x) + \mu^2 = 2\sigma^2$$

So, this is considered as a shifted exponential distribution by certain constant. Another thing we notice here is that, when we considered the form $\lambda e^{-\lambda t}$, where λ was a rate of the Poisson distribution, then the mean of this turnout to be $1/\lambda$. Since, there is a single parameter, if we write, say $1/\lambda$ by σ , then this will become $e^{-x/\sigma}$.

So, in the popular form with the shifted origin, we can consider the form of exponential distribution as, so we call it shifted exponential distribution. The density, we can write as $\frac{1}{\sigma} e^{-\frac{x-\mu}{\sigma}}$. So, here, x is greater than μ and of course, σ is positive. The advantage of this form is that, in place of time, if we are considering some other representation for x ; x need not be time all the time, so if it is something else, then μ can be negative also and then also, this distribution remains valid. So, you can consider, this is a generalization of the original exponential distribution.

If we consider this particular form, then you can see it is easier to calculate the moments of $x - \mu$ or $(x - \mu)/\sigma$. So, let us consider expectation of $(x - \mu)/\sigma$ to the power k which is, of course not necessarily the central moment because we have not shown that μ is the mean. However, for convenience we are considering this. So, it is equal to $\int_{\mu}^{\infty} (x - \mu)^k \frac{1}{\sigma} e^{-\frac{x-\mu}{\sigma}} dx$ from μ to infinity.

So, we may put $x - \mu$ by σ is equal to y , that is 1 by σ dx is equal to dy . Then, this is equal to 0 to infinity σ to the power k , y to the power k , e to the power minus y dy , that is equal to $\Gamma(k+1)$ or $k!$ σ^k to the power k .

So, in particular, if we consider expectation of $x - \mu$ to the power 1 , this is equal to σ . This means, mean of the shifted exponential distribution is $\mu + \sigma$. If we had considered λ there, then it would have been $\mu + 1/\lambda$. So, definitely we would be interested in the variance here. So, expectation of $x - \mu$ square is equal to $2\sigma^2$, which gives us expectation of x square minus 2μ expectation x plus μ square is equal to $2\sigma^2$.

(Refer Slide Time: 08:26)

The image shows a whiteboard with the following handwritten derivations:

$$\Rightarrow E(X^2) = 2\sigma^2 - \mu^2 + 2\mu(\mu + \sigma)$$

$$= 2\sigma^2 + 2\mu\sigma + \mu^2$$

$$\text{Var}(X) = E(X^2) - (E(X))^2 = 2\sigma^2 + 2\mu\sigma + \mu^2 - (\mu + \sigma)^2$$

$$= \sigma^2$$

$$M_X(u) = E(e^{uX})$$

$$= \int_{\mu}^{\infty} \frac{1}{\sigma} e^{ux} \cdot e^{-\frac{x-\mu}{\sigma}} dx$$

$$\text{Var}(X+c) = E\{X+c - E(X+c)\}^2$$

$$= E\{X - E(X)\}^2$$

$$= V(X)$$

So, further simplification gives, expectation of x square is equal to $2\sigma^2$ minus μ square plus 2μ expectation x is $\mu + \sigma$. So, this becomes equal to $2\sigma^2$ plus $2\mu\sigma$ minus μ square plus 2μ square. So, it becomes plus μ square. So, if you calculate the variance of x , that is equal to expectation x square minus expectation x whole square, that is equal to $2\sigma^2$ plus $2\mu\sigma$ plus μ square minus μ plus σ whole square, which is simply σ^2 .

So, variance has not changed by shifting. This is true because the fact that variance is independent of the shift in the origin. The variance of $x + c$ is expectation of $x + c$ minus expectation of $x + c$ whole square. So, this is equal to expectation of x minus expectation x whole square, which is the variance of x . So, the variance is unaffected by

the change in the origin. We can consider here, the moment generating function of this and also, the moment generating function of the original exponential distribution.

(Refer Slide Time: 11:12)

The image shows a handwritten derivation on a blue background. It starts with the integral definition of the moment generating function $M_X(u) = E[e^{ux}] = \int_0^{\infty} e^{ux} f(x) dx$. A substitution $y = \sigma x - \mu$ is used, where $\sigma = 1/\lambda$. The integral is transformed to $\int_0^{\infty} e^{u(\sigma y + \mu)} \sigma e^{-\sigma y} dy$. This simplifies to $e^{\mu u} \int_0^{\infty} e^{y(\sigma u - 1)} dy$. The integral is then evaluated as $\left[\frac{e^{y(\sigma u - 1)}}{\sigma u - 1} \right]_0^{\infty}$. The condition $\sigma u < 1$ (or $u < \frac{1}{\sigma} = \lambda$) is noted. The result is $-\frac{e^{\mu u}}{\sigma u - 1} = \frac{e^{\mu u}}{1 - \sigma u}$. For $u=0$, $M_X(u) = \frac{1}{1 - \sigma u} = \frac{1}{1 - u/\lambda} = \frac{\lambda}{\lambda - u}$, where $u < \lambda$.

So, let us look at this. Moment generating function at a point u , it is equal to expectation of e to the power ux . So, this is equal to $\int_0^{\infty} e^{ux} \sigma e^{-\sigma x} dx$. So, we can consider here, $x - \mu$ by σ is equal to y , the transformation that we made here. So, it is becoming 0 to infinity e to the power u and x becomes $\sigma y + \mu$, e to the power μu and e to the power $y(\sigma u - 1)$.

So, this can be simplified. It is equal to $e^{\mu u} \int_0^{\infty} e^{y(\sigma u - 1)} dy$. This becomes $e^{\mu u} \left[\frac{e^{y(\sigma u - 1)}}{\sigma u - 1} \right]_0^{\infty}$, provided $\sigma u < 1$ or $u < 1/\sigma$, which was basically λ from 0 to infinity. So, at infinity, this becomes 0 because I am taking σu to be less than 1 and at 0 , this becomes 1 . So, you are getting $e^{\mu u} / (\sigma u - 1)$.

In particular, if we had taken μ is equal to 0 , then this $M_X(u)$ would have been $1 / (\sigma u - 1)$. If we put $\sigma = 1/\lambda$, then this will become $1 / (u/\lambda - 1)$, that is, equal to $\lambda / (u - \lambda)$, where there is a minus sign here because when we put 0 , there is a minus sign here. So, this will become actually equal to $e^{\mu u} / (1 - \sigma u)$. So, this will become with a minus sign $1 / (1 - u/\lambda)$.

1 minus sigma u. So, minus is here. So, this is equal to lambda by lambda minus u for u less than lambda.

So, the moment generating function of an exponential distribution, when we consider the original form, that is waiting time starting from the 0, then mgf is lambda by lambda minus mu or 1 by 1 minus sigma mu sigma u. If we consider the starting from mu, then it is e to the power muu by 1 minus sigma. So, here, we can see the relationship between these exponential distributions.

(Refer Slide Time: 13:49)

Let $X \sim \text{Exp}(\mu, \sigma)$, $M_X(u) = \frac{e^{-\mu u}}{1 - \sigma u}$, $u < \frac{1}{\sigma}$

$Y = aX + b$

$$M_Y(u) = E e^{u(aX+b)} = e^{bu} M_X(au)$$

$$= e^{bu} \frac{e^{-a\mu u}}{1 - \sigma a u} = \frac{e^{(a\mu+b)u}}{1 - (a\sigma)u}$$

which mgf of $\text{Exp}(a\mu+b, a\sigma)$.

Theorem: If $X \sim \text{Exp}(\mu, \sigma)$ then $Y = aX + b, a > 0 \sim \text{Exp}(a\mu+b, a\sigma)$.

Linearity property of an Exponential Distribution.

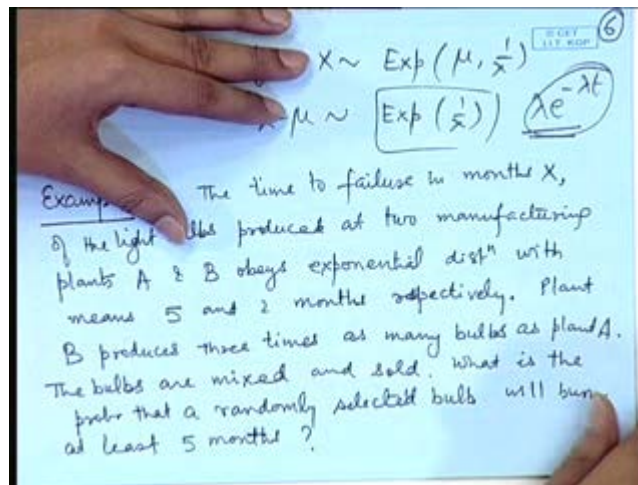
So, let X follow exponential distribution with parameters mu and sigma. Consider, the transformation Y is equal to, say ax plus b. Consider the moment generating function of this, then it is expectation of e to the power u into ax plus b. This is equal to e to the power bu and the moment generating function of x at the point au. So, what is the moment generating function of x at the point u, that we derived it as e to the power muu by 1 minus sigma u for u less than 1 by sigma.

So, if we make use of this term here, in place of u, we can put a mu here. So, it is equal to e to the power aumu divided by 1 minus sigma au. So, this we can adjust and write it as aumu plus bu. See this term was but here it is amu. So, amu, this actually u will interchange here. So, it becomes amu plus b and then, u divided by 1 minus a sigma mu. Compare this the term here, the moment generating function of x is e to the power muu by 1 minus sigma u. So, here, mu is replaced by amu plus b and sigma is replaced

by a sigma. The form is the same. At the same time, since we have replaced λ by λ/σ in the expression for M_X , this expression was valid for $\lambda/\sigma < 1$. So, here, λ/σ must be less than 1. That means λ is less than σ .

So, obviously, you can see by comparing this, that, this is mgf of, which is mgf of exponential distribution with parameter λ/σ and σ . So, this means, we have proved the following theorem. If X follows exponential λ/σ , then $Y = aX + b$, where $a > 0$ follows exponential distributions with parameter λ/σ . This is basically linearity property of an exponential distribution. That means any linear function of an exponential distribution is again having exponential distribution. So, this form is more general. This is a 2 parameter exponential distribution and it is more useful.

(Refer Slide Time: 17:32)



In particular, we will have, if X follows exponential, say λ , then, $X - \mu$ will follow exponential λ . That means the standard form, that is, $\lambda e^{-\lambda t}$. So, both the forms are useful. Let us take one example here for exponential distribution. The time to failure in months, so suppose, it is X of the light bulbs produced at two manufacturing plants, say A and B obeys exponential distribution with means 5 and 2 months respectively. Plant B produces 3 times as many bulbs as plant A . The bulbs are mixed, so they look indistinguishable to a naked eye and sold. What is the probability that a randomly selected bulb will burn at least 5 months?

So, let us consider here, the distribution of X is exponential, but it is not the same for all the bulbs. For a certain proportion of bulbs, the distribution is exponential with mean 5. That means, the density will be $\frac{1}{5} e^{-x/5}$ for $x > 0$ and for certain bulbs, the mean time is 2 months. So, the density will be $\frac{1}{2} e^{-x/2}$ for $x > 0$.

(Refer Slide Time: 20:55)

Handwritten notes on a whiteboard:

- $X \rightarrow$ life of bulb
- $P(X > 0) = e^{-\lambda x}$
- $X/A \sim \frac{1}{5} e^{-x/5}, x > 0$
- $X/B \sim \frac{1}{2} e^{-x/2}, x > 0$
- $P(X > 5) = P(X > 5 | A) P(A) + P(X > 5 | B) P(B)$
- $= e^{-5 \cdot \frac{1}{5}} \cdot \frac{1}{4} + e^{-5 \cdot \frac{1}{2}} \cdot \frac{3}{4} \approx 0.1534$

So, if you are X as the time or life of bulb, then X given, that it is produced by plant A, this has a density $\frac{1}{5} e^{-x/5}$ for $x > 0$ and of course, 0 otherwise. If we are considering the bulb produced by plant B, then the density function is $\frac{1}{2} e^{-x/2}$ for $x > 0$.

We are interested in a randomly selected bulb's life to be more than 5 months. So, here, we can apply the theorem of total probability. So, probability of X greater than 5, given that, it is produced by plant A is equal to the probability of being produced at plant A plus probability of X greater than 5, given that it is produced at plant B into probability of plant B. So, this is equal to, now if we are considering probability of X greater than 5 from this one, then consider the general formula probability of X greater than A is equal to $e^{-\lambda x}$.

So, here λ is equal to $\frac{1}{5}$ and A is equal to $\frac{1}{4}$. So, this becomes $e^{-5 \cdot \frac{1}{5}}$ into $\frac{1}{4}$ into probability of A. Now, what is probability of A? It is given that B produces 3 times as many bulbs as A. So, the probability of the bulb being selected from

A may be 1 by 4 and the probability of this may be 3 by 4. So, this is into 1 by 4 plus e to the power minus 5. Here, the parameter is 1 by 2, 3 by 4. So, this can be simplified and it is 0.1535.

So, probability that a randomly selected bulb is working beyond 5 months is only 0.15 approximately, which may look slightly surprising because you say that, there are 2 plants and from 1 of the plants, it is 2 months and from another plant it is 5 months. So, it should be nearly half, but it is not. So, because the plant B gives more supply compare to the plant A, and in the plant B, the probability of having more than 5 months is much smaller because it is e to the power minus 5 by 2 and here, it is e to the power minus 1. So, this number is, obviously larger than this number.

Now, consider that we are not interested in a single occurrence or a single failure or a single happening in a Poisson process. In place of that we are considering, so consider a Poisson process.

(Refer Slide Time: 24:22)

Consider a Poisson process $X(t)$ with rate λ .
 Let T_r denote the time of r^{th} occurrence.
 T_r ??

$P(T_r > t) = \begin{cases} P(X(t) \leq r-1), & t > 0 \\ 1, & t \leq 0 \end{cases}$

$= \begin{cases} \sum_{j=0}^{r-1} \frac{e^{-\lambda t} (\lambda t)^j}{j!}, & t > 0 \\ 1, & t \leq 0 \end{cases}$

Let us denote X_t with rate λ 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 r th occurrence. So, in place of the first occurrence, we are looking at a certain number of occurrences. As we have 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 a certain mishaps may close down the plant itself. A sequence of so many smaller kinds of events may lead to a major disaster. So, we may be interested in the waiting time for that, of course, considering the events to occur in a Poisson process. For example, a certain number of occurrences have occurred, say a certain number of people purchase certain tickets. Then, we may have to close down the window because the seats are full.

So, what is a distribution of T_r ? Once again, we can consider probability of T_r greater than t . So, starting from a time 0, consider time small t . If we say that r th 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 occurrences 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 λt . So, this is equal to $e^{-\lambda t}$ to the power λt $\lambda^j t^j$ by j factorial, summation j is equal to 0 to r minus 1. This is for t greater than 0 and it is 1 for t less than or equal to 0. 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 $e^{-\lambda t}$ to the power λt $\lambda^j t^j$ by j factorial for t greater than 0.

(Refer Slide Time: 27:26)

Handwritten notes on a whiteboard:

$$F_{T_r}(t) = 1 - P(T_r > t)$$

$$= \begin{cases} 0, & t \leq 0 \\ 1 - \sum_{j=0}^{r-1} \frac{e^{-\lambda t} (\lambda t)^j}{j!}, & t > 0 \end{cases}$$

The pdf of T_r is

$$f_{T_r}(t) = \frac{d}{dt} F_{T_r}(t) = 0, \quad t \leq 0$$

$$= -\frac{d}{dt} \left[e^{-\lambda t} + (\lambda t) e^{-\lambda t} + \frac{(\lambda t)^2 e^{-\lambda t}}{2!} + \dots + \frac{e^{-\lambda t} (\lambda t)^{r-1}}{(r-1)!} \right], \quad t > 0$$

So, 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 sum of a series. So, we have to 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 dt of e to the power minus λt plus λt into e to the power minus λt plus λt square e to the power minus λt by 2 factorial and so on, plus e to the power minus λt , λt to the power r minus 1 by r minus 1 factorial for t greater than 0.

So, here, you observe that when we differentiate here one term is there, but there after each term is a product of two terms involve in t . So, when we differentiate, we have to do by applying the formula of derivative of a product. So, if we expand the terms, the derivative of the first term will give us minus λe to the power minus λt and there is a minus sign here. Consequently, we will get it as λe to the power minus λt .

(Refer Slide Time: 29:41)

$$\begin{aligned}
&= \lambda e^{-\lambda t} - \lambda^2 t e^{-\lambda t} + \lambda^3 t^2 e^{-\lambda t} - \lambda^4 t^3 e^{-\lambda t} + \dots + \frac{\lambda^r t^{r-1} e^{-\lambda t}}{(r-1)!} \\
f_{T_r}(t) &= \frac{\lambda^r t^{r-1} e^{-\lambda t}}{\Gamma(r)} \quad t > 0, \quad \lambda > 0, \quad r > 0 \\
&\text{Gamma or Erlang's Distribution} \\
E(T_r^k) &= \int_0^{\infty} t^k \cdot \frac{\lambda^r t^{r-1} e^{-\lambda t}}{\Gamma(r)} dt \\
&= \frac{\lambda^r}{\Gamma(r)} \int_0^{\infty} t^{k+r-1} e^{-\lambda t} dt = \frac{\lambda^r}{\Gamma(r)} \cdot \frac{\Gamma(k+r)}{\lambda^{k+r}} =
\end{aligned}$$

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 stand cancelled out. The next term will be the derivative of this, so that will give minus lambda and minus is here, 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 you look at the derivative of this one, here t square is coming. So, 2t and this 2 will cancel out and there will be a minus sign. So, we will get with, because there as a minus sign outside, it becomes minus lambda square t e to the power minus lambda t. So, you can easily observe, that this is a telescopic sum.

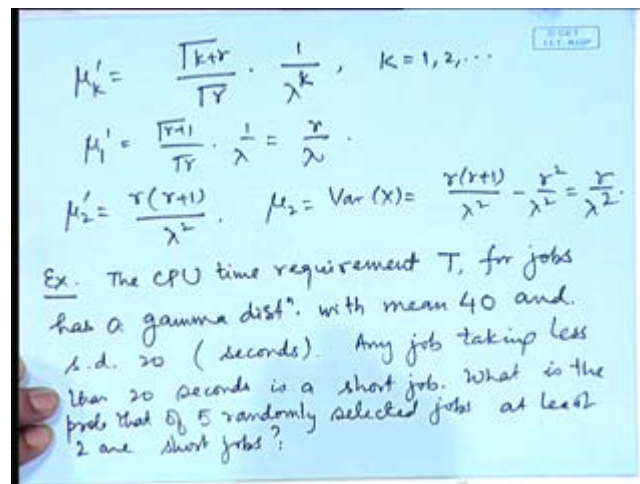
The final term will be plus, which will be contributed by this term that will become lambda to the power r t to the power r minus 1 by r minus 1 factorial e to the power minus lambda t. So, this will be the left out term. So, we are getting the probability density function of T r as lambda to the power r t to the power r 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 an integer because of 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 or Erlang's distribution.

So, this distribution arises as the distribution of the waiting 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 and you will get lambda e to the

power minus lambda t. So, 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 applications, one can see the characteristics of this distribution. If we consider a kth order non-central moment, that is μ_k' . So, μ_k' is equal to integral from 0 to infinity $t^k \lambda^r t^{r-1} e^{-\lambda t} dt$. So, obviously, this is a gamma function you can write it as $\lambda^r \int_0^\infty t^{k+r-1} e^{-\lambda t} dt$. So, it is $\lambda^r \Gamma(k+r) / \lambda^{k+r} = \Gamma(k+r) / \lambda^k$.

(Refer Slide Time: 33:38)



This is equal to, that is μ_k' is equal to $\Gamma(k+r) / \lambda^k$. Of course, this is valid for any k greater than 0, but we will be more concerned with the positive integral moments. So, μ_1' is equal to $\Gamma(r+1) / \lambda = r / \lambda$, that is, r by λ . So, if the waiting time for the first occurrence was average, waiting time for the first occurrence was $1 / \lambda$, then for the r th occurrence, it will be r / λ . See this is again because of the memoryless 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, waiting time will be $1 / \lambda$ and so on.

If we consider μ^2 prime, then this becomes r into $r + 1$ by λ square. Therefore, μ^2 , that is a variance of this distribution is equal to $r + 1$ by λ square minus r square by λ square, that is r by λ square which is again, you can see in the exponential distribution, the variance was 1 by λ square and if you are looking at the variance of the r th occurrence, then it becomes r by λ square.

Let us look at one application of this. The CPU time requirement T for jobs has a gamma distribution with mean 40 and 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?

(Refer Slide Time: 37:21)

$\frac{r}{\lambda} = 40, \quad \frac{r}{\lambda^2} = 400 \Rightarrow r = 4, \quad \lambda = \frac{1}{10}$

$P(T < 20) = \int_0^{20} \frac{1}{14 \cdot 10^4} e^{-t/10} \cdot t^3 dt$

$f_T(t) = \frac{1}{14 (10^4)} e^{-t/10} \cdot t^3, \quad t > 0$

$= 1 - \int_{20}^{\infty} \frac{1}{6 \cdot 10^4} e^{-t/10} t^3 dt$

$= 1 - \int_2^{\infty} \frac{1}{6} e^{-y} y^3 dy = 1 - \frac{19}{3} e^{-2} \approx 0.1429$

$\frac{t}{10} = y$
 $\frac{1}{10} dt = dy$

So, in order to answer this question, let us consider the setup. Here, mean is given to be 40, so that r by λ is 40 and standard deviation. So, here variance is r by λ square. So, r by λ square is 400. So, this is a 2 parameter distribution and we have now 2 equations. So, we can have r by λ is equal to 40, r by λ square is equal to 400. So, if we solve this, we get r is equal to 4 and so λ is equal to $1/10$. So, that is right. What is a probability of a short job? That means T is less than 20.

So, we need to consider the density function of T . f_T is equal to e to the power minus λt . So, that is t by 100, t to the power $r - 1$, that is, 3, then λ to the power r , that is $1/10$ 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 dt . This is the probability of a randomly selected job to be a short job.

So, if you 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. So, it is convenient if we consider this has 1 minus 20 to infinity, 1 by 6 into 10 to the power 4 , e to the power minus t by 10 , t cube dt . So, from the integrals integrand, we can observe that is convenient if we put t by 10 is equal to y , that is, 1 by $10dt$ is equal to dy . So, this becomes equal to 1 minus, this will be 2 to infinity 1 by 6 , e to the power minus y cube dy .

So, this is not a complete gamma function. This is an incomplete gamma function because integral is from 2 to infinity. 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 . So, 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 μ minus σ is actually much smaller.

(Refer Slide Time: 41:03)

$Z \rightarrow$ no of short jobs
 $Z \sim \text{Bin} \left(\underset{N}{5}, \underset{p}{0.1429} \right)$
 $P(Z \geq 2) = 1 - P(Z=0) - P(Z=1)$
 $= 1 - (1 - 0.1429)^5 - 5(1 - 0.1429)^4(0.1429)$
 ≈ 0.1519
 $M_{T_r}(u) = E(e^{uT_r}) = \int_0^{\infty} e^{ut} \cdot \frac{\lambda^r}{\Gamma(r)} e^{-\lambda t} t^{r-1} dt$

Now, we can consider the random variable, say Z is the number of short jobs out of 5 , then Z is following by 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 the 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 complementary event. So, this is equal to 0.1 minus 0.1429 to the power 5 minus 5 C 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 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 uT r, that is equal to integral e to the power ut, lambda to the power r by gamma r, e to the power minus lambda, t to the power r minus 1 dt from 0 to infinity.

So, here, if you see this term can be combined with this. So, it becomes directly a gamma function.

(Refer Slide Time: 43:19)

$$= \frac{\lambda^r}{\Gamma(r)} \int_0^\infty e^{-t(\lambda-u)} t^{r-1} dt$$

$$\left(\frac{\lambda}{\lambda-u}\right)^r, \quad u < \lambda.$$

Weibull Distribution

$$f(x) = \begin{cases} \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} & x > 0 \\ 0 & \text{ew.} \end{cases} \quad \alpha > 0, \beta > 0$$

$$F(x) = \begin{cases} \int_0^x f(t) dt & x > 0 \\ 0 & x \leq 0 \end{cases}$$

So, 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 moment generating function of the exponential distribution, where r was 1. So, it was lambda by lambda

minus u . So, 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 μ_k' . So, μ_3' , μ_4' can be calculated and then, μ_3 and μ_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 density function, it is $\lambda^r t^{r-1} e^{-\lambda t}$ divided by, of course $\Gamma(r)$.

So, 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. Then, in the beginning, since it is t to the power $r-1$, it may increase little bit, but there after this term will dominate. Therefore, the densities will always be positively skewed for various gamma distributions. Of course, it will depend upon, what is the value of r and what is the value of λ for different shapes, but whatever be the shapes, they will be positively skewed. The P equal of course, depend upon the value of the r and λ .

If we are considering the exponential distribution or the gamma distribution as the life of certain components, then another interesting distribution in the same direction is so-called Weibull distribution. The general form of a probability density function of a Weibull distribution is given by $\alpha \beta x^{\beta-1} e^{-\alpha x^\beta}$, where x is always positive and α and β are positive parameters.

So, quite naturally, one can see the form of the CDF here, because if you integrate from, since x is positive valued random variable, the integral will be from 0 to x $f(t) dt$ for x positive and it is 0, of course for x less than or equal to 0. So, if you consider this integrand, it looks like a derivative of $e^{-\alpha x^\beta}$. Therefore, when we calculate the CDF, this will be simply equal to $1 - e^{-\alpha x^\beta}$ for x positive and it is 0 for x less than or equal to 0.

(Refer Slide Time: 47:16)

$$f = \begin{cases} \alpha - e^{-\alpha x^\beta} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$M_k' = E(X^k) = \int_0^{\infty} \alpha \beta x^{\beta+k-1} e^{-\alpha x^\beta} dx$$

$$= \int_0^{\infty} \alpha y^{\frac{\beta+k}{\beta}} e^{-\alpha y} dy$$

$$= \frac{\alpha \frac{\Gamma(\frac{\beta+k}{\beta} + 1)}{\beta + 1}}{\alpha \frac{\beta}{\beta + 1}} = \frac{\Gamma(\frac{\beta+k}{\beta})}{\alpha^{\frac{\beta+k}{\beta}}}$$

$\alpha x^\beta = y$
 $\beta x^{\beta-1} dx = dy$

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. So, 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.

(Refer Slide Time: 49:43)

$$M_1' = E(X) = \frac{\alpha \frac{\Gamma(\frac{\beta+1}{\beta})}{\beta + 1}}{\alpha^{\frac{\beta+1}{\beta}}} = \alpha^{-\frac{1}{\beta}} \frac{\Gamma(\frac{\beta+1}{\beta})}{\beta}$$

$$M_2' = \alpha^{-\frac{2}{\beta}} \frac{\Gamma(\frac{\beta+2}{\beta})}{\beta}$$

$$M_2 = \alpha^{-\frac{2}{\beta}} \left[\frac{\Gamma(\frac{\beta+2}{\beta})}{\beta} - \left(\frac{\Gamma(\frac{\beta+1}{\beta})}{\beta} \right)^2 \right]$$

$T \rightarrow$ life of a system
 $P(T > t) = R(t) \rightarrow$ Reliability of the system at time t

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 we put x to the power beta is equal to y , that is βx to the power beta minus 1 is equal to dx is equal to dy , then this becomes integral from 0 to infinity $\alpha x^{\beta-1} dx$ to the power beta minus one term is combined here. So, you have x to the power k , that is y to the power k/β , e^{-y} to the power minus αy dy . So, this is simply a gamma function and the expression for this turns out to be $\alpha \Gamma(k/\beta + 1)$ divided by α to the power $k/\beta + 1$. That means $\Gamma(k/\beta)$ or you can write it as k/β plus β by β divided by α to the power k/β .

So, in particular, if I am looking at the mean, μ_1' is equal to expectation of x , that is equal to $\alpha^{-1/\beta} \Gamma(1 + 1/\beta)$ divided by $\alpha^{-1/\beta} \Gamma(1 + 1/\beta)$, that is equal to $\alpha^{-1/\beta} \Gamma(1 + 1/\beta)$ divided by $\alpha^{-1/\beta} \Gamma(1 + 1/\beta)$ and μ_2' will become $\alpha^{-2/\beta} \Gamma(2 + 1/\beta)$ divided by $\alpha^{-2/\beta} \Gamma(2 + 1/\beta)$. Therefore, the variance of the Weibull distribution will be, $\alpha^{-2/\beta} \Gamma(2 + 1/\beta)$ divided by $\alpha^{-2/\beta} \Gamma(2 + 1/\beta)$ minus $\alpha^{-2/\beta} \Gamma(1 + 1/\beta)$ 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 from calculators or computers. **Nowadays, this can be easily.**

However, this distribution has more importance in the reliability analysis of certain mechanical or electronic systems. So, let us define, what you mean by the reliability of a 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 has not failed 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.

(Refer Slide Time: 52:11)

Instantaneous Failure Rate of System at time t

$$\lim_{h \rightarrow 0} \frac{1}{h} P\left(\frac{t < T \leq t+h}{A} \mid \frac{T > t}{B}\right) = H(t)$$

\downarrow
hazard rate
at time t

$$= \lim_{h \rightarrow 0} \frac{1}{h} \cdot \frac{P(A \cap B)}{P(B)} = \lim_{h \rightarrow 0} \frac{P(t < T \leq t+h)}{h \cdot P(T > t)}$$

$$\lim_{h \rightarrow 0} \frac{\frac{F_T(t+h) - F_T(t)}{h}}{R(t)} = \frac{f_T(t)}{R(t)}$$

$$H(t) = \frac{f_T(t)}{1 - F_T(t)} = -\frac{d}{dt} \log(1 - F_T(t))$$

So, 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 .

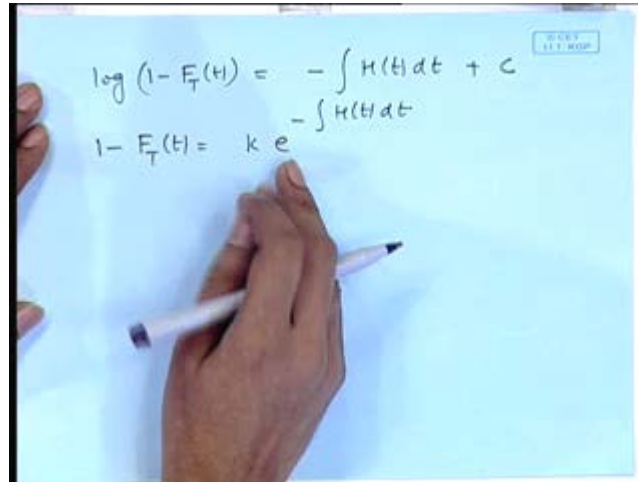
We also define, what is known as Instantaneous Failure Rate of System at time t . So, let us consider the interpretation of this. The system is functioning at time t and immediately after the time t ; it fails, that means in an interval from T to t plus h . So, if we are considering the rate, we consider it by divided by h and take limit as h tends to 0. Let us give some notations, say h of t ; we also call it hazard rate at time t . So, 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, given that the system is functioning at time t , what is the probability that it will fail immediately after that. Therefore, we want to calculate the rate. Then, we divide by the length of the interval and take the limiters. So, let us evaluate this.

Now, if we 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 can see that 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 we 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 $R(t)$ or you can consider it as f_T divided by $1 - F_T$. So, the hazard rate of a lifetime

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.

(Refer Slide Time: 55:33)



The image shows a hand holding a white marker writing on a whiteboard. The equations written are:

$$\log(1 - F_T(t)) = - \int H(t) dt + C$$
$$1 - F_T(t) = k e^{- \int H(t) dt}$$

So, we are calling it as $H(t)$. We can notice another relationship here. This is minus d by dt of $\log(1 - F_T)$. So, this looks like a first order differential equation. We can integrate it out and we will get, $\log(1 - F_T)$ as equal to and therefore, plus of course, the constant of integration. So, $1 - F_T$ becomes some k times into the power minus integral $H(t) dt$. 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 these quantities, that is, the reliability, the hazard rate, in context of the Weibull distribution, the exponential distribution and try to see what it signifies. So, in the next lecture, we will be covering these issues. Thank you.