EXPECTATIONS AND CONDITIONAL EXPECTATIONS

The joint density function of x and y is

$$f(x,y) = f(x|y)f(y) = f(y|x)f(x), \tag{1}$$

where

$$f(x) = \int_y f(x,y)dx \qquad \text{and} \qquad f(y) = \int_x f(x,y)dy \tag{2}$$

are the marginal distributions of x and y respectively and where

$$f(x|y) = \frac{f(y,x)}{f(y)} \qquad \text{and} \qquad f(y|x) = \frac{f(y,x)}{f(x)} \tag{3}$$

are the conditional distributions of x given y and of y given x.

The unconditional expectation of $y \sim f(y)$ is

$$E(y) = \int_y yf(y)dy. \tag{4}$$

The conditional expectation of y given x is

$$E(y|x) = \int_y yf(y|x)dy = \int_y y\frac{f(y,x)}{f(x)}dy. \tag{5}$$

The expectation of the conditional expectation is an unconditional expectation:

$$\begin{aligned} E\{E(y|x)\} &= \int_x \left\{ \int_y y\frac{f(y,x)}{f(x)}dy \right\} f(x)dx \\ &= \int_x \int_y yf(y,x)f(x)dydx \\ &= \int_y y\left\{ \int_x f(y,x)dx \right\} dy = \int_y f(y)dy = E(y). \end{aligned} \tag{6}$$

The conditional expectation of y given x is the minimum mean squared error prediction; and the error in predicting y is uncorrelated with x. The proof of this depends on showing that $E(\hat{y}x) = E(yx)$, where $\hat{y} = E(y|x)$:

$$\begin{aligned} E(\hat{y}x) &= \int_x xE(y|x)f(x)dx \\ &= \int_x x\left\{ \int_y y\frac{f(y,x)}{f(x)}dy \right\} f(x)dx \\ &= \int_x \int_y xyf(y,x)dydx = E(xy). \end{aligned} \tag{7}$$

The result can be expressed as $E\{(y-\hat{y})x\}=0$.

This result can be used in deriving expressions for the parameters α and β of a linear regression of the form

$$E(y|x)=\alpha+\beta x, \tag{8}$$

from which and unconditional expectation is derived in the form of

$$E(y)=\alpha+\beta E(x). \tag{9}$$

The orthogonality of the prediction error implies that

$$\begin{aligned} 0=E\{(y-\hat{y})x\} &= E\{(y-\alpha-\beta x)x\} \\ &= E(xy)-\alpha E(x)-\beta E(x^2). \end{aligned} \tag{10}$$

In order to eliminate $\alpha E(x)$ from this expression, equation (9) is multiplied by $E(x)$ and rearranged to give

$$\alpha E(x)=E(x)E(y)-\beta\{E(x)\}^2. \tag{11}$$

This substituted into (10) to give

$$E(xy)-E(x)E(y)=\beta\big[E(x^2)-\{E(x)\}^2\big], \tag{13}$$

whence

$$\beta=\frac{E(xy)-E(x)E(y)}{E(x^2)-\{E(x)\}^2}=\frac{C(x,y)}{V(x)}. \tag{14}$$

The expression

$$\alpha=E(y)-\beta E(x) \tag{15}$$

comes directly from (9).