Conditional Expectation and Variance

Definition 1. *Let (X,Y) be a random vector and $h:\mathbb{R}\longrightarrow\mathbb{R}$ be a function such that $h^{-1}(A)\in\mathbb{B}_{\mathbb{R}}$. Then*

(1) *the conditional expectation of $h(X)$, given Y, written as $E[h(X)|Y]$, is an random variable that takes the value $E[h(X)|Y=y]$, defined by*

$$E[h(X)|Y=y]=\begin{cases}\sum\limits_{x\in E_{X|Y=y}} h(x)P(X=x|Y=y), & \text{if } (X,Y) \text{ is of discrete type and } P(Y=y)>0\\ \int\limits_{-\infty}^{\infty} h(x)f_{X|Y}(x|y)\,dx, & \text{if } (X,Y) \text{ is of continuous type and } f_Y(y)>0\end{cases}$$

(2) *the conditional variance of $h(X)$, given Y, written as $Var[h(X)|Y]$, is an random variable that takes the value $Var[h(X)|Y=y]$, defined by*

$$\begin{aligned}Var[h(X)|Y=y]&=E[\big(h(X)-E[h(X)|Y=y]\big)^2|Y=y]\\&=E[\big(h(X)\big)^2|Y=y]-(E[h(X)|Y=y])^2\end{aligned}$$

Remark 2. (1) *For any constant c, $E[c|Y]=c$.*

(2) *Let $h_i:\mathbb{R}\longrightarrow\mathbb{R}$ be a function such that $h_i^{-1}(A)\in\mathbb{B}_{\mathbb{R}}$, for $i=1,2$. Then*

$$E[a_1h_1(X)+a_2h_2(X)|Y]=a_1E[h_1(X)|Y]+a_2E[h_2(X)|Y],$$

for any constants a_1,a_2.

(3) *If X and Y are independent, then*

$$E[h(X)|Y]=E(h(X)) \text{ and } Var[h(X)|Y]=Var(h(X)).$$

(4) *If $P(X\geq 0)=1$, then $E[X|Y]\geq 0$.*

(5) *If $P(X_1\geq X_2)=1$, then $E[X_1|Y]\geq E[X_1|Y]$.*

Theorem 3. (1) *Let $E(h(X))$ exist. Then*

$$E(h(X))=E(E[h(X)|Y]).$$

(2) ***The conditional Variance Formula:***

$$Var(h(X))=Var(E[h(X)|Y])+E(Var[h(X)|Y]).$$

Proof. Let (X,Y) be of the discrete type. Then

$$\begin{aligned}E(E[h(X)|Y])&=\sum_y E[h(X)|Y=y]P(Y=y)\\&=\sum_y\left[\sum_x h(x)P(X=x|Y=y)\right]P(Y=y)\\&=\sum_y\left[\sum_x h(x)P(X=x,Y=y)\right]\\&=\sum_x\left[\sum_y h(x)P(X=x,Y=y)\right]\\&=\sum_x h(x)P(X=x)\\&=E(h(X)).\end{aligned}$$

(2)

$$\begin{aligned} E(Var[h(X)|Y]) &= E(E[\big(h(X)\big)^2|Y] - (E[h(X)|Y])^2) \\ &= E(E[\big(h(X)\big)^2|Y]) - E((E[h(X)|Y])^2)) \\ &= E(\big(h(X)\big)^2) - E((E[h(X)|Y])^2)) \end{aligned}$$

and

$$\begin{aligned} Var(E[h(X)|Y]) &= E((E[h(X)|Y])^2)) - (E(E[h(X)|Y]))^2) \\ &= E((E[h(X)|Y])^2)) - (E(h(X)))^2 \end{aligned}$$

Thus $Var(E[h(X)|Y]) + E(Var[h(X)|Y]) = E(\big(h(X)\big)^2) - (E(h(X)))^2 = Var(h(X))$.

□

Example 4. *Let $\underline{Z} = (X, Y, Z)$ be a random vector with joint p.m.f.*

$$f(x,y,z) = \begin{cases} \frac{xyz}{72}, & \text{if } (x,y,z) \in \{1,2\} \times \{1,2,3\} \times \{1,3\} \\ 0, & \text{otherwise} \end{cases}$$

(1) *Let $Y_1 = 2X - Y + 3Z$ and $Y_2 = X - 2Y + Z$. Find the correlation coefficient between Y_1 and Y_2.*
(2) *For a fixed $y \in \{1, 2, 3\}$, find $E[Y_3|Y = y]$ and $Var[Y_3|Y = y]$, where $Y_3 = XZ$.*

Solution:

(1) By Example 10 of Lecture 16, we know that the marginal p.m.f. of X, Y and Z are

$$f_X(x) = \begin{cases} \frac{x}{3}, & \text{if } x \in \{1,2\} \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{y}{6}, & \text{if } y \in \{1,2,3\} \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_Z(z) = \begin{cases} \frac{z}{4}, & \text{if } y \in \{1,3\} \\ 0, & \text{otherwise} \end{cases}$$

respectively. Also X, Y, Z are independent. Therefore, $Cov(X, Y) = Cov(X, Z) = Cov(Y, Z) = 0$. Hence,

$$Cov(Y_1, Y_2) = 2Var(X) + 2Var(Y) + 3Var(Z);$$

$$Var(Y_1) = 4Var(X) + Var(Y) + 9Var(Z);$$

and

$$Var(Y_2) = Var(X) + 4Var(Y) + Var(Z).$$

By a simple calculation, we have

$$E(X) = \frac{5}{3}, E(Y) = \frac{7}{3} \text{ and } E(Z) = \frac{5}{2};$$
$$E(X^2) = 3, E(Y^2) = 6 \text{ and } E(Z^2) = 7;$$
$$Var(X) = \frac{2}{9}, Var(Y) = \frac{5}{9} \text{ and } Var(Z) = \frac{3}{4}$$

Therefore,

$$Cov(Y_1, Y_2) = \frac{137}{36}, Var(Y_1) = \frac{295}{36} \text{ and } Var(Y_2) = \frac{115}{36}.$$

Thus,

$$\rho(Y_1, Y_2) = \frac{Cov(Y_1, Y_2)}{\sqrt{Var(Y_1)Var(Y_2)}} = \frac{137}{\sqrt{295}\sqrt{115}}.$$

(2) As we know that X, Y, Z are independent, it follows that (X, Z) and Y are independent. Thus, $Y_3 = XZ$ and Y are independent. Therefore, $E[Y_3|Y = y] = E(Y_3) = E(X)E(Z) = \frac{25}{6}$ and

$$\begin{aligned} Var[Y_3|Y = y] &= Var(Y_3) \\ &= Var(E[XZ|Z]) + E(Var[XZ|Z]) \\ &= Var(ZE[X|Z]) + E(Z^2 Var[X|Z]) \\ &= Var(ZE(X)) + E(Z^2 Var(X)) \\ &= Var(\frac{5}{3}Z) + E(\frac{2}{9}Z^2) \\ &= \frac{25}{9}Var(Z) + \frac{2}{9}E(Z^2) \\ &= \frac{131}{36}. \end{aligned}$$

Example 5. *Let $\underline{Z} = (X, Y)$ be a random vector with joint p.d.f.*

$$f(x, y) = \begin{cases} 2, & \text{if } 0 < x < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

For a fixed $0 < x < 1$, find $E[Y|X = x]$ and $Var[Y|X = x]$, and for a fixed $0 < y < 1$, find $E[X|Y = y]$ and $Var[X|Y = y]$.

Solution: The marginal p.d.f. of X and Y are

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_x^1 2\, dy = \begin{cases} 2(1-x), & \text{if } 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_0^y 2\, dx = \begin{cases} 2y, & \text{if } 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

respectively. Hence, the conditional p.d.f. of Y, given $X = x$ and the conditional p.d.f. of X, given $Y = y$ are

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \begin{cases} \frac{1}{1-x}, & \text{if } x < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \begin{cases} \frac{1}{y}, & \text{if } 0 < x < y \\ 0, & \text{otherwise} \end{cases}$$

Thus,

$$E[Y|X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\, dy = \int_x^1 \frac{y}{1-x}\, dy = \frac{1+x}{2};$$

$$E[Y^2|X=x] = \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|x)\,dy = \int_{x}^{1} \frac{y^2}{1-x}\,dy = \frac{1+x+x^2}{3};$$

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx = \int_{0}^{y} \frac{x}{y}\,dy = \frac{y}{2};$$

$$E[X^2|Y=y] = \int_{-\infty}^{\infty} x^2 f_{X|Y}(x|y)\,dx = \int_{0}^{y} \frac{x^2}{y}\,dy = \frac{y^2}{3}.$$

Hence,

$$Var[Y|X=x] = E[Y^2|X=x] - (E[Y|X=x])^2 = \frac{1+x+x^2}{3} - \frac{1+2x+x^2}{4} = \frac{x^2-2x+1}{12};$$

and

$$Var[X|Y=y] = E[X^2|Y=y] - (E[X|Y=y])^2 = \frac{y^2}{3} - \frac{y^2}{4} = \frac{y^2}{12}.$$

Example 6. *Suppose that the expected number of accidents per week at an industrial plant is four. Suppose also that the numbers of workers injured in each accident are independent random variables with a common mean of* 2*. Assume also that the number of workers injured in each accident is independent of the number of accidents that occur. What is the expected number of injuries during a week?*

Solution: Let N denote the number of accidents and X_i the number of workers injured in the $i-$th accident, $i = 1, 2, \ldots$, then the total number of injuries can be expressed as $\sum_{i=1}^{N} X_i$. Now, $E(\sum_{i=1}^{N} X_i) = E(E[\sum_{i=1}^{N} X_i|N])$.

But $E[\sum_{i=1}^{N} X_i|N = n] = E[\sum_{i=1}^{n} X_i|N = n] = E(\sum_{i=1}^{n} X_i) = nE(X_i)$ (since X_i and N are independent, and X_i has common mean). Thus, $E[\sum_{i=1}^{N} X_i|N] = NE(X_i)$. Therefore, $E(\sum_{i=1}^{N} X_i) = E(NE(X_i)) = E(N)E(X_i) = 8$.