1. Introduction on Conditional Mean & Conditional Variance

Welcome to the series of E-learning modules on Conditional Mean and Conditional Variance.

By the end of this session, you will be able to:

- Explain conditional mean for discrete and continuous random variables
- Explain conditional variance for discrete and continuous random variables
- Explain the expectation of conditional mean
- Explain variance as the sum of expectation of conditional variance plus variance of conditional expectation
- Explain the independence of random variables

Let us define conditional mean and conditional variance for two variables separately.
The two variables are discrete random variables & continuous random variables.

First let us consider discrete random variables:
Conditional mean value or expectation of a continuous function g of X, Y, given that Y is equal to yj, is defined by
Expectation of g of X, Y given Y is equal to yj is equal to summation over I is equal to 1 to infinity g of xi, yj into probability of X is equal to xi, given Y is equal to yj.
Is equal to summation over I is equal to 1 to infinity, g of xi, yj into probability of X is equal to xi intersection Y is equal to yj whole divided by Probability of Y is equal to yj.

That is Expectation of g of X, Y given Y is equal to yj is nothing but the expectation of the function of g of X, yj of X with respect to the conditional distribution of X when Y is equal to yj.
In particular, the conditional expectation of discrete random variable X given Y is equal to yj is given by,
Expectation of X given Y is equal to yj is equal to summation over I is equal to 1 to infinity xi into Probability of X is equal to xi given Y is equal to yj.
The conditional variance of X given Y is equal to yj is likewise given by,
Variance of X given Y is equal to yj is equal to expectation of X minus Expectation of X given Y is equal to yj whole square given Y is equal to yj.

Similarly, we can write conditional mean and conditional variance of Y given X is equal to xi can be written as follows.
Expectation of Y given X is equal to xi is equal to summation over j is equal to 1 to infinity, yj into probability of Y is equal to Yj given X is equal to xi.
Variance of Y given X is equal to xi is equal to expectation of Y minus expectation of Y given X is equal xi whole square given X is equal xi.

Now consider the case of continuous random variables:
The conditional expectation of g of (X,Y) on the hypothesis Y is equal to y is defined by,
Expectation of g of X, Y given Y is equal to y is equal to integral from minus infinity to infinity, g of x, y into f X given Y of x given by dx is equal to integral from minus infinity to infinity, g of

x, y into f Xy of xy into dx whole divided by f x of y.
The conditional mean of X given Y is equal to y is defined by
Expectation of X given Y is equal to y is equal to integral from minus infinity to infinity x into f XY of x, y dx whole divided by f Y of y.

Similarly, Expectation of Y given X is equal to x is equal to integral from minus infinity to infinity y into f XY of x, y dy whole divided by f X of x.
The conditional variance of X may be defined as variance of X given Y is equal to y is equal to expectation of X minus expectation of X given Y is equal to Y whole square given Y is equal to y.

Similarly the conditional variance of Y is, variance of Y given X is equal to x is equal to Expectation of Y minus Expectation of Y given X is equal to x the whole square given X is equal to x.

2. Theorem - 1

Now consider the following theorems:
The expected value of X is equal to the expectation of the conditional expectation of X given Y.
Symbolically, Expectation of X is equal to Expectation of {Expectation of X given Y}.

Let us prove this theorem as follows.
If X and Y are discrete random variables, then
Expectation of {Expectation of X given Y} is equal to Expectation of summation over I, xi into Probability of X is equal to xi given Y is equal to yi.
Is equal to expectation of summation over I, xi into probability of X is equal to xi intersection Y is equal to yj divided by probability of Y is equal to yj.
Is equal to summation over j, summation over i xi into probability of X is equal to xi intersection Y is equal to yj divided by probability of Y is equal to yj into probability of Y is equal to yj.

Is equal to double summation over I & j, xi into probability of X is equal to xi intersection Y is equal to yj.
Is equal to summation over I, xi into summation over j, probability of X is equal to xi intersection Y is equal to yj.
Is equal to summation over I, xi into probability of X is equal to xi, is equal to expectation of X.
If X and Y are continuous random variables,
Expectation of Expectation of X given Y is equal to Expectation of integral from minus infinity to infinity x into f of x given y dx.

Is equal to expectation of integral from minus infinity to infinity x into f of x, y divided by f Y of y dx.
Is equal to integral from minus infinity to infinity of integral from minus infinity to infinity x into f of x, y divided by f Y of y dx into f Y of y dy.
Is equal to integral from minus infinity to infinity, x into integral from minus infinity to infinity f of x, y dy, dx.
Is equal to integral from minus infinity to infinity x into f X of x dx, is equal to Expectation of X.

3. Theorem - 2

Consider the second theorem, which states that,
The variance of X can be regarded as consisting of two parts, the expectation of the conditional variance and the variance of the conditional expectation.
Symbolically, Variance of X is equal to Expectation of Variance of X given Y plus Variance of Expectation of X given Y.

To prove this, consider
Expectation of Variance of X given Y plus variance of Expectation of X given Y.
Is equal to expectation of E of X square given Y minus E of X given Y the square, plus expectation of E of X given Y whole square minus expectation of E of X given Y the whole square.
Is equal to Expectation of E of X square given y minus Expectation of E of X given Y whole square plus Expectation of E of X given Y whole square minus Expectation of E of X given Y the whole square.
Cancelling the same terms with opposite signs and using the first theorem for the last term we get,
Expectation of E of X square given Y minus Expectation of X the whole square.

Now consider the following two cases.
First case: If X and Y are discrete random variables.
Expectation of E of X square given Y.
Is equal to Expectation of Summation over I xi square into Probability of X is equal to xi given Y is equal to yj.
Is equal to summation over j, summation over I, xi square into probability of X is equal to xi intersection Y is equal to yj, whole divided by probability of Y is equal to yj into Probability of Y is equal to Yj.
Is equal to summation over I, xi square summation over j Probability of X is equal to xi intersection Y is equal to yj,
Is equal to summation over I, xi square into probability of X is equal to xi, is equal to Expectation of X square.

Consider the second case, if X and Y are continuous random variables.
Expectation of E of X square given Y.
Is equal to Expectation of integral from minus infinity to infinity x square into f of x given y dx is equal to expectation of integral from minus infinity to infinity x into f of x, y divided by f Y of y dx.
Is equal to integral from minus infinity to infinity of integral from minus infinity to infinity x square into f of x, y divided by f Y of y dx into f Y of y dy.
Is equal to integral from minus infinity to infinity x square into integral from minus infinity to infinity f of x, y dy dx.
Is equal to integral from minus infinity to infinity x square into f X of x dx, is equal to Expectation of X square.
Therefore, Expectation of Variance of X given Y plus Variance of Expectation of X given Y is equal to Expectation of X square minus Expectation of X the whole square, which is equal to Variance of X.

4. Theorem - 3

Now consider the third theorem:
Let A and B be two mutually exclusive events, then
Expectation of X given A union B is equal to probability of A into Expectation of X given A plus probability of B into Expectation of X given B divided by probability of A union B.
Where by definition, expectation of X given A is equal to 1 divided by probability of A into summation over xi belongs to A, xi into Probability of X is equal to xi.

Let us prove the above theorem as follows.
Consider Expectation of X given A union B is equal to 1 divided by probability of A union B into summation over xi belongs to A union B, xi into probability of X is equal to xi.
Since A and B are mutually exclusive events
Summation over xi belongs to A union B, xi into Probability of X is equal to xi is equal to Summation over xi belongs to A , xi into Probability of X is equal to xi plus Summation over xi belongs to B, xi into Probability of X is equal to xi.

Therefore, expectation of X given A union B is equal to 1 divided by probability of A union B into Probability of A into expectation of X given A plus probability of B into Expectation of X given B.
Hence the proof.
Note that, if we put B is equal to A complement, in the above theorem, we get
Expectation of X given A union A complement is equal to 1 divided by probability of A union A complement into probability of A into expectation of X given A plus Probability of A complement into Expectation of X given A complement.
Implies Expectation of X is equal to probability of A into expectation of X given A plus Probability of A complement into Expectation of X given A complement.

Consider the following result.
Given two variables X and Y with joint density function f of x,y, prove that conditional mean of Y given X coincide with unconditional mean only if X and Y are independent.

Let us prove the result as follows.
Conditional mean of Y given X is given by
Expectation of Y given X is equal to integral from minus infinity to infinity y into f of y given x dy, where f of y given x is a conditional probability density function of Y given X is equal to x, which is given by f of y given x is equal to f of x, y divided by f of x.

Hence, Expectation of Y given X is equal to integral from minus infinity to infinity y into f of x, y divided by f of x dy. Name it as equation 1.
Unconditional mean of Y is given by,
Expectation of Y is equal to integral from minus infinity to infinity y into f of y dy. Name it as equation 2.
From (1) and (2) we conclude that the conditional mean of Y given X will coincide with unconditional mean of Y only if,
f of x, y divided by f of x is equal to f of y which implies f of x, y is equal to f of x into f of y.
That is X and Y are independent.

5. Illustrative Example

Now let us consider some illustrations.
Let f of x, y is equal to 21 into x square into y cube & zero less than x less than y less than 1 and zero elsewhere be the joint probability density function of X and Y.
Find the conditional mean and variance of X given Y is equal to y, where zero less than y less than 1.

Let us solve the problem as follows.
Given f of x, y is equal to 21 into x square into y cube, zero less than x less than y less than 1 and zero elsewhere.
Marginal probability density function of Y is given by,
f Y of y is equal to integral from zero to y, f of x, y dx is equal to 21 into y cube into integral from zero to y, x square dx is equal to 7 into y power 6.
Therefore, the conditional probability density function of X given Y is given by,
f of x given y is equal to f of x, y divided by f Y of y is equal to 21 into x square into y cube whole divided by 7 into y power 6 is equal to 3 into x square divided by y cube, where zero less than x less than y and zero less than y less than 1.

Conditional mean of X is,
Expectation of X given y is equal to y is equal to integral from zero to y, x into f of x given y dx. Is equal to 3 divided by y cube into integral from zero to y, x cube dx. On integrating and simplifying we get, 3 into y divided by 4, where zero less than y less than 1.
To find variance, first let us find the expectation of X square given Y is equal to y is equal to integral from zero to y, x square into f of x given y dx.
Is equal to 3 divided by y cube into integral from zero to y x power 4 dx is equal to 3 divided by y cube into y power 5 divided by 5 is equal to 3 divided by 5 into y square, where zero less than y less than 1.
Therefore, Variance of X given Y is equal to y is equal to Expectation of X square given Y is equal to y minus Expectation of X given Y is equal to y the whole square.
Is equal to 3 divided by 5 into y square minus 9 divided by 16 into y square is equal to 3 divided by 80 y square, where zero less than y less than 1.

Let X and Y be two random variable each taking three values minus 1, 0, 1 and having the following joint probability distribution, find Variance of Y given X is equal to minus1.

Figure 1

Y \ X	-1	0	1	Total
-1	0	0.1	0.1	**0.2**
0	0.2	0.2	0.2	**0.6**
1	0	0.1	0.1	**0.2**
Total	**0.2**	**0.4**	**0.4**	**1.0**

The table shows the values taken by the variables X and Y with their respective probabilities. That is when Y takes value minus 1, X take values minus 1, zero and 1 with respective probabilities zero, zero point 1, zero point 1.
When Y takes value zero, X take values minus 1, zero and 1 with probabilities zero point 2, zero point 2 and zero point 2 respectively and
When Y takes value 1, X take values, minus 1, zero and 1 with respective probabilities zero, zero point 1 and zero point 1.
The numbers written in bold gives the row totals and column total and the grand total adds up to 1.

We solve the above problem as follows.
We know that, Variance of Y given X is equal to minus 1 is equal to Expectation of Y square given X is equal to minus 1 minus Expectation of Y given x is equal to minus 1 the whole square. Name it as 1.
Expectation of Y given x is equal to minus 1 is equal to summation over y, y into Probability of Y is equal to y and X is equal to minus 1.
Is equal to minus 1 into zero plus zero into zero point 2 plus 1 into zero is equal to zero.
Expectation of Y square given x is equal to minus 1 is equal to summation over y, y square into Probability of Y is equal to y and X is equal to minus 1.
Is equal to minus 1 square into zero plus zero square into zero point 2 plus 1 square into zero is equal to zero.
Therefore, on substituting the values in 1, we get,
Variance of Y given X is equal to minus 1 is equal to zero.

Here's a summary of our learning in this session, where we have understood:

- The Conditional mean for discrete and continuous random variables
- The Conditional variance for discrete and continuous random variables
- The Expectation of conditional mean
- The Variance as the sum of expectation of conditional variance plus variance of conditional expectation
- The Independence of random variables