§ 1.5 Independence

Henceforth, unless otherwise stated, whenever a probability variable is considered, it is on a certain probability space (Ω, P). Two random variables are said to be **independent** when

$$P\{X = x, Y = y\} = P\{X = x\}P\{Y = y\}, \qquad x \in \Omega^X,\ y \in \Omega^Y$$

holds. In this instance,

$$\begin{aligned} P_{X=x}\{Y = y\} &= P\{Y = y\} \qquad (\text{when } P(X = x) > 0), \qquad y \in \Omega^Y, \\ P_{Y=y}\{X = x\} &= P\{X = x\} \qquad (\text{when } P(Y = y) > 0), \qquad x \in \Omega^X. \end{aligned}$$

Or either from these two conditions, the independence between X and Y emerges. These can also be described as, respectively,

$$\begin{aligned} &P_{X=x}\{Y = y\} \quad \text{is unrelated to } x \qquad (\text{when } P(X = x) > 0), \\ &P_{Y=y}\{X = x\} \quad \text{is unrelated to } y \qquad (\text{when } P(Y = y) > 0). \end{aligned}$$

The terminology independence is rooted in these facts.

Theorem 1.16 The following conditions are mutually equivalent:

(i) X and Y are independent,

(ii) Suppose N is the totality of all the values of x such that $P_{X=x}(Y = y) \neq P(Y = y)$. Then

$$P^X(N) = 0,$$

(iii) $P_X(Y = y) = P(Y = y)$ a.s.

Proof If X and Y are independent, as also noted above, for all $x \in N$,

$$P\{X = x\} = 0.$$

Hence

$$P^X(N) = P\{X \in N\} = \sum_{x \in N} P\{X = x\} = 0.$$

Thus (i) $\Longrightarrow$ (ii). Now let

$$M = \{\omega | P_{X(\omega)}(Y = y) \neq P\{Y = y\}\}.$$

If $\omega \in M$, then $X(\omega) \in N$. Hence $M \subset X^{-1}(N)$ (in fact = holds). Accordingly from (ii)

$$P(M) \leqq P\big(X^{-1}(N)\big) = P^X(N) = 0\,, \quad \text{i.e.} \quad P(M) = 0\,,$$

which means (iii). Thus (ii) $\Longrightarrow$ (iii). If (iii) is assumed, using the above M,

$$\begin{aligned} E\big(P_{X(\omega)}(Y = y), X^{-1}\{x\}\big) &= E\big(P_{X(\omega)}(Y = y), X^{-1}\{x\}\backslash M\big)\,, \\ E\big(P(Y = y), X^{-1}\{x\}\big) &= E\big(P(Y = y), X^{-1}\{x\}\backslash M\big)\,. \end{aligned}$$

Since on $X^{-1}\{x\}\backslash M)$, $P_{X(\omega)}(Y = y)$ and $P(Y = y)$ are equal, and

$$E\big(P_{X(\omega)}(Y = y), X^{-1}\{x\}\big) = E\big(P(Y = y), X^{-1}\{x\}\big)\,.$$

That is, from Theorem 1.13 and Theorem 1.3's (iii),

$$P(X = x, Y = y) = P(Y = y)P\big(X^{-1}\{x\}\big) = P(X = x)\,P(Y = y)\,,$$

and the independence between X and Y appears. Thus (iii) $\Longrightarrow$ (i). Finally the equivalence of (i), (ii), and (iii) is proven. ▮

Theorem 1.17 If X and Y are independent,

$$P(X \in E, Y \in F) = P(X \in E)\,P(Y \in F)\,.$$

Proof Since $\{\omega | X(\omega) \in E, Y(\omega) \in F\} = \sum_{x\in E}\sum_{y\in F}\{\omega | X(\omega) = y, Y(\omega) = y\}$,

$$P(X \in E, Y \in F) = \sum_{x\in E}\sum_{y\in F} P\{X = x\}\,P\{Y = y\} = P(E)\,P(F)\,. \text{ ▮}$$

For $i = 1, 2, \cdots, n$, $X_1, X_2, \cdots, X_n$ are independent is about the validity of

$$P\{X_1 = x_1, X_2 = x_2, \cdots X_n = x_n\} = P\{X_1 = x_1\}P\{X_2 = x_2\}\cdots P\{X_n = x_n\}\,, x_i \in \Omega_i^{X_i}\,.$$

In this instance, even if the order of $X_1, X_2, \cdots, X_n$ is exchanged, it is still independent. From the above equation, with respect to $x_n \in \Omega^{X_n}$, one obtains

$$P\{X_1 = x_1, X_2 = x_2, \cdots X_{n-1} = x_{n-1}\} = P\{X_1 = x_1\}P\{X_2 = x_2\}\cdots P\{X_{n-1} = x_{n-1}\}\,.$$

Hence, for a system of independent probability variables, its subsystem is also independent.

Theorem 1.18 Separate $\{X_1, X_2, \cdots, X_n\}$ into two subsystems

$$\{X_1, X_2, \cdots, X_r\} \qquad \text{and} \qquad \{X_{r+1}, X_{r+2}, \cdots, X_n\}\,.$$

For $X_1, X_2, \cdots, X_n$ to be independent, it is necessary and sufficient for the following two conditions to hold.

(i) Every subsystem is independent, i.e., $X_1, X_2, \cdots, X_r$ is independent, and $X_{r+1}, X_{r+2}, \cdots, X_n$ is independent,

(ii) Combined variables $(X_1, X_2, \cdots, X_r), (X_{r+1}, X_{r+2}, \cdots, X_n$ are independent.

proof It is easy to prove if the following points are noted:

$$\begin{aligned} &X_1 = x_1, \cdots, X_r = x_r, X_{r+1} = x_{r+1}, \cdots, X_n = x_n \\ &\quad \Longleftrightarrow (X_1, \cdots, X_r) = (x_1, \cdots, x_r)\,,\, (X_{r+1}, \cdots, X_n) = (x_{r+1}, \cdots, x_n)\,. \quad \blacksquare \end{aligned}$$

Theorem 1.19 If $X_1, X_2, \cdots, X_n$ are independent, then

$$P\{X_1 \in E_1, X_2 \in E_2, \cdots, X_n \in E_n\} = P\{X_1 \in E_1\}\, P\{X_2 \in E_2\} \cdots P\{X_n \in E_n\}\,.$$

Proof Theorem 1.17 is a special case of $n = 2$ of this theorem. From the previous theorem, $(X_1, X_2, \cdots, X_{n-1})$ is independent of X_n, and from Theorem 1.17,

$$\begin{aligned} \text{l.h.s} &= P\big\{(X_1, X_2, \cdots, X_{n-1}) \in E_1 \times E_2 \times \cdots \times E_{n-1}\,, X_n \in E_n\big\} \\ &= P\big\{(X_1, X_2, \cdots, X_{n-1}) \in E_1 \times E_2 \times \cdots \times E_{n-1}\big\}\, P\{X_n \in E_n\} \\ &= P\big\{X_1 \in E_1, X_2 \in E_2, \cdots, X_{n-1} \in E_{n-1}\big\}\, P\{X_n \in E_n\}\,. \end{aligned}$$

Repeat this argument, it is understood that in the end it equals the right hand side of the theorem's equation. $\blacksquare$

In the above theorem,

$$P^{(X_1, X_2, \cdots, X_n)}(E_1 \times E_2 \times \cdots \times E_n) = P^{X_1}(E_1)\, P^{X_2}(E_2) \cdots P^{X_n}(E_n),,$$

that is, it means

$$P^{(X_1, X_2, \cdots, X_n)} = P^{X_1} \times P^{X_2} \times \cdots \times P^{X_n} \qquad \text{(direct product measure)}\,.$$

Theorem 1.20 If $X_1, X_2, \cdots, X_n$ are independent, and $Y_1 = \varphi_1(X_1)$, $Y_2 = \varphi_2(X_2)$, $\cdots$, $Y_n = \varphi_n(X_n)$, then $Y_1, Y_2, \cdots, Y_n$ are also independent.

Proof Note the following and apply the previous theorem:

$$Y_1 = y_1, Y_2 = y_2, \cdots, Y_n = y_n \Longleftrightarrow X_1 \in \varphi_1^{-1}(y_1), X_1 \in \varphi_2^{-1}(y_2), \cdots, X_n \in \varphi_n^{-1}(y_n)\,. \quad ∎$$

The sets $A_1, A_2, \cdots, A_n$ are **independent** is about the independence of their indicator functions $1_{A_1}(\omega), 1_{A_2}(\omega), \cdot$ Since sets represent events, it is alright to think of them as the independence of events.

Theorem 1.21 The following three conditions are equivalent.

(i) $A_1, A_2, \cdots, A_n$ are independent,

(ii) When $A_i' = A_i$ or A_i^c, then

$$P\big(A_1' \cap A_2' \cap \cdots \cap A_n'\big) = P\big(A_1'\big)\,P\big(A_2'\big)\,\cdots\,P\big(A_n'\big)$$

always holds,

(iii) For any $k = 2, 3, \cdots, n$, and any $1 \leqq i_i < i_2 < \cdots < i_k \leqq n$,

$$P\big(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\big) = P\big(A_{i_1}\big)\,P\big(A_{i_2}\big)\,\cdots\,P\big(A_{i_k}\big)\,.$$

Proof Suppose $e_i(\omega)$ is the indicator function of A_i. (i) means that with respect to $a_i = 0, 1 (i = 1, 2, \cdots, n)$

$$P\big(e_1^{-1}(a_1) \cap e_2^{-1}(a_2) \cap \cdots \cap e_n^{-1}(a_n)\big) = P\big(e_1^{-1}(a_1)\big)\,P\big(e_2^{-1}(a_2)\big) \cdots P\big(e_n^{-1}(a_n)\big)$$

holds. Because in Accordance to $a_i = 1, 0$, $e_i^{-1}(a_i) = A_i, A_i^c$, the above equation is only a rewrite of (ii). Hence, (i) and (ii) are equivalent. If (i) holds, any subsystem $A_{i_1}, A_{i_2}, \cdots A_{i_k}$ of $A_1, A_2, \cdots A_n$ is also independent. It follows that from the already proven (i) $\Longleftrightarrow$ (ii), the equation of (iii) is obtained. Thus (i) $\Longrightarrow$ (iii) is proven. What remains to prove is (iii) $\Longrightarrow$ (ii). From the equation of (iii)

$$E(e_{i_1} e_{i_2} \cdots e_{i_k}) = E(e_{i_1}) E(e_{i_2}) \cdots E(e_{i_k})\,.$$

Now the expansion of $(a_1 + b_1\, t_1)(a_2 + b_2\, t_2) \cdots (a_n + b_n\, t_n)$ is denoted by

$$\sum_{k,\,\{i_\nu\}} c_{i_1 i_2 \cdots i_k}\, t_{i_1} t_{i_2} \cdots t_{i_k}\,.$$

Since

$$E\big(a_1 + b_1\, e_1)(a_2 + b_2\, e_2) \cdots (a_n + b_n\, e_n)\big) = \sum_{k,\,\{i_\nu\}} c_{i_1 i_2 \cdots i_k}\, E\big(e_{i_1} e_{i_2} \cdots e_{i_k}\big)\,,$$

$$E(a_1 + b_1\, e_1)\, E(a_2 + b_2\, e_2) \cdots E(a_n + b_n\, e_n) = \sum_{k,\,\{i_\nu\}} c_{i_1 i_2 \cdots i_k}\, E(e_{i_1}) E(e_{i_2}) \cdots E(e_{i_k})\,,$$

From the equation discussed above, it equals to the right hand side. Hence,

$$E\big(a_1 + b_1\, e_1)(a_2 + b_2\, e_2) \cdots (a_n + b_n\, e_n)\big) = E(a_1 + b_1\, e_1)\, E(a_2 + b_2\, e_2) \cdots E(a_n + b_n\, e_n)\,.$$

As $A_i' = A_i$ or A_i^c, let its indicator function be e_i'. Then

$$e_i' = e_i\,, \qquad \text{or} \qquad 1 - e_i\,.$$

Accordingly, e_i' takes the form of $a_i + b_i\, e_i$. Hence

$$E(e_1' e_2' \cdots e_n') = E(e_1') E(e_2') \cdots E(e_n')\,,$$

that is

$$P\big(A_1' \cap A_2' \cap \cdots \cap A_n'\big) = P\big(A_1'\big)\, P\big(A_2'\big) \cdots P\big(A_n'\big)\,.$$

As a result, (iii) $\Longrightarrow$ (ii) is proven. ▮

Suppose $\boldsymbol{T}$ is the direct union of $\boldsymbol{T}_1$, $\boldsymbol{T}_2$, $\cdots$ $\boldsymbol{T}_n$, and their probability spaces are, respectively,

$$(\Omega_1, P_1)\,, \quad (\Omega_2, P_2)\,, \quad \cdots\,, \quad (\Omega_n, P_n)\,, \quad (\Omega, P)\,.$$

Then

$$\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n\,, \qquad P = P_1 \times P_2 \times \cdots \times P_n\,.$$

When the result of $\boldsymbol{T}_i$ is ω_i $(i = 1, 2, \cdots, n)$, the result ω of $\boldsymbol{T}$ is given by $(\omega_1, \omega_2, \cdots \omega_n)$, and considering ω_i on (Ω, P),

$$\omega_i = \pi_i(\omega) \qquad (\pi_i \text{ is } i\text{-th image})$$

Accordingly, the probability variable $X_i(\omega_i)$ is on (Ω_i, P_i)

$$X_i(\omega_i) = X_i\big(\pi_i(\omega)\big) = (X_i \circ \pi)(\omega)$$

is on (Ω, P).

Theorem 1.22 On (Ω, P), $\omega_1, \omega_2, \cdots, \omega_n$ are independent. Hence $X_1(\omega_1), X_2(\omega_2), \cdots X_n(\omega_n)$ are also independent.

Proof Since $\omega_i = \pi_i(\omega)$,

$$P\{\pi_1(\omega) = a_1, \pi_2(\omega) = a_2, \cdots, \pi_n(\omega) = a_n\} = P\{(a_1, a_2, \cdots, a_n)\} = P_1\{a_1\}P_2\{a_2\}\cdots P_n\{a_n\},$$

$$P\{\pi_1(\omega) = a_1\} = P\big(\{a_1\} \times \Omega_2 \times \Omega_3 \times \cdots \times \Omega_n\big) = P_1\{a_1\}P_2(\Omega_2)P_3(\Omega_3)\cdots P_n(\Omega_n) = P_1\{a_1\}.$$

In the same way,

$$P\{\pi_i(\omega) = a_i\} = P_i\{a_i\},$$

hence

$$P\{\pi_1(\omega) = a_1, \pi_2(\omega) = a_2, \cdots, \pi_n(\omega) = a_n\} = P\{\pi_1(\omega) = a_1\}\, P\{\pi_2(\omega) = a_2\} \cdots P\{\pi_n(\omega) = a_n\}.$$

It means the independence of $\omega_1 = \pi_1(\omega)$, $\pi_2(\omega), \cdots, \pi_n(\omega)$. Accordingly, from Theorem 1.20, $X_1(\omega_1)$, $X_2(\omega_2), \cdots, X_n(\Omega_n)$ are also independent. ▌

Exercise 1.5 (i) Show that the number 2 and number 3 of a rolled dice are independent. (Apply Theorem 1.21 (i) $\Longleftrightarrow$ (iii))

(ii) A dice is rolled $m + n$ times. Using Theorems 1.22, 1.18, and 1.20, prove that the sum of the first m points and the product of the subsequent n points.

(iii) Prove that if X_1 and X_2 are independent. (X_1, X_2) and X_3 are independent, (X_1, X_2, X_3) and X_4 are independent, then X_1, X_2, X_3, X_4 are independent. Show also that the converse is true.

(iv) If A is independent of itself, show that $P(A) = 0$ or 1. (Note that $A \cap A = A$).

(v) If X is independent of itself, show that almost surely, it is equal to a constant.

[Hint] For any $a \in \Omega^X$, note that $P\{X = a\}$ is equal to 1 or 0.