



## Why

How should we modify probabilities, given that we know some aspect of the outcomes (i.e., that some event has occurred).

## Definition

Let  $\mathbf{P} : \mathcal{P}(\Omega) \rightarrow \mathbf{R}$  be a finite probability measure. Let  $A, B \subset \Omega$  and  $\mathbf{P}(B) \neq 0$ . The *conditional probability* of  $A$  *given*  $B$  is fraction of the probability of  $A \cap B$  over the probability of  $B$ .

## Notation

In a slightly slippery but universally standard notation, we denote the conditional probability of  $A$  given  $B$  by  $\mathbf{P}(A \mid B)$ . In other words, we define

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)},$$

for all  $A, B \subset \Omega$ , whenever  $\mathbf{P}(B) \neq 0$ .

For example, we can express the law of total probability (see Event Probabilities) as

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

where  $A_1, \dots, A_n$  partition  $\Omega$  and  $B \subset \Omega$  with  $\mathbf{P}(B) > 0$ .

## Conditional probability measure

Suppose  $B \subset \Omega$  and  $\mathbf{P}(B) > 0$ . Then (i)  $\mathbf{P}(A \mid B) \geq 0$  for all  $A \subset \Omega$ , since  $\mathbf{P}(A \cap B) \geq 0$ . Moreover, (ii)

$$\mathbf{P}(\Omega \mid B) = \mathbf{P}(\Omega \cap B) / \mathbf{P}(B) = \mathbf{P}(B) / \mathbf{P}(B) = 1$$

Similarly,  $\mathbf{P}(\emptyset \mid B) = \mathbf{P}(\emptyset \cap B) / \mathbf{P}(B) = 0 / \mathbf{P}(B) = 0$ .

Finally, (iii) if  $A \cap C = \emptyset$ , then

$$\begin{aligned}
\mathbf{P}(A \cap C \mid B) &= \mathbf{P}((A \cap C) \cap B) / \mathbf{P}(B) \\
&= \mathbf{P}((A \cap B) \cap (C \cap B)) / \mathbf{P}(B) \\
&\stackrel{(a)}{=} (\mathbf{P}(A \cap B) + \mathbf{P}(C \cap B)) / \mathbf{P}(B) \\
&= \mathbf{P}(A \cap B) / \mathbf{P}(B) + \mathbf{P}(C \cap B) / \mathbf{P}(B) \\
&= \mathbf{P}(A \mid B) + \mathbf{P}(C \mid B).
\end{aligned}$$

where (a) follows since  $A \cap B$  and  $C \cap B$  are disjoint.

Together, (i)-(iii) mean that  $\mathbf{P}(\cdot \mid B)$  is itself a probability measure on  $\Omega$ . We therefore refer to  $\mathbf{P}(\cdot \mid B)$  as a *conditional probability measure*.

### Induced conditional distribution

Therefore, we expect there to also correspond a new distribution on the set of outcomes. For  $\mathbf{P}_p$ , define  $q : \Omega \rightarrow \mathbf{R}$  by

$$q(\omega) = \begin{cases} \frac{p(\omega)}{\mathbf{P}(B)} & \text{if } \omega \in B \\ 0 & \text{otherwise.} \end{cases}$$

In this case  $\mathbf{P}_q(A) = \mathbf{P}_p(A \mid B)$ . We call  $q$  the *conditional distribution* induced by *conditioning on* the event  $B$ .

