Covariance and correlation.

- Covariance. Suppose that X and Y are random variables with $E(X) = \mu_X$ and $E(Y) = \mu_Y$. Then the covariance of X and Y, denoted by $Cov(X,Y)$, is the expectation of $(X-\mu_X)(Y-\mu_Y)$. That is,

$$Cov(X,Y) = E((X-\mu_X)(Y-\mu_Y)) = E(XY) - E(X)E(Y).$$

- Suppose that X and Y are discrete random variables. Then $Cov(X,Y)$ can be determined if $P((X,Y)=(x,y))$ is given for each (x,y).

- Example 1. Suppose that

$$P((X,Y)=(x,y)) = \begin{cases} 0.1 & \text{if } (x,y)=(1,-2); \\ 0.3 & \text{if } (x,y)=(2,-4); \\ 0.6 & \text{if } (x,y)=(3,-6); \\ 0 & \text{otherwise.} \end{cases}$$

  Find $E(X)$, $E(Y)$, $E(XY)$ and $Cov(X,Y)$.

  Sol.

$$E(X) = 1\times 0.1 + 2\times 0.3 + 3\times 0.6 = 2.5,$$

$$E(Y) = (-2)\times 0.1 + (-4)\times 0.3 + (-6)\times 0.6 = -5,$$

$$E(XY) = (1)(-2)\times 0.1 + (2)(-4)\times 0.3 + (3)(-6)\times 0.6 = -13.4,$$

  and

$$Cov(X,Y) = E(XY) - E(X)E(Y) = -13.4 - (2.5)\times(-5) = -0.9.$$

- If X and Y are independent, then $E(XY) = E(X)E(Y)$ and $Cov(X,Y) = 0$.

- Rules for covariance calculation.

  1. $Cov(X,X) = Var(X)$.
  2. $Cov(X,Y) = Cov(Y,X)$.
  3. Linearity (in one variable when the other variable is held fixed).

$$\begin{array}{l} Cov(aX,Y) = aCov(X,Y) = Cov(X,aY) \\ Cov(X_1+X_2,Y) = Cov(X_1,Y) + Cov(X_2,Y) \quad \text{and} \\ Cov(X,Y_1+Y_2) = Cov(X,Y_1) + Cov(X,Y_2). \end{array}$$

- $Var(aX+bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X,Y)$.

Example 2. Suppose that $Cov(X, Y) = -0.9$, $Var(X) = 0.45$ and $Var(Y) = 1.8$. Find $Var(2X + Y)$.

Sol. $Var(2X+Y) = 4Var(X) + 4Cov(X, Y) + Var(Y) = 4 \times 0.45 + 4 \times (-0.9) + 1.8 = 0$.

- Correlation. Suppose that $Var(X) > 0$ and $Var(Y) > 0$. Then the correlation (or correlation coefficient) between X and Y, denoted by $Corr(X, Y)$, is
$$\frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

- Example 3. Consider the (X, Y) in Example 1. Find $Corr(X, Y)$.

  Sol. From the solution for Example 1, we have $Cov(X, Y) = -0.9$, $E(X) = 2.5$ and $E(Y) = -5$. Therefore,

  $Var(X) = E(X^2) - (2.5)^2 = 1^2 \times 0.1 + 2^2 \times 0.3 + 3^2 \times 0.6 - (2.5)^2 = 0.45$,

  $Var(Y) = E(Y^2) - (-5)^2 = ((-2)\times 1)^2 \times 0.1 + ((-2)\times 2)^2 \times 0.3 + ((-2)\times 3)^2 \times 0.6 - (-5)^2 = 1.8$,

  and
$$Corr(X, Y) = \frac{-0.9}{\sqrt{0.45 \times 1.8}} = -1.$$

- Properties of correlation. Suppose that $Var(X) > 0$ and $Var(Y) > 0$.

  1. $-1 \le Corr(X, Y) \le 1$.
  2. $|Corr(X, Y)| = 1 \Leftrightarrow Y = a + bX$ for some a, b.
  3. Suppose that $Y = a + bX$ for some a, b (with probability one). Then
$$Corr(X, Y) = \begin{cases} 1 & \text{if } b > 0; \\ -1 & \text{if } b < 0. \end{cases}$$
  4. If X and Y are independent, then $Corr(X, Y) = 0$.

- Note that
$$Corr(X, Y) = 0 \not\Rightarrow \ X \text{ and } Y \text{ are independent}$$

Example 4. Suppose that $P(X = 0) = P(X = 1) = P(X = -1) = 1/3$ and $Y = X^2$. Then $Cov(X, Y) = 0$ but X and Y are not independent since $P((X, Y) = (0, 0)) \neq P(X = 0)P(Y = 0)$.

Proof of $-1 \le Corr(X,Y) \le 1$ assuming $Var(X) > 0$ and $Var(Y) > 0$.

Consider the problem of finding constants a and b so that

$$E(Y-a-bX)^2$$

is minimized. Note that

$$\begin{aligned}
&E(Y-a-bX)^2 \\
&= Var(Y-a-bX) + (E(Y-a-bX))^2 \\
&= Var(Y-bX) + (E(Y) - bE(X) - a)^2 \\
&= b^2 Var(X) - 2bCov(X,Y) + Var(Y) + (E(Y) - bE(X) - a)^2 \\
&= Var(X)\left(b - \frac{Cov(X,Y)}{Var(X)}\right)^2 + Var(Y) - Var(X)\left(\frac{Cov(X,Y)}{Var(X)}\right)^2 \\
&\quad + (E(Y) - bE(X) - a)^2,
\end{aligned}$$

where

$$\begin{aligned}
&Var(Y) - Var(X)\left(\frac{Cov(X,Y)}{Var(X)}\right)^2 \\
&= Var(Y)\left(1 - \frac{Cov(X,Y)^2}{Var(X)Var(Y)}\right) \\
&= Var(Y)\left(1 - (Corr(X,Y))^2\right),
\end{aligned}$$

so

$$\begin{aligned}
&E(Y-a-bX)^2 \\
&= Var(X)\left(b - \frac{Cov(X,Y)}{Var(X)}\right)^2 + (E(Y) - bE(X) - a)^2 \\
&\quad + Var(Y)\left(1 - (Corr(X,Y))^2\right). \qquad (1)
\end{aligned}$$

(1) implies that $E(Y-a-bX)^2$ is minimized when

$$b = \frac{Cov(X,Y)}{Var(X)} \text{ and } a = E(Y) - bE(X), \qquad (2)$$

and the minimum of $E(Y-a-bX)^2$ is

$$Var(Y)\left(1 - (Corr(X,Y))^2\right),$$

which must be nonnegative. Therefore,

$$1 - Corr(X,Y)^2 \ge 0$$

and $-1 \le Corr(X,Y) \le 1$.

- Remark. (1) also implies that when $|Corr(X,Y)| = 1$, $E(Y-a-bX)^2 = 0$ for the a and b in (2).