Lecture 8: Conditional Expectation & Covariance

24 March, 2018

Sunil Kumar Gauttam

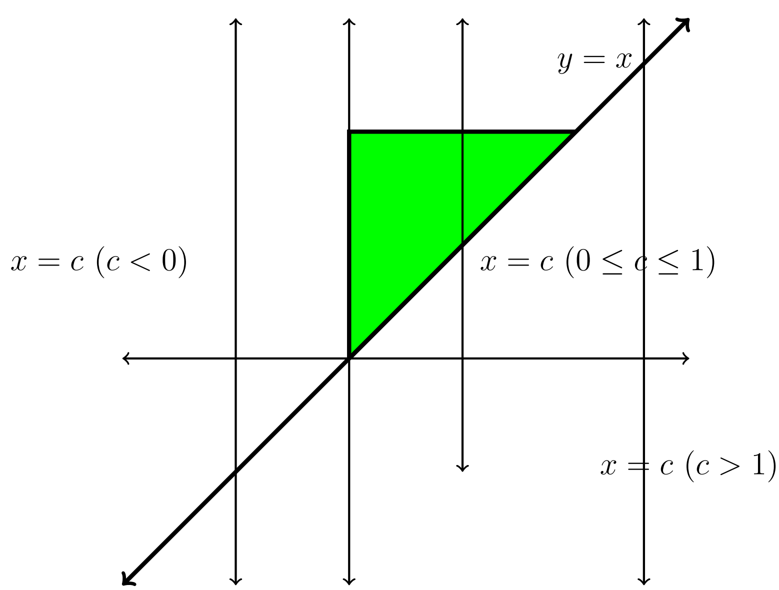
Department of Mathematics, LNMIIT

Example 8.1 *Let X, Y be continuous random variables with joint pdf given by*

$$f(x, y) = \begin{cases} 6(y - x) & ; \quad 0 \leq x \leq y \leq 1 \\ 0 & ; \quad \textit{otherwise} \end{cases}$$

Find $E[Y|X = x]$ and hence calculate EY.

Solution: In order to calculate $E[Y|X = x]$ we need to find $f_{Y|X}$, which is by definition equal to $\dfrac{f(x, y)}{f_X(x)}$.



Note that

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\
&= \begin{cases} \int_x^1 f(x,y)dy & ; \quad 0 \le x \le 1 \\ 0 & ; \quad \text{otherwise} \end{cases} \\
&= \int_x^1 6(y-x)dy \\
&= 6\left[\frac{y^2}{2} - xy\right]_x^1 \\
&= 6\left[\frac{x^2}{2} - x + \frac{1}{2}\right] \\
&= \begin{cases} 3(x-1)^2 & ; \quad 0 \le x \le 1 \\ 0 & ; \quad \text{otherwise} \end{cases}
\end{aligned}
$$

This implies that

$$
f_{Y|X}(y|x) = \begin{cases} \dfrac{2(y-x)}{(x-1)^2} & ; \quad 0 \le x \le y < 1 \\ 0 & ; \quad \text{otherwise} \end{cases}
$$

Hence $E[Y|X=x]$ would be non-zero only if $0 \le x < 1$.

$$
\begin{aligned}
E[Y|X=x] &= \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy \\
&= \int_x^1 y\frac{2(y-x)}{(x-1)^2}dy \\
&= \frac{2}{(x-1)^2}\int_x^1 (y^2 - xy)dy \\
&= \frac{2}{(x-1)^2}\left[\frac{y^3}{3} - x\frac{y^2}{2}\right]_x^1 \\
&= \frac{2}{(x-1)^2}\left[\frac{1}{3} - \frac{x}{2} + \frac{x^3}{6}\right] \\
&= \frac{2(x^3 - 3x + 2)}{6(x-1)^2} \\
&= \frac{x^2 + x - 2}{3(x-1)}
\end{aligned}
$$

Therefore

$$\begin{aligned} EY &= \int_{-\infty}^{\infty} E[Y|X=x] f_X(x) dx \\ &= \int_0^1 \frac{x^2+x-2}{3(x-1)} \times 3(x-1)^2 dx \\ &= \int_0^1 (x^2+x-2)(x-1) dx \\ &= \int_0^1 (x^3-3x+2) dx \\ &= \frac{3}{4} \end{aligned}$$

■

Theorem 8.2 *1. Let X and Y be discrete random variables with joint pmf f. If g is a function then*

$$E[g(X)|Y=y] = \sum_x g(x) f_{X|Y}(x|y),$$

provided $\sum_x |g(x)| f_{X|Y}(x|y) < \infty$.

2. Let X and Y be random variables with joint pdf f. If $g : \mathbb{R} \to \mathbb{R}$ is a Borel function then

$$E[g(X)|Y=y] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y) dx,$$

provided $\int_{-\infty}^{\infty} |g(x)| f_{X|Y}(x|y) dx < \infty$.

Covariance

Definition 8.3 *The covariance of two random variables X and Y, denoted by $\mathrm{cov}(X,Y)$ is defined by*

$$\mathrm{cov}(X,Y) = E\left[(X-EX)(Y-EY)\right]$$

When $\mathrm{cov}(X,Y)=0$, we say that X and Y are uncorrelated.

The covariance gives information about how random variable X and Y are linearly related. Intuitively, the covariance between X and Y indicates how the values of X and Y move relative to each other. If large values of X tend to happen with large values of Y, then the covariance is positive (or small values of X tend to happen with small values of Y) and we say X and Y are positively correlated. On the other hand, if X tends to be small when Y is large (or vice-versa), then the covariance is negative and we say X and Y are negatively correlated.

By appealing to the linearity of the expectation,

$$\begin{aligned}\text{cov}(X,Y) &= E\left[XY - XEY + YEX - EXEY\right] \\ &= E[XY] - EXEY + EYEX - EXEY \\ &= E[XY] - E[X]E[Y].\end{aligned}$$

If X and Y are independent then $E[XY] = EXEY$, therefore $\text{cov}(X,Y) = 0$. But converse is not true in general.

Example 8.4 *Let the joint probabilities of random variables X and Y are given by the following table.*

X \ Y	*-1*	*0*	*1*
-1	0	$\frac{1}{4}$	0
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{4}$	0

Then we have seen that X and Y are identically distributed and

$$P(X=-1) = P(X=1) = \frac{1}{4} \quad \textit{and } P(X=0) = \frac{1}{2}$$

Also, it is easy to see that $E[X] = E[Y] = 0$. Furthermore, random variable XY takes values $\{-1,0,1\}$ with the pmf

$$P(XY=1) = 0 = P(XY=-1) \textit{ and } P(XY=0) = 1.$$

Therefore, $E[XY] = 0$, which in turn implies $\text{cov}(X,Y) = 0$. However, X and Y are not independent since

$$P(X=-1, Y=-1) = 0 \neq \frac{1}{16} = P(X=-1)P(Y=-1)$$

Proposition 8.5 *For any random variable X, Y and Z, and any $a, b \in \mathbb{R}$,*

1. $\text{cov}(X,X) = \text{var}(X)$.
2. $\text{cov}(X,Y) = \text{var}(Y,X)$.
3. $\text{cov}(X, aY+b) = a\ \text{cov}(X,Y)$.

4. $\text{cov}(X, Y + Z) = \text{cov}(X, Y) + \text{cov}(X, Z)$*.*

Proof: All the proof follows from definition of covariance and linearity of expectation.

1. $$\text{cov}(X, X) = E[X^2] - EX\ EX = var(X)$$

2. immediate from definition.

3. $$\begin{aligned} \text{cov}(X, aY + b) &= E[X(aY + b)] - EXE[aY + b] = E[aXY + bX] - EX[aEY + b] \\ &= aE[XY] + bEX - aEXEY - bEX = a\left[E[XY] - EXEY\right] = \text{cov}(X, Y) \end{aligned}$$

4. $$\begin{aligned} \text{cov}(X, Y + Z) &= E[X(Y + Z)] - EXE(Y + Z) = E[XY + XZ] - EX[EY + EZ] \\ &= E[XY] + E[XZ] - EXEY - EXEZ = \text{cov}(X, Y) + \text{cov}(X, Z) \end{aligned}$$

■

Example 8.6 *Let X and Y be two independent $N(0, 1)$ random variables and $Z = 1 + X + XY^2, W = 1 + X$. Find $\text{cov}(Z, W)$.*

Solution:

$$\begin{aligned} \text{cov}(Z, W) &= \text{cov}(1 + X + XY^2, 1 + X) \\ &= \text{cov}(X + XY^2, X) \quad (\because \text{adding a constant to any rv does not affect the covariance}) \\ &= \text{cov}(X, X) + \text{cov}(XY^2, X) \\ &= var(X) + E[XY^2X] - E[XY^2]EX = 1 + E[X^2]E[Y^2] = 2 \end{aligned}$$

■

Example 8.7 *For any random variables X, Y, show that*

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y).$$

Solution:

$$\begin{aligned} \text{var}(X + Y) &= \text{cov}(X + Y, X + Y) = \text{cov}(X + Y, X) + \text{cov}(X + Y, Y) \\ &= \text{cov}(X, X) + \text{cov}(Y, X) + \text{cov}(X, Y) + \text{cov}(Y, Y) \\ &= \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y). \end{aligned}$$

■