PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

Week - 05

Lecture - 21

Applications of Poisson Distribution

So we can get X - 1, we can cancel. Instead of finding the expected value of X^2 directly, we will first find the expected value of X * (X - 1). Then what will be the benefit? If you find the expected value of X * (X - 1), this is nothing but the expected value of $X^2 - X$, which is the same as the expected value of X^2 minus the expected value of X. Then you can simplify it like this:

$E[X^2] = \mu_2' = E[X^2]$

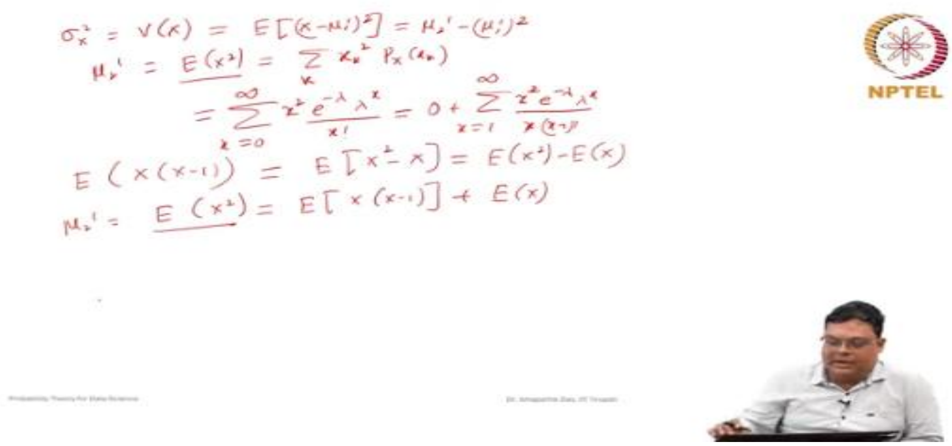
It is the same as the expected value of X * (X - 1) + E[X]. The expected value of X we have already computed here and that is λ. Now we will compute the expected value of X * (X - 1) and then we can find this value.

So let us first compute the expected value of X * (X - 1). The expected value of X * (X - 1) is the summation:

$E[X * (X - 1)] = \Sigma$ (from $x = 0$ to ∞) $x * (x - 1) * p(x)$

where p(x) is the probability mass function, which gives us:

$E[X * (X - 1)] = \Sigma$ (from $x = 0$ to ∞) $x * (x - 1) * e^{-\lambda} * (\lambda^x / x!)$



When $x = 0$ or $x = 1$, the term is 0, so we can start the summation from $x = 2$. We then rewrite $x * (x - 1)$ as $x * (x - 2) + 1$, which simplifies to $x * (x - 2) + x$. This gives us the summation from $x = 2$ to infinity of $(\lambda^x) / ((x - 2)!) * e^{(-\lambda)}$. By factoring out $\lambda^2 * e^{(-\lambda)}$, we get:

$\lambda^2 * e^{(-\lambda)}$ * summation from $r = 0$ to infinity of $(\lambda^r) / (r!)$

This summation is the known series expansion for e^λ, so the result becomes:

$\lambda^2 * e^{(-\lambda)} * e^\lambda$

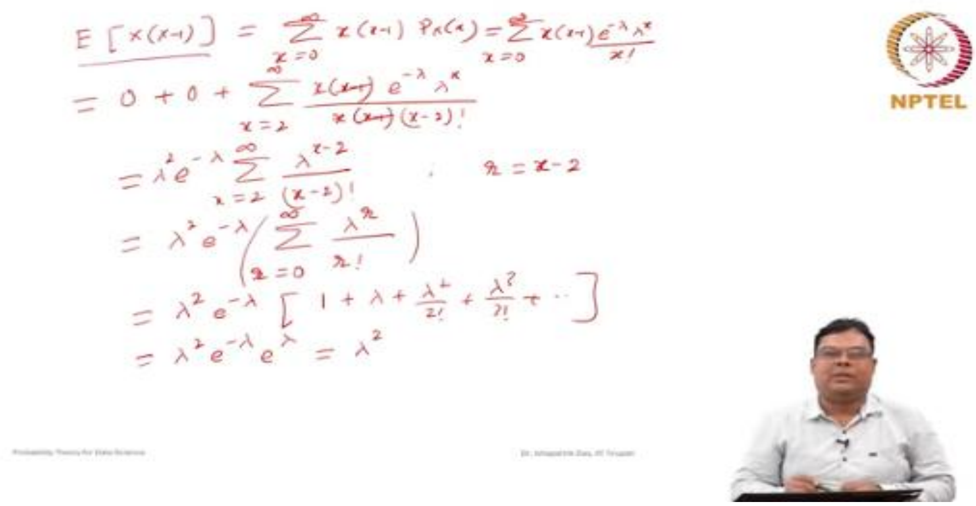
which simplifies to λ^2.

So the expected value of $x * (x - 1)$ is λ^2. We found that this is equal to λ^2, and the expected value of x is λ. So this gives us:

$\lambda * (\lambda + 1) = \mu_2 + \lambda$

where μ_2 is $\lambda^2 + \lambda$. Now we will find the variance. The variance σ_X^2 is the expected value of $(X - \mu_1)^2$, which is:

$\mu_2 - \mu_1^2$

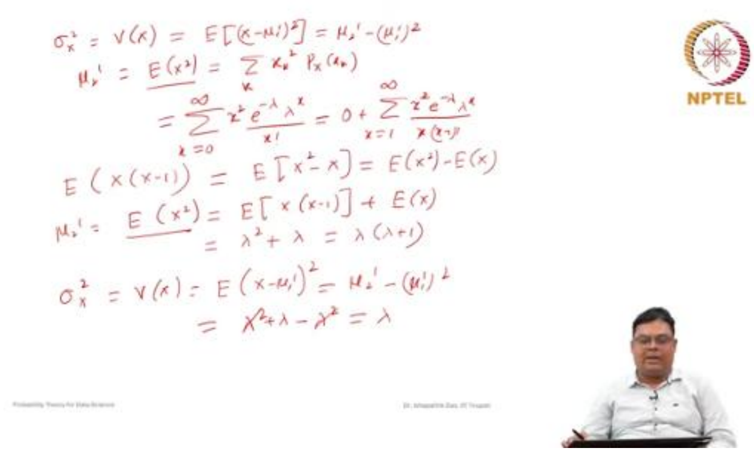


From earlier, we know $\mu_2 = \lambda^2 + \lambda$ and $\mu_1 = \lambda$. So the variance is:

$\lambda^2 + \lambda - \lambda^2$

which simplifies to λ. Therefore, the variance of the Poisson distribution is λ. We also found that the mean of the Poisson distribution is λ. So the mean and variance of the Poisson distribution are the same and the parameter is λ.

This is an example of a random variable where the mean and variance are equal. For random phenomena where the data shows these kinds of properties, such as being a discrete random variable with countably infinite values, we can model it using the Poisson distribution. Now we will discuss a numerical example of the Poisson distribution and how we can apply it. So, we can find some applications of the Poisson distribution. In real life, there are many applications of the Poisson distribution.

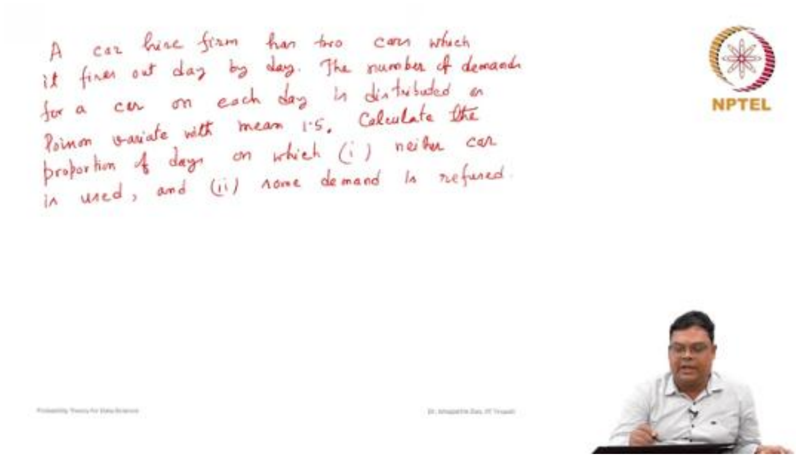


For example, the number of aircraft accidents or road accidents in a given time interval. Usually, insurance companies need this information. They want to know the probability of accidents happening because they provide insurance to people. They need to estimate how many accidents may happen, and based on that, they compute the premiums they should charge. This helps them make a profit, because common people might not know the probability of an accident and could be scared, thinking a huge loss could happen.

But the insurance company can estimate the number of accidents that might happen, and from that, they can calculate the potential loss and use this to make a profit. Another application is the number of arrivals at a service facility, like at an ATM, a bank, a railway station, or a petrol pump. Whenever there's a count of some kind of event, like accidents or service usage, this data can be modeled using the Poisson distribution. For example, a bank can use this information to decide where to establish ATMs or how many ATMs are needed. Similarly, it can be used for the number of patients in a doctor's clinic, the number of visitors to a mall or exhibition, or the number of services completed per unit time at a bank or any other service facility.

So, whenever you see something in real life that involves counting events like accidents or service usage in a discrete sense, you can use the Poisson distribution. Now we will discuss some of the numerical examples of the Poisson distribution, some of the numerical examples. So let us do some numerical examples using the Poisson distribution, how we can solve. So, a car hire firm has two cars. So let us do one problem here, a car hire firm has two cars which it rents out day by day.

The number of demands for a car on each day is distributed as a Poisson variate with a mean of 1.5. Calculate the proportion of days on which neither car is used and some demand is refused. So you can see this problem, a car hire firm has two cars which it rents out day by day. The number of demands for a car on each day is distributed as a Poisson distribution. So, let X be the random variable defined as the number of demands for a car on each day.

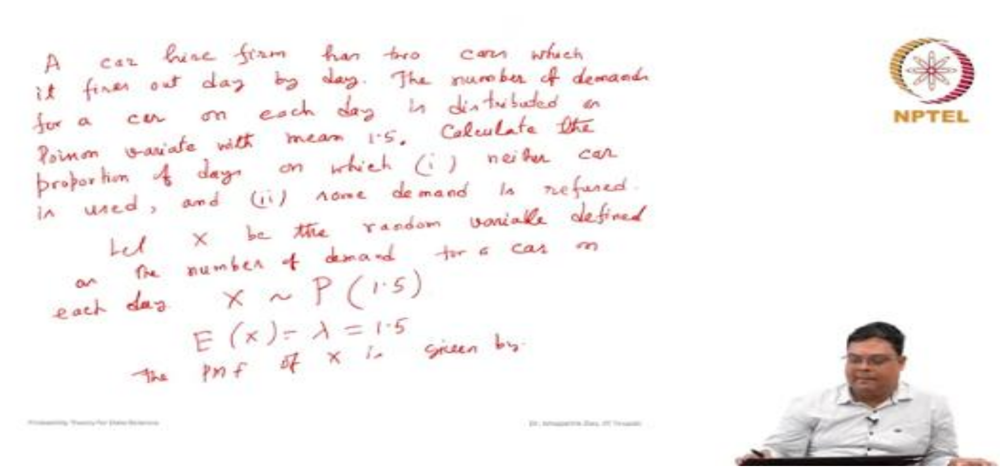


It is given that X follows a Poisson distribution with parameter λ. So, we write it as X ~ Poisson(λ). The expected value of X, we know, is equal to λ whenever X is a Poisson random variable. Now, this λ mean is given as 1.5. So, we write it as Poisson(1.5). Now, we need to calculate the proportion of days.

So, what will be the probability mass function here? The probability mass function of X is given by:

$P(X = x) = (e^{-\lambda} * \lambda^x) / x!$, for x = 0, 1, 2, ...

and 0 otherwise. So, this is the general probability mass function we discussed. Here, it is given that λ = 1.5. So, that is why it is known to us.



Now, calculate the proportion of days on which neither car is used. So, that means you have to find the probability that there is no demand. So, that is the probability that $X = 0$. So, that is the first probability we have to compute, nothing but $P(X = 0)$. So, then you can have this $\lambda = 1.5$.

We will use $e^{(-\lambda)}$, this probability is nothing but λ^X, so $X = 0$ here, so that is $e^{(-\lambda)}$, where $\lambda = 1.5$, so $e^{(-1.5)}$. So, this is the probability, the proportion of days, that means the probability that neither of the cars is used.

And the next is, the next question is to find the probability that some demand is refused. So, when will the demand be refused? So, demand will be refused when, because there are only two cars.

So, demand will be refused if the demand is greater than 2. So, up to 2, then there will be some cars. But the demand is refused means we have to find the probability that $X > 2$. So, $X > 2$ will be nothing but the summation from $X = 3$ and greater, basically $X \geq 3$. So, now instead of finding, because you have to take the infinite sum from 3 to infinity, that can also be done, but it is easy to find the complement of that.

So, $1 - P(X > 2)$, that means $P(X \leq 2)$. So, this is equal to:

$1 - P(X \leq 2)$

This is nothing but:

1 - summation from $X = 0$ to 2, $P(X = x)$

So, this is nothing but:

1 - summation from $X = 0$ to 2, $P(X = x)$ = 1 - summation from $X = 0$ to 2, $e^{(-\lambda)} \lambda^X / X!$

So, then this is nothing but:

$1 - (e^{(-1.5)} + e^{(-1.5)} * 1.5 + e^{(-1.5)} * (1.5^2) / 2!)$

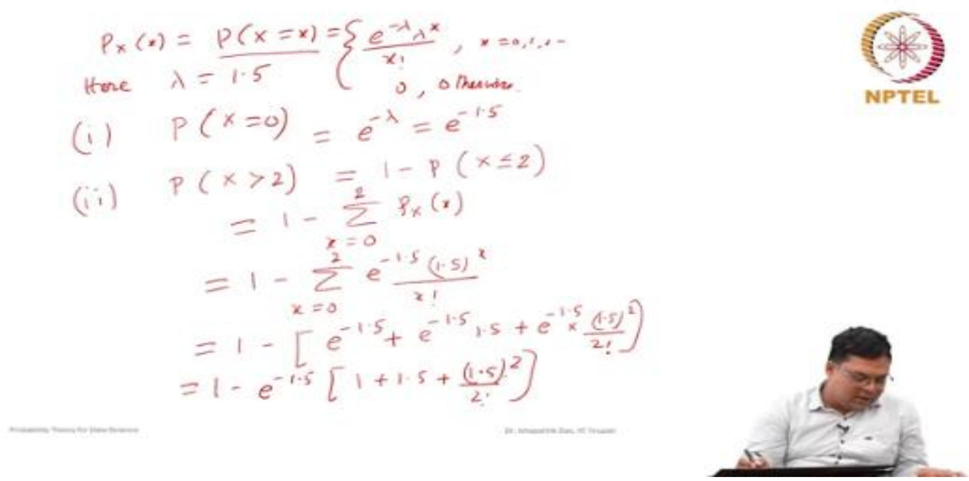
So, $X = 0$, $X = 1$, and $X = 2$, and the expression becomes:

$1 - (e^{(-1.5)} + e^{(-1.5)} * 1.5 + e^{(-1.5)} * (1.5^2) / 2!)$

So, then you have to just use the calculator to compute it:

$e^{(-1.5)} * 1 + 1.5 + (1.5^2) / 2!$

Let's solve this numerical problem. I'll provide the answer here so you can calculate it yourself and check if it's close to 0.19126. This is just one example, and we'll go through more to get comfortable with the Poisson distribution. So, suppose X is a Poisson random variable.



We're given that the probability of X = 2 is 9 times the probability of X = 4, plus 90 times the probability of X = 6. We need to find the value of λ, which is the mean of this distribution.

For a Poisson distribution with parameter λ, the probability of X = x is calculated by:

P(X = x) = (e^(-λ) * λ^x) / x!

This formula applies for values of x = 0, 1, 2, ... To find the probability that X = 2, we substitute x = 2, giving us:

P(X = 2) = (e^(-λ) * λ^2) / 2!

We're told that the probability of X = 2 is 9 times the probability of X = 4, plus 90 times the probability of X = 6. Now, we substitute the expressions for these probabilities:

P(X = 4) = (e^(-λ) * λ^4) / 4!, P(X = 6) = (e^(-λ) * λ^6) / 6!

Substituting these values into the equation, we get:

(e^(-λ) * λ^2) / 2! = 9 * (e^(-λ) * λ^4) / 4! + 90 * (e^(-λ) * λ^6) / 6!

Since $e^{-\lambda}$ is present in each term, we can cancel it out. This simplifies to:

$$(\lambda^2) / 2 = 9 * (\lambda^4) / 24 + 90 * (\lambda^6) / 720$$

Clearing the denominators further simplifies this to:

$$12 * \lambda^2 = \lambda^4 + 15 * \lambda^6$$

Finally, we continue simplifying by factoring and canceling terms until we isolate λ, giving us the mean of this Poisson random variable. So, then we get this equation. I am just writing this equation:

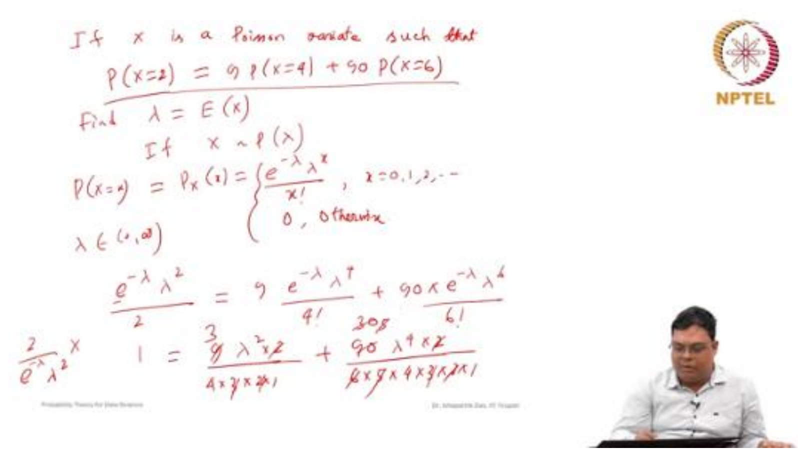
$$12 * \lambda^2 = \lambda^4 + 15 * \lambda^6$$

Suppose we multiply both sides by 4. Then this is nothing but:

$$4 * 12 * \lambda^2 = 4 * (\lambda^4 + 15 * \lambda^6)$$

which simplifies to:

$$48 * \lambda^2 = 4 * \lambda^4 + 60 * \lambda^6$$



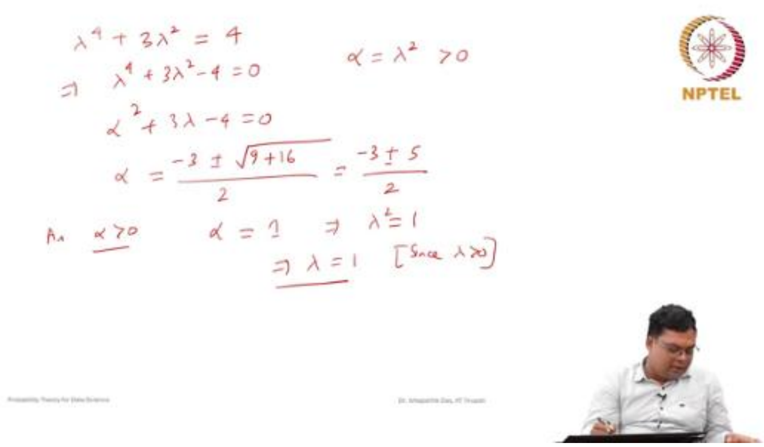
We get λ^4. We multiplied both sides by 4. $\lambda^4 + 3\lambda^2 = 4$. We multiplied both sides by 4. This is 4. So, which implies $\lambda^4 + 3\lambda^2 - 4 = 0$.

Suppose λ is a positive number. So, let us denote another number, suppose α, where $\alpha = \lambda^2$. This is a positive number. Then this equation will convert to a quadratic equation, because it's a fourth power. I want to simplify it so that we have something more familiar.

If you write λ^2 as α, then it becomes $\alpha^2 + 3\alpha - 4 = 0$. From here, you can get the roots of this equation. This is $(-b \pm \sqrt{b^2 - 4ac})$. What we got from this solution is $-3 \pm \sqrt{9 - 4 * 1 * -4}$. This gives 9 + 16, which is 25.

So, 4 * 4 = 16, divided by 2 * 1 = 2. So, this is $-3 \pm \sqrt{25}$, which is 5, divided by 2. So, we get two solutions: -3 + 5 = 2, and 2 ÷ 2 = 1. This implies $\lambda^2 = 1$. Since λ is positive, $\lambda = 1$.

So, from here we can solve for the value of λ. This is just another problem, really. This is how to find λ. So, let us consider another problem. Suppose X and Y are independent Poisson variates. It is given that the probability of X = 1 is the same as the probability of X = 2.



Also, it is given that P(Y = 2) = P(Y = 3). They are independent Poisson variates, but the parameters are not known. λ_1 and λ_2 are the two unknown parameters. So, we are asked to find the variance of X - 2Y. Since X and Y are independent, the variance of X - 2Y will be the variance of X + 4 times the variance of Y.

We will use this formula later when we discuss bivariate random variables and covariance in more detail. Now, let X be a Poisson random variable with parameter λ_1, and Y be a Poisson random variable with parameter λ_2. The probability mass function for X is:

$P(X = x) = (e^{-\lambda_1} * \lambda_1^x) / x!$ for x = 0, 1, 2, ...

Similarly, for Y:

$P(Y = y) = (e^{-\lambda_2} * \lambda_2^y) / y!$ for y = 0, 1, 2, ...

Now, it is given that $P(X = 1) = P(X = 2)$. From the probability mass function, we know that:

$P(X = 1) = e^{(-\lambda_1)} * \lambda_1 / 1!$

$P(X = 2) = e^{(-\lambda_1)} * \lambda_1^2 / 2!$

So, equating these two probabilities, we get:

$e^{(-\lambda_1)} * \lambda_1 / 1! = e^{(-\lambda_1)} * \lambda_1^2 / 2!$

Which simplifies to:

$\lambda_1 = \lambda_1^2 / 2$.

Solving this, we get:

$\lambda_1^2 = 2$, which implies $\lambda_1 = \sqrt{2}$.

Next, it is given that $P(Y = 2) = P(Y = 3)$. Using the probability mass function for Y, we have:

$P(Y = 2) = e^{(-\lambda_2)} * \lambda_2^2 / 2!$

$P(Y = 3) = e^{(-\lambda_2)} * \lambda_2^3 / 3!$

Equating these two probabilities, we get:

$e^{(-\lambda_2)} * \lambda_2^2 / 2! = e^{(-\lambda_2)} * \lambda_2^3 / 3!$

Simplifying this, we get:

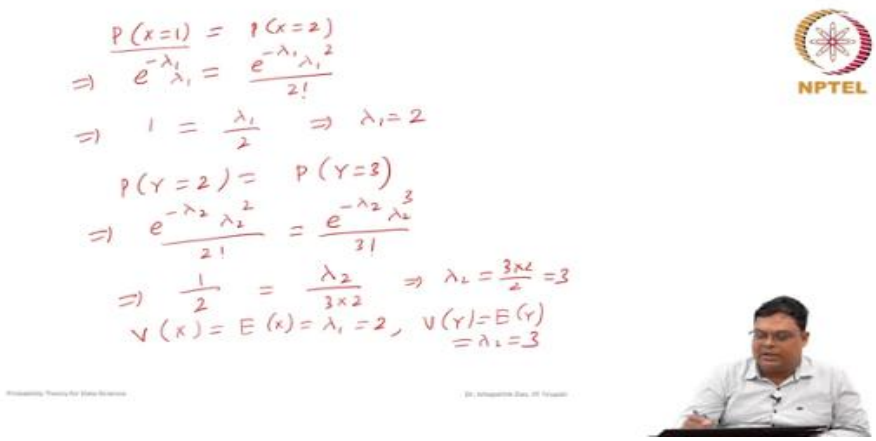
$\lambda_2^2 / 2! = \lambda_2^3 / 3!$

Solving for λ_2, we get:

$\lambda_2 = 3$.

Now that we have $\lambda_1 = \sqrt{2}$ and $\lambda_2 = 3$, we can find the variance. The variance of a Poisson random variable is equal to its mean. Therefore, the variance of X is $\lambda_1 = \sqrt{2}$, and the variance of Y is $\lambda_2 = 3$.

Now, it is asked to find the variance of X - 2Y. So, the variance of X - 2Y. We will discuss this later. But let us remember that for any two random variables, the variance of X - 2Y is the variance of X + the variance of 2Y, because they are independent. There is another term, covariance, but since X and Y are independent, this term is 0.



So, because X and Y are independent random variables, this formula works. Actually, we will discuss this formula later when we talk about bivariate random variables, but for now, we are using it. This is nothing but the variance of X + 4 times the variance of Y. The constant comes outside as the square, so this becomes:

$\text{Var}(X - 2Y) = \text{Var}(X) + 4 \times \text{Var}(Y)$

Since X and Y are independent, $\text{Cov}(X, Y) = 0$. The variance of X is λ_1, which is 2, plus 4 times the variance of Y, which is λ_2, and $\lambda_2 = 3$.

Therefore, we have:

$\text{Var}(X - 2Y) = \lambda_1 + 4 \times \lambda_2 = 2 + 4 \times 3 = 2 + 12 = 14$

Thus, the variance of X - 2Y is 14. This is the answer. If there is any mistake, please check and see whether this answer is correct, or if there is a different answer. These are some of the examples we discussed.

$$V(X-2Y) = V(X) + V(2Y)$$

$$= V(X) + 4\,V(Y)$$

[Since X and Y are independent random variable.]

$$= 2 + 4 \times 3$$

$$= 2 + 12 = 14.$$