

2.1: Conditional Probability and Bayes' Rule

In many situations, additional information about the result of a probability experiment is known (or at least assumed to be known) and given that information the probability of some other event is desired. For this scenario, we compute what is referred to as *conditional probability*.

Definition 2.1.1

For events A and B , with $P(B) > 0$, the **conditional probability** of A given B , denoted $P(A | B)$, is given by

$$P(A | B) = \frac{P(A \cap B)}{P(B)}.$$

In computing a conditional probability we *assume* that we know the outcome of the experiment is in event B and then, given that additional information, we calculate the probability that the outcome is also in event A . This is useful in practice given that partial information about the outcome of an experiment is often known, as the next example demonstrates.

Example 2.1.1

Continuing in the context of [Example 1.2.1](#), where we considered tossing a fair coin twice, define D to be the event that at least one tails is recorded:

$$D = \{ht, th, tt\}$$

Let's calculate the conditional probability of A given D , i.e., the probability that at least one heads is recorded (event A) assuming that at least one tails is recorded (event D). Recalling that outcomes in this sample space are equally likely, we apply the definition of conditional probability ([Definition 2.1.1](#)) and find

$$P(A | D) = \frac{P(A \cap D)}{P(D)} = \frac{P(\{ht, th\})}{P(\{ht, th, tt\})} = \frac{(2/4)}{(3/4)} = \frac{2}{3} \approx 0.67.$$

Note that in [Example 1.2.1](#) we found the **un**-conditional probability of A to be $P(A) = 0.75$. So, knowing that at least one tails was recorded, i.e., assuming event D occurred, the conditional probability of A given D decreased. This is because, if event D occurs, then the outcome hh in A cannot occur, thereby decreasing the chances that event A occurs.

Exercise 2.1.1

Suppose we randomly draw a card from a [standard deck of 52 playing cards](#).

- If we know that the card is a King, what is the probability that the card is a club?
- If we instead know that the card is black, what is the probability that the card is a club?

Answer

In order to compute the necessary probabilities, first note that the sample space is given by the set of cards in a standard deck of playing cards. So the number of outcomes in the sample space is 52. Next, note that the outcomes are equally likely, since we are *randomly* drawing the card from the deck.

For part (a), we are looking for the conditional probability that the randomly selected card is club, given that it is a King. If we let C denote the event that the card is a club and K the event that it is a King, then we are looking to compute

$$P(C | K) = \frac{P(C \cap K)}{P(K)}. \quad (2.1.1)$$

To compute these probabilities, we count the number of outcomes in the following events:

of outcomes in C = # of clubs in standard deck = 13

of outcomes in K = # of Kings in standard deck = 4

of outcomes in $C \cap K$ = # of King of clubs in standard deck = 1

The probabilities in Equation 2.1.1 are then given by dividing the counts of outcomes in each event by the total number of outcomes in the sample space (by the boxed Equation 1.3.2 in Section 1.3):

$$P(C | K) = \frac{P(C \cap K)}{P(K)} = \frac{(1/52)}{(4/52)} = \frac{1}{4} = 0.25.$$

For part (b), we are looking for the conditional probability that the randomly selected card is club, given that it is instead black. If we let B denote the event that the card is black, then we are looking to compute

$$P(C | B) = \frac{P(C \cap B)}{P(B)}. \quad (2.1.2)$$

To compute these probabilities, we count the number of outcomes in the following events:

of outcomes in B = # of black cards in standard deck = 26

of outcomes in $C \cap B$ = # of black clubs in standard deck = 13

The probabilities in Equation 2.1.2 are then given by dividing the counts of outcomes in each event by the total number of outcomes in the sample space:

$$P(C | B) = \frac{P(C \cap B)}{P(B)} = \frac{(13/52)}{(26/52)} = \frac{13}{26} = 0.5.$$

Remark: Exercise 2.1.1 demonstrates the following fact. For sample spaces with equally likely outcomes, conditional probabilities are calculated using

$$P(A | B) = \frac{\text{number of outcomes in } A \cap B}{\text{number of outcomes in } B}. \quad (2.1.3)$$

In other words, if we know that the outcome of the probability experiment is in the event B , then we restrict our focus to the outcomes in that event that are also in A . We can think of this as event B taking the place of the sample space, since we know the outcome must lie in that event.

Properties of Conditional Probability

As with unconditional probability, we also have some useful properties for conditional probabilities. The first property below, referred to as the *Multiplication Law*, is simply a rearrangement of the probabilities used to define conditional probability. The Multiplication Law provides a way for computing the probability of an intersection of events when the conditional probabilities are known.

Multiplication Law

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

The next two properties are useful when a *partition* of the sample space exists, where a partition is a way of dividing up the outcomes in the sample space into non-overlapping sets. A partition is formally defined in the *Law of Total Probability* below. In many cases, when a partition exists, it is easy to compute the conditional probability of an event in the sample space given an event in the partition. The Law of Total Probability then provides a way of using those conditional probabilities of an event, given the partition to compute the unconditional probability of the event. Following the Law of Total Probability, we state *Bayes' Rule*, which is really just an application of the Multiplication Law. Bayes' Rule is used to calculate what are informally referred to as "reverse conditional probabilities", which are the conditional probabilities of an event in a partition of the sample space, given any other event.

Law of Total Probability

Suppose events B_1, B_2, \dots, B_n , satisfy the following:

1. $\Omega = B_1 \cup B_2 \cup \dots \cup B_n$
2. $B_i \cap B_j = \emptyset$, for every $i \neq j$
3. $P(B_i) > 0$, for $i = 1, \dots, n$

We say that the events B_1, B_2, \dots, B_n , **partition** the sample space Ω . Then for any event A , we can write

$$P(A) = P(A | B_1)P(B_1) + \dots + P(A | B_n)P(B_n).$$

Bayes' Rule

Let B_1, B_2, \dots, B_n , partition the sample space Ω and let A be an event with $P(A) > 0$. Then, for $j = 1, \dots, n$, we have

$$P(B_j | A) = \frac{P(A | B_j)P(B_j)}{P(A)}.$$

A common application of the Law of Total Probability and Bayes' Rule is in the context of medical diagnostic testing.

Example 2.1.2

Consider a test that can diagnose kidney cancer. The **test correctly detects when a patient has cancer 90% of the time**. Also, **if a person does not have cancer, the test correctly indicates so 99.9% of the time**. Finally, suppose it is known that **1 in every 10,000 individuals has kidney cancer**. We find the probability that a patient has kidney cancer, given that the test indicates she does.

First, note that we are finding a conditional probability. If we let A denote the event that the patient tests positive for cancer, and we let B_1 denote the event that the patient actually has cancer, then we want

$$P(B_1 | A).$$

If we let $B_2 = B_1^c$, then we have a partition of all patients (which is the sample space) given by B_1 and B_2 .

In the first paragraph of this example, we are given the following probabilities:

$$\text{test correctly detects cancer 90\% of time: } P(A | B_1) = 0.9$$

$$\text{test correctly detects no cancer 99.9\% of time: } P(A^c | B_2) = 0.999 \Rightarrow P(A | B_2) = 1 - P(A^c | B_2) = 0.001$$

$$\text{1 in every 10,000 individuals has cancer: } P(B_1) = 0.0001 \Rightarrow P(B_2) = 1 - P(B_1) = 0.9999$$

Since we have a partition of the sample space, we apply the Law of Total Probability to find $P(A)$:

$$P(A) = P(A | B_1)P(B_1) + P(A | B_2)P(B_2) = (0.9)(0.0001) + (0.001)(0.9999) = 0.0010899$$

Next, we apply Bayes' Rule to find the desired conditional probability:

$$P(B_1 | A) = \frac{P(A | B_1)P(B_1)}{P(A)} = \frac{(0.9)(0.0001)}{0.0010899} \approx 0.08$$

This implies that only about 8% of patients that test positive under this particular test actually have kidney cancer, which is not very good.

This page titled [2.1: Conditional Probability and Bayes' Rule](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Kristin Kuter](#).