PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

Week - 06

Lecture - 27

Properties of Normal Distributions

Let X be a normally distributed random variable with mean μ and variance σ^2, then the probability density function is given by $f_X(x) = (1 / \sqrt{2\pi\sigma}) * e^{-(x - \mu)^2 / (2\sigma^2)}$.

What is the mean (expected value) of X? Let X be a normally distributed random variable with mean μ and variance σ^2. The probability density function is given by $f_X(x) = (1 / \sqrt{2\pi\sigma}) * e^{-(x - \mu)^2 / (2\sigma^2)}$.

What is the mean, or expected value, of X? $\mu' = \int$ from $-\infty$ to $+\infty$ of $x * f_X(x)\, dx$.

So, this is the formula we know. This is nothing but $\int$ from $-\infty$ to $+\infty$ of $x * f_X(x)$. Then, we put $(1 / \sqrt{2\pi\sigma}) * e^{-(x - \mu)^2 / (2\sigma^2)}\, dx$.

Now, we have to solve this integral. So, first, we will take a similar type of transformation.

The transformation is where we take $z = (x - \mu) / \sigma$. So, we have $z = (x - \mu) / \sigma$. So, $x = \sigma z + \mu$, and $dx = \sigma\, dz$. Now, as $x \to -\infty$, $z \to -\infty$ because σ is positive. So, as $x \to +\infty$, $z \to +\infty$.

Then, what is x? $x = \sigma z + \mu$, and $(1 / \sqrt{2\pi\sigma})$ is constant. $e^{-(x - \mu) / \sigma}$ is z, so $(x - \mu)^2 / \sigma^2 = z^2 / 2$. Then, $dz = \sigma\, dz$. So, the σ cancels out here.

Now, what have we found? We need to find these things. So, it is nothing but $(1 / \sqrt{2\pi\sigma})$. This is $(1 / \sqrt{2\pi})$, and the integral goes from $-\infty$ to $+\infty$. Then we have σ, σz, so we take σ outside. This is $\sigma / \sqrt{2\pi}$, times $z * e^{-z^2 / 2}\, dz$.

The remaining part is from $-\infty$ to $+\infty$, $\mu * (1 / \sqrt{2\pi}) * e^{-z^2 / 2}\, dz$.

So, now, what we have is the simplification we found because it is $\sigma z + \mu$. So, $\sigma * z * e^{-z^2 / 2}\, dz$, minus $e^{+\infty}$, $(1 / \sqrt{2\pi})$.

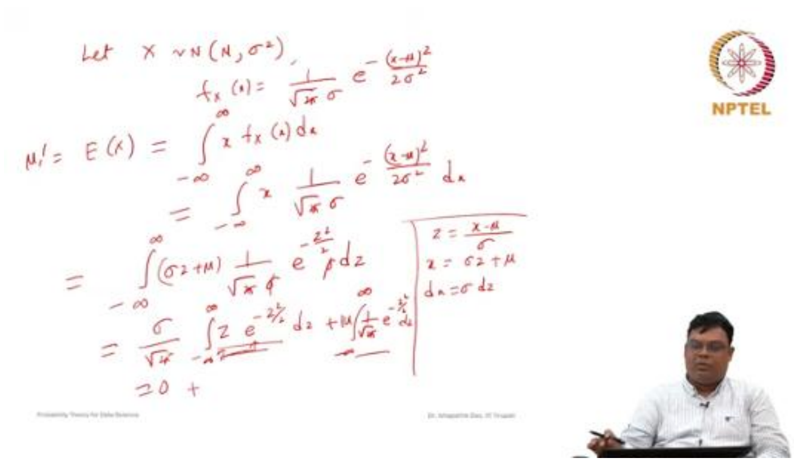
Now, you can see that this is an even function, and this is an odd function. So, the product of an odd and even function is an odd function. So, that's why it is an odd function.

For this odd function, we discussed that the integral will be equal to 0. Specifically, this integral from $-\infty$ to $+\infty$ of $f_X(x)\,dx$ will be equal to 0. So, what did we find? This value will be 0.What will be the value for this function?

So, let us see what the value will be. We have already computed that whenever X is normally distributed with mean μ and variance σ^2, and $f_X(x)$ is the density function, the integral from $-\infty$ to $+\infty$ of $f_X(x)\,dx = 1$.

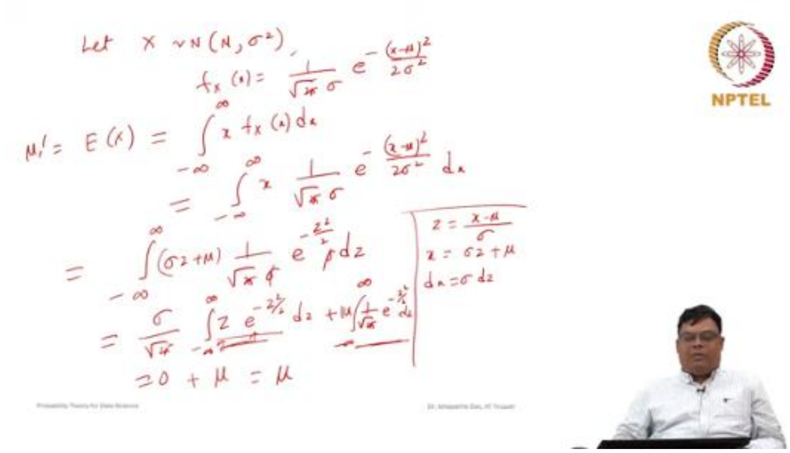
This implies that $\int$ from $-\infty$ to $+\infty$ of $f_X(x) = (1 / \sqrt{(2\pi\sigma)}) * e^{(-(x - \mu)^2 / (2\sigma^2))}\,dx = 1$.

Note that this is true for any real number μ and any value of $\sigma > 0$. For example, if $\mu = 0$ and $\sigma = 1$, it will still hold true.



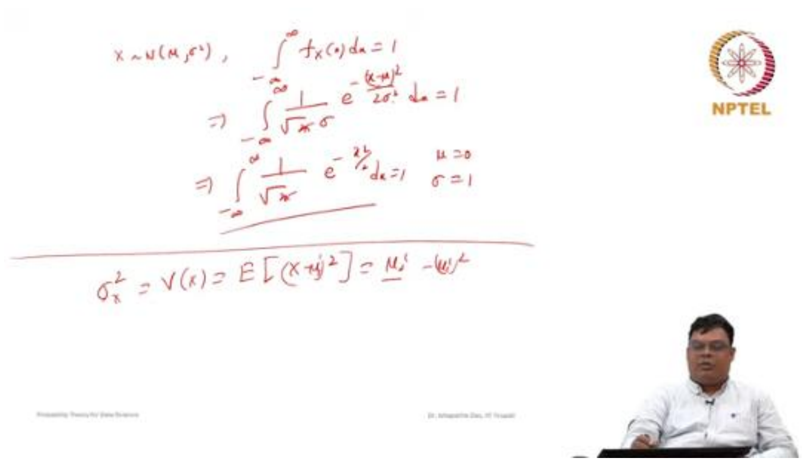
In that case, we find that $(1 / \sqrt{(2\pi)}) * e^{(-x^2 / 2)}\,dx = 1$. This is simply the density of the standard normal variate. Here, you can see that this is the integral from $-\infty$ to $+\infty$ of $(1 / \sqrt{(2\pi)}) * e^{(-z^2 / 2)}\,dz$. So, that is why this value is equal to 1. Hence, this is simply $\mu * 1$, which equals μ.

So, the value is μ. Therefore, for a normally distributed X with mean μ and variance σ^2, the expected value of X is μ. μ is the parameter that represents the mean of this random variable. Next, we will find the variance of this random variable. For convenience, we often use the variance formula.



You can remember that the variance of X, denoted as σ^2_X, is simply the variance of X. We use the expected value of $(X - \mu)^2$, μ_1'. Actually, μ_1' is nothing but μ. We found that this parameter is μ. So, this is $\mu_2' - \mu_1'^2$.

You can either find μ_2' and then subtract it, but since this density function contains terms like $(x - \mu)^2$, it's easier this way. So, that's why it may be convenient to use this formula directly. Let's find that. σ^2_X, the variance of X, is nothing but the expected value of $(X - \mu_1')^2$. Now, μ_1' is μ, as we've already found. Now, we want to find the expected value of $(X - \mu)^2$.



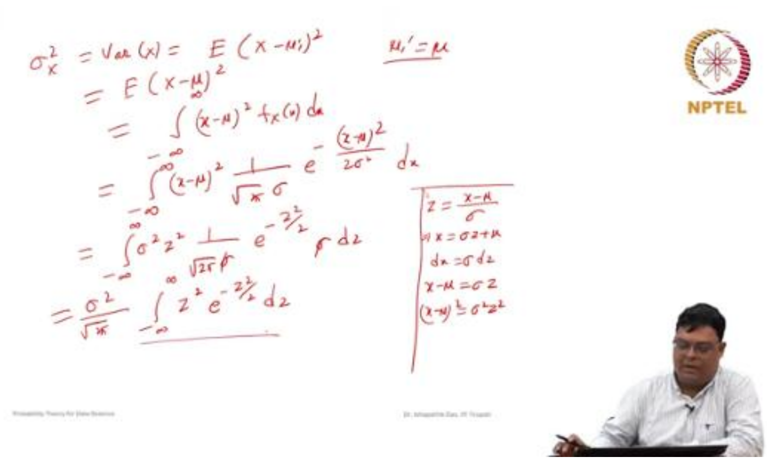
So, what is the formula? It is nothing but the integral from $-\infty$ to $+\infty$ of $(x - \mu)^2 * f_X(x)$ dx. This is the same as the integral from $-\infty$ to $+\infty$ of $(x - \mu)^2 * (1 / \sqrt{(2\pi\sigma)}) * e^{(- (x - \mu)^2 / 2\sigma^2)}$ dx. This may be helpful.

Let's see; if it's not helpful, then we will use the previous formula. First, we will find μ_2' because in the density function, there is also the $(x - \mu)^2$ term. That's why I thought this integration might be helpful if we can find it directly. So, let's do the same transformation again. The transformation is $z = (x - \mu) / \sigma$, which implies that $x = \sigma z + \mu$.

Then, $dx = \sigma\, dz$. So, this is equal to the integral from $-\infty$ to $+\infty$ because z goes to $-\infty$ as x goes to $-\infty$, and z goes to $+\infty$ as x goes to $+\infty$. Now, $(x - \mu)^2$ is simply σz. So, $(x - \mu)^2 = \sigma^2 z^2$. So, this is nothing but $\sigma^2 z^2 * (1 / \sqrt{(2\pi\sigma)})$.

Then, we have $(1 / \sqrt{(2\pi\sigma)}) * e^{(- (\sigma * (x - \mu)) / \sigma)^2 / 2}$, and then $dx = \sigma\, dz$. So, the σ terms cancel out. Now, what have we found? It is nothing but $\sigma^2 * \sigma^2 / \sqrt{(2\pi)} * \int$ from $-\infty$ to $+\infty$ of $z^2 * e^{(-z^2 / 2)}$ dz.

So, what we have found is $\sigma^4 / \sqrt{(2\pi)} * \int$ from $-\infty$ to $+\infty$ of $z^2 * e^{(-z^2 / 2)}$ dz. Now, we want to find the value of this integral. What will the value be? Let us find that out. So, the integral from $-\infty$ to $+\infty$ of $z^2 * e^{(-z^2 / 2)}$ dz. Now, you can see that this is an even function, and this is also an even function. When multiplied, the result will be an even function.



This can be written as $2 * \int$ from 0 to ∞ of $z^2 * e^{(-z^2 / 2)}$ dz. Now, how will we do this integration? Again, we will use the gamma function, with the help of gamma. So, $\Gamma(\alpha)$ is nothing but ∫ from 0 to ∞ of $x^{(\alpha - 1)} * e^{(-x)}$ dx. Now, here you can see that $e^{(-z^2 / 2)}$ is present, so we will take this.

So, all these formulas we know: $\Gamma(\alpha + 1) = \alpha * \Gamma(\alpha)$, and for an integer natural number l, $\Gamma(l + 1) = l!$. Also, $\Gamma(1/2) = \sqrt{\pi}$. We will use these results here. So, now, to get it into a gamma function form, what we will do is take the transformation $z^2 / 2 = t$. So, $z^2 = 2t$, and then $z = \sqrt{(2t)}$.

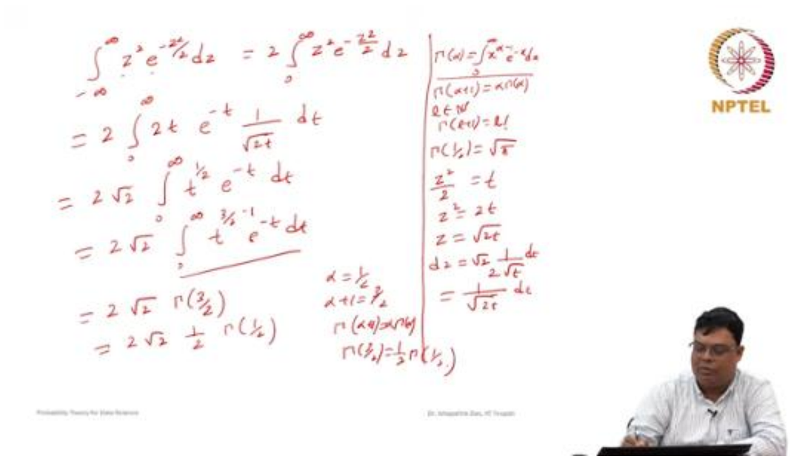
Therefore, $dz = (1 / \sqrt{2}) * (1 / \sqrt{t})$ dt. So, we get $1 / \sqrt{2}$. So, this is $1 / \sqrt{(2t)}$, and this is dt. I think this is fine. Now, we will substitute, and the limits do not change. When $z \to 0$, $t \to 0$, and when $z \to \infty$, $t \to \infty$. So, this becomes 2 * ∫ from 0 to ∞. Now, $z^2 = 2t$. So, that is why, from here, $z^2 = 2t$. This is nothing but 2t.

$e^{(-z^2 / 2)}$ becomes $e^{(-t)}$, and then $dz = (1 / \sqrt{(2t)})$ dt. So, that's the integration we are finding. Now, you can see that this is coming out as a gamma function form with $e^{(-t)}$. Now, here, $2 * 2 / \sqrt{2}$ becomes $2\sqrt{2}$. Then, $\sqrt{2} * \sqrt{2}$ simplifies further.

This leads to the integral from 0 to ∞ of t * (1 - 1/2), which equals 1/2. So, t = 1. What we need to find is something - 1, which results in 2. Finally, we have 1/2 because t is divided by $\sqrt{2} * \sqrt{t}$. So, this is nothing but $t * (1 / 2) * e^{(-t)}$ dt.

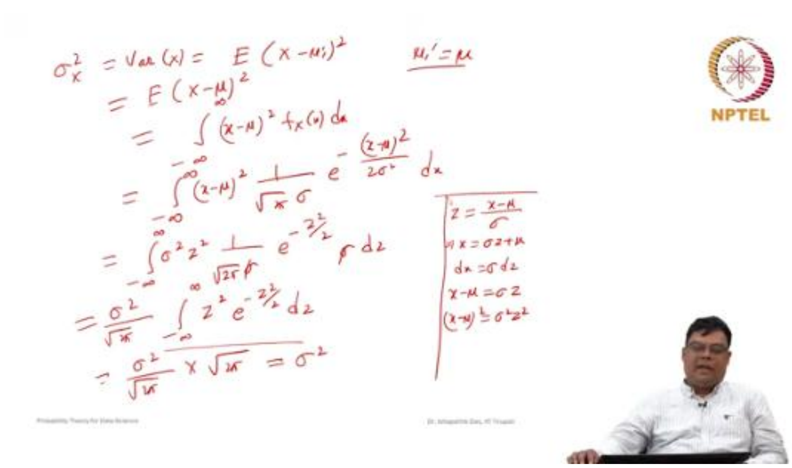
We have to write this in a form like something α - 1. This becomes $2\sqrt{2}$ * ∫ from 0 to ∞ of $t^{(1/2)}$, which we can write as $t^{(3/2 - 1)}$, so that it matches the desired form with $e^{(-t)}$ dt. Now, comparing this with the gamma function, this is nothing but $\Gamma(3/2)$. Therefore, the result is $2\sqrt{2} * \Gamma(3/2)$. Now we will use this formula: $\Gamma(\alpha + 1) = \alpha * \Gamma(\alpha)$.

Here, 3/2 can be written as 1 + 1/2. That is why we have $2\sqrt{2} * (1 / 2) * \Gamma(1/2)$, because $\alpha = 3/2$. So, $\alpha + 1$ becomes 3/2, and $\alpha = 1/2$. So, $\Gamma(\alpha + 1) = \alpha * \Gamma(\alpha)$. So, $\Gamma(3/2)$ is nothing but $(1 / 2) * \Gamma(1/2)$.



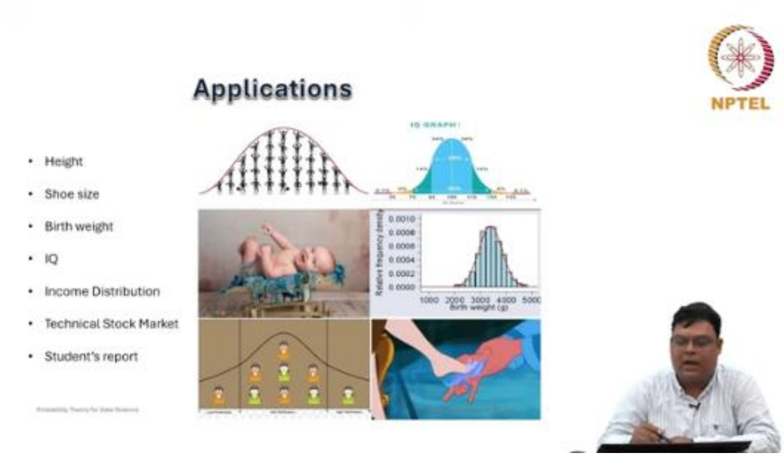
We have used that here, so 1 / 2 is canceled, leaving $\sqrt{2} * \Gamma(1 / 2)$. Since $\Gamma(1 / 2) = \sqrt{\pi}$, this simplifies to $\sqrt{(2\pi)}$. Hence, you see that from here, this value we have computed is nothing but $\sqrt{(2\pi)}$. So, this is $\sigma^2 / \sqrt{(2\pi)} * \sqrt{(2\pi)}$. This simplifies to σ^2.

So, hence we learn about the normal distribution function, its probability density function, and we check that it is a density function, integrating to 1. We also found its mean and variance. The parameters μ and σ^2 are very significant. μ is nothing but the mean of this random variable, and σ^2 is nothing but the variance of this random variable. Next, we will discuss some important properties of the random variable.



Let us also discuss some applications of the normal distribution. There are many applications of the normal distribution, as most data sets found in nature follow a normal distribution. For example, if you look at the height distribution, you'll see that the frequency or probability is less for heights < 5 feet. Similarly, for heights > 6 feet, the frequency or probability will also be less. But in the middle, around 5.5 or 5.7 feet, the number or proportion of people will be higher.

So, it looks like it increases and then decreases, showing symmetry around some point. Similarly, if you look at graphs for IQ, shoe size, birth weight, income distribution, stock market data, or student reports—like when grading students and analyzing their performance based on marks—you will see that, when the data is large, it tends to follow a normal distribution. So, this is very important because many distributions encountered in practice, such as the Poisson distribution, binomial distribution, hypergeometric distribution, and exponential distribution, can be approximated by a normal distribution. Even if a variable is not normally distributed, it can often be transformed into a normal form through a simple transformation. Additionally, many distributions of sample statistics are important.



For example, later on, we will discuss the sampling distribution in statistical inference, which will be very important. So, for sampling distributions, we will look at the distribution of $\bar{X}$ (the sample mean) and the distribution of the sample variance (S^2). Most of the sample variances will follow the square of a normal distribution, which is known as the chi-square distribution. So, these are all normally distributed random variables when the sample size is large. Now, one important property is that if $X_1, X_2, ..., X_n$ are independently normally distributed random variables, then their linear combination will also be a normally distributed random variable.

So, let us write that $X_1, X_2, \ldots, X_n$ are independent random variables, with each X_i having a normal distribution with mean μ_i and variance σ_i^2. Now, if you take the sum of the linear combination, like $Y = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_n X_n$. So, we have taken n random variables, which are independent. Now, this will also have a normal distribution with mean μ_γ and variance σ_γ^2. So, we will not focus on that proof right now.

Later, we will learn about other important topics, such as the moment-generating function. The moment-generating function is simply the transformation of a random variable. We will learn how to find the distribution of a transformation of a random variable, and then we can determine this distribution. Now, if we assume that the linear combination follows a normal distribution, then the mean will be μ_γ and the variance will be σ_γ^2. Then you can easily find the mean and σ_γ^2.

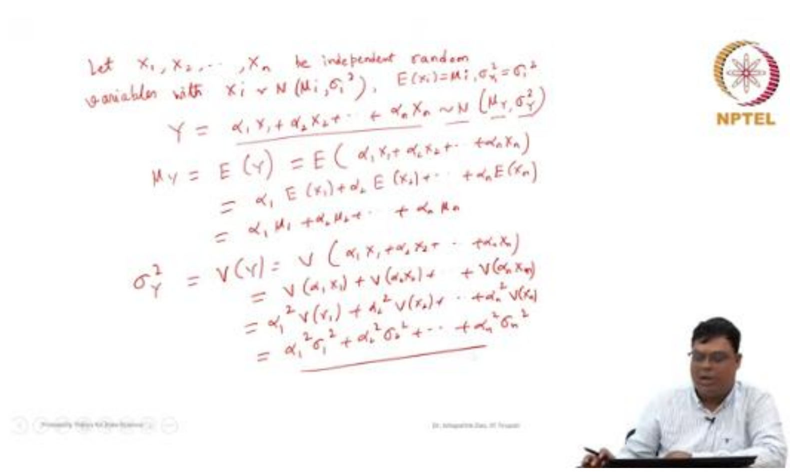
The mean of Y, μ_γ, is nothing but the expected value of Y. The expected value of Y is $\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_n X_n$. We have discussed that some of the properties involve linear transformation and expansion. This is nothing but the expected value being a linear transformation. So, this can be written as $\alpha_1 E[X_1] + \alpha_2 E[X_2] + ... + \alpha n E[Xn]$.

Then, because the expected value of X_1 and the expected value of X_2 are already known, since Xi is normally distributed, we can use that information. The expected value of Xi is nothing but μ_i, and the variance of Xi is σ_i^2. So, we will use that here. This gives $\alpha_1\mu_1 + \alpha_2\mu_2 + ... + \alpha n\mu n$. This is the mean of this normal distribution.

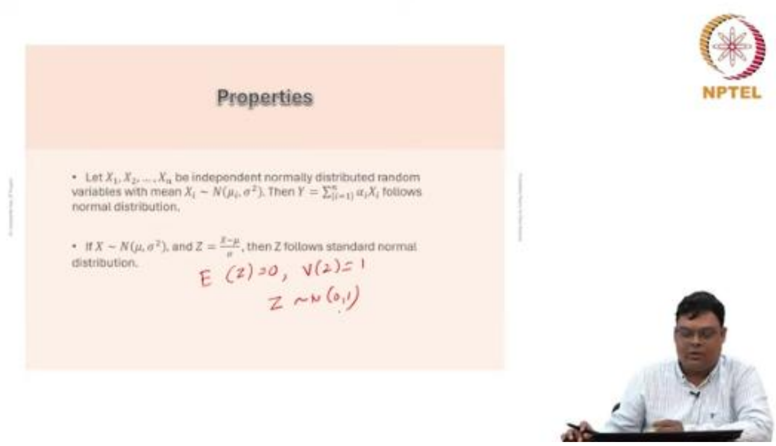
Now, what will be the variance of this normal distribution Y? σ_γ^2 will be the variance of Y. This is nothing but the variance of $\alpha_1 X_1 + \alpha_2 X_2 + ... + \alpha nXn$. By another property, since these are independent random variables, the variance can be written as the variance of $\alpha_1 X_1$ + the variance of $\alpha_2 X_2$ + ... + the variance of αnXn. Because they are independently distributed random variables, that is why you can write this.

Otherwise, there would be some covariance terms, but we will discuss that whenever we learn about bivariate random variables. Here, because of independence, this is the formula. Now, as we have already discussed, the variance of a constant multiplied by a random variable will involve the square of the constant. This results in $\alpha_1^2 * Var(X_1) + \alpha_2^2 * Var(X_2) + ... + \alpha n^2 * Var(Xn)$. The variance of X_1 is nothing but σ_1^2.

This gives $\alpha_1^2 * \sigma_1^2 + \alpha_2^2 * \sigma_2^2 + ... + \alpha n^2 * \sigma n^2$. This represents the variance of this random variable, which is one of its properties. Now, from these properties, we can see that this is a general property. Next, it's very important. The next property is that if X is normally distributed with mean μ and variance σ^2, then if you take this transformation, $Z = (X - \mu) / \sigma$, Z follows a standard normal distribution.



This means that the expected value of Z is 0, and the variance of Z is 1. So, Z follows a normal distribution with mean 0 and variance 1, or $Z \sim N(0, 1)$. So, if $X \sim N(\mu, \sigma^2)$, and you take the transformation $Z = (X - \mu) / \sigma$, it's still a linear transformation. Because it's $(1/\sigma) * (X - \mu)$, it's a linear transformation. As we discussed, any linear transformation of a normal distribution is again a normal distribution.



So, we can say that Z is also a normal distribution. Now, we want to find what the mean and variance will be. So, it will be a normal distribution with mean μ_Z and variance σ^2_Z. Now, we want to find what the mean of this random variable and the variance of this random variable will be. The mean of this random variable, μ_Z, is nothing but the expected value of Z. So, this is the expected value of $(X - \mu) / \sigma$.

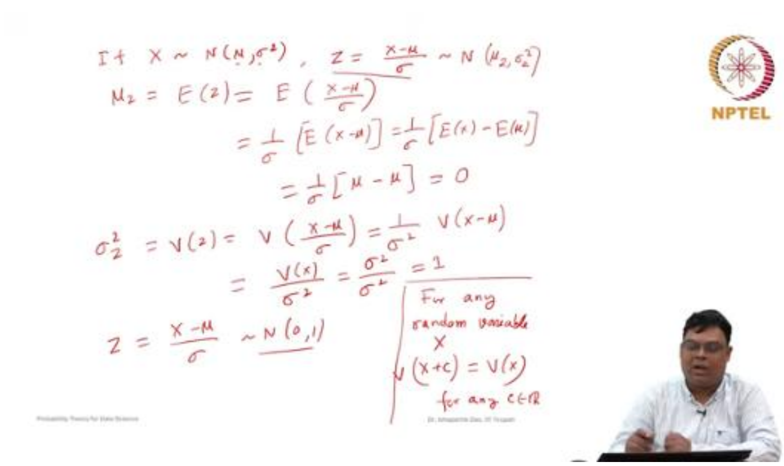
So, since σ is a constant, we have $(1/\sigma) * E(X - \mu)$. Now, $E(X - \mu)$ is $E(X) - E(\mu)$. Since μ is a constant, we have already discussed that $E(X)$ is nothing but μ. And since μ is a constant, this is again μ. Therefore, this simplifies to 0.

Now, what will be σ^2_Z? What will be the variance of Z? So, this is the variance of Z. This is nothing but the variance of $(X - \mu) / \sigma$. Since σ is a constant, we have already discussed that it will be $(1/\sigma^2) * Var(X - \mu)$. Now, this is another exercise.

You may observe that for any random variable, another property is that if you take the transformation X + c, where c is a constant, the variance does not change. For any $c \in \mathbb{R}$, $Var(X + c) = Var(X)$. So, that is why μ is a real number. So, it is X - c. This is nothing but $Var(X) / \sigma^2$.

Now, what is the variance of X? The variance of X is simply the parameter σ^2. So, this becomes σ^2 / σ^2, which equals 1. Therefore, if you take the transformation where $Z = (X - \mu) / \sigma$, it results in a standard normal distribution, which is written as $N(0, 1)$. This is the standard normal variate.

Now, suppose you want to find the probability distribution function or any probability related to a normal distribution function, but the integration is not tractable. For example, suppose $X \sim N(\mu, \sigma^2)$. Now, you want to find the probability that $X > c_1$ and $X \leq c_2$. So, what will the probability be? This probability will be the integral from c_1 to c_2 of f(x) dx.



So, c_1 and c_2 can be anything. For example, you can take c_1 as $-\infty$, and c_2 can be any real number, or it can tend to ∞ as well. So, both of them can be real numbers, any values like that. So, then this probability will be nothing but the integral from c_1 to c_2 of $(1 / \sqrt{(2\pi\sigma)}) * e^{(-(x - \mu)^2 / (2\sigma^2))}$ dx. Now, you can see that this is intractable because the integration is not straightforward.

We cannot compute this integration. So, up to this point, we can do the calculations, but after that, we can't find any closed-form value as a real number. So, we can either do it numerically or apply a transformation. For instance, if we take the transformation, where $z = (x - \mu) / \sigma$, we know that this transformation results in a normal distribution with mean 0 and variance 1. So, here you can apply this transformation.

Let $x = \sigma z + \mu$. Then, $dx = \sigma\, dz$. Now, what happens to the limits? The limits change as follows: when $x = c_1$, $z = (c_1 - \mu) / \sigma$. Similarly, when $x = c_2$, $z = (c_2 - \mu) / \sigma$. So, the assumption here is that $c_1 < c_2$. The expression becomes $(1 / \sqrt{(2\pi\sigma)}) * e^{(-z^2 / 2\sigma)}$ dz. The σ cancels out. So, we need to evaluate this integral, which is from $(c_1 - \mu) / \sigma$ to $(c_2 - \mu) / \sigma$, with the integrand $(1 / \sqrt{(2\pi)}) * e^{(-z^2 / 2)}$ dz. So, this is actually the density of a standard normal variate.