§ 1.4 Conditional Probability

Let (Ω, P) be the probability space, A, B any subsets of Ω. The **(conditional) probability** of B given A is defined as

$$P(A \cap B) / P(A).$$

It is denoted by $P(B|A)$ or $P_A(B)$. When $P(A) = 0$, though $P_A(B)$ is meaningless, it is convenient to write

$$P_A(B) = \begin{cases} 1, & \omega_0 \in B, \\ 0, & \omega_0 \notin B. \end{cases}$$

Here, ω_0 is any arbitrary fixed point in Ω.

Theorem 1.12 (i) $P(A \cap B) = P(A)\, P_A(B)$

(ii) When $P_A(B)$ is viewed as a function of B with A fixed, it is a probability measure on Ω. (In this sense, the set function P_A is the **(conditional) probability law** given A.

Proof (i) When $P(A) > 0$, it is straightforward from the definition. Even when $P(A) = 0$, it is understandable from the above constraint that the theorem holds. ▮

To repeat the above theorem's (i),

$$P\big(A_1 \cap A_2 \cap A_3\big) = P(A_1)\, P_{A_1}(A_2)\, P_{A_1 \cap A_2}(A_3),$$

and more generally,

$$P\big(A_1 \cap A_2 \cap \cdots \cap A_n\big) = P(A_1)\, P_{A_1}(A_2) \cdots P_{A_1 \cap A_2 \cap \cdots \cap A_{n-1}}(A_n)$$

is obtained.

Let $P_A(B)$, $A = X^{-1}\{x\}$ $(x \in \Omega^X)$, and define $P_{X^{-1}\{x\}}(B)$. Since $X^{-1}\{x\} = \{\omega | X(\omega) = x\}$, $P_{X^{-1}\{x\}}(B)$ is written as

$$P_{X=x}(B) \qquad \text{or} \qquad P(B|X = x).$$

This is called "the conditional probability of B given the condition of $X = X$." $P_{X=x}(B)$ is a function of $x \in \Omega^X$ and $B \subset \Omega$. Next, when x is fixed, $P_{X=x}(B)$ is a probability measure on Ω as a function of B. When x is rewritten as $X(\omega)$, that is, when

$$P_{X=x}(B)|_{x=X(\omega)}$$

is expressed as $P_X(B)$ or $P(B|X)$, it is called the "conditional probability of B given that the value of $X(\omega)$ is known." When B is fixed, $P_X(B)$ is a function of $X(\omega)$ hence a function of ω, and can be considered as the probability variable on (Ω, P).

Theorem 1.13

$$P\big(X^{-1}(F)\cap B\big)=\sum_{x\in F}P\{X=x\}\,P_{X=x}(B)=E^X\big(P_{X=x}(B),F\big)=E\big(P_X(B),X^{-1}(F)\big)\,.$$

Proof From $X^{-1}(F)=\sum_{x\in F}X^{-1}\{x\}$, the first equality is evident. With $P\{X=x\}=P^X\{x\}$, the second equality emerges. To show the third equality, denote $P_{X=x}(B)$ by $\varphi_B(x)$,

$$E^X\big(P_{X=x}(B),F\big)=E^X\big(\varphi_B,F\big)=E^X\big(\varphi_B\,1_F\big)\,,$$
$$E\big(P_X(B),X^{-1}(F)\big)=E\big(\varphi_B(X(\omega)),X^{-1}(F)\big)=E\big(\varphi_B(X(\omega))\,1_F(X(\omega))\big)\,,$$

is obtained. From Theorem 1.9, it is easy to understand that these two equations are equal. ▮

Since P_A is a probability law on Ω, with respect to a function $Y(\omega)$ on Ω, the average value E_AY of $Y(\omega)$ with respect to P_A can be defined by

$$E_A(Y)=\sum_{\omega\in\Omega}Y(\omega)P_A\{\omega\}\,.$$

Directly defined by

$$E_A(Y)=\begin{cases}E(Y,A)/P(A)\,, & P(A)>0\,,\\ Y(\omega_0)\,, & P(A)=0\end{cases}$$

yields the same thing also. Here ω_0 is that which is fixed when $P(A)=0$ in the definition of $P_A(B)$. Accordingly,

$$E(Y,A)=P(A)\,E_A(Y)\,.$$

Same as the definition of $P_{X=x}(B),P_X(B)$ by $P_A(B)$, it is possible to define $E_{X=x}(Y),E_X(Y)$ with $E_A(Y)$. Corresponding to Theorem 1.13, the following theorem is obtained:

Theorem 1.14

$$E\big(Y,X^{-1}(F)\big)=\sum_{x\in F}P\{X=x\}\,E_{X=x}(Y)=E^X\big(E_{X=x}(Y),F\big)=E\big(E_X(Y),X^{-1}(F)\big)\,.$$

Proof Same as Theorem 1.13. ▮

At times, $E_A(Y)$, $E_{X=x}(Y)$, and $E_X(Y)$ are also written as, respectively,

$$E(Y|A)\,,\qquad E(Y|X=x)\,,\qquad E(Y|X)\,.$$

Let $B=\{\omega|Y\in F\}=Y^{-1}(F)$ in $P_A(B)$ and define $P_A\{Y\in F\}$. As a function of F, it is a probability measure. Since this probability measure is the probability law of $Y(\omega)$ when $Y(\omega)$ is seen as the probability variable on (Ω, P_A), it can be written as $(P_A)^Y$. By the same token, one can consider $P_{X=x}\{Y \in F\}$, $P_X\{y \in F\}$, $(P_{X=x})^Y$, and $(P_X)^Y$.

The conditional probability has a close relationship with the tree union discussed in the previous section. Consider the tree union

$$\begin{tikzcd}[row sep=small] & \boldsymbol{T}_{a_1} \\ \boldsymbol{T}_0 \arrow[ur] \arrow[r] \arrow[ddr] & \boldsymbol{T}_{a_2} \\ & \vdots \\ & \boldsymbol{T}_{a_n} \end{tikzcd}$$

Let the probability space of $\boldsymbol{T}_0$ be

$$\big(\Omega_0 \equiv \{a_1, a_2, \cdots, a_n\}, P_0\big)\,,$$

and that of $\boldsymbol{T}_a$ be (Ω_a, P_a). The probability space (Ω, P) of this tree union is given by

$$\Omega = \{(a,b) | a \in \Omega_0, b \in \Omega_a\}\,,$$
$$P\big\{(a,b)\big\} = P_0\{a\}\, P_a\{b\}\,.$$

The result of the first trial $\boldsymbol{T}_0$ of $\boldsymbol{T}$, and the result of the next trial $\big($one amongst $\boldsymbol{T}_{a_1}, \boldsymbol{T}_{a_2}, \cdots, \boldsymbol{T}_{a_n}\big)$ are the probability variables on (Ω, P). They are denoted by $X(\omega)$ and $Y(\omega)$. That is

$$X(\omega) = \pi_1(\omega)\,, \qquad Y(\omega) = \pi_2(\omega) \qquad (\pi_i \text{ is image})\,.$$

Thus, $Y(\omega)$ is a probability variable that takes value in one of $\bigcup_{i=1}^{n} \Omega_{a_i}$, and

$$P\big\{(a,b)\big\} = P\{X = a, Y = b\} = P\{X = a\} P_{X=a}\{Y = b\}\,,$$
$$P\{X = a\} = \sum_{b \in \Omega_a} P\{X = a, Y = b\} = \sum_{b \in \Omega_a} P\big\{(a,b)\big\} = \sum_{b \in \Omega_a} P_0\{a\} P_a\{b\} = P_0\{a\}\,.$$

From these emerge

$$P_{X=a}\{Y = b\} = \frac{P\big\{(a,b)\big\}}{P\{X = a\}} = \frac{P_0\{a\}\, P_a\{b\}}{P_0\{a\}} = P_a\{b\}\,.$$

Hence, the following theorem is obtained.

Theorem 1.15 $\quad P_{X=a}\{Y = b\} = P_a\{b\}\,.$

Similarly, for tree union with many periods, denote the results of the first, second, $\cdots$ by, respectively, $X, Y, Z, \cdots$, and

$$P_{X=a}\{Y=b\} = P_a\{b\}\,,$$
$$P_{X=a,\,Y=b}\{Z=c\} = P_{a,b}\{c\}$$
$$\cdots\cdots\cdots\cdots\cdots$$

In the case of direct union, $P_{X=a}\{Y=b\}$ is unrelated to a, and $P_{X=a,\,Y=b}\{Z=c\}$ is unrelated to a, b.

Exercise 1.4 When $\Omega = A_1 + A_2 + \cdots + A_n$, prove the following equation

$$P_B(A_i) = \frac{P(A_i)\,P_{A_i}(B)}{\sum_{k=1}^{n} P(A_k)\,P_{A_k}(B)} \qquad (\textbf{Bayes' Theorem})$$

Christopher Ting
Translation Version 1.0
Date: 10 May, 2017