$$\begin{aligned}\mathrm{Var}(Z) &= \mathrm{Cov}(Z, Z) \\ &= \mathrm{Cov}(X+Y, X+Y) \\ &= \mathrm{Cov}(X, X) + \mathrm{Cov}(X, Y) + \mathrm{Cov}(Y, X) + \mathrm{Cov}(Y, Y) \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y).\end{aligned}$$

More generally, for $a, b \in \mathbb{R}$, we conclude:

$$\boxed{\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X, Y) \qquad (5.21)}$$

Correlation Coefficient:

The **correlation coefficient**, denoted by ρ_{XY} or $\rho(X, Y)$, is obtained by normalizing the covariance. In particular, we define the correlation coefficient of two random variables X and Y as the covariance of the standardized versions of X and Y. Define the standardized versions of X and Y as

$$U = \frac{X - EX}{\sigma_X}, \quad V = \frac{Y - EY}{\sigma_Y} \qquad (5.22)$$

Then,

$$\begin{aligned}\rho_{XY} = \mathrm{Cov}(U, V) &= \mathrm{Cov}\left(\frac{X - EX}{\sigma_X}, \frac{Y - EY}{\sigma_Y}\right) \\ &= \mathrm{Cov}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) && \text{(by Item 5 of Lemma 5.3)} \\ &= \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}.\end{aligned}$$

$$\boxed{\rho_{XY} = \rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var(X)}\ \mathrm{Var(Y)}}} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}}$$

A nice thing about the correlation coefficient is that it is always between -1 and 1. This is an immediate result of Cauchy-Schwarz inequality that is discussed in Section 6.2.4. One way to prove that $-1 \leq \rho \leq 1$ is to use the following inequality:

$$\alpha\beta \leq \frac{\alpha^2+\beta^2}{2}, \text{for } \alpha, \beta \in \mathbb{R}.$$

This is because $(\alpha-\beta)^2 \geq 0$. The equality holds only if $\alpha = \beta$. From this, we can conclude that for any two random variables U and V,

$$E[UV] \leq \frac{EU^2+EV^2}{2},$$

with equality only if $U = V$ with probability one. Now, let U and V be the standardized versions of X and Y as defined in Equation 5.22. Then, by definition $\rho_{XY} = \text{Cov}(U,V) = EUV$. But since $EU^2 = EV^2 = 1$, we conclude

$$\rho_{XY} = E[UV] \leq \frac{EU^2+EV^2}{2} = 1,$$

with equality only if $U = V$. That is,

$$\frac{Y-EY}{\sigma_Y} = \frac{X-EX}{\sigma_X},$$

which implies

$$\begin{aligned} Y &= \frac{\sigma_Y}{\sigma_X}X + \left(EY - \frac{\sigma_Y}{\sigma_X}EX\right) \\ &= aX+b, \text{ where } a \text{ and } b \text{ are constants.} \end{aligned}$$

Replacing X by $-X$, we conclude that

$$\rho(-X,Y) \leq 1.$$

But $\rho(-X,Y) = -\rho(X,Y)$, thus we conclude $\rho(X,Y) \geq -1$. Thus, we can summarize some properties of the correlation coefficient as follows.

Properties of the correlation coefficient:

1. $-1 \leq \rho(X,Y) \leq 1$;
2. if $\rho(X,Y) = 1$, then $Y = aX + b$, where $a > 0$;
3. if $\rho(X,Y) = -1$, then $Y = aX + b$, where $a < 0$;
4. $\rho(aX + b, cY + d) = \rho(X,Y)$ for $a, c > 0$.

Definition 5.2
Consider two random variables X and Y:
- If $\rho(X,Y) = 0$, we say that X and Y are **uncorrelated**.
- If $\rho(X,Y) > 0$, we say that X and Y are **positively** correlated.
- If $\rho(X,Y) < 0$, we say that X and Y are **negatively** correlated.

Note that as we discussed previously, two independent random variables are always uncorrelated, but the converse is not necessarily true. That is, if X and Y are uncorrelated, then X and Y may or may not be independent. Also, note that if X and Y are uncorrelated from Equation 5.21, we conclude that $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

If X and Y are uncorrelated, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

More generally, if $X_1, X_2, \ldots, X_n$ are pairwise uncorrelated, i.e., $\rho(X_i, X_j) = 0$ when $i \neq j$, then

$$\mathrm{Var}(X_1 + X_2 + \ldots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n).$$

Note that if X and Y are independent, then they are uncorrelated, and so $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. This is a fact that we stated previously in Chapter 3, and now we could easily prove using covariance.

Example 5.34

Let X and Y be as in Example 5.24 in Section 5.2.3, i.e., suppose that we choose a point (X, Y) uniformly at random in the unit disc

$$D = \{(x, y) | x^2 + y^2 \leq 1\}.$$

Are X and Y uncorrelated?

Solution

We need to check whether $\textrm{Cov}(X, Y) = 0$. First note that, in Example 5.24 of Section 5.2.3, we found out that X and Y are not independent and in fact, we found that

$$X|Y \quad \sim \quad Uniform(-\sqrt{1-Y^2}, \sqrt{1-Y^2}).$$

Now let's find $\textrm{Cov}(X, Y) = EXY - EXEY$. We have

$$\begin{aligned} EX &= E[E[X|Y]] && \big(\text{law of iterated expectations (Equation 5.17)}\big) \\ &= E[0] = 0 && \big(\text{since } X|Y \quad \sim \quad Uniform(-\sqrt{1-Y^2}, \sqrt{1-Y^2})\big). \end{aligned}$$

Also, we have

$$\begin{aligned} E[XY] &= E[E[XY|Y]] && \big(\text{law of iterated expectations (Equation 5.17)}\big) \\ &= E[YE[X|Y]] && \big(\text{Equation 5.6}\big) \\ &= E[Y \cdot 0] = 0. \end{aligned}$$

Thus,

$$\textrm{Cov}(X, Y) = E[XY] - EXEY = 0.$$

Thus, X and Y are uncorrelated.
