**Example 1.** An urn contains 100 marbles: 60 blue marbles and 40 red marbles. A marble is drawn from the urn, what is the **probability** that the marble is blue?

**Assumption:** Each marble is just as likely to be drawn as any other.

**Question:** Why is this a fair assumption?

**Answer:** The probability of an event reflects what we know (and don't know) about the event. With the information we have, the simplest assumption to make is that all the marbles are equally likely to be drawn... Because we have no reason to conclude otherwise. This is called the *principle of insufficient reason* (or the *principle of indifference*).

With more information or more experience with this urn of marbles, we may eventually choose to reconsider this assumption.

**Conclusion:** There are 60 blue marbles and 100 marbles total, all of which are equally likely to be drawn, the probability of drawing a blue marble is equal to the proportion of blue marbles in the urn, which is $60/100 = 60\%$.

**Shorthand:** I'll write $P(E)$ to denote the probability of the event $E$.

Since probability can be viewed as a proportion, we observe that

(*) If an event $E$ is certain to occur, then $P(E) = 100\%$.

(*) If $E$ is certain to ***not*** occur, then $P(E) = 0\%$.

(*) In all other cases, $0\% < P(E) < 100\%$.

$\Rightarrow$ The closer $P(E)$ is to 100% the more certain we are that $E$ will occur.

$\Rightarrow$ The closer $P(E)$ is to 0% the more certain we are that $E$ will not occur.

**Observation:** the event "*E does* ***not*** *occur*" is called the *complement* of $E$. Whenever we perform an experiment or procedure that might lead to the occurrence of $E$, then either $E$ or *not* $E$ must occur so their relative frequencies must add up to 100%. I.e.,

(*) $P(\text{not } E) = 100\% - P(E)$.

**Example 2.** Suppose that the marbles in Example 1 are marked with the letter $A$ or $B$. Specifically,

- 40 blue marbles are marked with an $A$ and 20 blue marbles are marked with a $B$.
- 10 red marbles are marked with an $A$ and 30 red marbles are marked with a $B$.

A marble is drawn *at random* from the urn. What is the probability that it is marked with an $A$?

$\Rightarrow$ There are 50 marbles marked with an $A$, so $P(A) = 50/100 = 50\%$.

A marble is drawn from the urn, and it is observed to be blue. What is the probability that it is marked with an $A$?

We now have more information: the marble is known to be blue. We can ignore the red marbles and imagine that the marble was drawn from an urn of 60 blue marbles, 40 of which are marked with an $A$.

$\Rightarrow$ $P(A \text{ given that the marble is blue}) = 40/60 \approx 66.67\%$.

**Definition.** The probability of event $E$ given that we know that event $F$ has occurred is called the ***conditional probability of E given F***. We use the shorthand $P(E|F)$ for this conditional probability.

E.g., in Example 2, we found that $P(A|$ blue marble$) \approx 66.67\%$.

(*) The probability of $E$ (without any additional information) is sometimes called the *unconditional probability* of $E$.

**Example 3.** A marble is drawn from the urn in Example 2 and you are told that it is marked with an $A$. What is the probability that the marble is blue?

$\Rightarrow$ There are 50 marbles marked with an $A$ and 40 of these are blue, so $P($blue marble $|A) = 40/50 = 80\%$.

**Example 4.** What is the probability that a marble drawn from the (same) urn is red, if we know that it is marked with a $B$?

$\Rightarrow$ There are 50 marbles marked with a $B$ of which 30 are red, so $P($red marble$|B) = 30/50 = 60\%$.

**Example 5.** What is the probability that a marble drawn at random from the (same) urn is red ***and*** marked with a $B$?

$\Rightarrow$ There are 100 marbles in the urn and 30 of them are both red and marked with a $B$, so $P(\text{red } \mathit{and} \ B) = 30/100 = 30\%$.

Or...

$\Rightarrow$ 40% of the marbles in the urn are red, and of these red marbles, 75% are marked with a $B$, so

$$P(\text{red } \mathit{and} \ B) = (40\%) \times (75\%) = 30\%.$$

In other words

$$P(\text{red } \mathit{and} \ B) = P(\text{red}) \times P(B|\text{red}).$$

**The Multiplication rule.** *Given two events $E$ and $F$*

$$P(E \text{ and } F) = P(E) \times P(F|E)$$

Also

$$P(E \text{ and } F) = P(F) \times P(E|F)$$

**Example 6.** Three cards are dealt from the top of a well-shuffled deck. What is the probability that the first card is a King and the third card is a 7?

$\Rightarrow$ The probability that the first card is a King is $4/52$ and the probability that the third card is a 7 given that the first card is a King is $4/51$, so

$$\begin{aligned} &P(\text{first card King } \textit{and} \text{ third card 7}) \\ &\qquad = P(\text{first card King}) \times P(\text{third card 7}|\text{first card King}) \\ &\qquad = \frac{4}{52} \cdot \frac{4}{51} = \frac{16}{2652} \approx 0.6\% \end{aligned}$$

**Observation:** The second card is unknown and so for all intents and purposes it is just another card in the deck.

**Example 7.** A card is dealt from the top of a well-shuffled deck, then it is replaced, the deck is reshuffled and another card is dealt. What is the probability that the second card is a 7 given that the first card is a King?

$\Rightarrow$ Since the first card was replaced (and the deck was reshuffled) before the second card was dealt, the nature of the first card doesn't provide any new information about the nature of the second card, so

$$P(\text{second card } 7|\text{first card King}) = P(\text{second card } 7) = \frac{4}{52} \approx 7.7\%.$$

**Definition.** If $P(E|F) = P(E)$, then the events $E$ and $F$ are said to be (statistically) ***independent***.

**Comment:** *Independence* is not the same as unrelated. Two events can be closely related, but statistically independent.

**Example 8.** A box contains 200 tickets...

- 120 of the tickets are marked with an $X$ and 80 tickets are marked with a $Y$.
- Of the $X$-tickets, 30 are also marked with an $A$ and the other 90 are marked with a $U$.
- Of the $Y$-tickets, 20 are also marked with an $A$ and the other 60 are marked with a $U$.

One ticket is drawn at random from the box...

(*) There are 200 tickets overall and $50 = 30 + 20$ of them are marked with an $A$, so $P(A) = 25\%$.

(*) There are 120 $X$-tickets and 30 of them are marked with an $A$, so $P(A|X) = 30/120 = 25\%$.

(*) The events *"ticket is marked with an A"* and *"ticket is marked with an X"* are *independent* (but not unrelated).

**Comments.**

(1) If the events $E$ and $F$ are independent, then $P(E|F) = P(E)$ (and $P(F|E) = P(F)$). So, if $E$ and $F$ are independent events, then the multiplication rule reduces to

$$P(E \text{ and } F) = P(E)P(F).$$

In fact, this formula can be used as the definition of independence.

(2) Independence is often an assumption that we make about the events that we are considering. This is an assumption that makes computing probability easier.

(3) But — the assumption of independence must be justified in one way or another, and reevaluated if subsequent results point in another direction. One of the easiest ways to misuse probability is to make an unjustified assumption of independence. See the example in Section 13.5 of the textbook.