



## Why

Since one and only one outcome occurs, given a distribution on outcomes, we define the probability of a set of outcomes as the sum of their probabilities.

## Definition

Suppose  $p$  is a distribution on  $\Omega$ . For any event  $A \subset \Omega$ , we call the value  $\sum_{a \in A} p(a)$  the *event probability*. We refer to the probability of  $A$ . The probability  $A$  is the sum of the probabilities of its outcomes.

We can define a function  $\mathbf{P} : \mathcal{P}(\Omega) \rightarrow \mathbf{R}$  by  $\mathbf{P}(A) = \sum_{a \in A} p(a)$ . We call  $\mathbf{P}$  the *event probability function* (or *the probability measure*) induced by  $p$ . Since  $\mathbf{P}$  depends on the sample space  $\Omega$  and the distribution  $p$ , we occasionally denote this dependence by  $\mathbf{P}_{\Omega,p}$  or  $\mathbf{P}_p$ .

## Example: die

Define  $p : \{1, \dots, 6\} \rightarrow \mathbf{R}$  by  $p(\omega) = 1/6$  for  $\omega = 1, \dots, 6$ . Define the event  $E = \{2, 4, 6\}$ . Then

$$\mathbf{P}(E) = \sum_{\omega \in E} p(\omega) = p(2) + p(4) + p(6) = 1/2.$$

## Properties of $\mathbf{P}$

As a result of the conditions on  $p$ ,  $\mathbf{P}$  satisfies

1.  $\mathbf{P}(A) \geq 0$  for all  $A \subset \Omega$ ;
2.  $\mathbf{P}(\Omega) = 1$  (and  $\mathbf{P}(\emptyset) = 0$ );
3.  $\mathbf{P}(A) + \mathbf{P}(B)$  for all  $A, B \subset \Omega$  and  $A \cap B = \emptyset$ . This statement follows from the more general identity

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$$

for  $A, B \subset \Omega$ , by using  $\mathbf{P}(\emptyset) = 0$  of (2) above.

Do all such  $\mathbf{P}$  satisfying (1)-(3) have a corresponding underlying probability distribution? In other words, suppose  $f : \mathcal{P}(\Omega) \rightarrow \mathbf{R}$  satisfies (1)-(3). These three conditions are sometimes called the *axioms of probability for finite sets*.

Define  $q : \Omega \rightarrow \mathbf{R}$  by  $q(\omega) = f(\{\omega\})$ . If  $f$  satisfies the axioms, then  $q$  is a probability distribution. For this reason we call any function satisfying (i)-(iii) an *event probability function* (or a *probability measure*).

### **Probability by cases**

Let  $\mathbf{P}$  be a probability event function. Suppose  $A_1, \dots, A_n$  partition  $\Omega$ . Then for any  $B \subset \Omega$ ,

$$\mathbf{P}(B) = \sum_{i=1}^n \mathbf{P}(A_i \cap B).$$

Some authors call this the *law of total probability*.

