

Introduction to Probability

The **sample space**, S , for a situation or experiment is the set of all possible basic outcomes.

For example, if an ordinary die is thrown once then

$$S = \{1, 2, 3, 4, 5, 6\}$$

which is the set of all the possible numbers that could be thrown.

An **event** is any set of possible basic outcomes. For example:

$$A: \text{“Throwing an even number”} = \{2, 4, 6\}$$

$$B: \text{“Throwing a number greater than 4”} = \{5, 6\}$$

are examples of events.

A set of events is **mutually exclusive** if no two can occur at the same time.

For example, $\{1, 3\}$, $\{2, 4\}$ and $\{6\}$ are three mutually exclusive events.

A set of events is **exhaustive** if at least one of them is bound to occur.

For example, $\{1, 2, 3\}$, $\{3, 4, 5\}$ and $\{5, 6\}$ are three exhaustive events.

The **probability** $P(E)$ of an event E

is an indication of how likely that event is to happen.

Probabilities have the following properties:

- For any event E : $0 \leq P(E) \leq 1$.
- If an event is impossible then its probability is 0 .
- If an event is certain then its probability is 1 .
- If A and B are *mutually exclusive* events then: $P(A \text{ or } B) = P(A) + P(B)$
(where “ A or B ” is the event that one or the other, or both occur).
- The sum of the probabilities of a *mutually exclusive and exhaustive* set of events is 1 .

If A and B are events, then the event “ A and B ” (that both occur) is called the **joint event**.

The **conditional probability** of event A , given that event B occurs, $P(A | B)$, is defined by:

$$P(A | B) = \frac{P(A \text{ and } B)}{P(B)}$$

It follows that the joint probability

$$P(A \text{ and } B) = P(A | B) \times P(B)$$

If A is any event and $\{E_1, E_2, \dots, E_n\}$ are *mutually exclusive and exhaustive* events, then we have the **generalized addition law**:

$$\begin{aligned} P(A) &= P(A \text{ and } E_1) + P(A \text{ and } E_2) + \dots + P(A \text{ and } E_n) \\ &= P(A | E_1) \times P(E_1) + P(A | E_2) \times P(E_2) + \dots + P(A | E_n) \times P(E_n) \end{aligned}$$

Two events are said to be **independent** if the occurrence of one has no effect on the probability of the occurrence of the other.

If events A and B are independent, then:

$$P(A | B) = P(A)$$

and so

$$P(A \text{ and } B) = P(A) \times P(B)$$

Three are three alternative interpretations of probability:

Classical (or theoretical) probability:

Based on simple games of chance involving symmetric objects such as fair coins, dice and packs of cards, in which basic outcomes are equally likely.

For example, for a fair die:

$$P(1) = P(2) = \dots = P(6) = 1/6 .$$

In such simple cases, with finite sample spaces (i.e., with a finite number of basic outcomes), the probability of an event E is simply:

$$P(E) = \frac{\text{number of basic outcomes in } E}{\text{number of basic outcomes in sample space}}$$

For example, if E is the event of throwing a number greater than 4 with a fair die, then $P(E) = 2/6 = 1/3$.

Long-term frequency:

Based on observing n repeated trials of an experiment and counting the number of times m that a particular event E occurs.

The relative frequency, m/n , with which the event occurs is an estimate of the probability of E .

As n is increased, this ratio becomes a more and more accurate estimate of the probability.

$$P(E) = \lim_{n \rightarrow \infty} \frac{m}{n}$$

This interpretation applies to a far wider range of phenomena than does classical probability, for example to industrial processes and to life insurance.

Subjective probability:

A measure of the **degree of belief** one has that an event will occur.

It is a personal judgment, based on all relevant information one has at the time, such as a bookie's odds in a horse race.

This interpretation is applicable to the widest range of phenomena as it neither requires symmetry (as in classical probability) nor repeatability of identical trials (as in the frequentist approach).

It is the most appropriate interpretation in the area of management decision making.

Bayes' Theorem

If A and B are any two events then:

$$P(A \text{ and } B) = P(A | B) \times P(B)$$

Similarly:

$$P(B \text{ and } A) = P(B | A) \times P(A)$$

But the joint event “ A and B ” is the same as “ B and A ” and so we can equate the probabilities.

Therefore:

$$P(A | B) \times P(B) = P(B | A) \times P(A)$$

and so:
$$P(A | B) = \frac{P(B | A)}{P(B)} \times P(A)$$

Bayes' theorem can be seen as relating $P(A | B)$, the conditional probability of A given B , to the absolute probability $P(A)$.

If $\{A_1, A_2, \dots, A_n\}$ are mutually exclusive and exhaustive, then for any one of the A_i :

$$P(A_i | B) = \frac{P(B | A_i) \times P(A_i)}{P(B | A_1) \times P(A_1) + \dots + P(B | A_n) \times P(A_n)}$$

We often use Bayes' Theorem when we want to find the probability of a particular state of affairs, in the light of observations or experiments that have been made.

If A_1, A_2, \dots, A_n are alternative states and observation B has been made, the last equation shows how to relate:

the conditional probability that state A_i is the true state, given observation B
to

the “absolute” probability of A_i , estimated before B had been observed.

Bayes' Theorem provides the mechanism for updating our estimate as to the chance of A_i being the true state, in the light of new information.

$P(A_i)$ is sometimes called the **prior probability** of A_i

prior to

collecting any extra relevant information or

making any extra observations related to the possible occurrence of A_i .

$P(A_i | B)$ is sometimes called the **posterior probability** of A_i

posterior (i.e., after)

observing that B occurred.

$P(B | A_i)$ is sometimes called the **likelihood** of B

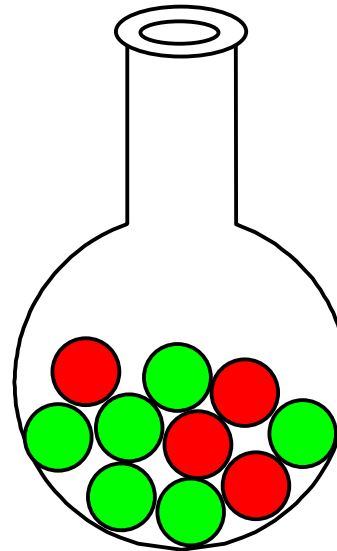
the probability of observing B , in state A_i .

Example:

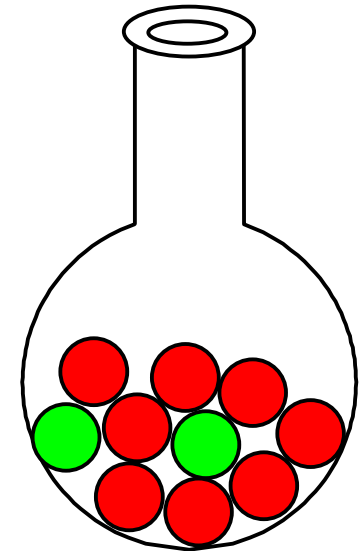
Two opaque jars *A* and *B* each contain ten balls.

Jar *A* contains
four red balls and six green balls.

Jar *B* contains
eight red balls and two green balls.



A



B

One of the two jars is chosen at random.

What is the probability that jar *A* was chosen?

In the absence of any further information, the Laplace criterion gives us the answer $1/2$.

i.e., the prior probability is: $P(A) = 1/2$.

Now suppose that some information is collected to help in deciding which jar was chosen.

A ball is taken at random from the jar.

It turns out to be a red ball.

What effect does this have on the assessment of the probability that jar A had been chosen?

If jar A had been chosen, the probability that a red ball would be withdrawn is $4/10$.

If jar B had been chosen, the probability that a red ball would be withdrawn is $8/10$.

The *likelihoods* of a red ball being withdrawn are:

$$P(\text{red} \mid A) = 4/10,$$

$$P(\text{red} \mid B) = 8/10.$$

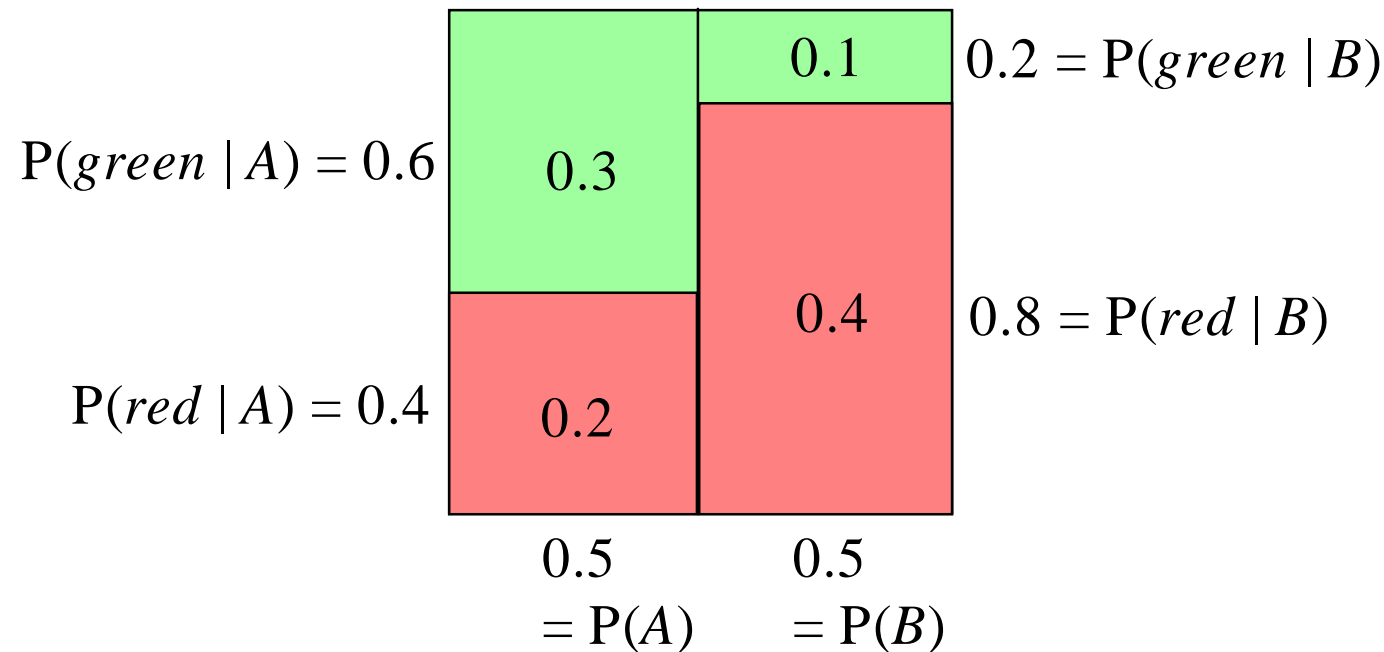
But we are interested in the *posterior probability* $P(A \mid red)$.

Using Bayes' Theorem:

$$\begin{aligned} P(A \mid red) &= \frac{P(red \mid A) \times P(A)}{P(red \mid A) \times P(A) + P(red \mid B) \times P(B)} \\ &= \frac{4/10 \times 1/2}{(4/10 \times 1/2) + (8/10 \times 1/2)} \\ &= \frac{1/5}{1/5 + 2/5} \\ &= 1/3 \end{aligned}$$

It is sometimes easier to employ a diagrammatic method based on Bayes' theorem, which involves:

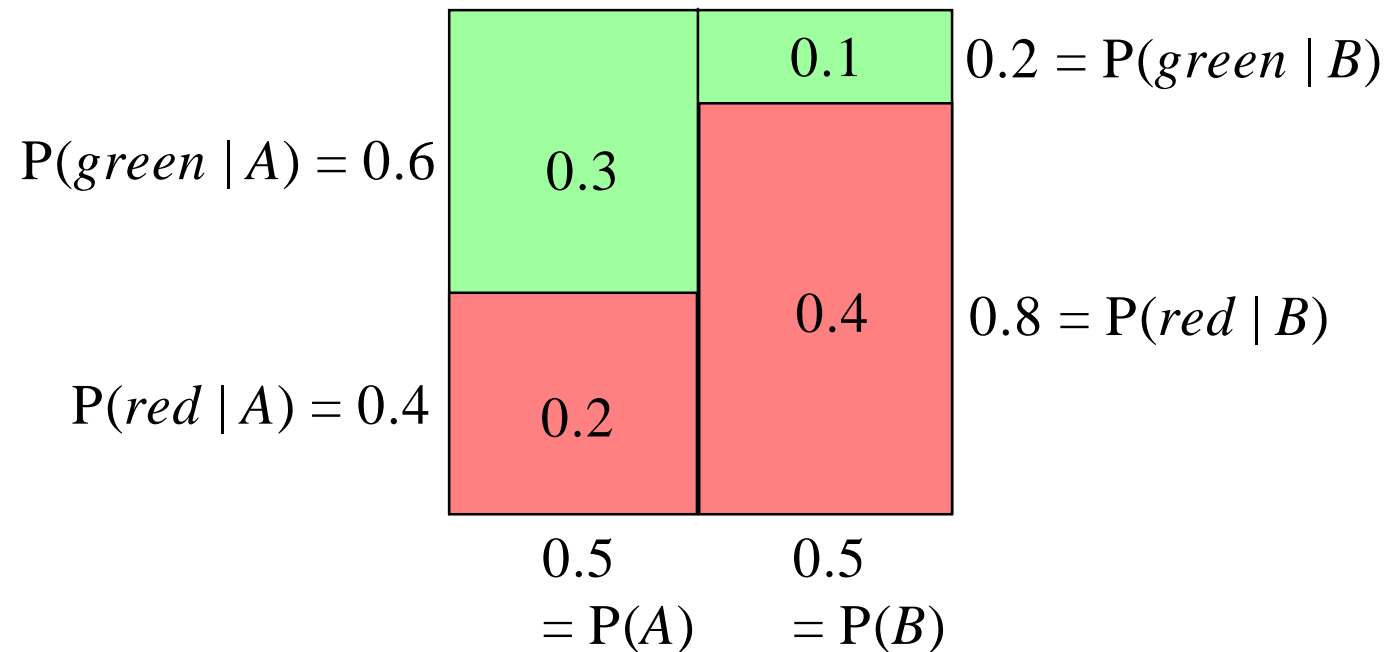
- drawing a unit square,
- marking the a priori probabilities along the base,
- dividing the square into corresponding vertical rectangles,
- dividing each rectangle according to the likelihood values, and
- calculating the area of each of the resulting smaller rectangles.



The two original vertical rectangles correspond to the two events: “jar *A* was picked” and “jar *B* was picked”.

Each of the smaller rectangles represents a joint event, such as “jar *A* was picked **and** a *red* ball was drawn from it”.

The areas of the rectangles are the probabilities of these different events.



The rectangles shaded red represent the two joint events which involve drawing a red ball. The overall probability of such a result is:

$$\begin{aligned}P(\textit{red}) &= P(A \text{ and } \textit{red}) + P(B \text{ and } \textit{red}) \\&= P(\textit{red} \mid A) \times P(A) + P(\textit{red} \mid B) \times P(B) \\&= 0.4 \times 0.5 + 0.8 \times 0.5 \\&= 0.2 + 0.4 \\&= 0.6\end{aligned}$$

The regions shaded green represent the joint events that involve drawing a green ball:

$$\begin{aligned}P(\textit{green}) &= P(A \text{ and } \textit{green}) + P(B \text{ and } \textit{green}) \\&= P(\textit{green} \mid A) \times P(A) + P(\textit{green} \mid B) \times P(B) \\&= 0.6 \times 0.5 + 0.2 \times 0.5 \\&= 0.3 + 0.1 \\&= 0.4\end{aligned}$$

The posterior probabilities are:

$$P(A \mid red) = \frac{P(A \text{ and } red)}{P(red)} = \frac{P(red \mid A) \times P(A)}{P(red)} = \frac{0.2}{0.6} = 1/3$$

$$P(B \mid red) = \frac{P(B \text{ and } red)}{P(red)} = \frac{P(red \mid B) \times P(B)}{P(red)} = \frac{0.4}{0.6} = 2/3$$

$$P(A \mid green) = \frac{P(A \text{ and } green)}{P(green)} = \frac{P(green \mid A) \times P(A)}{P(green)} = \frac{0.3}{0.4} = 3/4$$

$$P(B \mid green) = \frac{P(B \text{ and } green)}{P(green)} = \frac{P(green \mid B) \times P(B)}{P(green)} = \frac{0.1}{0.4} = 1/4$$

Drawing a red ball:

decreases the probability of jar A from $1/2$ to $1/3$ and
increases the probability of jar B from $1/2$ to $2/3$.

Drawing a green ball:

increases the probability of jar A from $1/2$ to $3/4$ and
decreases the probability of jar B from $1/2$ to $1/4$.