8

Conditional Expectations

If there is partial information on the outcome of a random experiment, the probabilities for the possible events may change. The concept of conditional probabilities and conditional expectations formalises the corresponding calculus.

8.1 Elementary Conditional Probabilities

Example 8.1. We throw a die and consider the events

$$A := \{\text{the face shows three or smaller}\},$$
$$B := \{\text{the face shows an odd number}\}.$$

Clearly, $\mathbf{P}[A] = \frac{1}{2}$ and $\mathbf{P}[B] = \frac{1}{2}$. However, what is the probability that B occurs if we already know that A occurs?

We model the experiment on the probability space $(\Omega, \mathcal{A}, \mathbf{P})$, where $\Omega = \{1, \ldots, 6\}$, $\mathcal{A} = 2^\Omega$ and $\mathbf{P}$ is the uniform distribution on Ω. Then

$$A = \{1, 2, 3\} \quad \text{and} \quad B = \{1, 3, 5\}.$$

If we know that A has occurred, it is plausible to assume the uniform distribution on the remaining possible outcomes; that is, on $\{1, 2, 3\}$. Thus we define a new probability measure $\mathbf{P}_A$ on $(A, 2^A)$ by

$$\mathbf{P}_A[C] = \frac{\#C}{\#A} \quad \text{for } C \subset A.$$

By assigning the points in $\Omega \setminus A$ probability zero (since they are impossible if A has occurred), we can extend $\mathbf{P}_A$ to a measure on Ω:

$$\mathbf{P}[C\,|\,A] := \mathbf{P}_A[C \cap A] = \frac{\#(C \cap A)}{\#A} \quad \text{for } C \subset \Omega.$$

In this way, we get $\mathbf{P}[B\,|\,A] = \dfrac{\#\{1, 3\}}{\#\{1, 2, 3\}} = \dfrac{2}{3}$. ◇

Motivated by this example, we make the following definition.

Definition 8.2 (Conditional probability). *Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space and $A \in \mathcal{A}$. We define the* ***conditional probability given*** *A for any $B \in \mathcal{A}$ by*

$$\mathbf{P}[B\,|\,A] = \begin{cases} \dfrac{\mathbf{P}[A \cap B]}{\mathbf{P}[A]}, & \text{if } \mathbf{P}[A] > 0, \\ 0, & \text{else}. \end{cases} \tag{8.1}$$

Remark 8.3. The specification in (8.1) for the case $\mathbf{P}[A] = 0$ is arbitrary and is of no importance. ◇

Theorem 8.4. *If $\mathbf{P}[A] > 0$, then $\mathbf{P}[\,\cdot\,|\,A]$ is a probability measure on $(\Omega, \mathcal{A})$.*

Proof. This is obvious. □

Theorem 8.5. *Let $A, B \in \mathcal{A}$ with $\mathbf{P}[A], \mathbf{P}[B] > 0$. Then*

$$A, B \text{ are independent} \iff \mathbf{P}[B\,|\,A] = \mathbf{P}[B] \iff \mathbf{P}[A\,|\,B] = \mathbf{P}[A].$$

Proof. This is trivial! □

Theorem 8.6 (Summation formula). *Let I be a countable set and let $(B_i)_{i \in I}$ be pairwise disjoint sets with $\mathbf{P}\left[\biguplus_{i \in I} B_i\right] = 1$. Then, for any $A \in \mathcal{A}$,*

$$\mathbf{P}[A] = \sum_{i \in I} \mathbf{P}[A\,|\,B_i]\,\mathbf{P}[B_i]. \tag{8.2}$$

Proof. Due to the σ-additivity of $\mathbf{P}$, we have

$$\mathbf{P}[A] = \mathbf{P}\left[\biguplus_{i \in I}(A \cap B_i)\right] = \sum_{i \in I} \mathbf{P}[A \cap B_i] = \sum_{i \in I} \mathbf{P}[A\,|\,B_i]\mathbf{P}[B_i]. \qquad \square$$

Theorem 8.7 (Bayes' formula). *Let I be a countable set and let $(B_i)_{i \in I}$ be pairwise disjoint sets with $\mathbf{P}\left[\biguplus_{i \in I} B_i\right] = 1$. Then, for any $A \in \mathcal{A}$ with $\mathbf{P}[A] > 0$ and any $k \in I$,*

$$\mathbf{P}[B_k\,|\,A] = \frac{\mathbf{P}[A\,|\,B_k]\,\mathbf{P}[B_k]}{\sum_{i \in I} \mathbf{P}[A\,|\,B_i]\,\mathbf{P}[B_i]}. \tag{8.3}$$