3

Conditional probability and independence

Knowing that an event has occurred sometimes forces us to reassess the probability of another event; the new probability is the *conditional* probability. If the conditional probability equals what the probability was before, the events involved are called *independent*. Often, conditional probabilities and independence are needed if we want to compute probabilities, and in many other situations they simplify the work.

3.1 Conditional probability

In the previous chapter we encountered the events L, "born in a long month," and R, "born in a month with the letter r." Their probabilities are easy to compute: since $L = \{\text{Jan, Mar, May, Jul, Aug, Oct, Dec}\}$ and $R = \{\text{Jan, Feb, Mar, Apr, Sep, Oct, Nov, Dec}\}$, one finds

$$\mathrm{P}(L) = \frac{7}{12} \quad \text{and} \quad \mathrm{P}(R) = \frac{8}{12}.$$

Now suppose that it is *known* about the person we meet in the street that he was born in a "long month," and we wonder whether he was born in a "month with the letter r." The information given excludes five outcomes of our sample space: it cannot be February, April, June, September, or November. Seven possible outcomes are left, of which only four—those in $R \cap L = \{\text{Jan, Mar, Oct, Dec}\}$—are favorable, so we reassess the probability as $4/7$. We call this the *conditional probability of R given L*, and we write:

$$\mathrm{P}(R \,|\, L) = \frac{4}{7}.$$

This is not the same as $\mathrm{P}(R \cap L)$, which is $1/3$. Also note that $\mathrm{P}(R \,|\, L)$ is the proportion that $\mathrm{P}(R \cap L)$ is of $\mathrm{P}(L)$.

Quick exercise 3.1 Let $N = R^c$ be the event "born in a month without r." What is the conditional probability $\mathrm{P}(N \mid L)$?

Recalling the three envelopes on our doormat, consider the events "envelope 1 is the middle one" (call this event A) and "envelope 2 is the middle one" (B). Then $\mathrm{P}(A) = \mathrm{P}(213 \text{ or } 312) = 1/3$; by symmetry, the same is found for $\mathrm{P}(B)$. We say that the envelopes are in order if their order is either 123 or 321. Suppose we know that they are *not* in order, but otherwise we do not know anything; what are the probabilities of A and B, given this information?

Let C be the event that the envelopes are not in order, so: $C = \{123, 321\}^c = \{132, 213, 231, 312\}$. We ask for the probabilities of A and B, given that C occurs. Event C consists of four elements, two of which also belong to A: $A \cap C = \{213, 312\}$, so $\mathrm{P}(A \mid C) = 1/2$. The probability of $A \cap C$ is half of $\mathrm{P}(C)$. No element of C also belongs to B, so $\mathrm{P}(B \mid C) = 0$.

Quick exercise 3.2 Calculate $\mathrm{P}(C \mid A)$ and $\mathrm{P}(C^c \mid A \cup B)$.

In general, computing the probability of an event A, given that an event C occurs, means finding which fraction of the probability of C is also in the event A.

Definition. The *conditional probability of A given C* is given by:

$$\mathrm{P}(A \mid C) = \frac{\mathrm{P}(A \cap C)}{\mathrm{P}(C)},$$

provided $\mathrm{P}(C) > 0$.

Quick exercise 3.3 Show that $\mathrm{P}(A \mid C) + \mathrm{P}(A^c \mid C) = 1$.

This exercise shows that the rule $\mathrm{P}(A^c) = 1 - \mathrm{P}(A)$ also holds for conditional probabilities. In fact, even more is true: if we have a fixed conditioning event C and define $Q(A) = \mathrm{P}(A \mid C)$ for events $A \subset \Omega$, then Q is a probability function and hence satisfies all the rules as described in Chapter 2. The definition of conditional probability agrees with our intuition and it also works in situations where computing probabilities by counting outcomes does not.

A chemical reactor: residence times

Consider a continuously stirred reactor vessel where a chemical reaction takes place. On one side fluid or gas flows in, mixes with whatever is already present in the vessel, and eventually flows out on the other side. One characteristic of each particular reaction setup is the so-called residence time distribution, which tells us how long particles stay inside the vessel before moving on. We consider a continuously stirred tank: the contents of the vessel are perfectly mixed at all times.

Let R_t denote the event "the particle has a residence time longer than t seconds." In Section 5.3 we will see how continuous stirring determines the probabilities; here we just use that in a particular continuously stirred tank, R_t has probability e^{-t}. So:

$$\begin{aligned} \mathrm{P}(R_3) &= \mathrm{e}^{-3} = 0.04978\ldots \\ \mathrm{P}(R_4) &= \mathrm{e}^{-4} = 0.01831\ldots. \end{aligned}$$

We can use the definition of conditional probability to find the probability that a particle that has stayed more than 3 seconds will stay more than 4:

$$\mathrm{P}(R_4 \,|\, R_3) = \frac{\mathrm{P}(R_4 \cap R_3)}{\mathrm{P}(R_3)} = \frac{\mathrm{P}(R_4)}{\mathrm{P}(R_3)} = \frac{\mathrm{e}^{-4}}{\mathrm{e}^{-3}} = \mathrm{e}^{-1} = 0.36787\ldots.$$

Quick exercise 3.4 Calculate $\mathrm{P}(R_3 \,|\, R_4^c)$.

For more details on the subject of residence time distributions see, for example, the book on reaction engineering by Fogler ([11]).

3.2 The multiplication rule

From the definition of conditional probability we derive a useful rule by multiplying left and right by $\mathrm{P}(C)$.

The multiplication rule. For any events A and C:

$$\mathrm{P}(A \cap C) = \mathrm{P}(A \,|\, C) \cdot \mathrm{P}(C)\,.$$

Computing the probability of $A \cap C$ can hence be decomposed into two parts, computing $\mathrm{P}(C)$ and $\mathrm{P}(A \,|\, C)$ separately, which is often easier than computing $\mathrm{P}(A \cap C)$ directly.

The probability of no coincident birthdays

Suppose you meet two arbitrarily chosen people. What is the probability their birthdays are different? Let B_2 denote the event that this happens. Whatever the birthday of the first person is, there is only one day the second person cannot "pick" as birthday, so:

$$\mathrm{P}(B_2) = 1 - \frac{1}{365}.$$

When the same question is asked with *three* people, conditional probabilities become helpful. The event B_3 can be seen as the intersection of the event B_2, "the first two have different birthdays," with event A_3 "the third person has a birthday that does not coincide with that of one of the first two persons." Using the multiplication rule:

$$\mathrm{P}(B_3) = \mathrm{P}(A_3 \cap B_2) = \mathrm{P}(A_3 \,|\, B_2)\mathrm{P}(B_2)\,.$$

The conditional probability $\mathrm{P}(A_3 \,|\, B_2)$ is the probability that, when two days are already marked on the calendar, a day picked at random is not marked, or

$$\mathrm{P}(A_3 \,|\, B_2) = 1 - \frac{2}{365},$$

and so

$$\mathrm{P}(B_3) = \mathrm{P}(A_3 \,|\, B_2)\mathrm{P}(B_2) = \left(1 - \frac{2}{365}\right) \cdot \left(1 - \frac{1}{365}\right) = 0.9918.$$

We are already halfway to solving the general question: in a group of n arbitrarily chosen people, what is the probability there are no coincident birthdays? The event B_n of no coincident birthdays among the n persons is the same as: "the birthdays of the first $n-1$ persons are different" (the event B_{n-1}) *and* "the birthday of the nth person does not coincide with a birthday of any of the first $n-1$ persons" (the event A_n), that is,

$$B_n = A_n \cap B_{n-1}.$$

Applying the multiplication rule yields:

$$\mathrm{P}(B_n) = \mathrm{P}(A_n \,|\, B_{n-1}) \cdot \mathrm{P}(B_{n-1}) = \left(1 - \frac{n-1}{365}\right) \cdot \mathrm{P}(B_{n-1})$$

as person n should avoid $n-1$ days. Applying the same step to $\mathrm{P}(B_{n-1})$, $\mathrm{P}(B_{n-2})$, etc., we find:

$$\begin{aligned}
\mathrm{P}(B_n) &= \left(1 - \frac{n-1}{365}\right) \cdot \mathrm{P}(A_{n-1} \,|\, B_{n-2}) \cdot \mathrm{P}(B_{n-2}) \\
&= \left(1 - \frac{n-1}{365}\right) \cdot \left(1 - \frac{n-2}{365}\right) \cdot \mathrm{P}(B_{n-2}) \\
&\;\;\vdots \\
&= \left(1 - \frac{n-1}{365}\right) \cdots \left(1 - \frac{2}{365}\right) \cdot \mathrm{P}(B_2) \\
&= \left(1 - \frac{n-1}{365}\right) \cdots \left(1 - \frac{2}{365}\right) \cdot \left(1 - \frac{1}{365}\right).
\end{aligned}$$

This can be used to compute the probability for arbitrary n. For example, we find: $\mathrm{P}(B_{22}) = 0.5243$ and $\mathrm{P}(B_{23}) = 0.4927$. In Figure 3.1 the probability

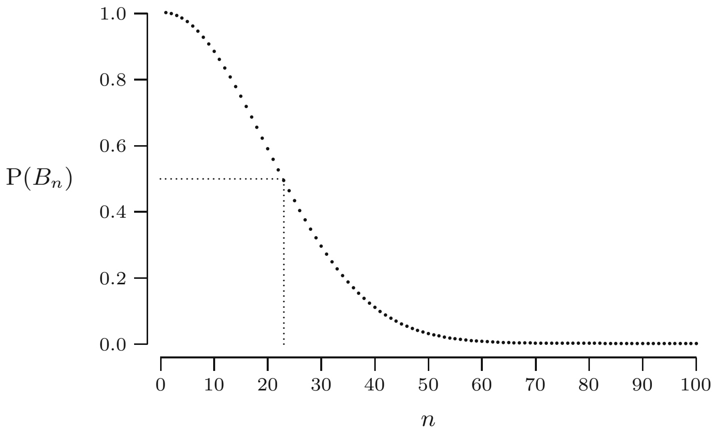


Fig. 3.1. The probability $\mathrm{P}(B_n)$ of no coincident birthdays for $n = 1, \ldots, 100$.

$\mathrm{P}(B_n)$ is plotted for $n = 1, \ldots, 100$, with dotted lines drawn at $n = 23$ and at probability 0.5. It may be hard to believe, but with just 23 people the probability of all birthdays being different is less than 50%!

QUICK EXERCISE 3.5 Compute the probability that three arbitrary people are born in different months. Can you give the formula for n people?

It matters how one conditions

Conditioning can help to make computations easier, but it matters how it is applied. To compute $\mathrm{P}(A \cap C)$ we may condition on C to get

$$\mathrm{P}(A \cap C) = \mathrm{P}(A \mid C) \cdot \mathrm{P}(C)\,;$$

or we may condition on A and get

$$\mathrm{P}(A \cap C) = \mathrm{P}(C \mid A) \cdot \mathrm{P}(A)\,.$$

Both ways are valid, but often *one* of $\mathrm{P}(A \mid C)$ and $\mathrm{P}(C \mid A)$ is easy and the other is not. For example, in the birthday example one could have tried:

$$\mathrm{P}(B_3) = \mathrm{P}(A_3 \cap B_2) = \mathrm{P}(B_2 \mid A_3)\mathrm{P}(A_3)\,,$$

but just trying to understand the conditional probability $\mathrm{P}(B_2 \mid A_3)$ already is confusing:

> The probability that the first two persons' birthdays differ given that the third person's birthday does not coincide with the birthday of one of the first two ...?

Conditioning should lead to easier probabilities; if not, it is probably the wrong approach.

3.3 The law of total probability and Bayes' rule

We will now discuss two important rules that help probability computations by means of conditional probabilities. We introduce both of them in the next example.

Testing for mad cow disease

In early 2001 the European Commission introduced massive testing of cattle to determine infection with the transmissible form of *Bovine Spongiform Encephalopathy* (BSE) or "mad cow disease." As no test is 100% accurate, most tests have the problem of false positives and false negatives. A *false positive* means that according to the test the cow is infected, but in actuality it is not. A *false negative* means an infected cow is not detected by the test.

Imagine we test a cow. Let B denote the event "the cow has BSE" and T the event "the test comes up positive" (this is test jargon for: according to the test we should believe the cow is infected with BSE). One can "test the test" by analyzing samples from cows that are known to be infected or known to be healthy and so determine the effectiveness of the test. The European Commission had this done for four tests in 1999 (see [19]) and for several more later. The results for what the report calls Test A may be summarized as follows: an infected cow has a 70% chance of testing positive, and a healthy cow just 10%; in formulas:

$$\begin{aligned} \mathrm{P}(T \mid B) &= 0.70, \\ \mathrm{P}(T \mid B^c) &= 0.10. \end{aligned}$$

Suppose we want to determine the probability $\mathrm{P}(T)$ that an arbitrary cow tests positive. The tested cow is either infected or it is not: event T occurs in combination with B or with B^c (there are no other possibilities). In terms of events

$$T = (T \cap B) \cup (T \cap B^c),$$

so that

$$\mathrm{P}(T) = \mathrm{P}(T \cap B) + \mathrm{P}(T \cap B^c)\,,$$

because $T \cap B$ and $T \cap B^c$ are disjoint. Next, apply the multiplication rule (in such a way that the known conditional probabilities appear!):

$$\begin{aligned} \mathrm{P}(T \cap B) &= \mathrm{P}(T \mid B) \cdot \mathrm{P}(B) \\ \mathrm{P}(T \cap B^c) &= \mathrm{P}(T \mid B^c) \cdot \mathrm{P}(B^c) \end{aligned} \tag{3.1}$$

so that

$$\mathrm{P}(T) = \mathrm{P}(T \mid B) \cdot \mathrm{P}(B) + \mathrm{P}(T \mid B^c) \cdot \mathrm{P}(B^c)\,. \tag{3.2}$$

This is an application of the law of total probability: computing a probability through conditioning on several disjoint events that make up the whole sample space (in this case two). Suppose[1] $P(B) = 0.02$; then from the last equation we conclude: $P(T) = 0.02 \cdot 0.70 + (1 - 0.02) \cdot 0.10 = 0.112$.

QUICK EXERCISE 3.6 Calculate $P(T)$ when $P(T \mid B) = 0.99$ and $P(T \mid B^c) = 0.05$.

Following is a general statement of the law.

THE LAW OF TOTAL PROBABILITY. Suppose $C_1, C_2, \ldots, C_m$ are disjoint events such that $C_1 \cup C_2 \cup \cdots \cup C_m = \Omega$. The probability of an arbitrary event A can be expressed as:

$$P(A) = P(A \mid C_1)P(C_1) + P(A \mid C_2)P(C_2) + \cdots + P(A \mid C_m)P(C_m).$$

Figure 3.2 illustrates the law for $m = 5$. The event A is the disjoint union of $A \cap C_i$, for $i = 1, \ldots, 5$, so $P(A) = P(A \cap C_1) + \cdots + P(A \cap C_5)$, and for each i the multiplication rule states $P(A \cap C_i) = P(A \mid C_i) \cdot P(C_i)$.

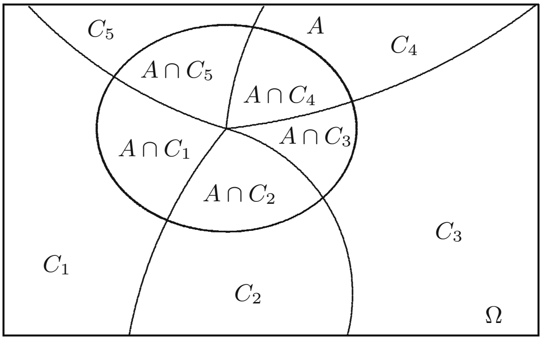


Fig. 3.2. The law of total probability (illustration for $m = 5$).

In the BSE example, we have just two mutually exclusive events: substitute $m = 2$, $C_1 = B$, $C_2 = B^c$, and $A = T$ to obtain (3.2).

Another, perhaps more pertinent, question about the BSE test is the following: suppose my cow tests positive; what is the probability it really has BSE? Translated, this asks for the value of $P(B \mid T)$. The information we were given is $P(T \mid B)$, a conditional probability, but the wrong one. We would like to switch T and B.

Start with the definition of conditional probability and then use equations (3.1) and (3.2):

[1] We choose this probability for the sake of the calculations that follow. The true value is unknown and varies from country to country. The BSE risk for the Netherlands for 2003 was estimated to be $P(B) \approx 0.000013$.

$$\mathrm{P}(B \,|\, T) = \frac{\mathrm{P}(T \cap B)}{\mathrm{P}(T)} = \frac{\mathrm{P}(T \,|\, B) \cdot \mathrm{P}(B)}{\mathrm{P}(T \,|\, B) \cdot \mathrm{P}(B) + \mathrm{P}(T \,|\, B^c) \cdot \mathrm{P}(B^c)}.$$

So with $\mathrm{P}(B) = 0.02$ we find

$$\mathrm{P}(B \,|\, T) = \frac{0.70 \cdot 0.02}{0.70 \cdot 0.02 + 0.10 \cdot (1 - 0.02)} = 0.125,$$

and by a similar calculation: $\mathrm{P}(B \,|\, T^c) = 0.0068$. These probabilities reflect that this Test A is not a very good test; a perfect test would result in $\mathrm{P}(B \,|\, T) = 1$ and $\mathrm{P}(B \,|\, T^c) = 0$. In Exercise 3.4 we redo this calculation, replacing $\mathrm{P}(B) = 0.02$ with a more realistic number.

What we have just seen is known as Bayes' rule, after the English clergyman Thomas Bayes who derived this in the 18th century. The general statement follows.

> BAYES' RULE. Suppose the events $C_1, C_2, \ldots, C_m$ are disjoint and $C_1 \cup C_2 \cup \cdots \cup C_m = \Omega$. The conditional probability of C_i, given an arbitrary event A, can be expressed as:
>
> $$\mathrm{P}(C_i \,|\, A) = \frac{\mathrm{P}(A \,|\, C_i) \cdot \mathrm{P}(C_i)}{\mathrm{P}(A \,|\, C_1)\mathrm{P}(C_1) + \mathrm{P}(A \,|\, C_2)\mathrm{P}(C_2) + \cdots + \mathrm{P}(A \,|\, C_m)\mathrm{P}(C_m)}.$$

This is the traditional form of Bayes' formula. It follows from

$$\mathrm{P}(C_i \,|\, A) = \frac{\mathrm{P}(A \,|\, C_i) \cdot \mathrm{P}(C_i)}{\mathrm{P}(A)} \tag{3.3}$$

in combination with the law of total probability applied to $\mathrm{P}(A)$ in the denominator. Purists would refer to (3.3) as Bayes' rule, and perhaps they are right.

QUICK EXERCISE 3.7 Calculate $\mathrm{P}(B \,|\, T)$ and $\mathrm{P}(B \,|\, T^c)$ if $\mathrm{P}(T \,|\, B) = 0.99$ and $\mathrm{P}(T \,|\, B^c) = 0.05$.

3.4 Independence

Consider three probabilities from the previous section:

$$\mathrm{P}(B) = 0.02,$$
$$\mathrm{P}(B \,|\, T) = 0.125,$$
$$\mathrm{P}(B \,|\, T^c) = 0.0068.$$

If we know nothing about a cow, we would say that there is a 2% chance it is infected. However, if we know it tested positive, we can say there is a 12.5%

chance the cow is infected. On the other hand, if it tested negative, there is only a 0.68% chance. We see that the two events are related in some way: the probability of B *depends* on whether T occurs.

Imagine the opposite: the test is useless. Whether the cow is infected is unrelated to the outcome of the test, and knowing the outcome of the test does not change our probability of B: $\mathrm{P}(B \mid T) = \mathrm{P}(B)$. In this case we would call B independent of T.

DEFINITION. An event A is called *independent of* B if

$$\mathrm{P}(A \mid B) = \mathrm{P}(A)\,.$$

From this simple definition many statements can be derived. For example, because $\mathrm{P}(A^c \mid B) = 1 - \mathrm{P}(A \mid B)$ and $1 - \mathrm{P}(A) = \mathrm{P}(A^c)$, we conclude:

$$A \text{ independent of } B \quad \Leftrightarrow \quad A^c \text{ independent of } B. \tag{3.4}$$

By application of the multiplication rule, if A is independent of B, then $\mathrm{P}(A \cap B) = \mathrm{P}(A \mid B)\mathrm{P}(B) = \mathrm{P}(A)\,\mathrm{P}(B)$. On the other hand, if $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$, then $\mathrm{P}(A \mid B) = \mathrm{P}(A)$ follows from the definition of independence. This shows:

$$A \text{ independent of } B \quad \Leftrightarrow \quad \mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)\,.$$

Finally, by definition of conditional probability, if A is independent of B, then

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

that is, B is independent of A. This works in reverse, too, so we have:

$$A \text{ independent of } B \quad \Leftrightarrow \quad B \text{ independent of } A. \tag{3.5}$$

This statement says that in fact, independence is *a mutual property*. Therefore, the expressions "A is independent of B" and "A and B are independent" are used interchangeably. From the three $\Leftrightarrow$-statements it follows that there are in fact 12 ways to show that A and B are independent; and if they are, there are 12 ways to use that.

INDEPENDENCE. To show that A and B are independent it suffices to prove *just one* of the following:

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

where A may be replaced by A^c and B replaced by B^c, or both. If one of these statements holds, *all* of them are true. If two events are not independent, they are called *dependent*.

Recall the birthday events L "born in a long month" and R "born in a month with the letter r." Let H be the event "born in the first half of the year," so $\mathrm{P}(H) = 1/2$. Also, $\mathrm{P}(H \mid R) = 1/2$. So H and R are independent, and we conclude, for example, $\mathrm{P}(R^c \mid H^c) = \mathrm{P}(R^c) = 1 - 8/12 = 1/3$.

We know that $\mathrm{P}(L \cap H) = 1/4$ and $\mathrm{P}(L) = 7/12$. Checking $1/2 \times 7/12 \neq 1/4$, you conclude that L and H are dependent.

QUICK EXERCISE 3.8 Derive the statement "R^c is independent of H^c" from "H is independent of R" using rules (3.4) and (3.5).

Since the words dependence and independence have several meanings, one sometimes uses the terms *stochastic* or *statistical* dependence and independence to avoid ambiguity.

Remark 3.1 (Physical and stochastic independence). Stochastic dependence or independence can sometimes be established by inspecting whether there is any physical dependence present. The following statements may be made.
If events have to do with processes or experiments that have no physical connection, they are always stochastically independent. If they are connected to the same physical process, then, as a rule, they are stochastically dependent, but stochastic independence is possible in exceptional cases. The events H and R are an example.

Independence of two or more events

When more than two events are involved we need a more elaborate definition of independence. The reason behind this is explained by an example following the definition.

INDEPENDENCE OF TWO OR MORE EVENTS. Events A_1, A_2, ..., A_m are called independent if

$$\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_m) = \mathrm{P}(A_1)\,\mathrm{P}(A_2) \cdots \mathrm{P}(A_m)$$

and this statement *also* holds when any number of the events A_1, ..., A_m are replaced by their complements throughout the formula.

You see that we need to check 2^m equations to establish the independence of m events. In fact, $m + 1$ of those equations are redundant, but we chose this version of the definition because it is easier.

The reason we need to do so much more checking to establish independence for multiple events is that there are subtle ways in which events may depend on each other. Consider the question:

> Is independence for three events A, B, and C the same as: A and B are independent; B and C are independent; and A and C are independent?

The answer is "No," as the following example shows. Perform two independent tosses of a coin. Let A be the event "heads on toss 1," B the event "heads on toss 2," and C "the two tosses are equal."

First, get the probabilities. Of course, $\mathrm{P}(A) = \mathrm{P}(B) = 1/2$, but also

$$\mathrm{P}(C) = \mathrm{P}(A \cap B) + \mathrm{P}(A^c \cap B^c) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

What about independence? Events A and B are independent by assumption, so check the independence of A and C. Given that the first toss is heads (A occurs), C occurs if and only if the second toss is heads as well (B occurs), so

$$\mathrm{P}(C \mid A) = \mathrm{P}(B \mid A) = \mathrm{P}(B) = \frac{1}{2} = \mathrm{P}(C)\,.$$

By symmetry, also $\mathrm{P}(C \mid B) = \mathrm{P}(C)$, so all pairs taken from A, B, C are independent: the three are called *pairwise independent*. Checking the full conditions for independence, we find, for example:

$$\mathrm{P}(A \cap B \cap C) = \mathrm{P}(A \cap B) = \frac{1}{4}, \quad \text{whereas} \quad \mathrm{P}(A)\,\mathrm{P}(B)\,\mathrm{P}(C) = \frac{1}{8},$$

and

$$\mathrm{P}(A \cap B \cap C^c) = \mathrm{P}(\emptyset) = 0, \quad \text{whereas} \quad \mathrm{P}(A)\,\mathrm{P}(B)\,\mathrm{P}(C^c) = \frac{1}{8}.$$

The reason for this is clear: whether C occurs follows deterministically from the outcomes of tosses 1 and 2.

3.5 Solutions to the quick exercises

3.1 $N = \{\text{May, Jun, Jul, Aug}\}$, $L = \{\text{Jan, Mar, May, Jul, Aug, Oct, Dec}\}$, and $N \cap L = \{\text{May, Jul, Aug}\}$. Three out of seven outcomes of L belong to N as well, so $\mathrm{P}(N \mid L) = 3/7$.

3.2 The event A is contained in C. So when A occurs, C also occurs; therefore $\mathrm{P}(C \mid A) = 1$.

Since $C^c = \{123, 321\}$ and $A \cup B = \{123, 321, 312, 213\}$, one can see that two of the four outcomes of $A \cup B$ belong to C^c as well, so $\mathrm{P}(C^c \mid A \cup B) = 1/2$.

3.3 Using the definition we find:

$$\mathrm{P}(A \mid C) + \mathrm{P}(A^c \mid C) = \frac{\mathrm{P}(A \cap C)}{\mathrm{P}(C)} + \frac{\mathrm{P}(A^c \cap C)}{\mathrm{P}(C)} = 1,$$

because C can be split into disjoint parts $A \cap C$ and $A^c \cap C$ and therefore

$$\mathrm{P}(A \cap C) + \mathrm{P}(A^c \cap C) = \mathrm{P}(C)\,.$$

3.4 This asks for the probability that the particle stays more than 3 seconds, given that it does not stay longer than 4 seconds, so 4 or less. From the definition:

$$\mathrm{P}(R_3 \,|\, R_4^c) = \frac{\mathrm{P}(R_3 \cap R_4^c)}{\mathrm{P}(R_4^c)}.$$

The event $R_3 \cap R_4^c$ describes: longer than 3 but not longer than 4 seconds. Furthermore, R_3 is the disjoint union of the events $R_3 \cap R_4^c$ and $R_3 \cap R_4 = R_4$, so $\mathrm{P}(R_3 \cap R_4^c) = \mathrm{P}(R_3) - \mathrm{P}(R_4) = \mathrm{e}^{-3} - \mathrm{e}^{-4}$. Using the complement rule: $\mathrm{P}(R_4^c) = 1 - \mathrm{P}(R_4) = 1 - \mathrm{e}^{-4}$. Together:

$$\mathrm{P}(R_3 \,|\, R_4^c) = \frac{\mathrm{e}^{-3} - \mathrm{e}^{-4}}{1 - \mathrm{e}^{-4}} = \frac{0.0315}{0.9817} = 0.0321.$$

3.5 Instead of a calendar of 365 days, we have one with just 12 months. Let C_n be the event n arbitrary persons have different months of birth. Then

$$\mathrm{P}(C_3) = \left(1 - \frac{2}{12}\right) \cdot \left(1 - \frac{1}{12}\right) = \frac{55}{72} = 0.7639$$

and it is no surprise that this is much smaller than $\mathrm{P}(B_3)$. The general formula is

$$\mathrm{P}(C_n) = \left(1 - \frac{n-1}{12}\right) \cdots \left(1 - \frac{2}{12}\right) \cdot \left(1 - \frac{1}{12}\right).$$

Note that it is correct even if n is 13 or more, in which case $\mathrm{P}(C_n) = 0$.

3.6 Repeating the calculation we find:

$$\mathrm{P}(T \cap B) = 0.99 \cdot 0.02 = 0.0198$$
$$\mathrm{P}(T \cap B^c) = 0.05 \cdot 0.98 = 0.0490$$

so $\mathrm{P}(T) = \mathrm{P}(T \cap B) + \mathrm{P}(T \cap B^c) = 0.0198 + 0.0490 = 0.0688$.

3.7 In the solution to Quick exercise 3.5 we already found $\mathrm{P}(T \cap B) = 0.0198$ and $\mathrm{P}(T) = 0.0688$, so

$$\mathrm{P}(B \,|\, T) = \frac{\mathrm{P}(T \cap B)}{\mathrm{P}(T)} = \frac{0.0198}{0.0688} = 0.2878.$$

Further, $\mathrm{P}(T^c) = 1 - 0.0688 = 0.9312$ and $\mathrm{P}(T^c \,|\, B) = 1 - \mathrm{P}(T \,|\, B) = 0.01$. So, $\mathrm{P}(B \cap T^c) = 0.01 \cdot 0.02 = 0.0002$ and

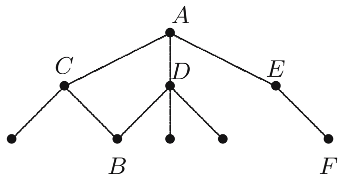
$$\mathrm{P}(B \,|\, T^c) = \frac{0.0002}{0.9312} = 0.00021.$$

3.8 It takes three steps of applying (3.4) and (3.5):

$$\begin{aligned}
H \text{ independent of } R \quad &\Leftrightarrow \quad H^c \text{ independent of } R \quad \text{by (3.4)}\\
H^c \text{ independent of } R \quad &\Leftrightarrow \quad R \text{ independent of } H^c \quad \text{by (3.5)}\\
R \text{ independent of } H^c \quad &\Leftrightarrow \quad R^c \text{ independent of } H^c \quad \text{by (3.4).}
\end{aligned}$$

3.6 Exercises

3.1 ⊞ Your lecturer wants to walk from A to B (see the map). To do so, he first randomly selects one of the paths to C, D, or E. Next he selects randomly one of the possible paths at that moment (so if he first selected the path to E, he can either select the path to A or the path to F), etc. What is the probability that he will reach B after two selections?



3.2 ⊞ A fair die is thrown twice. A is the event "sum of the throws equals 4," B is "at least one of the throws is a 3."

a. Calculate $\mathrm{P}(A \mid B)$.

b. Are A and B independent events?

3.3 ⊞ We draw two cards from a regular deck of 52. Let S_1 be the event "the first one is a spade," and S_2 "the second one is a spade."

a. Compute $\mathrm{P}(S_1)$, $\mathrm{P}(S_2 \mid S_1)$, and $\mathrm{P}(S_2 \mid S_1^c)$.

b. Compute $\mathrm{P}(S_2)$ by conditioning on whether the first card is a spade.

3.4 ⊡ A Dutch cow is tested for BSE, using Test A as described in Section 3.3, with $\mathrm{P}(T \mid B) = 0.70$ and $\mathrm{P}(T \mid B^c) = 0.10$. Assume that the BSE risk for the Netherlands is the same as in 2003, when it was estimated to be $\mathrm{P}(B) = 1.3 \cdot 10^{-5}$. Compute $\mathrm{P}(B \mid T)$ and $\mathrm{P}(B \mid T^c)$.

3.5 A ball is drawn at random from an urn containing one red and one white ball. If the white ball is drawn, it is put back into the urn. If the red ball is drawn, it is returned to the urn together with two more red balls. Then a second draw is made. What is the probability a red ball was drawn on *both* the first and the second draws?

3.6 We choose a month of the year, in such a manner that each month has the same probability. Find out whether the following events are independent:

a. the events "outcome is an even numbered month" (i.e., February, April, June, etc.) and "outcome is in the first half of the year."

b. the events "outcome is an even numbered month" (i.e., February, April, June, etc.) and "outcome is a summer month" (i.e., June, July, August).

3.7 ⊞ Calculate

a. $P(A \cup B)$ if it is given that $P(A) = 1/3$ and $P(B \mid A^c) = 1/4$.

b. $P(B)$ if it is given that $P(A \cup B) = 2/3$ and $P(A^c \mid B^c) = 1/2$.

3.8 ⊞ Spaceman Spiff's spacecraft has a warning light that is supposed to switch on when the freem blasters are overheated. Let W be the event "the warning light is switched on" and F "the freem blasters are overheated." Suppose the probability of freem blaster overheating $P(F)$ is 0.1, that the light is switched on when they actually *are* overheated is 0.99, and that there is a 2% chance that it comes on when nothing is wrong: $P(W \mid F^c) = 0.02$.

a. Determine the probability that the warning light is switched on.

b. Determine the conditional probability that the freem blasters are overheated, given that the warning light is on.

3.9 ⊡ A certain grapefruit variety is grown in two regions in southern Spain. Both areas get infested from time to time with parasites that damage the crop. Let A be the event that region R_1 is infested with parasites and B that region R_2 is infested. Suppose $P(A) = 3/4$, $P(B) = 2/5$ and $P(A \cup B) = 4/5$. If the food inspection detects the parasite in a ship carrying grapefruits from R_1, what is the probability region R_2 is infested as well?

3.10 A student takes a multiple-choice exam. Suppose for each question he either knows the answer or gambles and chooses an option at random. Further suppose that if he knows the answer, the probability of a correct answer is 1, and if he gambles this probability is 1/4. To pass, students need to answer at least 60% of the questions correctly. The student has "studied for a minimal pass," i.e., with probability 0.6 he knows the answer to a question. Given that he answers a question correctly, what is the probability that he actually *knows* the answer?

3.11 A breath analyzer, used by the police to test whether drivers exceed the legal limit set for the blood alcohol percentage while driving, is known to satisfy

$$P(A \mid B) = P(A^c \mid B^c) = p,$$

where A is the event "breath analyzer indicates that legal limit is exceeded" and B "driver's blood alcohol percentage exceeds legal limit." On Saturday night about 5% of the drivers are known to exceed the limit.

a. Describe in words the meaning of $P(B^c \mid A)$.

b. Determine $P(B^c \mid A)$ if $p = 0.95$.

c. How big should p be so that $P(B \mid A) = 0.9$?

3.12 The events A, B, and C satisfy: $P(A \mid B \cap C) = 1/4$, $P(B \mid C) = 1/3$, and $P(C) = 1/2$. Calculate $P(A^c \cap B \cap C)$.

3.13 In Exercise 2.12 we computed the probability of a "dream draw" in the UEFA playoffs lottery by counting outcomes. Recall that there were ten teams in the lottery, five considered "strong" and five considered "weak." Introduce events D_i, "the ith pair drawn is a dream combination," where a "dream combination" is a pair of a strong team with a weak team, and $i = 1, \ldots, 5$.

a. Compute $\mathrm{P}(D_1)$.

b. Compute $\mathrm{P}(D_2 \mid D_1)$ and $\mathrm{P}(D_1 \cap D_2)$.

c. Compute $\mathrm{P}(D_3 \mid D_1 \cap D_2)$ and $\mathrm{P}(D_1 \cap D_2 \cap D_3)$.

d. Continue the procedure to obtain the probability of a "dream draw": $\mathrm{P}(D_1 \cap \cdots \cap D_5)$.

3.14 Recall the Monty Hall problem from Section 1.3. Let R be the event "the prize is behind the door you chose initially," and W the event "you win the prize by switching doors."

a. Compute $\mathrm{P}(W \mid R)$ and $\mathrm{P}(W \mid R^c)$.

b. Compute $\mathrm{P}(W)$ using the law of total probability.

3.15 Two independent events A and B are given, and $\mathrm{P}(B \mid A \cup B) = 2/3$, $\mathrm{P}(A \mid B) = 1/2$. What is $\mathrm{P}(B)$?

3.16 You are diagnosed with an uncommon disease. You know that there only is a 1% chance of getting it. Use the letter D for the event "you have the disease" and T for "the test says so." It is known that the test is imperfect: $\mathrm{P}(T \mid D) = 0.98$ and $\mathrm{P}(T^c \mid D^c) = 0.95$.

a. Given that you test positive, what is the probability that you really *have* the disease?

b. You obtain a second opinion: an independent repetition of the test. You test positive again. Given this, what is the probability that you really *have* the disease?

3.17 You and I play a tennis match. It is deuce, which means if you win the next two rallies, you win the game; if I win both rallies, I win the game; if we each win one rally, it is deuce again. Suppose the outcome of a rally is independent of other rallies, and you win a rally with probability p. Let W be the event "you win the game," G "the game ends after the next two rallies," and D "it becomes deuce again."

a. Determine $\mathrm{P}(W \mid G)$.

b. Show that $\mathrm{P}(W) = p^2 + 2p(1-p)\mathrm{P}(W \mid D)$ and use $\mathrm{P}(W) = \mathrm{P}(W \mid D)$ (why is this so?) to determine $\mathrm{P}(W)$.

c. Explain why the answers are the same.

3.18 Suppose A and B are events with $0 < \mathrm{P}(A) < 1$ and $0 < \mathrm{P}(B) < 1$.

a. If A and B are disjoint, can they be independent?
b. If A and B are independent, can they be disjoint?
c. If $A \subset B$, can A and B be independent?
d. If A and B are independent, can A and $A \cup B$ be independent?