Chapter 4

Probability Theory

Probability theory is a branch of mathematics that is an essential component of statistics. It originally evolved from efforts to understand the odds and probabilities involved in games of chance, called classical probability theory (Weatherford 1982). The modern theory is developed from a small number of *a priori* axioms (like other mathematical theories) from which the rest the theory is deduced, including the behavior of probabilities and various rules for calculating them (Kolmogorov 1951, Weatherford 1982). While theoretical in origin, probability theory has proven to be spectacularly useful because it provides explanations for many natural processes, as well as the mathematical underpinnings for an enormous range of statistical procedures in the sciences.

4.1 Probability theory

4.1.1 Events

We can develop many elements of probability theory using a simple example, a single throw of a dice cube. If we throw the cube once, there are six possible outcomes corresponding to $1, 2, \ldots$, or 6 spots appearing on the cube. We call the possible outcomes of this single throw of a dice cube a **sample space** S, commonly written using set notation as

$$S = \{1, 2, 3, 4, 5, 6\} \tag{4.1}$$

We now define as **events** various subsets of the elements in S. **Simple events** contain exactly one element of S. For S defined above, the simple events are $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}$, and $\{6\}$. More specifically, the event $\{2\}$ signifies that a single throw of a dice cube showed two spots. More complex events contain more than one element of S. For example, consider the event A that the number of spots is odd, meaning that either one, three, or five spots showed on the dice cube after a single throw. We would write this event as the set $A = \{1,\ 3,\ 5\}$. Another possible event B is that the number of spots is less than or equal to three, or $B = \{1, 2, 3\}$. An event C such that number of spots is even would be written as $C = \{2, 4, 6\}$. Technically, both S itself and the empty set $\phi = \{\}$ are also possible events. S would always happen no matter the outcome of the throw, because some number of spots always occurs. The event corresponding to the empty set $\phi = \{\}$ would never happen because some number of spots always occurs after a throw. Sometimes certain events are subsets of other ones. For example, the event A defined above is a subset of S because every element of A is contained in S. This is written as $A \subset S$ using set notation.

4.1.2 Union, intersection, and complement of events

We now consider various combinations of events, again using our dice example. The **union** of two events A and B is defined to be the set containing all the simple events in A and B. The union is written using the notation $A \cup B$. For example, consider two of the events defined above for the dice example, $A = \{1, 3, 5\}$ (the number of spots is odd) and $B = \{1, 2, 3\}$ (the number of spots is less than or equal to three). We have

$$A \cup B = \{1, 3, 5\} \cup \{1, 2, 3\} = \{1, 2, 3, 5\}. \tag{4.2}$$

The union of two events can also be visualized using Venn diagrams, with the events A and B represented by circles and the shaded area their union (Fig. 4.1). The rectangle labeled S represents the entire sample space.

The **intersection** of two events A and B is defined to be the set containing simple events present in both A and B. The intersection is written using the notation $A \cap B$ or just AB. For example, consider the events A and B from the dice example. We have

$$A \cap B = \{1, 3, 5\} \cap \{1, 2, 3\} = \{1, 3\}. \tag{4.3}$$

The intersection of these two events is shown by the shaded area in Fig. 4.2. It is possible to have the intersection of two events be the empty set ϕ.

Consider the events A (spots is odd) and C (spots even) for the dice example. We have

$$A \cap C = \{1, 3, 5\} \cap \{2, 4, 6\} = \{\} = \phi. \tag{4.4}$$

Fig. 4.3 shows this outcome, with no shaded area because the intersection is empty. When the intersection of two events is the empty set, we say the two events are **mutually exclusive**. This means either one or the other event has occurred – it is impossible for them to happen at the same time.

The **complement** of an event A is the set of simple events in S remaining after we subtract those in A, typically written as A^c. For the event $A = \{1, 3, 5\}$ from the dice example, we have $A^c = \{2, 4, 6\}$, the simple events remaining in S after we substract those in A. Using set notation, $A^c = S - A = \{1, 2, 3, 4, 5, 6\} - \{1, 3, 5\} = \{2, 4, 6\}$. We can also represent A^c in a diagram with the shaded area representing all outcomes outside of A (Fig. 4.4). Complements of events frequently arise in the use of statistical tables.

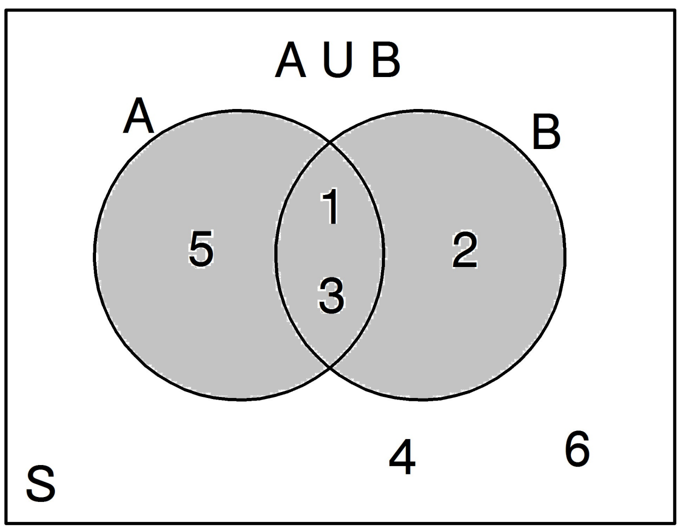

Figure 4.1: $A \cup B = \{1, 2, 3, 5\}$.

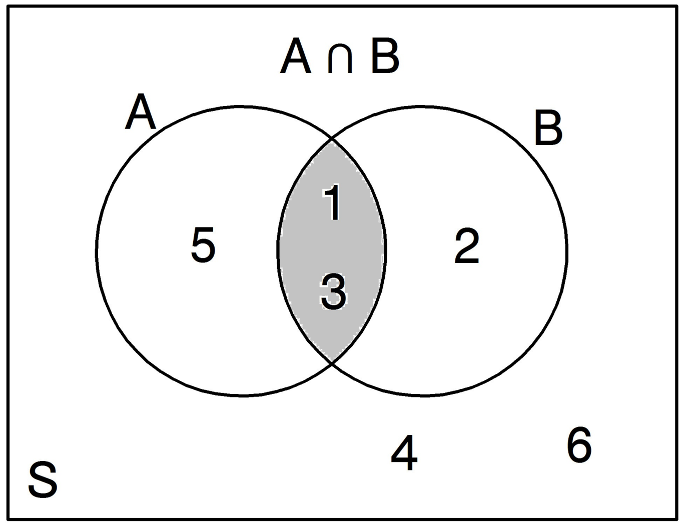

Figure 4.2: $A \cap B = \{1, 3\}$.

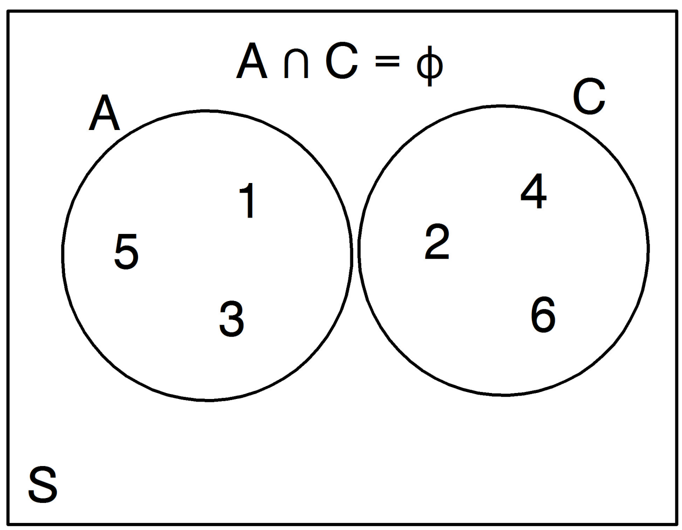

Figure 4.3: $A \cap C = \phi$.

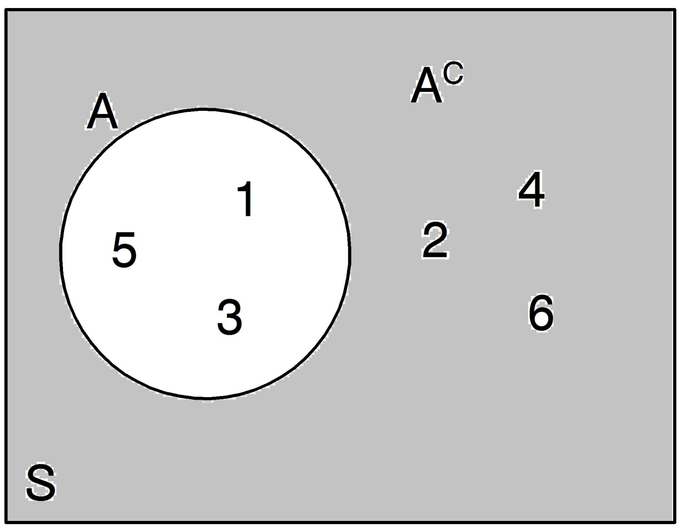

Figure 4.4: $A^C = \{2, 4, 6\}$.

4.1.3 Probability distributions

We now describe how probabilities are attached to events using a probability distribution, which can be mathematically defined based on certain axioms (Mood et al. 1974). Here, we simply list a number of properties of probability distributions that are useful to the practicing statistician. Let the symbol $P[A]$ stand for the probability of some event A. For a sample space S and any two events A and B, we have

1. $P[A] \geq 0$.
2. $P[S] = 1$.
3. $P[\phi] = 0$, where ϕ is the empty set.
4. $P[A \cup B] = P[A] + P[B] - P[A \cap B]$.
5. $P[A^c] = 1 - P[A]$.

While we have listed some of the properties of a probability distribution, we have not actually defined one yet. Recall the dice example, in which a single dice cube is thrown and the number of spots observed. If the cube is fair, then it is reasonable to assume that each number is equally likely to occur, and there are six possible numbers, so we assign a probability of 1/6 to each number. In particular, we have

$$P[\{1\}] = P[\{2\}] = P[\{3\}] = P[\{4\}] = P[\{5\}] = P[\{6\}] = 1/6 \qquad (4.5)$$

How should we assign probabilities to events like $A = \{1, 3, 5\}$? We define these events to have a probability equal to the sum of the probabilities for each simple event within them. For example, we have

$$\begin{aligned} P[A] &= P[\{1,3,5\}] = P[\{1\}] + P[\{3\}] + P[\{5\}] && (4.6) \\ &= 1/6 + 1/6 + 1/6 = 3/6 = 1/2. && (4.7) \end{aligned}$$

This result also makes intuitive sense for the event A, because we would expect the dice cube to produce an odd number of spots half of the time. We can view this probability distribution as a model of the dice cube's behavior, which would be accurate if the dice cube is fair. This is a common task faced by a statistician in analyzing a problem – determine an appropriate probability distribution to describe a particular type of data.

We will now calculate the probabilities for certain events to illustrate how this probability distribution can be used. Recall that the sample space for this distribution is $S = \{1, 2, 3, 4, 5, 6\}$. Suppose we have three events, namely $A = \{1, 3, 5\}$ (an odd number of spots), $B = \{1, 2, 3\}$ (less than or equal to three spots), and $C = \{2, 4, 6\}$ (an even number of spots).

We have already illustrated how to find the probability for A. For B, we have

$$P[B] = P[\{1, 2, 3\}] = P[\{1\}] + P[\{2\}] + P[\{3\}] \tag{4.8}$$
$$= 1/6 + 1/6 + 1/6 = 3/6 = 1/2. \tag{4.9}$$

For C the probability is

$$P[C] = P[\{2, 4, 6\}] = P[\{2\}] + P[\{4\}] + P[\{6\}] \tag{4.10}$$
$$= 1/6 + 1/6 + 1/6 = 3/6 = 1/2. \tag{4.11}$$

For the sample space S, which is also an event, we have

$$P[S] = P[\{1, 2, 3, 4, 5, 6\}] \tag{4.12}$$
$$= P[\{1\}] + P[\{2\}] + P[\{3\}] + P[\{4\}] + P[\{5\}] + P[\{6\}] \tag{4.13}$$
$$= 1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 6/6 = 1. \tag{4.14}$$

We also have $P[\{\}] = P[\phi] = 0$ because it is impossible to have no spots showing on the dice cube.

What is the probability for $A \cap B$? We have

$$P[A \cap B] = P[\{1, 3, 5\} \cap \{1, 2, 3\}] = P[\{1, 3\}] \tag{4.15}$$
$$= P[\{1\}] + P[\{3\}] = 1/6 + 1/6 = 1/3. \tag{4.16}$$

For $A \cup B$ we can calculate the probability in two ways. We can directly find it as follows. We have

$$P[A \cup B] = P[\{1, 3, 5\} \cup \{1, 2, 3\}] = P[\{1, 2, 3, 5\}] \tag{4.17}$$
$$= P[\{1\}] + P[\{2\}] + P[\{3\}] + P[\{5\}] \tag{4.18}$$
$$= 1/6 + 1/6 + 1/6 + 1/6 = 2/3. \tag{4.19}$$

We can also use the formula listed in Property 4 to find this probability. We have

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] \tag{4.20}$$
$$= 1/2 + 1/2 - 1/3 = 2/3. \tag{4.21}$$

We obtain the same answer as by direct calculation.

We can understand how the Property 4 formula works by considering the diagram for $A \cup B$ (Fig. 4.1). Suppose that the shaded area for event A represents the probability for A, and similarly for event B. If we add $P[A]$ and $P[B]$ together, this would actually be greater than $P[A \cup B]$ because it counts the area of the intersection ($A \cap B$) twice. This explains why we need to subtract $P[A \cap B]$ in Property 4 to obtain $P[A \cup B]$.

We now find the probability for A^c. We can directly calculate it by finding the probability for $A^c = S - A = \{1, 2, 3, 4, 5, 6\} - \{1, 3, 5\} = \{2, 4, 6\}$, so $P[A^c] = P[\{2, 4, 6\}] = 1/2$. Alternately, by Property 5 above,

$$P[A^c] = 1 - P[A] = 1 - 1/2 = 1/2. \tag{4.22}$$

Property 5 can also be explained by a diagram. The rectangle in Fig. 4.4 represents the sample space S, and by Property 2 we have $P[S] = 1$. If the circle for event A represents $P[A]$, then clearly $P[A^c] = 1 - P[A]$.

4.1.4 Probability spaces

The combination of a sample space S, a collection of all possible events on the sample space (A, B, S, ϕ, etc.), and a probability distribution is called a **probability space**.

4.1.5 Independence of events

Independence of events is an important concept in statistics, and basically implies that an event A has no effect on whether B occurs, and vice versa. In terms of probabilities, two events A and B are defined to be independent if

$$P[A \cap B] = P[A]P[B]. \tag{4.23}$$

Are the events $A = \{1, 3, 5\}$ and $B = \{1, 2, 3\}$ defined for the dice cube example independent? We have

$$P[A \cap B] = P[\{1, 3, 5\} \cap \{1, 2, 3\}] = P[\{1, 3\}] = 1/3. \tag{4.24}$$

However,

$$P[A]P[B] = 1/2 \times 1/2 = 1/4. \tag{4.25}$$

This implies that A and B are not independent because $P[A \cap B] \neq P[A]P[B]$. To see why this happens, observe that when the number of spots is less than or equal to three (B occurs), the number of spots is more likely to be odd (A occurs) because two of the three outcomes in B are odd.

We now work an example where the two events are independent. Suppose that $D = \{1, 2, 3, 4\}$, the event that the number of spots is less than or equal to four. Are A and D independent? We have

$$P[A \cap D] = P[\{1,3,5\} \cap \{1,2,3,4\}] \quad (4.26)$$
$$= P[\{1,3\}] = 1/6 + 1/6 = 1/3, \quad (4.27)$$

and

$$P[A]P[D] = 1/2 \times P[\{1,2,3,4\}] \quad (4.28)$$
$$= 1/2 \times (1/6 + 1/6 + 1/6 + 1/6) \quad (4.29)$$
$$= 1/2 \times 2/3 = 1/3. \quad (4.30)$$

This implies that A and D are independent because $P[A \cap D] = P[A]P[D]$. This outcome seems reasonable – when the number of spots is less than or equal to four (D occurs), the probability of the number of spots being odd is still equal to 1/2 because half of the outcomes in D are odd.

4.1.6 Conditional probability

Suppose that an event B has already happened, so that we have some information on a particular system or situation. Could this affect the probability that some other event A would occur? This is the idea behind **conditional probability**, an important concept in statistics that is related to independence. The conditional probability of an event A, given that B has occurred, is given by the formula

$$P[A|B] = \frac{P[A \cap B]}{P[B]}. \quad (4.31)$$

The notation '$A|B$' is read as A given B. For the dice cube example, what is the conditional probability of $A = \{1, 3, 5\}$ given that $B = \{1, 2, 3\}$ has occurred? We have

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{1/3}{1/2} = 2/3. \quad (4.32)$$

Note that the $P[A|B] > P[A]$ because $2/3 > 1/2$. Thus, if B has occurred it is more likely that A occurs, because two of three outcomes in B are odd.

If two events are independent, implying that $P[A \cap B] = P[A]P[B]$, then we have

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[A]P[B]}{P[B]} = P[A]. \qquad (4.33)$$

Thus, if two events are independent then the fact that B has occurred does not alter the probability for A. We can illustrate this for the dice cube example using the events $A = \{1, 3, 5\}$ and $D = \{1, 2, 3, 4\}$, which we earlier showed to be independent. We have

$$P[A|D] = \frac{P[A \cap D]}{P[D]} = \frac{P[A]P[D]}{P[D]} \qquad (4.34)$$

$$= \frac{1/3}{2/3} = 1/2 = P[A]. \qquad (4.35)$$

Thus, if D has occurred it has no effect on the probability of A occurring. This follows because half the events in D are odd, and so the probability of obtaining an odd number (1/2) is exactly the same as the original probability.

4.1.7 A biological probability distribution

We now examine a more biological example involving the infection of amphibians by the chytrid fungus *Batrachochytrium dendrobatidis*, which appears responsible for the decline of amphibians in some regions (Lips et al. 2006). Certain amphibian species appear less susceptible than others by virtue of natural immunity or their ecological traits (Lips et al. 2003), and we would expect infection rates to therefore vary among species. Suppose we know that at a particular location the amphibians can be classified into three common species (A, B, and C) that can also be divided into infected and uninfected individuals, with the frequency of individuals in each category having the distribution given in Table 4.1. In practice, we would need to estimate these proportions, but we will assume they are already known.

Table 4.1: Proportions of individuals from three amphibian species (A, B, and C), classified as infected (Yes) or free of chytrid fungus (No).

		Species		
		A	B	C
Infected	No	0.25	0.2	0.15
	Yes	0.25	0.1	0.05

Suppose we now sample a single individual from this location. The sample space would be $S = \{$A-Yes, A-No, B-Yes, B-No, C-Yes, C-No$\}$. Here 'A-No' stands for an amphibian of species A that is free of fungus, and is one of six simple events. The probability of sampling an A-No individual would be $P[\text{A-No}] = 0.25$, with the probabilities for other simple events given by the entries in Table 4.1. Note that $P[S] = P[\{$A-No, A-Yes, B-No, B-Yes, C-No, C-Yes$\} = 0.25 + 0.25 + 0.2 + 0.1 + 0.15 + 0.05 = 1$ as is necessary for a probability distribution.

We now calculate the probabilities for certain events. Suppose that A is the event that species A is sampled, implying that $A = \{$A-No, A-Yes$\}$. We have

$$P[A] = P[\{\text{A-No}, \text{A-Yes}\}] \quad (4.36)$$

$$= P[\{\text{A-No}\}] + \text{P}[\{\text{A-Yes}\}] \quad (4.37)$$

$$= 0.25 + 0.25 = 0.5 \quad (4.38)$$

Thus, we would expect half the amphibians sampled to be species A. Suppose we also want to find the probability for A^c. By Property 5 above, we have

$$P[A^c] = 1 - P[A] = 1 - 0.5 = 0.5 \quad (4.39)$$

Now let B be the event that species B is sampled, so that $B = \{$B-No, B-Yes$\}$ and $P[B] = P[\{\text{B-No}, \text{B-Yes}\}] = \text{P}[\{\text{B-No}\}] + \text{P}[\{\text{B-Yes}\}] = 0.2 + 0.1 = 0.3$. What is the probability for $A \cap B$? We see that A and B share no simple events, so $P[A \cap B] = P[\{\}] = P[\phi] = 0$. The two events are therefore mutually exclusive, which is not surprising because the sampled amphibian can only be species A or B, not both.

What happens for $A \cup B$? We can directly calculate this probability by finding the simple events in $A \cup B$. We have

$$\begin{aligned} P[A \cup B] &= P[\{\text{A-No}, \text{A-Yes}\} \cup \{\text{B-No}, \text{B-Yes}\}] && (4.40) \\ &= P[\{\text{A-No}, \text{A-Yes}, \text{B-No}, \text{B-Yes}\}] && (4.41) \\ &= P[\{\text{A-No}\}] + \text{P}[\{\text{A-Yes}\}] + \text{P}[\{\text{B-No}\}] + \text{P}[\{\text{B-Yes}\}] && (4.42) \\ &= 0.25 + 0.25 + 0.20 + 0.10 = 0.80. && (4.43) \end{aligned}$$

An alternate way to calculate this probability uses Property 4 listed above. In particular,

$$\begin{aligned} P[A \cup B] &= P[A] + P[B] - P[A \cap B] && (4.44) \\ &= 0.5 + 0.3 - 0 = 0.8, && (4.45) \end{aligned}$$

the same answer as before.

We now define an event I which stands for infected amphibians, meaning $I = \{\text{A-Yes, B-Yes, C-Yes}\}$. We have

$$\begin{aligned} P[I] &= P[\{\text{A-Yes}, \text{B-Yes}, \text{C-Yes}\}] && (4.46) \\ &= P[\{\text{A-Yes}\}] + \text{P}[\{\text{B-Yes}\}] + \text{P}[\{\text{C-Yes}\}] && (4.47) \\ &= 0.25 + 0.1 + 0.05 = 0.4 && (4.48) \end{aligned}$$

This means that the overall probability of sampling an infected animal is 0.4. Suppose that we already know the sampled amphibian is species C. What is the probability that it is infected given it is species C, or $P[I|C]$? We have

$$\begin{aligned} P[I|C] &= \frac{P[I \cap C]}{P[C]} = \frac{P[\{\text{A-Yes}, \text{B-Yes}, \text{C-Yes}\} \cap \{\text{C-No}, \text{C-Yes}\}]}{P[\{\text{C-No}, \text{C-Yes}\}]} && (4.49) \\ &= \frac{P[\{\text{C-Yes}\}]}{P[\{\text{C-No}, \text{C-Yes}\}]} && (4.50) \\ &= \frac{0.05}{0.2} = 0.25. && (4.51) \end{aligned}$$

Thus, if an individual of species C has been sampled the probability of it being infected is 0.25. We can also see this by examining the column for species C in Table 4.1, where the proportion of infected animals is $0.05/(0.15 + 0.05) = 0.25$.

4.1.8 Bayes theorem

Another use of conditional probability involves Bayes Theorem, named for the Reverend Thomas Bayes, an eighteenth century clergyman who first derived the theorem. The theorem is often used in the interpretation of medical tests as well as the field of Bayesian statistics (Ellison 1996).

Recall the example above involving amphibians and their infection by chytrid fungus. Let D be the event an amphibian actually has the disease while D^C implies they are disease-free. Now suppose a particular test is used to determine if a sampled amphibian has the disease. Let T be the event the amphibian tests positive for the disease, while T^c means the amphibian tests negative. The test is less than perfect, however, and sometimes gives a positive result when the amphibian is disease-free (a false positive) and a negative one when it is diseased (a false negative). What we would like to calculate is the probability that an amphibian actually has the disease given that it tests positive, or $P[D|T]$. This is called the **positive predictive value** of the test.

What is known for the test is the probability of testing positive for amphibians with the disease, $P[T|D]$, called the **sensitivity** of the test. This would be determined by testing a large number of amphibians that are known to have the disease by other means, and finding the proportion that test positive. Also known is the probability of testing negative for amphibians that are disease-free, $P[T^c|D^c]$, called the **specificity** of the test. We will also need an estimate of the probability that an amphibian has the disease in the population as a whole, $P[D]$, called the **prevalence** of the disease.

To find $P[D|T]$, we begin by using the definition of conditional probability:

$$P[D|T] = \frac{P[D \cap T]}{P[T]} = \frac{P[T \cap D]}{P[T]} \tag{4.52}$$

We can also write

$$P[T|D] = \frac{P[T \cap D]}{P[D]}, \tag{4.53}$$

which implies that $P[T \cap D] = P[T|D]P[D]$. Inserting this result into Eq. 4.52, we obtain

$$P[D|T] = \frac{P[T|D]P[D]}{P[T]}. \tag{4.54}$$

We are nearly there, except that we need to express $P[T]$ in terms of known quantities. The event T is made up of two mutually exclusive groups, amphibians that test positive and have the disease ($T \cap D$), and ones that test positive that are disease-free ($T \cap D^c$). From above, we have $P[T \cap D] = P[T|D]P[D]$ and can similarly show that $P[T \cap D^c] = P[T|D^c]P[D^c]$. Because the two groups are mutually exclusive, we can write $P[T]$ as the sum of the probabilities for each group:

$$P[T] = P[T \cap D] + P[T \cap D^c] = P[T|D]P[D] + P[T|D^c]P[D^c]. \tag{4.55}$$

Substituting this quantity into Eq. 4.52, we obtain Bayes' theorem:

$$P[D|T] = \frac{P[T|D]P[D]}{P[T|D]P[D] + P[T|D^c]P[D^c]}. \tag{4.56}$$

Because $P[T|D^c] = 1 - P[T^c|D^c]$ and $P[D^c] = 1 - P[D]$, we can also write Bayes' theorem as

$$P[D|T] = \frac{P[T|D]P[D]}{P[T|D]P[D] + (1 - P[T^c|D^c])(1 - P[D])}. \tag{4.57}$$

or

$$P[D|T] = \frac{\text{sensitivity} \times \text{prevalence}}{\text{sensitivity} \times \text{prevalence} + (1 - \text{specificity}) \times (1 - \text{prevalence})}. \tag{4.58}$$

We can thus express the theorem in terms of the sensitivity and specificity of the test, and the overall prevalence of the disease, which are known quantities.

Bayes theorem – sample calculation

Suppose that the test for amphibian disease has a high sensitivity ($P[T|D] = 0.95$) as well as a high specificity ($P[T^c|D^c] = 0.90$). A particular amphibian population has a fairly high prevalence of the disease ($P[D] = 0.25$, implying 25% are infected). What is the probability that an animal that tests positive from this population has the disease, $P[D|T]$? Inserting these quantities in Bayes theorem (Eq. 4.57 or 4.58), we obtain

$$P[D|T] = \frac{P[T|D]P[D]}{P[T|D]P[D] + (1 - P[T^c|D^c])(1 - P[D])} \tag{4.59}$$

$$= \frac{0.95 \times 0.25}{0.95 \times 0.25 + (1 - 0.9) \times (1 - 0.25)} \tag{4.60}$$

$$= \frac{0.2375}{0.2375 + 0.075} \tag{4.61}$$

$$= 0.76 \tag{4.62}$$

So, the probability that an animal that tests positive actually has the disease is 0.76. We now examine what happens if the prevalence of the disease is lower, say $P[D = 0.05]$, implying only 5% are infected. We have

$$P[D|T] = \frac{P[T|D]P[D]}{P[T|D]P[D] + (1 - P[T^c|D^c])(1 - P[D])} \tag{4.63}$$

$$= \frac{0.95 \times 0.05}{0.95 \times 0.05 + (1 - 0.9) \times (1 - 0.05)} \tag{4.64}$$

$$= \frac{0.0475}{0.0475 + 0.095} \tag{4.65}$$

$$= 0.3333 \tag{4.66}$$

Now the probability that the animal has the disease is only 0.3333, despite using exactly the same sensitivity and specificity values for the test. What has happened here?

The explanation is that when prevalence is low, the majority of positive test results are actually false positives, in which disease-free animals test positive. This is reflected in the denominator of Eq. 4.63, where the term 0.095 (the probability of testing positive and being disease-free) is actually larger than the term .0475 (the probability of testing positive and having the disease). To fix this problem it would be helpful to have a test with higher specificity to reduce the incidence of false positives.

Bayesian statistics

Another type of probability theory, called subjective or Bayesian probability theory, equates probability with a degree of belief on the part of the analyst (Weatherford 1982). This theory makes use of Bayes theorem but with a different interpretation of the probabilities. Suppose that $P[D]$ is the belief by an investigator that a particular animal has the disease before the test, rather than the prevalence (frequency of the disease) in the amphibian population. The value of $P[D|T]$ calculated using Bayes' theorem now represents the investigator's belief that the animal has the disease after observing a positive test result. These two probabilities are simple examples of the prior and posterior distributions used in Bayesian statistics. See Ellison et al. (1996) and Ellison (2004) for a summary of arguments for Bayesian statistics, which is based on this interpretation of probability as belief. Dennis (1996) and Lele and Dennis (2009) provide arguments against Bayesian statistics and in favor of ‘frequentist’ statistics, the kind of statistics based on the form of probability developed in this chapter. While frequentist statistics has its problems, it remains the most commonly used method in many fields and has a long successful record in science.

4.2 References

Dennis, B. (1996) Discussion: should ecologists become Bayesians? *Ecological Applications* 6: 1095-1103.

Ellison, A. M. (1996) An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecological Applications* 6: 1036-1046.

Ellison, A. M. (2004) Bayesian inference in ecology. *Ecology Letters* 7: 509-520.

Lele, S. R., and B. Dennis (2009) Bayesian methods for hierarchical models: Are ecologists making a Faustian bargain? *Ecological Applications* 19: 581-584.

Lips, K. R., F. Brem, R. Brenes, J. D. Reeve, R. A. Alford, J. Voyles, C. Carey, L. Livo, A. P. Pessier, and J. P. Collins (2006) Emerging infectious disease and the loss of biodiversity in a Neotropical amphibian community. *Proceedings of the National Academy of Sciences* 103: 3165-3170.

Lips, K. R., J. D. Reeve, and L. R. Witters (2003) Ecological traits predicting amphibian population declines in Central America. *Conservation Biology* 17: 1078-1088.

Kolmogorov, A. (1951) *Foundations of the Theory of Probability*. Chelsea, New York, NY.

Mood, A. M., Graybill, F. A. & Boes, D. C. (1974) *Introduction to the Theory of Statistics*. McGraw-Hill, Inc., New York, NY.

Mettlin, C., Littrup, P. J., Kane, R. A., Murphy, G. P, Lee, F., Chesley, A., Badalament, R. & Mostofi, F. K. (1994) Relative sensitivity and specificity of serum prostate specific antigen (PSA) level compared with age-referenced PSA, PSA density, and PSA change. *Cancer* 74: 1615-1620.

Weatherford, R. (1982) *Philosophical Foundations of Probability Theory*. Routledge & Kegan Paul Ltd., Boston, MA.

4.3 Problems

1. Suppose you have a loaded dice cube, such that $P[\{1\}] = 0.1, P[\{2\}] = 0.1, P[\{3\}] = 0.1, P[\{4\}] = 0.2, P[\{5\}] = 0.2$, and $P[\{6\}] = 0.3$. The cube is tossed a single time and the number of spots observed. Answer the following questions. Note that 'and' denotes an intersection of events, 'or' the union of events, and 'given' a conditional probability.

   (a) What is the probability that the number is even?

   (b) What is the probability that the number is odd and ≥ 3?

   (c) Are the events odd and ≥ 3 independent?

   (d) What is the probability that the number is odd or ≥ 3?

   (e) What is the probability that the number is odd, given that it is ≥ 3?

2. The PSA (prostate specific antigen) test is used to screen older men for prostate cancer. This test has a sensitivity of 0.90 and specificity of 0.719 (Mettlin et al. 1994). Assuming a prevalence of 0.1, find the probability that an individual with a positive test has cancer. Show your calculations.

3. Suppose you know that a particular animal population consists of 40% juveniles and 60% adults, and have a sample of two animals selected at random from the population.

   (a) What is the sample space for this scenario?

   (b) In a sample of two animals, what is the probability of obtaining two juveniles in a row?

   (c) What is the probability of obtaining one adult and one juvenile, in that order?

   (d) What is the probability of obtaining one juvenile and one adult, in that order?

   (e) What is the probability of obtaining two adults in a row?