



## Why

Given that we know that one event has occurred, we want language for what the new probabilities should be.<sup>1</sup>

## 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

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

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

## Induced conditional distribution

Conditioning on an event  $B$  induces 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  $B$ .

---

<sup>1</sup>Future editions will improve.

### Total probability

Using the notation  $\mathbf{P}(\cdot \mid \cdot)$ , we can express the law of total probability for  $\mathbf{P}(B)$ ,  $B \subset \Omega$ , 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  $\Omega$ .

