4.5 Covariance and Correlation

In earlier sections, we have discussed the absence or presence of a relationship between two random variables, Independence or nonindependence. But if there is a relationship, the relationship may be strong or weak. In this section, we discuss two numerical measures of the strength of a relationship between two random variables, the covariance and correlation.

Throughout this section, we will use the notation $EX = \mu_X$, $EY = \mu_Y$, $\mathrm{Var}X = \sigma_X^2$, and $\mathrm{Var}Y = \sigma_Y^2$.

Definition 4.5.1 The covariance of X and Y is the number defined by

$$\mathrm{Cov}(X, Y) = E((X - \mu_X)(Y - \mu_Y)).$$

Definition 4.5.2 The correlation of X and Y is the number defined by

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

The value ρ_{XY} is also called the correlation coefficient.

Theorem 4.5.3 For any random variables X and Y,

$$\mathrm{Cov}(X, Y) = EXY - \mu_X \mu_Y.$$

Theorem 4.5.5 If X and Y are independent random variables, then $\mathrm{Cov}(X, Y) = 0$ and $\rho_{XY} = 0$.

Theorem 4.5.6 If X and Y are any two random variables and a and b are any two constants, then

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

If X and Y are independent random variables, then

$$\mathrm{Var}(AX + bY) = a^2 \mathrm{Var}X + b^2 \mathrm{Var}Y.$$

Theorem 4.5.7 For any random variables X and Y,

a. $-1 \le \rho_{XY} \le 1$.

b. $|\rho_{XY}| = 1$ if and only if there exist numbers $a \neq 0$ and b such that $P(Y = aX + b) = 1$. If $\rho_{XY} = 1$, then $a > 0$, and if $\rho_{XY} = -1$, then $a < 0$.

PROOF: Consider the function $h(t)$ defined by

$$\begin{aligned} h(t) &= E((X-\mu_X)t + (Y-\mu_Y))^2 \\ &= t^2\sigma_X^2 + 2t\text{Cov}(X,Y) + \sigma_Y^2. \end{aligned}$$

Since $h(t) \ge 0$ and it is quadratic function,

$$(2\text{Cov}(X,Y))^2 - 4\sigma_X^2\sigma_Y^2 \le 0.$$

This is equivalent to

$$-\sigma_X\sigma_Y \le \text{Cov}(X,Y) \le \sigma_X\sigma_Y.$$

That is,

$$-1 \le \rho_{XY} \le 1.$$

Also, $|\rho_{XY}| = 1$ if and only if the discriminant is equal to 0, that is, if and only if $h(t)$ has a single root. But since $((X-\mu_X)t + (Y-\mu_Y))^2 \ge 0$, $h(t) = 0$ if and only if

$$P((X-\mu_X)t + (Y-\mu_Y) = 0) = 1.$$

This $P(Y = aX + b) = 1$ with $a = -t$ and $b = \mu_X t + \mu_Y$, where t is the root of $h(t)$. Using the quadratic formula, we see that this root is $t = -\text{Cov}(X,Y)/\sigma_X^2$. Thus $a = -t$ has the same sign as ρ_{XY}, proving the final assertion. □

Example 4.5.8 (Correlation-I) Let X have a uniform(0,1) distribution and Z have a uniform(0,0.1) distribution. Suppose X and Z are independent. Let $Y = X + Z$ and consider the random vector (X,Y). The joint pdf of (X,Y) is

$$f(x,y) = 10, \quad 0 < x < 1, \quad x < y < x + 0.1$$

Note $f(x,y)$ can be obtained from the relationship $f(x,y) = f(y|x)f(x)$. Then

$$\begin{aligned} \text{Cov}(X,Y) &= EXY = -(EX)(EY) \\ &= EX(X+Z) - (EX)(E(X+Z)) \\ &= \sigma_X^2 = \frac{1}{12} \end{aligned}$$

The variance of Y is $\sigma_Y^2 = \text{Var}X + \text{Var}Z = \frac{1}{12} + \frac{1}{1200}$. Thus

$$\rho_{XY} = \frac{1/12}{\sqrt{1/12}\sqrt{1/12 + 1/1200}} = \sqrt{\frac{100}{101}}.$$

The next example illustrates that there may be a strong relationship between X and Y, but if the relationship is not linear, the correlation may be small.

Example 4.5.9 (Correlation-II) Let $X \sim Unif(-1, 1)$, $Z \sim Unif(0, 0.1)$, and X and Z be independent. Let $Y = X^2 + Z$ and consider the random vector (X, Y). Since given $X = x$, $Y \sim Unif(x^2, x^2 + 0.1)$. The joint pdf of X and Y is

$$f(x, y) = 5, \quad -1 < x < 1, \quad x^2 < y < x^2 + \frac{1}{10}.$$

$$\begin{aligned}
\text{Cov}(X, Y) &= E(X(X^2 + Z)) - (EX)(E(X^2 + Z)) \\
&= EX^3 + EXZ - 0E(X^2 + Z) \\
&= 0
\end{aligned}$$

Thus, $\rho_{XY} = \text{Cov}(X, Y)/(\sigma_X \sigma_Y) = 0$.

Definition 4.5.10 Let $-\infty < \mu_X < \infty$, $-\infty < \mu_Y < \infty$, $0 < \sigma_X$, $0 < \sigma_Y$, and $-1 < \rho < 1$ be five real numbers. The bivariate normal pdf with means μ_X and μ_Y, variances σ_X^2 and σ_Y^2, and correlation ρ is the bivariate pdf given by

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_Y\sqrt{1-\rho^2}} \exp\Big\{-\frac{1}{2(1-\rho^2)}\Big((\frac{x-\mu_X}{\sigma_X})^2 - 2\rho(\frac{x-\mu_X}{\sigma_X})(\frac{y-\mu_Y}{\sigma_Y}) + (\frac{y-\mu_Y}{\sigma_Y})^2\Big)\Big\}$$

for $-\infty < x < \infty$ and $-\infty < y < \infty$.

The many nice properties of this distribution include these:

a. The marginal distribution of X is $N(\mu_X, \sigma_X^2)$.

b. The marginal distribution of Y is $N(\mu_Y, \sigma_Y^2)$.

c. The correlation between X and Y is $\rho_{XY} = \rho$.

d. For any constants a and b, the distribution of $aX + bY$ is $N(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y)$.

Assuming (a) and (b) are true, we will prove (c). Let

$$s = (\frac{x-\mu_X}{\sigma_X})(\frac{y-\mu_Y}{\sigma_Y}) \quad \text{and} \quad t = (\frac{x-\mu_X}{\sigma_X}).$$

Then $x = \sigma_X t + \mu_X$, $y = (\sigma_Y s/t) + \mu_Y$, and the Jacobian of the transformation is $J = \sigma_X \sigma_Y / t$. With this change of variables, we obtain

$$\begin{aligned}
\rho_{XY} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s f(\sigma_X t + \mu_X, \frac{\sigma_Y s}{t} + \mu_Y) |\frac{\sigma_X \sigma_Y}{t}| ds dt \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})^{-1} \exp\big(-\frac{1}{2(1-\rho)^2}(t^2 - 2\rho s + (\frac{s}{t})^2)\big)\frac{\sigma_X\sigma_Y}{|t|} ds dt \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2}) dt \int_{-\infty}^{\infty} \frac{s}{\sqrt{2\pi}\sqrt{(1-\rho^2)t^2}} \exp\big(-\frac{(s-\rho t^2)^2}{2(1-\rho^2)t^2}\big) ds
\end{aligned}$$

The inner integral is ES, where S is a normal random variable with $ES = \rho t^2$ and $\text{Var}S = (1-\rho^2)t^2$. Thus,

$$\rho_{XY} = \int_{-\infty}^{\infty} \frac{\rho t^2}{\sqrt{2\pi}} \exp\{-t^2/2\} dt = \rho.$$