Equivalence of Statistical Independence and No-Correlation for a Pair of Random Variables Taking Two Values

Toru Ohira
Graduate School of Mathematics, Nagoya University, Japan

March 5, 2018

Abstract

It is well known that when a pair of random variables is statistically independent, it has no-correlation (zero covariance), and that the converse is not true (e.g. [1]). However, if both of these random variables take only two values, no-correlation entails statistical independence. We provide here a general proof.

It is well known and can be simply proven that when two random variables are statistically independent, they are not correlated. The converse is not true in general (e.g. [1]). We can have a pair of random variables which is not correlated but not statistically independent, and such examples can be easily constructed as well.

In this note, however, we show that when both of these random variables take only two values, statistical independence and no-correlation become equivalent. In other words, the proposed theorem means that one cannot have an uncorrelated pair of random variables with two distinct values which is not statistically independent.

Consider two random variables X and Y, such that they both take only two distinct finite values (x_1, x_2) and (y_1, y_2). Denote the joint probability distribution for these variables as $P(X : Y)$, and assume it is given by

$$p(x_i : y_j) \equiv P(X = x_i : Y = y_j) = p_{ij}, \quad (i, j \in \{1, 2\}) \tag{1}$$

Then, the probability distributions $P(X)$ for X and $P(Y)$ for Y are simply expressed as follows.

$$p(x_i) \equiv P(X = x_i) = p_{i1} + p_{i2}, \quad p(y_j) \equiv P(Y = y_j) = p_{1j} + p_{2j}. \tag{2}$$

By the requirement that both X, Y take only two values,

$$p(x_1) + p(x_2) = p(y_1) + p(y_2) = 1. \tag{3}$$

Y \ X	x_1	x_2	
y_1	p_{11}	p_{21}	$p(y_1)$
y_2	p_{12}	p_{22}	$p(y_2)$
	$p(x_1)$	$p(x_2)$	1

These relations can be summarized in the following table.

The statistical independence of X and Y is defined as

$$P(X : Y) = P(X)P(Y). \tag{4}$$

Also, with the definition of expectation values as

$$E[X] = \sum_i p(x_i)x_i, \quad E[Y] = \sum_i p(y_i)y_i, \quad E[XY] = \sum_{i,j} p(x_i : y_j)x_i y_j, \tag{5}$$

we define that X and Y are not correlated when their covariance is zero.

$$Cov[X, Y] \equiv E[XY] - E[X]E[Y] = 0, \tag{6}$$

or equivalently,

$$E[XY] = E[X]E[Y]. \tag{7}$$

We can see from the above definition that given (4), (7) follows. Our main statement here is that the converse is true, i.e., (4) and (7) are equivalent when both of these random variable take two finite distinct values. (In passing, we note that if either (or both) X or Y takes more than two values, one can easily create examples showing this equivalence does not hold.)

Theorem

When both random variables X and Y take two distinct finite values as set up above, and

$$E[XY] = E[X]E[Y], \tag{8}$$

then

$$P(X : Y) = P(X)P(Y). \tag{9}$$

proof

Let us define all the relevant probabilities with three parameters using relations (1), (2) and (3). We set

$$\alpha = p_{11}, \quad u = p(x_1), \quad v = p(y_1). \tag{10}$$

Y \ X	x_1	x_2	
y_1	α	v-α	v
y_2	u-α	1+α -u-v	1- v
	u	1- u	1

Then, other relevant probabilities can be expressed as summarized in the following table.

By definition of the expectation values, we have the following

$$\begin{aligned} E[X] &= ux_1 + (1-u)x_2, \\ E[Y] &= vy_1 + (1-v)y_2, \\ E[XY] &= \alpha x_1 y_1 + (u-\alpha)x_1 y_2 + (v-\alpha)x_2 y_1 + (1-v-u+\alpha)x_2 y_2. \end{aligned}$$

The condition (8) is now used together with above so that we obtain

$$\begin{aligned} 0 &= E[XY] - E[X]E[Y] \\ &= \{\alpha x_1 y_1 + (u-\alpha)x_1 y_2 + (v-\alpha)x_2 y_1 + (1-v-u+\alpha)x_2 y_2\} \\ &\quad -\{ux_1 + (1-u)x_2\}\{vy_1 + (1-v)y_2\} \\ &= (\alpha - uv)(x_1 - x_2)(y_1 - y_2). \end{aligned}$$

By the assumption that both of these stochastic variables take two distinct values ($x_1 \neq x_2, y_1 \neq y_2$), this leads to

$$\alpha - uv = p_{11} - p(x_1)p(y_1) = 0, \tag{11}$$

from which one can deduce that

$$P(X:Y) = P(X)P(Y). \tag{12}$$

References

[1] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol.1, John Wiley & Sons, New York, 1957.