PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

Week - 09

Lecture - 45

Conditional Mean and Variance for Continuous Random Variables

Now, with respect to the conditional probability mass function, we will find the conditional mean of Y given X = 2. It is defined as the expected value of Y given X, where $x_i = 2$. So, we have $\mu_\gamma \mid 2$. By definition, this is $\Sigma (y_j * P(Y = y_j \mid X = 2))$ for $y_j \in \{1, 2\}$.

In this case, only this variable is represented, where y_j can take the values 1 or 2. That's why we have:

$1 * P(Y = 1 \mid X = 2) + 2 * P(Y = 2 \mid X = 2)$.

If we plug in the values, we get:
$1 * (5/11) + 2 * (6/11)$.

This gives us:

$5/11 + 12/11 = 17/11$.

Now, to find the variance, we will use the simplified formula. The conditional variance is defined by the expected value:

$\sigma^2_\gamma \mid 2 = E(Y^2 \mid X = 2) - (\mu_\gamma \mid 2)^2$.

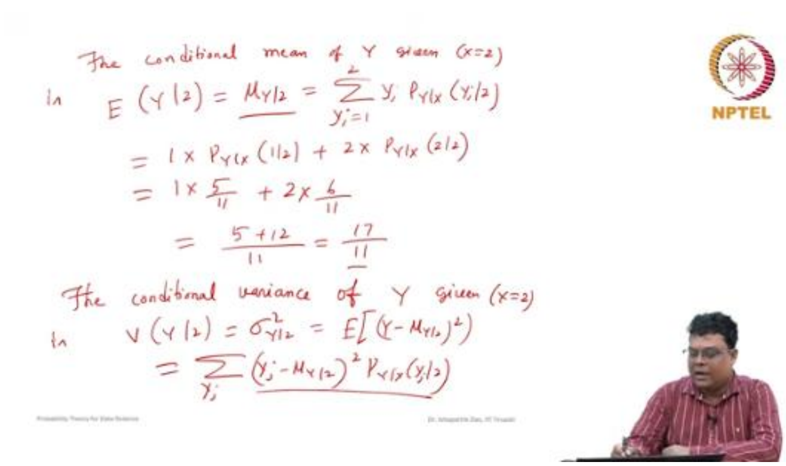
The formula becomes:

$\Sigma (y_j - \mu_\gamma \mid 2)^2 * P(Y = y_j \mid X = 2)$.

Using this formula is complicated because you have to subtract all possible values from 17/11 (which is μ_Y | 2), square the result, and then calculate the variance. To simplify, we use a simplified formula. This formula involves finding the expected value of Y^2 | 2, and then subtracting the square of the expected value of Y | 2, which we have already found.

So, we will use the simplified formula:

$E(Y^2 | X = 2) = (1^2 * P(Y = 1 | X = 2)) + (2^2 * P(Y = 2 | X = 2))$.



Substituting the values: For $y_j = 1$:
$1^2 = 1$, and $P(Y = 1 | X = 2) = (4 + 1) / 11 = 5 / 11$.
So, this term is:
$1 * (5 / 11) = 5 / 11$.

For $y_j = 2$:
$2^2 = 4$, and $P(Y = 2 | X = 2) = (4 + 2) / 11 = 6 / 11$.
So, this term is:
$4 * (6 / 11) = 24 / 11$.

Now, adding the two terms together:
$5 / 11 + 24 / 11 = 29 / 11$.
Thus, the expected value of Y^2 given X = 2 is:
$E(Y^2 | X = 2) = 29 / 11$.

Now, we can find the variance of Y given X = 2. The formula for variance is:
Variance = $E(Y^2 \mid X = 2) - (E(Y \mid X = 2))^2$.

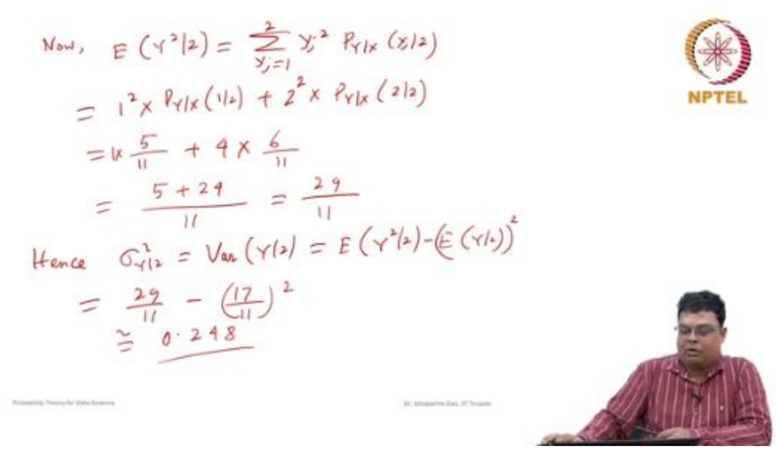
We have already calculated $E(Y \mid X = 2) = 17 / 11$, so:
Variance = $29 / 11 - (17 / 11)^2$.

First, compute the square of 17 / 11:
$(17 / 11)^2 = 289 / 121$.

Now, subtract:
29 / 11 - 289 / 121 = (29 * 11) / 121 - 289 / 121 = 319 / 121 - 289 / 121 = 30 / 121.

So, the variance is approximately:
Variance ≈ 30 / 121 ≈ 0.248.

This is the method to compute the conditional mean and conditional variance. Once you understand the process, you can perform the necessary calculations using a calculator. I hope you have understood it, so for finding the conditional mean and conditional variance.



You first need to find the conditional probability mass function. Based on that, you can compute the conditional mean and conditional variance. In this question, we completed the calculation of the conditional mean and conditional variance of Y given X = 2. The conditional mean is 17/11, and the conditional variance is calculated as $29/11 - (17/11)^2$, which is approximately 0.248. We have now completed the discrete part for finding the conditional mean and conditional variance.

Next, we will discuss the case for continuous random variables. Specifically, we will look at how to find the conditional mean and conditional variance for continuous bivariate random variables. If X and Y are continuous random variables, then X and Y are a bivariate continuous random variable with a joint probability density function. The marginal probability density function of X is denoted by $f_X(x)$, and the marginal probability density function of Y is denoted by $f_Y(y)$. The conditional probability density function of Y given X = x is equal to the joint probability density function $f_{XY}(x, y)$ divided by the marginal probability density function $f_X(x)$, provided that $f_X(x) \neq 0$.

This calculation is only valid for those values of x where $f_X(x) \neq 0$.

Now we will define the conditional mean and conditional variance. The conditional mean of Y given X = x is defined as the expected value of Y given X = x. This is denoted as $\mu_{Y|x}$. Since we are dealing with continuous variables, we use integration instead of summation, which is used for discrete random variables.

The conditional mean is calculated as:

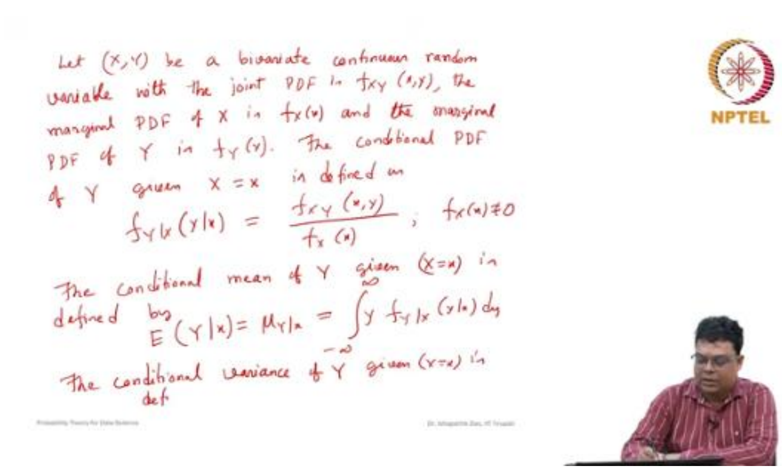
$\mu_{Y|x} = \int y * f_{Y|x}(y \mid x)\, dy$

The conditional variance of Y given X is defined as the expected value of $(Y - \mu_{Y|x})^2$. This is expressed as:

$Var(Y \mid X = x) = \int (y - \mu_{Y|x})^2 * f_{Y|x}(y \mid x)\, dy$

This calculation can be a bit complicated, so we use a simplified formula. The simplified formula is:

$Var(Y \mid X = x) = E(Y^2 \mid X = x) - (E(Y \mid X = x))^2$



We already defined the expected value of Y given X as $\mu_{Y|X}$. The expected value of Y^2 given X is calculated as the integral from negative infinity to positive infinity of y^2, multiplied by the conditional probability density function of Y given X. Using this formula, we obtain the conditional mean and variance for continuous random variables.

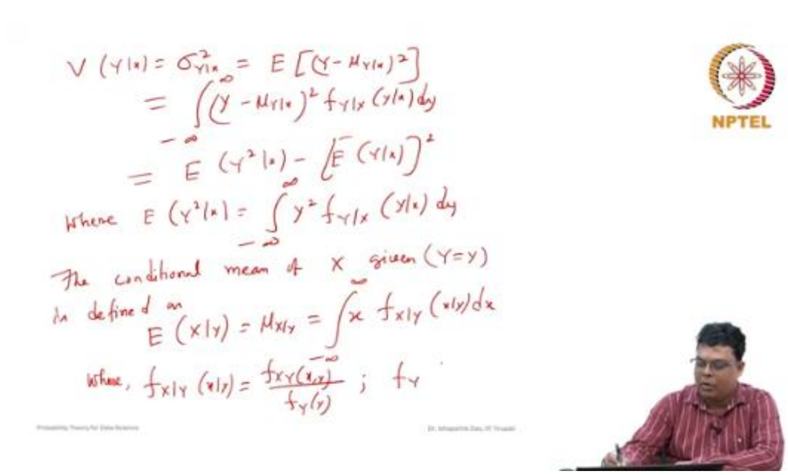
For the continuous bivariate random variable (X, Y), with the joint probability density function $f_{XY}(x, y)$, the conditional mean of Y given X is defined by $\mu_{Y|X}$, and the conditional variance is represented by the expected value of Y^2 given X minus the square of the expected value of Y given X.

If you want to find the conditional mean and variance of X given Y, you simply swap the positions of X and Y.
The conditional mean of X given Y is the expected value of X given Y, which is calculated by integrating x multiplied by the conditional probability density function of X given Y.

The conditional probability density function of X given Y is the joint probability density function divided by the marginal probability density function of Y:
$f_{X|Y}(x \mid y) = f_{XY}(x, y) / f_Y(y)$

The conditional variance can be found only if the conditional probability density function is not equal to zero. If it equals zero, the variance will be undefined.

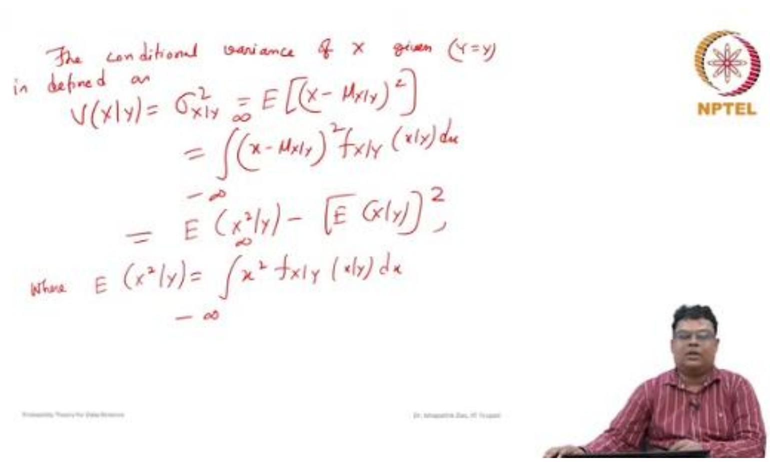
Now, how do we compute the conditional variance? The conditional variance of X given Y is defined as the variance of X given Y, denoted as $\sigma^2_{X|Y}$.

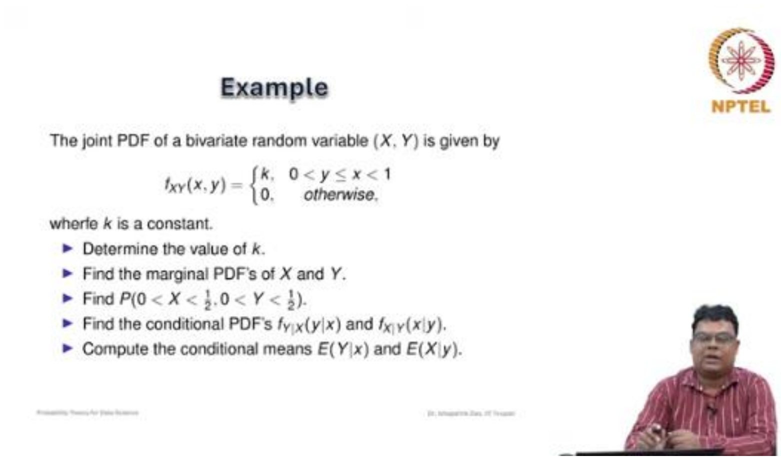


This is calculated as the expected value of $(X - \mu_{x|\gamma})^2$. It can be expressed as the integral of $(X - \mu_{x|\gamma})^2$ multiplied by the conditional probability density function of X given Y. Since this computation can be complex, we use a simplified formula: the conditional variance is the expected value of X^2 given Y minus the square of the expected value of X given Y. The expected value of X given Y is denoted as $\mu_{x|\gamma}$, which is the conditional mean, and the expected value of X^2 given Y is the integral of X^2 multiplied by the conditional density function of X given Y.

For computation, we will use this formula, which can be more convenient for certain density functions. Based on my observations, this formula is typically easier to apply when finding the conditional variance.

Let us go through a numerical example for conditional mean and conditional variance. This example has been discussed before, so it serves as a revision. You can review it again if needed. In this example, we need to find the conditional probability density function (PDF) in order to compute the conditional mean and conditional variance. Therefore, we also need to compute the marginal functions.

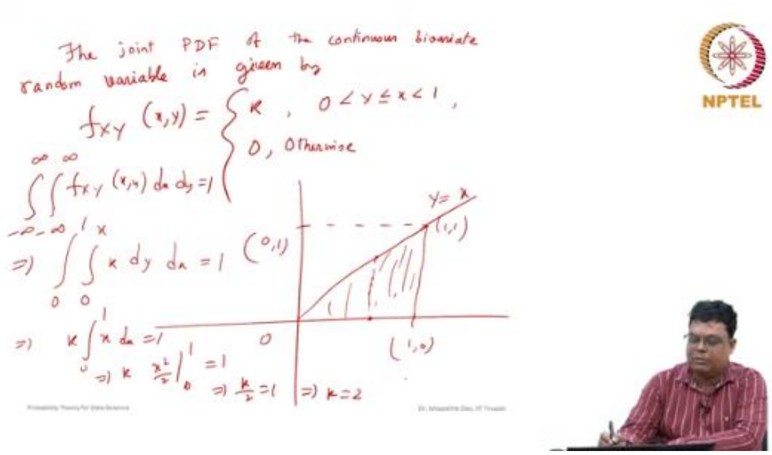




We previously computed these, but we will do it again for clarity. The joint probability density function of a bivariate random variable (X, Y) is given by a uniform distribution, which is constant when $0 \leq y \leq x \leq 1$ and 0 otherwise. This means that the value of the joint probability density function is non-zero only in the region where $0 \leq y \leq x \leq 1$. If you visualize this on a graph, with the axes representing x and y, the region of interest is the area below the line $y = x$ from (0, 0) to (1, 1). This region is defined by the condition $y \leq x$.

Now, to determine the value of k, we need to ensure that the total probability integrates to 1. To do this, we integrate the joint PDF over the entire region. The integration will be non-zero only in the region where $0 \leq y \leq x \leq 1$. We start by changing the order of integration to simplify the process. This allows us to integrate over y first, from 0 to x, and then over x from 0 to 1.

After performing the integration, we obtain the result $k * (x^2)/2$, evaluated from 0 to 1. This simplifies to $k/2$, and since the total probability must be 1, we set $k/2 = 1$, which gives $k = 2$. Thus, the value of k is 2. So, I quickly completed this because we have already done it before. You can go through it again if needed.



This integration is very simple. The value of k becomes 2 because the area under the curve is half of the total area, which is 1/2. This is a triangle with an area of (1/2) * 1 * 1. Therefore, to make the total area equal to 1, you must multiply the constant by 2, which gives k = 2. Now, let's quickly find the marginal probability density function of X.

The marginal probability density function of X is obtained by integrating the joint density function with respect to y. The joint density is non-zero only when x is between 0 and 1. Outside this range, the joint density becomes zero. For x between 0 and 1, the limit for y is from 0 to x because the function is non-zero in this region. Since we have found that k = 2, the integration becomes 2 * x.

Therefore, the marginal probability density function of X is:
$f(x) = 2x$ for $0 \leq x \leq 1$
$f(x) = 0$ otherwise.

Next, let's find the marginal probability density function of Y. This is calculated by integrating the joint density function with respect to x.

The joint density is non-zero only when y is between 0 and 1. For a particular value of y, the lower limit for x is y, and the upper limit is 1. So, the integration becomes 2 * (1 - y). Therefore, the marginal probability density function of Y is:

$f(y) = 2(1 - y)$ for $0 \leq y \leq 1$
$f(y) = 0$ otherwise.

We have already computed these values. Now, the next question is how to find the probability when 0 < x < 1/2 and 0 < y < 1/2.