PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

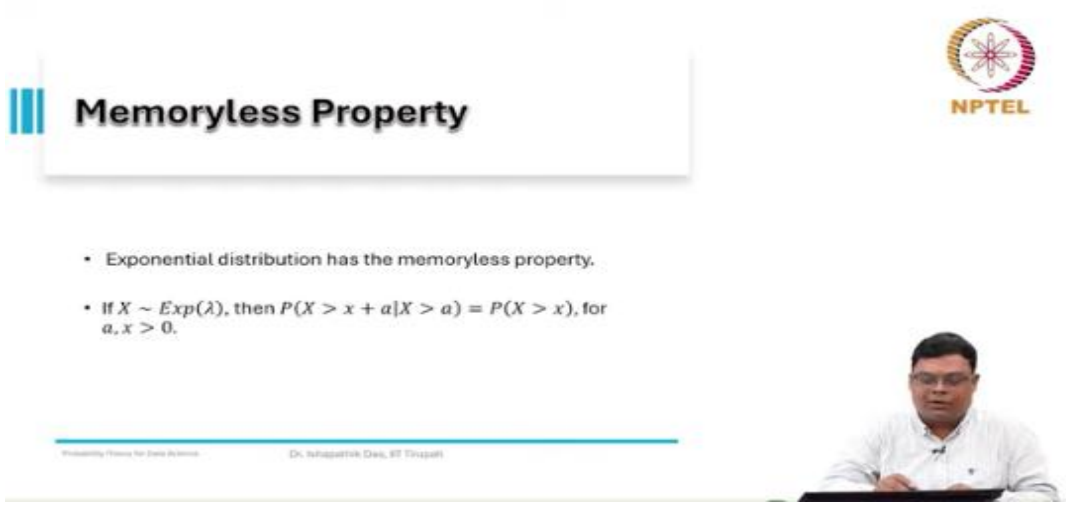
Indian Institute of Technology Tirupati

Week - 05

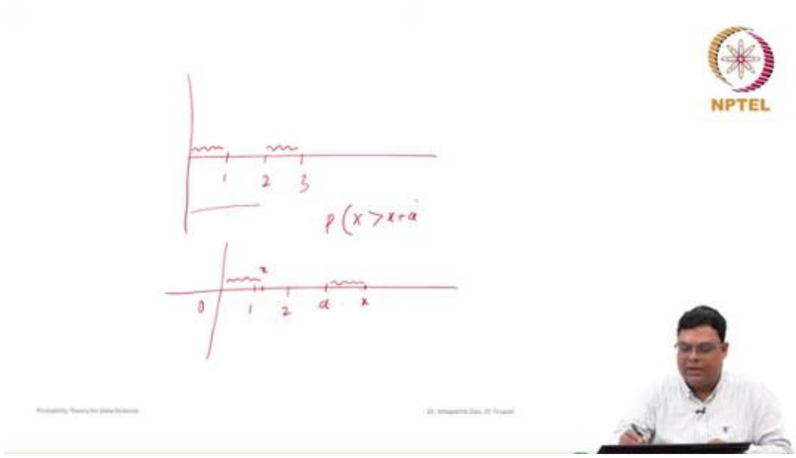
Lecture - 25

Memoryless Property and Gamma Distribution

The exponential distribution has a special property known as the memoryless property. It's typically used to model lifetime data. For instance, if we buy an electronic product and want to know its lifetime, it could be 1 year, 2 years, or more. Now, suppose the product has already lasted for 2 years, and we want to know the probability that it will last for another year. The probability of it lasting for an additional year is the same as if we were asking about the probability of it lasting 1 year from the beginning. This means that the random variable X, or the phenomenon we're observing, doesn't "remember" that it's already lasted for 2 years. It's like forgetting, because usually, if something has already worked for 2 years, we assume the probability of it working for another year is lower than when it was new. But with this kind of random variable, it doesn't take the past into account. This property is known as the memoryless property. So, the exponential distribution has the memoryless property.



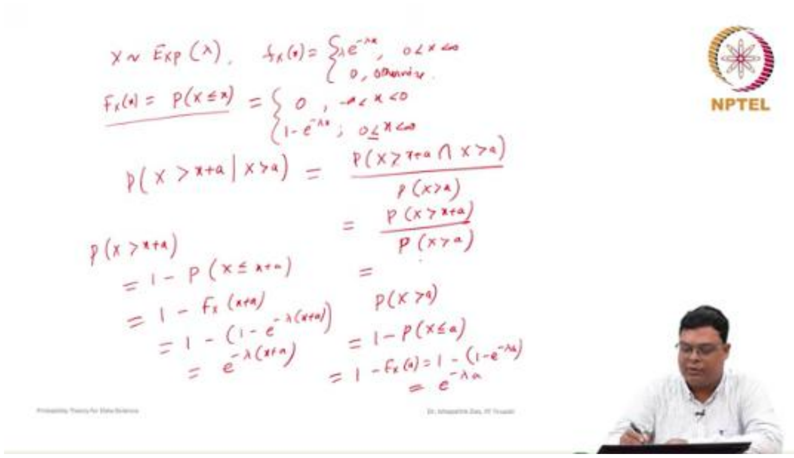
If X has an exponential distribution with parameter lambda, then the probability that X > x + a, given that X > a, is the same as the probability that X > x for any real number x and small a. This means that for any real number x, the probability that X > x + a is the same as the probability that X > x, even when a is a small value. Exponential distribution has the memoryless property. If X follows an exponential distribution with parameter lambda, then the probability that X > x + a, given that X > a, is the same as the probability that X > x for any real numbers x and a. This means that the probability of survival for a certain additional time x is independent of the time that has already passed a.



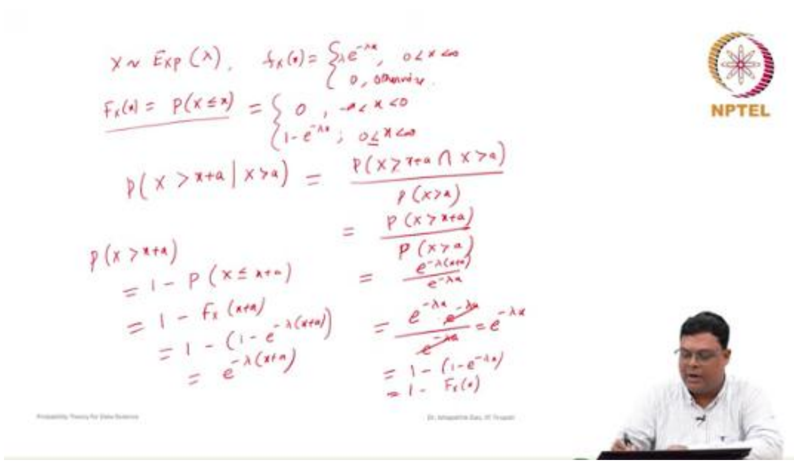
For example, if an electronic product has already lasted for a certain period, say 2 years, the probability that it will last for another year is the same as if it were brand new. This property is known as the memoryless property. To verify that the exponential distribution exhibits this property, we recall that when X follows an exponential distribution with parameter λ, the probability density function is e^(-λ x) for x > 0, and 0 otherwise. From this, we can derive the cumulative distribution function, which is 1 - e^(-λ x) for x ≥ 0. Now, let's calculate the probability that X > x + a, given that X > a.

This is a conditional probability, expressed as the intersection of the events "X > x + a" and "X > a", divided by the probability that X > a. Since "X > x + a" implies "X > a", the intersection is just the event "X > x + a." Therefore, we have the probability that X > x + a, divided by the probability that X > a. To calculate these probabilities, we use the fact that the probability that X > x + a is 1 - CDF(x + a), which simplifies to e^(-λ(x + a)). Similarly, the probability that X > a is 1 - CDF(a), which simplifies to e^(-λ a).

So, this is e^(-λ(x + a)) / e^(-λ a). Let’s erase everything we’ve written before because we’ve already computed this value, and we’ll use it. It simplifies to e^(-λ a). Now, this is (e^(-λ x) * e^(-λ a)) / e^(-λ a). The e^(-λ a) terms cancel out, so we’re left with e^(-λ x).



And what is $e^{(-\lambda x)}$? This can be represented as $1 - (1 - e^{(-\lambda x)})$, which is the same as $1 - F_X(x)$. So, this is basically the probability that $X \leq x$, which is $1 - P(X > x)$. This is the relationship we obtained. We can see that the exponential distribution has this memoryless property.

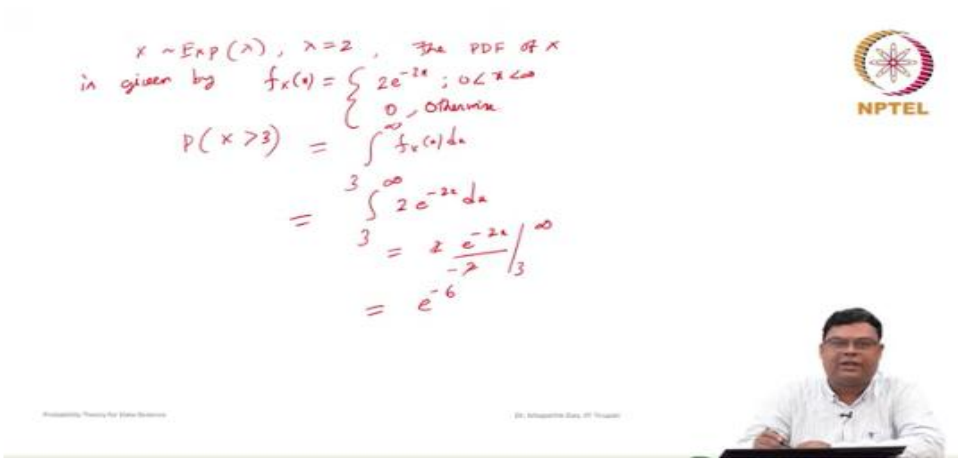


That's why we found that the probability that $X > x + a$, given that $X > a$, is the same as the probability that $X > x$. This is why the exponential distribution is called memoryless. Let's discuss a numerical example. Suppose, during the rainy season on a tropical island, the length of the shower has an exponential distribution with a rate parameter $\lambda = 2$, with time measured in minutes. So, $\lambda = 2$.

The question is: what is the probability that the shower will last more than 3 minutes? We know that X follows an exponential distribution with $\lambda = 2$. The probability density function is given by $\lambda e^{(-\lambda x)}$, where $x > 0$ and $x < \infty$, and 0 otherwise. Since $\lambda = 2$, this becomes 2 e^(-2x) for x > 0. Now, we want to find the probability that the shower will last more than 3 minutes.

We need to find the probability that X > 3. This is the same as the integral from 3 to infinity of the probability density function. Alternatively, we can use the complement and apply the cumulative distribution function. So, we have the probability that X > 3, which is the integral from 3 to infinity of 2 e^(-2x). This simplifies to -e^(-2x), evaluated from 3 to infinity.

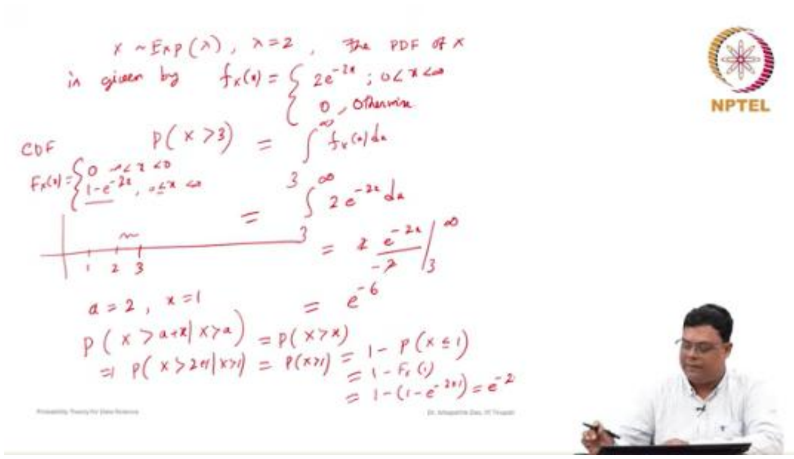
When x → ∞, this becomes 0, and when x = 3, we get e^(-6). So, the probability is e^(-6). So, this is the probability that X > 3. The question is asking what is the probability that the shower will last more than 3 minutes. If the shower has already lasted for 2 minutes, what is the probability that it will last at least 1 more minute?



Here, it is given that the shower has already lasted 2 minutes, so a = 2. The question is asking for the probability that it will last at least 1 more minute, which means the probability that X > 2 + 1. This is the probability that X > 3. By the memoryless property, we know that the probability is the same as if we were asking about X > 1. So, the probability that X > 2 + 1, given that X > 2, is the same as the probability that X > 1.

Therefore, the probability that X > 3, given X > 2, is the same as the probability that X > 1. We can find this probability using the cumulative distribution function (CDF). The CDF is 0 for values of X < 0, and 1 - e^(-λ x) for values of X ≥ 0. Since λ = 2, the CDF is 1 - e^(-2x). To find the probability that X > 1, we substitute x = 1 into the CDF formula: 1 - e^(-2).

This simplifies to e^(-2). So, the probability that X > 1 is e^(-2). So, this is the answer. This is one example we discussed for the exponential distribution, and I hope you have understood its properties. We have also discussed the gamma function, which will be useful as we now move on to the gamma distribution, another important continuous random variable that is widely used.



A continuous random variable X is said to follow a gamma distribution with parameters α > 0 and λ > 0, denoted as X ~ Gamma(α, λ), if its probability density function is given by:

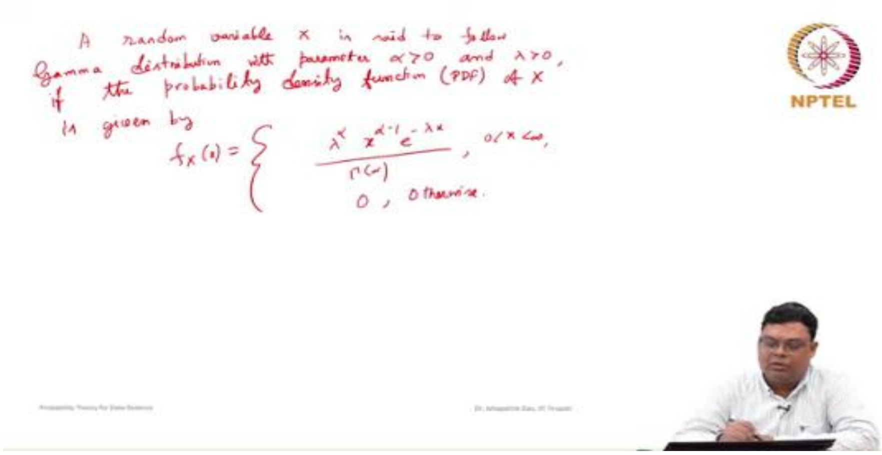
f_X(x) = (x^(α - 1) * e^(-λ x) * λ^α) / Γ(α)

where Γ(α) is the gamma function we've already discussed. While there are different variations of this form, we will mostly use this standard form. This variation can also be derived by reparameterizing with k = α and θ = 1 / λ. You can also see curves for different values of k and θ in the gamma density function and its cumulative distribution function. Now, let's look at the important properties of the gamma distribution.

A random variable X follows a gamma distribution with parameters α and λ, where both parameters are greater than zero. These parameters are real and positive numbers. The probability density function of X is given by:

f_X(x) = (x^(α - 1) * e^(-λ x) * λ^α) / Γ(α) for x > 0, and 0 otherwise.

So, this is the probability density function. Now, we want to check whether it follows all the properties of a density function.



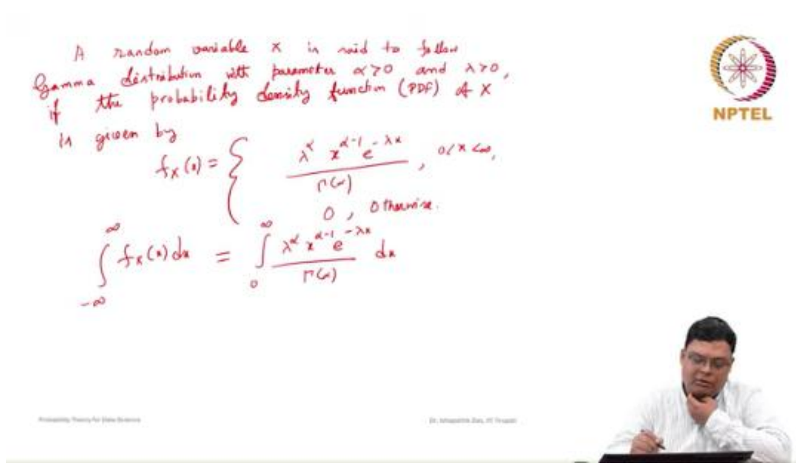
As you can see, it is always greater than or equal to zero, so that's fine. The next thing we need to verify is whether the integral of $f(x)\,dx$ equals 1. When you do this integration, you'll see that this density function is non-zero over the interval from 0 to infinity. So, this becomes:

$$\int_0^\infty \frac{\lambda^\alpha x^{\alpha - 1} e^{-\lambda x}}{\Gamma(\alpha)}\, dx$$

Now, how can we solve this integration?
We will use the gamma function formula we discussed earlier. You can go through it again if needed. The gamma function $\Gamma(\alpha)$ is defined as:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha - 1} e^{-x}\, dx$$

We will use all these properties here. Now, suppose we want to do this integration.



First, recall the gamma α formula. $\Gamma(\alpha)$ is the integral from 0 to infinity of $x^{(\alpha - 1)} * e^{(-x)}\, dx$. Now, we have $e^{(-\lambda x)}$ in the equation, but the gamma function has $e^{(-x)}$. So, we will take a transformation. Let $z = \lambda x$, meaning $x = z / \lambda$. This gives us $dx = dz / \lambda$.

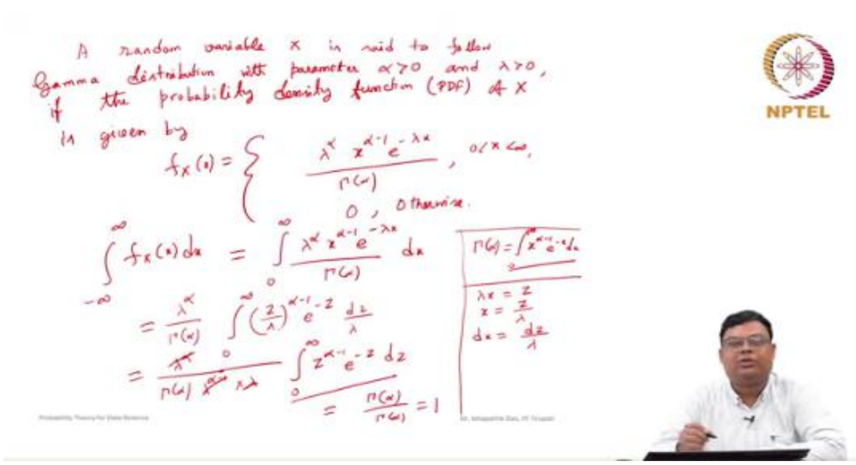
When x goes to 0, z goes to 0, and when x goes to infinity, z goes to infinity as well. Now, we can take the constant λ^α outside the integral, and we get $\Gamma(\alpha)$ in the denominator. The integral becomes:

$(\lambda^\alpha) / \Gamma(\alpha)$

Now, substitute x with z / λ, which gives us $z^{(\alpha - 1)} * e^{(-z)}$, and dx becomes dz / λ. The λ terms cancel out, and we are left with:

$\int$ from 0 to infinity of $z^{(\alpha - 1)} * e^{(-z)}\, dz$

This matches the $\Gamma(\alpha)$ formula, so we end up with $\Gamma(\alpha) / \Gamma(\alpha) = 1$. So, it is a probability density function. Now, we will discuss some of the important characteristics. One of the important characteristics is the moments. So, we will compute the first-order moments.



What is μ_1'? μ_1' is the expected value of X. The expected value of X is defined by the integral from minus infinity to plus infinity of $x\, f_x(x)\, dx$. Since the function is non-zero only when x is between 0 and infinity, the integral becomes from 0 to infinity. The formula for $f_x(x)$ is:

$f_x(x) = (\lambda^\alpha * x^{(\alpha - 1)} * e^{(-\lambda x)}) / \Gamma(\alpha)$

So, how can we find it? Now, if you want to find the variance, you'll need μ_2' as well. Typically, we compute μ_r' for r = 1, 2, and so on. μ_r' is:

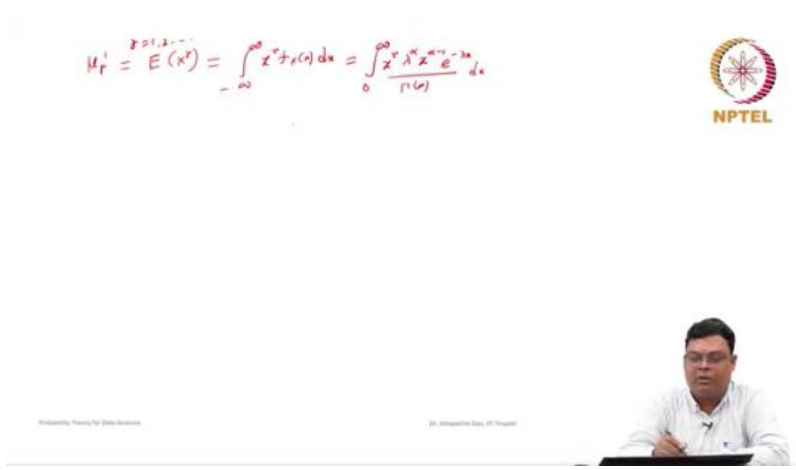
μ_r' = integral from 0 to infinity of x^p * $f_x(x)$ dx

This will be useful because, if we simplify it, we can find μ_1' when r = 1, μ_2' when r = 2, and so on.

So, we simplify these terms. Since (λ^α) / Γ(α) is a constant, it comes outside the integral. The integral becomes from 0 to infinity, and we have:

integral from 0 to infinity of x^(r + α - 1) * e^(-λ x) dx

Γ(α) stays outside. The integral is with respect to dx.



Now, we want to write this in the form of a gamma function. We know that the gamma function is defined as:

Γ(α) = $\int_0^\infty$ x^(α - 1) e^(-x) dx

Since we have e^(-λx) here, we apply the transformation λx = z. So, x becomes z / λ, and dx becomes dz / λ. The limits don't change. As x approaches 0, z also approaches 0, and as x approaches infinity, z also approaches infinity. Now, we replace x with z / λ. This gives us:

z^(r + α - 1) e^(-z) dz / λ

This leads to λ^α / Γ(α). Notice that the λ term becomes λ^(r + α). So, it will be:

$\int_0^\infty z^{(r + \alpha - 1)} e^{(-z)} dz$

Now, comparing these two, we have:

$\int_0^\infty x^{(\alpha - 1)} e^{(-x)} dx$

and

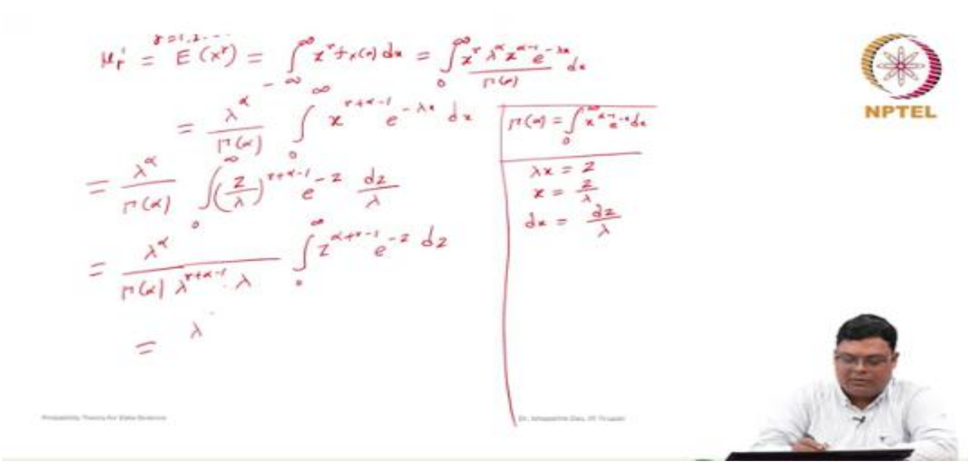
$\int_0^\infty z^{(\alpha + r - 1)} e^{(-z)} dz$

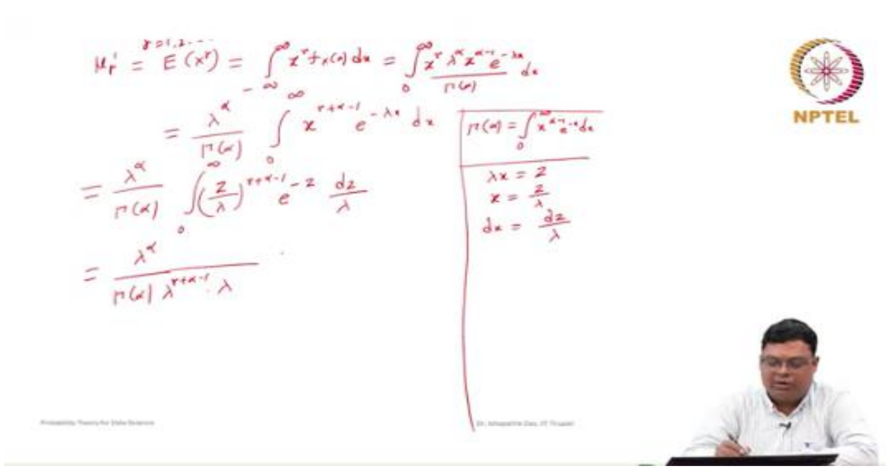
This is nothing but $\Gamma(\alpha + r)$. So, when you simplify these things, it becomes:

$\lambda^{(r + \alpha)} / \Gamma(\alpha)$

So, it actually simplifies to:

$1 / \Gamma(\alpha) * \Gamma(\alpha + r)$





You can see here $\Gamma(\alpha)$, then λ will be raised to the power of $r + \alpha - 1 + 1$, and then $- \alpha$. So, this cancels out and we're left with λ^r. This is nothing but $\Gamma(\alpha + r) / \Gamma(\alpha) * \lambda^r$ for different values of r. Now, what we found is that μ_r', which is the expected value of x^r, is coming out as $\Gamma(\alpha + r) / \Gamma(\alpha) * \lambda^r$. For r = 1, 2, and other integer values, when r = 1, we get μ_1' as $\Gamma(\alpha + 1) / \Gamma(\alpha) * \lambda$. We just discussed the formula, and $\Gamma(\alpha + 1)$ is $\Gamma(\alpha) * \alpha$, so it simplifies to α / λ.

Now, for r = 2, μ_2' is $\Gamma(\alpha + 2) / \Gamma(\alpha) * \lambda^2$. So, after simplification, this is $\alpha * (\alpha + 1) / \lambda^2$. Now, for variance, the variance of the gamma random variable is computed as $\mu_2' - (\mu_1')^2$. μ_2' is $\alpha * (\alpha + 1) / \lambda^2$ and μ_1' is α / λ^2. Simplifying, we get $\alpha * (\alpha + 1) / \lambda^2 - (\alpha^2 / \lambda^2)$. This simplifies to α / λ^2. So, the variance of the gamma random variable is α / λ^2. The mean of the gamma random variable is α / λ.