# Formula of Total Probability, Bayes' Rule, and Applications

Recall that for any event $A$, the pair of events $A$ and $\overline{A}$ has an intersection that is empty, whereas the union $A \cup \overline{A}$ represents the total population of interest. In fact, this pair of events $\left\{A, \overline{A}\right\}$ is a special case of a *partition* of the sample space, hereinafter denoted by $S$.

**1.** *Partition of Sample Space and Formula of Total Probability.*

**Definition of Partition.** A collection of events $\{ S_1 , S_2 , \cdots, S_n \}$ of a certain sample space (or population) $S$ is called a **partition** if

(i) $S_1 , S_2 , \cdots, S_n$ are *mutually exclusive events*;

(ii) $S_1 \cup S_2 \cup \cdots \cup S_n = S$.

**Illustrative Example.** In the nineteenth century G. Mendel conducted a famous experiment that led to the first announcement of elementary genetic principles. He bred hybrid strains of peas and simultaneously observed the color (green or yellow) and smoothness (round or wrinkled) of the offspring peas. If $S$ denotes the set of all peas involved in the pea-breeding experiment, and

$S_1$ denotes the subpopulation of round and green peas;

$S_2$ denotes the subpopulation of round and yellow peas;

$S_3$ denotes the subpopulation of wrinkled and green peas;

$S_4$ denotes the subpopulation of wrinkled and yellow peas;

then $\{ S_1 , S_2 , S_3 , S_4 \}$ represents a partition of $S$.

**Formula of Total Probability.**

Assume that the set of events $\{ S_1 , S_2 , \cdots, S_n \}$ constitutes a partition of the sample space $S$. Assume that for every $i$, $1 \le i \le n$,

$$P(S_i) > 0.$$

Then for any event $A$, we have

$$P(A) = \sum_{i=1}^{n} P(S_i) \cdot P(A \mid S_i). \tag{1}$$

**Proof.** It follows from the multiplication law of probability that for every $i$, $1 \le i \le n$,

$$P(S_i) \cdot P(A \mid S_i) = P(AS_i).$$

On the other hand, since the events $S_1$, $S_2$, $\cdots$, $S_n$ are mutually exclusive, we have that the events $AS_1$, $AS_2$, $\cdots$ $AS_n$ are also mutually exclusive. In addition note that

(2) $$AS_1 \cup AS_2 \cup \cdots \cup AS_n = A.$$

It would be instructive if the student tries to verify (2) by use of a Venn diagram. Then an application of the addition law of probability to (2) gives (1).

**Illustrative Example.** A diagnostic test for a certain disease is known to be 95% accurate, i.e., if a person has the disease, the test will detect it with probability 0.95. Also, if the person does not have the disease, the test will report that they do not have it with the same probability 0.95. In addition, it is known from previous data that only 1% of the population has this particular disease. What is the probability that a particular person chosen at random will be tested positive?

**Solution.** Let

$T^+$ denote the event that a person is tested positive;
$T^-$ denote the event that a person is tested negatively;
$D$ denote the event that a person has the disease.

Then it follows from the above stated conditions that

$$P(D) = 0.01, \quad P(\overline{D}) = 0.99,$$

$$P(T^+ \mid D) = 0.95, \text{ and } P(T^- \mid \overline{D}) = 0.95.$$

In particular, since

$$\underbrace{P(T^- \mid \overline{D})}_{0.95} + P(T^+ \mid \overline{D}) = 1,$$

we have that

$$P(T^+ \mid \overline{D}) = 0.05.$$

Now we apply the formula of total probability from (1) with $A = T^+$, $n = 2$, $S_1 = D$, and $S_2 = \overline{D}$, to obtain

$$\begin{aligned} P\left(T^{+}\right) &= P\left(T^{+} \mid D\right) \cdot P(D) + P\left(T^{+} \mid \overline{D}\right) \cdot P\left(\overline{D}\right) \\ &= (0.95) \cdot (0.01) + (0.05) \cdot (0.99) \\ &\approx 0.059. \end{aligned}$$

**2.** *Bayes' Rule.*

This important rule enables one to compute a conditional probability when the original condition now becomes the event of interest.

Assume that the set $\{S_1, S_2, \cdots, S_n\}$ constitutes a partition of the sample space $S$. Assume that for each $i$, $1 \le i \le n$,

$$P(S_i) > 0.$$

Fix any event $A$. Then for any given $j$, $1 \le j \le n$,

$$P\left(S_j \mid A\right) = \frac{P\left(S_j\right) \cdot P\left(A \mid S_j\right)}{\sum_{i=1}^{n} P(S_i) \cdot P(A \mid S_i)}. \tag{3}$$

The key thing to note about Bayes' theorem is that the information that will be given in a problem will be the conditional probabilities $P(A \mid S_i)$, $1 \le i \le n$, that appear on the right-hand side of the equation, whereas what is sought is one of the conditional probabilities, $P\left(S_j \mid A\right)$, where the events $S_j$ and $A$ are "reversed" from the given information. I.e., given that $A$ occurred, what is the probability it happened "through $S_j$".

**Proof of Bayes' theorem.** Note that by the multiplication law of probability, the numerator of the fraction on the right-hand side of (3) can be rewritten as

$$P\left(S_j\right) \cdot P\left(A \mid S_j\right) = P\left(AS_j\right).$$

At the same time, by the formula of total probability (formula (1) above), the denominator of the fraction on the right-hand side of (3) is equal to

$$\sum_{i=1}^{n} P(S_i) \cdot P(A \mid S_i) = P(A).$$

Hence, the right-hand side of (3) is equal to

$$\frac{P\left(AS_j\right)}{P(A)} = P\left(S_j \mid A\right)$$

by the definition of conditional probability.

**3.** *Examples*.

**Illustrative Example 1.** It is quite common that different illnesses produce similar or even identical symptoms. Suppose that any one of the illnesses $X$, $Y$, or $Z$ lead to the same set of symptoms, hereafter denoted as $U$. For simplicity assume that the illnesses $X$, $Y$, and $Z$ are mutually exclusive and that there are no other illnesses leading to the same set of symptoms. Suppose the probabilities of contracting these three illnesses are:

$$P(X) = 0.03, \qquad P(Y) = 0.01, \qquad P(Z) = 0.02,$$

and that the chances of developing the set of symptoms $U$, given a specific illness are:

$$P(U \mid X) = 0.85, \qquad P(U \mid Y) = 0.92 \qquad P(U \mid Z) = 0.80.$$

If a sick person develops the set of symptoms $U$, what are the chances he or she has illness $X$ ?

**Solution.** First note that the set of events $X$, $Y$, and $Z$ together do not represent a partition. Therefore, define $H$ to be the event of not suffering from any of $X$, $Y$, or $Z$, i.e., the complement of the union of $X$, $Y$, and $Z$,

$$H = \overline{X \cup Y \cup Z}.$$

Then we have

$$\begin{aligned} P(H) &= 1 - P(X) - P(Y) - P(Z) \\ &= 1 - 0.03 - 0.01 - 0.02 \\ &= 0.94. \end{aligned}$$

However,

$$P(U \mid H) = 0.$$

Applying Bayes' rule yields that the conditional probability that given the symptoms $U$, that a person indeed has the illness $X$, viz., $P(X \mid U)$, is

$$\begin{aligned} P(X \mid U) &= \frac{P(U \mid X)\cdot P(X)}{P(U \mid X)\cdot P(X) + P(U \mid Y)\cdot P(Y) + P(U \mid Z)\cdot P(Z) + P(U \mid H)\cdot P(H)} \\ &= \frac{(0.85)\cdot(0.03)}{(0.85)\cdot(0.03) + (0.92)\cdot(0.01) + (0.80)\cdot(0.02) + 0} \\ &= 0.5029. \end{aligned}$$

Note that the data given at the outset of the problem above involved the conditional probabilities of "$U$ given $X$, $U$ given $Y$, $U$ given Z, and $U$ given $H$", but what was sought was the conditional probability of "$X$ given $U$", which involved the reverse of the conditions of the given data in the problem. This is the prototypical situation for the application of Bayes' theorem.

**Illustrative Example 2.** In this example we consider a situation somewhat like the earlier example above on pp. 2 and 3 of this insert following the formula of total probability. Suppose we are concerned with medically testing for leukemia. Let

$T^+$ denote the event that the test is positive, suggesting the person has leukemia;

$T^-$ denote the event that the test is negative, suggesting the person does not have leukemia;

$L$ denote the event that the person tested has leukemia;

$\bar{L}$ denote the event that the person tested does not have leukemia.

It is the case that the medical test for leukemia is not perfectly accurate. Most of the time, if one has leukemia, the test will be positive. Past records indicate that $P\left(T^+ \mid L\right) = 0.98$. Similarly, if one does not have leukemia, the test is usually negative. Again, it is known that $P\left(T^- \mid \bar{L}\right) = 0.99$. All this notwithstanding, there are people who sometimes test positively, but do not , in fact, have the disease; and some who test negatively, but do indeed have the disease. If we also know that $P(L) = 0.000001$, find:

(a) the probability $P\left(\bar{L} \mid T^+\right)$ of a false positive test;

and

(b) the probability $P\left(L \mid T^-\right)$ of a false negative test.

**Solution. (a)** Note again that in this problem we are given the conditional probabilities of "$T^+$ given $L$ and $T^-$ given $\bar{L}$", but are asked to find the conditional probabilities that have the $T^+$, $T^-$ events and the $L$, $\bar{L}$ events "reversed". Hence we employ Bayes' rule. This yields

$$P\left(\bar{L} \mid T^+\right) = \frac{P\left(T^+ \mid \bar{L}\right) \cdot P\left(\bar{L}\right)}{P\left(T^+ \mid \bar{L}\right) \cdot P\left(\bar{L}\right) + P\left(T^+ \mid L\right) \cdot P(L)}.$$

Since

$$P\left(\bar{L}\right) = 1 - P(L) = 0.999999,$$

and

$$P\left(T^{+} \mid \overline{L}\right) = 1 - P\left(T^{-} \mid \overline{L}\right) = 0.01,$$

we obtain

$$P\left(\overline{L} \mid T^{+}\right) = \frac{(0.01)\cdot(0.999999)}{(0.01)\cdot(0.999999) + (0.98)\cdot(0.000001)} \approx 0.99991 .$$

In particular, this implies that

$$P\left(L \mid T^{+}\right) = 1 - P\left(\overline{L} \mid T^{+}\right) \approx 0.00009.$$

**(b)** Try to compute $P\left(L \mid T^{-}\right)$ in an analogous manner as an exercise.

(**Answer**: $2.020204\times10^{-8}$.)