1. Some examples

Example 16.1 (Example 14.2, continued)**.** We find that

$$\mathrm{E}(XY) = \left(1 \times 1 \times \frac{2}{36} \right) = \frac{2}{36}.$$

Also,

$$\mathrm{E}X = \mathrm{E}Y = \left(1 \times \frac{10}{36} \right) + \left(2 \times \frac{1}{36} \right) = \frac{12}{36}.$$

Therefore,

$$\mathrm{Cov}(X, Y) = \frac{2}{36} - \left(\frac{12}{36} \times \frac{12}{36} \right) = -\frac{72}{1296} = -\frac{1}{18}.$$

The *correlation* between X and Y is the quantity,

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\ \mathrm{Var}(Y)}}. \tag{14}$$

Example 16.2 (Example 14.2, continued)**.** Note that

$$\mathrm{E}(X^2) = \mathrm{E}(Y^2) = \left(1^2 \times \frac{10}{36} \right) + \left(2^2 \times \frac{1}{36} \right) = \frac{14}{36}.$$

Therefore,

$$\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{14}{36} - \left(\frac{12}{36} \right)^2 = \frac{360}{1296} = \frac{5}{13}.$$

Therefore, the correlation between X and Y is

$$\rho(X, Y) = -\frac{1/18}{\sqrt{\left(\frac{5}{13} \right) \left(\frac{5}{13} \right)}} = -\frac{13}{90}.$$

2. Correlation and independence

The following is a variant of the Cauchy–Schwarz inequality. I will not prove it, but it would be nice to know the following.

Theorem 16.3. *If* $E(X^2)$ *and* $E(Y^2)$ *are finite, then* $-1 \le \rho(X,Y) \le 1$.

We say that X and Y are *uncorrelated* if $\rho(X,Y) = 0$; equivalently, if $\mathrm{Cov}(X,Y) = 0$. A significant property of uncorrelated random variables is that $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$; see Theorem 15.4(2).

Theorem 16.4. *If* X *and* Y *are independent [with joint mass function f], then they are uncorrelated.*

Proof. It suffices to prove that $E(XY) = E(X)E(Y)$. But

$$E(XY) = \sum_x \sum_y xy f(x,y) = \sum_x \sum_y xy f_X(x) f_Y(y)$$
$$= \sum_x x f_X(x) \sum_y y f_Y(y) = E(X)E(Y),$$

as planned.   □

Example 16.5 (A counter example). Sadly, it is only too common that people some times think that the converse to Theorem 16.4 is also true. So let us dispel this with a counterexample: Let Y and Z be two independent random variables such that $Z = \pm 1$ with probability 1/2 each; and $Y = 1$ or 2 with probability 1/2 each. Define $X = YZ$. Then, I claim that X and Y are uncorrelated but not independent.

First, note that $X = \pm 1$ and ± 2, with probability 1/4 each. Therefore, $E(X) = 0$. Also, $XY = Y^2 Z = \pm 1$ and ± 4 with probability 1/4 each. Therefore, again, $E(XY) = 0$. It follows that

$$\mathrm{Cov}(X,Y) = \underbrace{E(XY)}_{0} - \underbrace{E(X)}_{0} E(Y) = 0.$$

Thus, X and Y are uncorrelated. But they are not independent. Intuitively speaking, this is clear because $|X| = Y$. Here is one way to logically justify our claim:

$$P\{X = 1, Y = 2\} = 0 \ne \frac{1}{8} = P\{X = 1\}P\{Y = 2\}.$$

Example 16.6 (Binomials). Let $X = \mathrm{Bin}(n,p)$ denote the total number of successes in n independent success/failure trials, where $P\{\text{success per trial}\} =$

p. Define I_j to be one if the jth trial leads to a success; else $I_j = 0$. The key observation is that

$$X = I_1 + \cdots + I_n.$$

Note that $\mathrm{E}(I_j) = 1 \times p = p$ and $\mathrm{E}(I_j^2) = \mathrm{E}(I_j) = p$, whence $\mathrm{Var}(I_j) = p - p^2 = pq$. Therefore,

$$\mathrm{E}(X) = \sum_{j=1}^{n} \mathrm{E}(I_j) = np \quad \text{and} \quad \mathrm{Var}(X) = \sum_{j=1}^{n} \mathrm{Var}(I_j) = npq.$$

3. The law of large numbers

Theorem 16.7. *Suppose $X_1, X_2, \ldots, X_n$ are independent, all with the same mean μ and variance $\sigma^2 < \infty$. Then for all $\epsilon > 0$, however small,*

$$\lim_{n \to \infty} \mathrm{P}\left\{ \left| \frac{X_1 + \cdots + X_n}{n} - \mu \right| \geq \epsilon \right\} = 0. \tag{15}$$

Lemma 16.8. *Suppose $X_1, X_2, \ldots, X_n$ are independent, all with the same mean μ and variance $\sigma^2 < \infty$. Then:*

$$\mathrm{E}\left(\frac{X_1 + \cdots + X_n}{n} \right) = \mu$$

$$\mathrm{Var}\left(\frac{X_1 + \cdots + X_n}{n} \right) = \frac{\sigma^2}{n}.$$

Proof of Theorem 16.7. Recall *Chebyshev's inequality*: For all random variables Z with $\mathrm{E}(Z^2) < \infty$, and all $\epsilon > 0$,

$$\mathrm{P}\left\{ |Z - \mathrm{E}Z| \geq \epsilon \right\} \leq \frac{\mathrm{Var}(Z)}{\epsilon^2}.$$

We apply this with $Z = (X_1 + \cdots + X_n)/n$, and then use use Lemma 16.8 to find that for all $\epsilon > 0$,

$$\mathrm{P}\left\{ \left| \frac{X_1 + \cdots + X_n}{n} - \mu \right| \geq \epsilon \right\} \leq \frac{\sigma^2}{n\epsilon^2}.$$

Let $n \nearrow \infty$ to finish. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$