Chapter 1

Sample Spaces and Probability

1.1 DISCRETE SAMPLE SPACES

Probability theory deals with situations in which there is an element of randomness or chance. Some models of the physical world are deterministic, that is, they predict exactly what will happen under certain circumstances. For example, if an object is dropped from a height and given no initial velocity, its distance, \( s \), from the starting point is given by

\[
s = \frac{1}{2} g \cdot t^2,
\]

where \( g \) is the acceleration due to gravity and \( t \) is the time. If one tried to apply the formula in a practical situation, one would not find very satisfactory results. The problem is that the formula applies only in a vacuum and ignores the shape of the object and the resistance of the air as well as other factors. Although some of these factors can be determined, we generally combine them and say that the result has a random or chance component. Our model then becomes

\[
s = \frac{1}{2} g \cdot t^2 + \varepsilon,
\]

where \( \varepsilon \) denotes the random component of the model. In contrast with the deterministic model, this model is stochastic.

Science often considers stochastic models; in formulating new models, the scientist may try to determine the contributions of both deterministic and random components of the model in predicting accurate results.

The mathematical theory of probability arose in consideration of games of chance, but, as the above-mentioned example shows, it is now widely used in far more practical and applied situations. We encounter other circumstances frequently in everyday life in which we presume that some random factors are at work. Here are some simple examples. What is the chance I will find that all eight traffic lights I pass through on my way to work are green? What are my chances for winning a lottery? I have a ten-volume encyclopedia that I have packed in separate boxes. If the boxes become mixed up and I draw the volumes out at random, what is the chance that my encyclopedia will be in order? My desk lamp has a bulb that is “guaranteed” to last 5000 hours. It has been used for 3000 hours. What is the chance that I must replace it before 2000 more hours are used? Each of these situations involves a random event whose specific outcome is unpredictable in advance.

Probability theory has become important because of the wide variety of practical problems it solves and its role in science. It is also the basis of the statistical analysis of data that is widely used in industry and in experimentation. Consider some examples. A manufacturer of television sets may know that 1% of the television sets manufactured have defects of some kind. What is the chance that a shipment of 200 sets a dealer has received contains 2% defective sets? Solving problems such as these has become important to manufacturers who are anxious to produce high quality products, and indeed such considerations play a central role in what has become known in manufacturing as statistical process control.

© 2015 John Wiley & Sons, Inc. Published 2015 by John Wiley & Sons, Inc.
Sample surveys, in which only a portion of a population or reference set is investigated, have become commonplace. A recent survey, for example, showed that two-thirds of welfare recipients in the United States were not old enough to vote. But surely we do not know that exactly two-thirds of all welfare recipients were not old enough to vote; there is some uncertainty, largely dependent on the size of the sample investigated as well as the manner in which the survey was conducted, connected with this result. How is this uncertainty calculated?

As a final example, consider a scientific investigation into say the relationship between temperature, a catalyst, and pressure in creating a chemical compound. A scientist can only carry out a few experiments in which several combinations of temperatures, amount of catalyst, and level of pressure are investigated. Furthermore, there is an element of randomness (largely due to other, unmeasured, factors) that influence the amount of compound produced. How is the scientist to determine which combination of factors maximizes the amount of chemical compound? We will encounter many of these examples in this book.

In some situations, we could measure all the forces involved and predict the outcome precisely but very often choose not to do so. In the traffic light example, we could, by knowledge of the timing of the lights, my speed, and the traffic pattern, predict precisely the color of each light as I approach it. While this is possible, it is probably not worth the effort, so we combine all the forces involved and call the result “chance.” So “chance” as we use it does not imply any new or unknown physical forces; it is simply an umbrella under which we put forces we choose not to measure.

How do we then measure the probability of events such as those described earlier? How do we determine how likely such events are? Such probability problems may be puzzling to us since we lack a framework in which to solve them. We lack a strategy for dealing with the randomness involved in these situations. A sensible way to begin is to consider all the possibilities that could occur. Such a list, or set, is called a sample space.

We begin here with some situations that are admittedly much simpler than some of those described earlier; more complex problems will also be encountered in this book. We will consider situations that we call experiments. These are situations that can be repeated under identical circumstances. Those of interest to us will involve some randomness so that the outcomes cannot be precisely predicted in advance. As examples, consider the following:

- Two people are chosen at random from a group of five people.
- Choose one of two brands of breakfast cereal at random.
- Throw two fair dice.
- Take an actuarial examination until it is passed for the first time.
- Any laboratory experiment.

Clearly, the first four of these experiments involve random factors. Laboratory experiments involve random factors as well and we would probably choose not to measure all the factors so as to be able to predict the exact outcome in advance.

Once the conditions for the experiment are set, and we are assured that these conditions can be repeated exactly, we can form the sample space, which we define as follows:

**Definition** A sample space is a set of all the possible outcomes from an experiment.
1.1 Discrete Sample Spaces

Example 1.1.1
The sample spaces for the first four experiments mentioned above are as follows:

(a) (Choose two people at random from a group of five people.) Denoting the five people as $A, B, C, D,$ and $E$, we find, if we disregard the order in which the persons are chosen, that there are ten possible samples of two people:

$$S = \{AB, AC, AD, AE, BC, BD, BE, CD, CE, DE\}.$$ 

This set, $S$, then comprises the sample space for the experiment.

If we consider the choice of people as random, we might expect that each of these ten samples occurs about 10% of the time. Further, we see that any particular person, say $B$, occurs in exactly four of the samples, so we say the probability that any particular person is in the sample is $\frac{4}{10} = \frac{2}{5}$. The reader may be interested to show that if three people were selected from a group of five people, then the probability a particular person is in the sample is $\frac{3}{5}$. Here, there is a pattern that we can establish with some results to be developed later in this chapter.

(b) (Choose one of two brands of breakfast cereal at random.) Denote the brands as $K$ and $P$. We take the sample space as

$$S = \{K, P\},$$

where the set $S$ contains each of the elementary outcomes, $K$ and $P$.

(c) (Toss two fair dice.) In contrast with the first two examples, we might consider several different sample spaces. Suppose first that we distinguish the two dice by color, say one is red and the other is green. Then we could write the result of a toss as an ordered pair indicating the outcome on each die, giving say the result on the red die first and the result on the green die second. Let a sample space be

$$S_1 = \{(1,1), (1,2), \ldots, (1,6), (2,1), (2,2), \ldots, (2,6), \ldots, (6,6)\}.$$ 

It is useful to see this sample space as a geometric space as in Figure 1.1.

Note that the 36 dots represent the only possible outcomes from the experiment. The sample space is not continuous in any sense in this case and may differ from our notions of a geometric space.

We could also describe all the possible outcomes from the experiment by the set

$$S_2 = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

since one of these sums must occur when the two dice are thrown.

<table>
<thead>
<tr>
<th>Second die</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

Figure 1.1 Sample space for tossing two dice.
Which sample space should be chosen? Note that each point in \( S_2 \) represents at least one point in \( S_1 \). So, while we might consider each of the 36 points in \( S_1 \) to occur with equal frequency if we threw the dice a large number of times, we would not consider that to be true if we chose sample space \( S_2 \). A sum of 7, for example, occurs on 6 of the points in \( S_1 \) while a sum of 2 occurs at only one point in \( S_1 \). The choice of sample space is largely dependent on what sort of outcomes are of interest when the experiment is performed. It is not uncommon for an experiment to admit more than one sample space. We generally select the sample space most convenient for the analysis of the probabilities involved in the problem.

We continue now with further examples of experiments involving randomness.

(d) (Take an actuarial examination until it is passed for the first time.) Letting \( P \) and \( F \) denote passing and failing the examination, respectively, we note that the sample space here is infinite:

\[
S = \{P, FP, FFP, FFFF, \ldots\}
\]

However, \( S \) here is a countably infinite sample space since its elements can be counted in the sense that they can be placed in a one-to-one correspondence with the set of natural numbers \( \{1, 2, 3, 4, \ldots\} \) as follows:

\[
\begin{align*}
P & \leftrightarrow 1 \\
FP & \leftrightarrow 2 \\
FFP & \leftrightarrow 3 \\
& \quad \cdot \\
& \quad \cdot \\
& \quad \cdot
\end{align*}
\]

The rule for the one-to-one correspondence is as follows: given an entry in the left column, the corresponding entry in the right column is the number of the attempt on which the examination is passed; given an entry in the right column, say \( n \), consider \( n - 1 \)’s followed by \( P \) to construct the corresponding entry in the left column. Hence, the correspondence with the set of natural numbers is one-to-one. Such sets are called countable or denumerable. We will consider countably infinite sets in much the same way that we will consider finite sets. In the next chapter, we will encounter infinite sets that are not countable.

(e) Sample spaces for laboratory experiments are usually difficult to enumerate and may involve a combination of finite and infinite factors.

Example 1.1.2

As a more difficult example, consider observing single births in a hospital until two girls are born in a row.

The sample space now is a bit more challenging to write down than the sample spaces for the situations considered in Example 1.1.1.
1.1 Discrete Sample Spaces

For convenience, we write the points, showing the births in order and grouped by the total number of births.

<table>
<thead>
<tr>
<th>Number of Births</th>
<th>Sample Points</th>
<th>Number of Sample Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>GG</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>BGG</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>BBGG</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>GBGG</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>BBBGG</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>BGBGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GBBGG</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>BBBBGG</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>BBGBGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BGBBGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GBBGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GBGBGG</td>
<td></td>
</tr>
</tbody>
</table>

and so on. We note that the number of sample points as we have grouped them follows the sequence 1, 1, 2, 4, 6, ..., which we recognize as the beginning of the Fibonacci sequence. The Fibonacci sequence is found by starting with the sequence 1, 1. Subsequent entries are found by adding the two immediately preceding entries. However, we only have evidence that the Fibonacci sequence applies to a few of the groups of points in the sample space. We will have to establish the general pattern in this example before concluding that the Fibonacci sequence does indeed give the number of sample points in the sample space. The reader may wish to do that before reading the following paragraphs!

Here is the reason the Fibonacci sequence occurs: consider a sequence of B's and G's in which GG occurs for the first time at the \( n \)th birth. Let \( a_n \) denote the number of ways in which this can occur. If GG occurs for the first time on the \( n \)th birth, there are two possibilities for the beginning of the sequence. These possibilities are mutually exclusive, that is, they cannot occur together.

One possibility is that the sequence begins with a B and is followed for the first time by the occurrence of GG in \( n - 1 \) births. Since we are requiring the sequence GG to occur for the first time at the \( n - 1 \)st birth, this can occur in \( a_{n-1} \) ways.

The other possibility for the beginning of the sequence is that the sequence begins with G, which must then be followed by B (else the pattern GG will occur in two births) and then the pattern GG occurs in \( n - 2 \) births. This can occur in \( a_{n-2} \) ways. Since the sequence begins either with B or G, it follows that

\[
a_n = a_{n-1} + a_{n-2}, \quad n \geq 4,
\]

where \( a_2 = a_3 = 1 \),

which describes the Fibonacci sequence.

The sequences for which GG occurs for the first time in 7 births can then be found by writing B followed by the sequences for 6 births and by writing GB followed by GG in 5 births:
Formulas such as (1.1) often describe a problem in a very succinct manner; they are called *recursions* because they describe one value of a function, here $a_n$, in terms of other values of the same function; in addition, they are easily programmed. Computer algebra systems are especially helpful in giving large number of terms determined by recursions. One can find, for example, that there are 46,368 ways for the sequence $GG$ to occur for the first time on the 25th birth. It is difficult to imagine determining this number without the use of a computer.

**EXERCISES 1.1**

1. Show the sample space when 3 people are selected from a group of 5 people. Verify the fact that any particular person in the selected group is 3/5.

2. In Example 1.1.2, show all the sample points where the births of two girls in a row occur in 8 or 9 births.

3. An experiment consists of drawing two numbered balls from a box of balls numbered from 1 to 9. Describe the sample space if
   (a) the first ball is not replaced before the second is drawn.
   (b) the first ball is replaced before the second is drawn.

4. In the diagram below, A, B, and C are switches that may be closed (current flows through the switch) or open (current cannot flow through the switch). Show the sample space indicating all the possible positions of the switches in the circuit.
1.2 Events; Axioms of Probability

5. Items being produced on an assembly line can be good (G) or not meeting specifications (N). Show the sample space for the next five items produced by the assembly line.

6. A student decides to take an actuarial examination until it is passed, but will attempt the test at most five times. Show the sample space.

7. In the World Series, games are played until one of the teams has won four games. Show all the points in the sample space in which the American League (A) wins the series over the National League (N) in at most six games.

8. We are interested in the sequence of male and female births in five-child families. Show the sample space.

9. Twelve chips numbered 1 through 12 are mixed in a bowl. Two chips are drawn successively and without replacement. Show the sample space for the experiment.

10. An assembly line is observed until items of both types—good (G) items and items not meeting specification (N)—are observed. Show the sample space.

11. Two numbers are chosen without replacement from the set \(\{2, 3, 4, 5, 6, 7\}\), with the additional restriction that the second number chosen must be smaller than the first. Describe an appropriate sample space for the experiment.

12. Computer chips coming off an assembly line are marked defective (D) or nondefective (N). The chips are tested and their condition listed. This is continued until two consecutive defectives are produced or until four chips have been tested, whichever occurs first. Show a sample space for the experiment.

13. A coin is tossed five times and a running count of the heads and tails is kept (so the number of heads and the number of tails tossed so far is recorded at each toss). Show all the sample points where the heads count always exceeds the tails count.

14. A sample space consists of all the linear arrangements of the integers 1, 2, 3, 4, and 5. (These linear arrangements are called permutations).

   (a) Use your computer algebra system to list all the sample points.

   (b) If the sample points are equally likely, what is the probability that the number 3 is in the third position?

   (c) What is the probability that none of the integers occupies its natural position?

1.2 EVENTS; AXIOMS OF PROBABILITY

After establishing a sample space, we are often interested in particular points, or sets of points, in that sample space. Consider the following examples:

(a) An item is selected at random from a production line. We are interested in the selection of a good item.

(b) Two dice are tossed. We are interested in the occurrence of a sum of 5.

(c) Births are observed until a girl is born. We are interested in this occurring in an even number of births.

Let us begin by defining an event.

**Definition** An event is a subset of a sample space.

Events then contain one or more elementary outcomes in the sample space.
Chapter 1  Sample Spaces and Probability

In the earlier examples, “a good item is selected,” “the sum is 5,” and “an even number of births was observed” can be described by subsets of the appropriate sample space and are, therefore, events.

We say that an event occurs if any of the elementary outcomes contained in the event occurs.

We will be interested in the relative frequency with which these events occur. In example (a), we would most likely say, if 99% of the items produced in the production line are good, then a good item will be selected about 99% of the time the experiment is performed, but we would expect some variation from this figure. In example (b), such a calculation is more complex since the event “the sum of the spots showing on the dice is 5” comprises several more elementary events. If the sample space distinguishing a red and a green die is

\[ S = \{(1, 1), (1, 2), \ldots, (1, 6), (2, 1), \ldots, (6, 6)\}, \]

then the points where the sum is 5 are

\[(1, 4), (2, 3), (3, 2), (4, 1).\]

If the dice are fair, then each of the 36 points in \( S \) occurs about \(1/36\) of the time, so we conclude that the sum of the spots showing 5 occurs about \(4 \cdot \frac{1}{36} = \frac{1}{9}\) of the time.

In example (c), observing births until a girl is born, the event “an even number of births is observed” is much more complex than examples (a) and (b) since there is an infinity of possibilities. How are we to judge the frequency of occurrence of each one? We cannot answer this question at this time, but we will consider it later.

Now we consider a structure so that we can deal with such questions, as well as many others far more complex than those considered so far. We start with some assumptions about any sample space.

**Axioms of Probability**

We consider the long-range relative frequency or probability of an event in a sample space. If we perform an experiment 120 times and an event, \( A \), occurs 30 times, then we say that the relative frequency of \( A \) is \(30/120 = 1/4\). In general, if in \( n \) trials an event \( A \) occurs \( n(A) \) times, then we say that the relative frequency of \( A \) is \(\frac{n(A)}{n}\). Of course, if we perform the experiment another \( n \) times, we do not expect \( A \) to occur exactly the same number of times as before, giving another relative frequency for the event \( A \). We do expect these variable ratios representing relative frequencies to settle down in some manner as \( n \) grows large. If \( A \) is an event, we denote this limiting relative frequency by the probability of \( A \) and denote this by \( P(A) \).

**Definition**  If \( A \) is an event, then the probability of \( A \) is

\[ P(A) = \lim_{n \to \infty} \frac{n(A)}{n}. \]

We assume at this point that the limit exists. We will discuss this in detail in Chapter 4.

In considering events, it is most convenient to use the language and notation of sets where the following notations are common:
1.2 Events; Axioms of Probability

The **union** of sets \( A \) and \( B \) is denoted by \( A \cup B \) where

\[
A \cup B = \{ x \mid x \in A \text{ or } x \in B \},
\]

where the word "or" is used in the inclusive sense, that is, an element in both sets \( A \) and \( B \) is included in the union of the sets.

The **intersection** of sets \( A \) and \( B \) is denoted by \( A \cap B \) where

\[
A \cap B = \{ x \mid x \in A \text{ and } x \in B \}.
\]

We will consider the following as axiomatic or self-evident:

1. \( P(A) \geq 0 \), where \( A \) is an event,
2. \( P(S) = 1 \), where \( S \) is the sample space, and
3. If \( A_1, A_2, \ldots \) are disjoint or mutually exclusive, that is, they have no sample points in common, then

\[
P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).
\]

Axioms of probability, of course, should reflect our common intuition about the occurrence of events. Since an event cannot occur with a negative relative frequency, (1) is evident. Since something must occur when the experiment is done and since \( S \) denotes the entire sample space, \( S \) must occur with relative frequency 1, hence assumption (2). Now suppose \( A \) and \( B \) are events with no sample points in common. We can illustrate events in a graphic manner by drawing a rectangle that represents all the points in \( S \); events are subsets of this sample space. A diagram showing the event \( A \), that is, the set of all elements of \( S \) that are in the event \( A \), is shown in Figure 1.2. Illustrations of sets and their relationships with each other are called **Venn diagrams**.

The event \( A \) or \( B \) consists of all points in \( A \) or in \( B \) and so its relative frequency is the sum of the relative frequencies of \( A \) or \( B \). This is assumption (3). Figure 1.3 shows a Venn diagram illustrating the disjoint events \( A \) and \( B \).

![Figure 1.2 Venn diagram showing the event A](image)

![Figure 1.3 Venn diagram showing disjoint events A and B](image)
1.3 PROBABILITY THEOREMS

In the above-mentioned example (b), we considered the event that the sum was 5 when two dice were thrown. This event in turn comprises elementary events

\[(1, 4), (2, 3), (3, 2), (4, 1)\]

each of which had probability \(\frac{1}{36}\). Since the events \((1, 4), (2, 3), (3, 2), \) and \((4, 1)\) are disjoint, axiom (3) shows that the probability of the event that the sum is 5 is the sum of the probabilities of these four elementary events or

\[\frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{4}{36} = \frac{1}{9}.\]

Assumption (3) shows that if \(A\) is an event that comprises elementary disjoint events \(a_1, a_2, a_3, ..., a_n\), then

**Theorem 1:**

\[P(A) = \sum_{i=1}^{n} P(a_i).\]

This fact is often used in the establishment of the theorems we consider in this section. Although we will not do so, all of them can be explained using Theorem 1.

What can we say about \(P(A \cup B)\) if \(A\) and \(B\) have sample points in common? If we find

\[P(A) + P(B)\]

we will have counted the points in the intersection \(A \cap B\) twice, as shown in Figure 1.4. So the intersection must be subtracted once giving

**Theorem 2:**

\[P(A \cup B) = P(A) + P(B) - P(A \cap B).\]

We call this the **addition theorem** (for two events).

**Example 1.3.1**

Choose a card from a well-shuffled deck of cards. Let \(A\) be the event “the selected card is a heart,” and let \(B\) be the event “the selected card is a face card.” Let the sample space \(S\)
1.3 Probability Theorems

consist of one point for each of the 52 cards. If the deck is really well shuffled, each point in \( S \) can be presumed to have probability \( \frac{1}{52} \). The event \( A \) contains 13 points and the event \( B \) contains 12 points, so \( P(A) = \frac{13}{52} \) and \( P(B) = \frac{12}{52} \). But the events \( A \) and \( B \) have three sample points in common, those for the King, Queen, and Jack of Hearts. The event \( A \cup B \) is then the event “the selected card is a Heart or a face card,” and its probability is

\[
P(A \cup B) = P(A) + P(B) - P(A \cap B)
\]

\[
= \frac{13}{52} + \frac{12}{52} - \frac{3}{52} = \frac{22}{52} = \frac{11}{26}.
\]

It is also easy to see by direct counting that the event “the selected card is a Heart or a face card” contains exactly 22 points in the sample space of 52 points.

How can the addition theorem for two events be extended to three or more events? First, consider events \( A, B, \) and \( C \) in a sample space \( S \). By adding and subtracting probabilities, the reader may be able to see that

**Theorem 3:**

\[
P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C),
\]

but we offer another proof as well. This proof will be based on the fact that a correct expression for \( P(A \cup B \cup C) \) must count each sample point in the event \( A \cup B \cup C \) once and only once. The Venn diagram in Figure 1.5 shows that \( S \) comprises 8 disjoint regions labeled as

0: points outside \( A \cup B \cup C \) (1 region)
1: points in \( A, B, \) or \( C \) alone (3 regions)
2: points in exactly two of the events (3 regions)
3: points in \( A \cap B \cap C \) (1 region).

![Figure 1.5 Venn diagram showing events, A, B, and C.](image)

Now we show that the right-hand side of Theorem 3 counts each point in the event \( A \cup B \cup C \) once and only once. By symmetry, we can consider only four cases:

Case 1. Suppose a point is in event \( A \) only. Then its probability is counted only once, in \( P(A) \), on the right-hand side of Theorem 3.

Case 2. Suppose a point is in \( A \cap B \) only. Then its probability is counted in \( P(A) \), \( P(B) \) and in \( P(A \cap B) \), a net count of one on the right-hand side in Theorem 3.
12 Chapter 1 Sample Spaces and Probability

Case 3. Suppose a point is in $A \cap B \cap C$. Then its probability is counted in each term on the right-hand side of Theorem 3, yielding a net count of 1.

Case 4. If a point is outside $A \cup B \cup C$, then it is not counted on the right-hand side in Theorem 3.

So Theorem 3 must be correct since it counts each point in $A \cup B \cup C$ exactly once and never counts any point outside the event $A \cup B \cup C$. This proof uses a combinatorial principle, that of inclusion and exclusion, a principle used in other ways as well in the field of combinatorics. We will make some use of this principle in the remainder of the book. Theorem 2 is of course a special case of Theorem 3.

We would like to extend Theorem 3 to $n$ events, but this requires some combinatorial facts that will be developed later and so we postpone this extension until they are established.

Example 1.3.2

A card is again drawn from a well-shuffled deck. Consider the events

$A$: the card shows an even number (2, 4, 6, 8, or 10),

$B$: the card is a Heart, and

$C$: the card is black.

We use a sample space containing one point for each of the 52 cards in the deck.

Then $P(A) = \frac{20}{52}$, $P(B) = \frac{13}{52}$, $P(C) = \frac{26}{52}$, $P(A \cap B) = \frac{5}{52}$, $P(A \cap C) = \frac{10}{52}$, $P(B \cap C) = 0$,

and $P(A \cap B \cap C) = 0$, so by Theorem 3,

$$P(A \cup B \cup C) = \frac{20}{52} + \frac{13}{52} + \frac{26}{52} - \frac{5}{52} - \frac{10}{52} = \frac{44}{52} = \frac{11}{13}.$$  

We will show one more fact in this section. Consider $S$ and an event $A$ in $S$. Denoting the set of points where the event $A$ does not occur by $\overline{A}$, it is clear that the events $\overline{A}$ and $A$ are disjoint. So, by Theorem 2, $P(A \cup \overline{A}) = P(A) + P(\overline{A}) = 1$, which is most often written as

Theorem 4:

$$P(\overline{A}) = 1 - P(A).$$

Example 1.3.3

Throw a pair of fair dice. What is the probability that the dice show different numbers? Here, it is convenient to let $A$ be the event “the dice show different numbers.” Referring to the sample space shown in Figure 1.1, we compute $P(\overline{A})$ since

$$P(\overline{A}) = P(\text{the dice show the same numbers}) = \frac{6}{36} = \frac{1}{6}.$$  

So $P(A) = 1 - \frac{6}{36} = \frac{5}{6}$. 

This is easier than counting the 30 sample points out of 36 for which the dice show different numbers.

The theorems we have developed so far appear to be fairly simple; the difficulty arises in applying them.

**EXERCISES 1.3**

1. Verify the probabilities in Example 1.3.2 by specifying the relevant sample points.
2. A fair coin is tossed until a head appears. Find the probability this occurs in four or fewer tosses.
3. A fair coin is tossed five times. Find the probability of obtaining
   (a) exactly three heads.
   (b) at most three heads.
4. A manufacturer of pickup trucks is required to recall all the trucks manufactured in a given year for the repair of possible defects in the steering column and defects in the brake linings. Dealers have been notified that 3% of the trucks have defective steering only, and that 6% of the trucks have defective brake linings only. If 87% of the trucks have neither defect, what percentage of the trucks have both defects?
5. A hat contains tags numbered 1, 2, 3, 4, and 5. A tag is drawn from the hat and it is replaced, then a second tag is drawn. Assume that the points in the sample space are equally likely.
   (a) Show the sample space.
   (b) Find the probability that the number on the second tag exceeds the number on the first tag.
   (c) Find the probability that the first tag has a prime number and the second tag has an even number. The number 1 is not considered to be a prime number.
6. A fair coin is tossed four times.
   (a) Show a sample space for the experiment, showing each possible sequence of tosses.
   (b) Suppose the sample points are equally likely and that a running count is made of the number of heads and the number of tails tossed. What is the probability the heads count always exceeds the tails count?
   (c) If the last toss is a tail, what is the probability an even number of heads was tossed?
7. In a sample space of two events is it possible to have \( P(A) = 1/2, P(A \cap B) = 1/3 \) and \( P(B) = 1/4 \)?
8. If \( A \) and \( B \) are events in a sample space of two events, explain why \( P(A \cap \overline{B}) \geq P(A) - P(B) \).
9. In testing the water supply for various cities in a state for two kinds of impurities commonly found in water, it was found that 20% of the water supplies had neither sort of impurity, 40% had an impurity of type A, and 50% had an impurity of type B. If a city is chosen at random, what is the probability its water supply has exactly one type of impurity?
10. A die is loaded so that the probability a face turns up is proportional to the number on that face. If the die is thrown, what is the probability an even number occurs?
11. Show that \( P(A \cap B) = P(B) - P(A \cap B) \).
Chapter 1 Sample Spaces and Probability

12. (a) Explain why \( P(A \cup B) \leq P(A) + P(B) \).

(b) Explain why \( P(A \cup B \cup C) \leq P(A) + P(B) + P(C) \).

13. Find a formula for \( P(A \text{ or } B) \) using the word “or” in an exclusive sense: that is, \( A \) or \( B \) means that event \( A \) occurs or event \( B \) occurs, but not both.

14. The entering class in an engineering college has 34% who intend to major in Mechanical Engineering, 33% who indicate an interest in taking advanced courses in Mathematics as part of their major field of study, and 28% who intend to major in Electrical Engineering, while 23% have other interests. In addition, 59% are known to major in Mechanical Engineering or take advanced Mathematics while 51% intend to major in Electrical Engineering or take advanced Mathematics. Assuming that a student can major in only one field, what percent of the class intends to major in Mechanical Engineering or in Electrical Engineering, but shows no interest in advanced Mathematics?

1.4 CONDITIONAL PROBABILITY AND INDEPENDENCE

Example 1.4.1

Suppose a card is drawn from a well-shuffled deck of 52 cards. What is the probability that the card is a Jack? If the sample space consists of a point for each card in the deck, the answer to the question is \( \frac{4}{52} \) since there are four Jacks in the deck.

Now suppose the person choosing the card gives us some additional information. Specifically, suppose we are told that the drawn card is a face card. Now what is the probability that the card is a Jack? An appropriate sample space for the experiment becomes the set of 12 points consisting of all the possible face cards that could be selected:

\[ \{JH, QH, KH, JD, QD, KD, JS, QS, KS, JC, QC, KC\} \]

Considering each of these 12 outcomes to be equally likely, the probability the chosen card is a Jack is now \( \frac{4}{12} \). The given additional information that the card is a face card has altered the probability of the event in question. Generally, such additional information, or conditions, has the effect of changing the probability of an event as the conditions change. Specifically, the conditions often reduce the sample space and, hence, alter the probabilities on those points that satisfy the conditions.

Let us denote by

\[ A : \text{ the event “the chosen card is a Jack”} \]

and

\[ B : \text{ the event “the chosen card is a face card”} \]

Further, we will use the notation \( P(A|B) \) to denote the probability of the event \( A \), given that the event \( B \) has occurred. We call \( P(A|B) \) the conditional probability of \( A \) given \( B \).

In this example, we see that \( P(A|B) = \frac{4}{12} \).

Now we can establish a general result by reasoning as follows. Suppose the event \( B \) has occurred; while this reduces the sample space to those points in \( B \), we cannot presume that the probability of the set of points in \( B \) is 1. However, if the probability of each point in \( B \) is divided by \( P(B) \), then the set of points in \( B \) has probability 1 and can therefore serve as
1.4 Conditional Probability and Independence

A sample space. This division by a constant also preserves the relative probabilities of the points in the original sample space; if one point in the original sample space was $k$ times as probable as another, it is still $k$ times as probable as the other point in the new sample space. Clearly, $P(A|B)$ accounts for the points in $A \cap B$ in the new sample space. We have found that

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

where we have presumed of course that $P(B) \neq 0$.

In the earlier example, $P(A \cap B) = \frac{4}{52}$ and $P(B) = \frac{12}{52}$ so $P(A|B) = \frac{4}{12}$ as before.

In this example, $P(A \cap B)$ reduces to $P(A)$, but this will not always be the case. We can also write this result as

$$P(A \cap B) = P(B) \cdot P(A|B), \text{ or, interchanging } A \text{ and } B,$$

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

We call this result the multiplication theorem.

**Example 1.4.2**

A box of transistors has four good transistors mixed up with two bad transistors. A production worker, in order to sample the product, chooses two transistors at random, the first chosen transistor not being replaced before the second transistor is chosen. What is the probability that both transistors are good?

If the events are

$$A : \text{ the first transistor chosen is good}$$

and

$$B : \text{ the second transistor chosen is good},$$

then we want $P(A \cap B)$.

Now $P(A) = \frac{4}{6}$ while $P(B|A) = \frac{3}{5}$ since the box, after the first good transistor is drawn, contains five transistors, three of which are good transistors. So the probability that both chosen transistors are good is

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

$$P(A \cap B) = \frac{4}{6} \cdot \frac{3}{5} = \frac{2}{5}.$$

by the multiplication theorem.

**Example 1.4.3**

In the context of the earlier example, what is the probability the second transistor chosen is good?
We need \( P(B) \). Now \( B \) can occur in two mutually exclusive ways: the first transistor is good and the second transistor is also good, or the first transistor is bad and the second transistor is good. So,
\[
P(B) = P(A \cap B) + P(A^c \cap B) = P(A) \cdot P(B|A) + P(A^c) \cdot P(B|A^c)
\]
\[
P(B) = \frac{4}{6} \cdot \frac{3}{5} + \frac{2}{6} \cdot \frac{4}{5} = \frac{2}{3}.
\]
We used the fact in this example that 
\[
P(B) = P(A) \cdot P(B|A) + P(A^c) \cdot P(B|A^c)
\]
since \( B \) occurs when either \( A \) or \( A^c \) occurs.

This result can be generalized. Suppose the sample space consists of disjoint events so that
\[
S = A_1 \cup A_2 \cup \cdots \cup A_n,
\]
where \( A_i \) and \( A_j \) have no sample points in common if \( i \neq j, i, j = 1, 2, \ldots, n \).

Then if \( B \) is an event,
\[
P(B) = P((A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B))
\]
\[
= P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)
\]
\[
= P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \cdots + P(A_n) \cdot P(B|A_n).
\]

We have then

**Theorem: (Law of Total Probability):** If \( S = A_1 \cup A_2 \cup \cdots \cup A_n \) where \( A_i \) and \( A_j \) have no sample points in common if \( i \neq j, i, j = 1, 2, \ldots, n \), then if \( B \) is an event,
\[
P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \cdots + P(A_n) \cdot P(B|A_n)
\]
or
\[
P(B) = \sum_{i=1}^{n} P(A_i) \cdot P(B|A_i).
\]

**Example 1.4.4**

A supplier purchases 10% of its parts from factory A, 20% of its parts from factory B, and the remainder of its parts from factory C. Out of which, 3% of A’s parts are defective; 2% of B’s parts are defective, and 1/2% of C’s parts are defective. What is the probability a randomly selected part is defective?

Let \( P(A) \) denote the probability the part is from factory A and define \( P(B) \) and \( P(C) \) similarly. Let \( P(D) \) denote the probability an item is defective. Then, from the law of total probability,
\[
P(D) = P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C)
\]
so
\[
P(D) = (0.10) \cdot (0.03) + (0.20) \cdot (0.02) + (0.70) \cdot (0.005) = 0.0105.
\]
1.4 Conditional Probability and Independence

So 1.05% of the items are defective.

We will encounter other uses of the law of total probability in the following examples.

**Example 1.4.5**

Suppose, in the context of the previous example, we are given that the second chosen transistor is good. What is the probability the first was also good?

Using the events $A$ and $B$ in the previous example, we want to find $P(A|B)$. That is

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

From the previous example, $P(A \cap B) = \frac{4}{6} \cdot \frac{3}{5} = \frac{2}{5}$, and we found in Example 1.4.3 that $P(B) = \frac{2}{3}$, so

$$P(A|B) = \frac{3}{5}.$$ 

When the earlier results are combined, we see that

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\overline{A}) \cdot P(B|\overline{A})} \quad (1.2)$$

This result is sometimes known as Bayes’ theorem.

The theorem can easily be extended to three or more mutually disjoint events.

**Theorem (Bayes’ theorem):** If $S = A_1 \cup A_2 \cup \cdots \cup A_n$ where $A_i$ and $A_j$ have no sample points in common if $i \neq j$ then, if $B$ is an event,

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i) \cdot P(B|A_i)}{P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \cdots + P(A_n) \cdot P(B|A_n)}$$

and

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{i=1}^{n} P(A_i) \cdot P(B|A_i)}.$$ 

Rather than remember this result, it is useful to look at Bayes’ theorem in a geometric way; it is not nearly as difficult as it may appear. This will first be illustrated using the current example.

Draw a square of side 1; as shown in Figure 1.6, divide the horizontal axis proportional to $P(A)$ and $P(\overline{A})$—in this case (returning to the context of Example 1.4.5) in the proportions 4/6 to 2/6. Along the vertical axis the conditional probabilities are shown. The vertical axis shows $P(B|A) = \frac{3}{5}$ and $P(B|\overline{A}) = \frac{4}{5}$, respectively.

The shaded area above $P(A)$ then shows $P(A) \cdot P(B|A)$. The total shaded area then shows $P(B)P(B) = \frac{4}{6} \cdot \frac{3}{5} + \frac{2}{6} \cdot \frac{4}{5} = \frac{2}{3}$. The doubly shaded region is the proportion of the
Chapter 1 Sample Spaces and Probability

P(B|A) = \frac{3}{5}

P(B|A) = \frac{4}{5}

P(A) = \frac{4}{6}

P(\overline{A}) = \frac{2}{6}

Figure 1.6 Diagram for Example 1.4.5.

Figure 1.7 A geometric view of Bayes’ theorem.

shaded area arising from the occurrence of A, which is P(A|B). We see that this is

\frac{4}{6} \cdot \frac{3}{5} + \frac{2}{6} \cdot \frac{4}{5} = \frac{3}{5}

yielding the same result found using Bayes’ theorem.

Figure 1.7 shows a geometric view of the general situation.

Bayes’ theorem then simply involves the calculation of areas of rectangles.

Example 1.4.6

According to the New York Times (September 5, 1987), a test for the presence of the HIV virus exists that gives a positive result (indicating the virus) with certainty if a patient actually has the virus. However, associated with this test, as with most tests, there is a false positive rate, that is, the test will sometimes indicate the presence of the virus in patients actually free of the virus. This test has a false
positive rate of 1 in 20,000. So the test would appear to be very sensitive. Assuming now that 1 person in 10,000 is actually HIV positive, what proportion of patients for whom the test indicates HIV actually have the HIV virus? The answer may be surprising.

A picture (greatly exaggerated so that the relevant areas can be seen) is shown in Figure 1.8.

Define the events as

\[ A : \text{patient has AIDS,} \]
\[ T^+ : \text{test indicates patient has AIDS.} \]

Then \[ P(A) = 0.0001; P(T^+|A) = 1; P(T^+|\bar{A}) = \frac{1}{20,000} \] from the data given. We are interested in \[ P(A|T^+) \]. So, from Figure 1.8, we see that

\[
P(A|T^+) = \frac{(0.0001) \cdot 1}{(0.0001) \cdot 1 + (0.9999) \cdot \frac{1}{20,000}} \quad \text{or} \quad P(A|T^+) = \frac{20,000}{29,999}
\]

We could also of course apply Bayes’ theorem to find that

\[
P(A|T^+) = \frac{P(A \cap T^+)}{P(T^+)} = \frac{P(A \cap T^+)}{P(A \cap T^+) + P(\bar{A}) \cdot P(T^+|\bar{A})}
\]

\[
= \frac{P(A) \cdot P(T^+|A)}{P(A) \cdot P(T^+|A) + P(\bar{A}) \cdot P(T^+|\bar{A})}
\]

\[
= \frac{(0.0001) \cdot 1}{(0.0001) \cdot 1 + (0.9999) \cdot \frac{1}{20,000}}
\]

\[
= \frac{20,000}{29,999}
\]

giving the same result as that found using simple geometry.
At first glance, the test would appear to be very sensitive due to its small false positive rate, but only two-thirds of those people testing positive would actually have the virus, showing that widespread use of the test, while detecting many cases of HIV, would also falsely detect the virus in about one-third of the population who test positive. This risk may be unacceptably high.

A graph of $P(A|T^+)$ (shown in Figure 1.9) shows that this probability is highly dependent on $P(A)$.

The graph shows that $P(A|T^+)$ increases as $P(A)$ increases, and that $P(A|T^+)$ is very large even for small values of $P(A)$. For example, if we desire $P(A|T^+)$ to be $\geq 0.9$, then we must have $P(A) \geq 0.0045$.

The sensitivity of the test may incorrectly be associated with $P(T^+|A)$. The patient, however, is concerned with $P(A|T^+)$. This example shows how easy it is to confuse $P(A|T^+)$ with $P(T^+|A)$.

Let us generalize the HIV example to a more general medical test in this way: assume a test has a probability $p$ of indicating a disease among patients actually having the disease; assume also that the test indicates the presence of the disease with probability $1 - p$ among patients not having the disease. Finally, suppose the incidence rate of the disease is $r$.

If $T^+$ denotes that the test indicates the disease, and if $A$ denotes the occurrence of the disease, then

$$P(A|T^+) = \frac{r \cdot p}{r \cdot p + (1 - r) \cdot (1 - p)}.$$

For example, if $p = 0.95$ and $r = 0.005$ (indicating that the test is 95% accurate on both those who have the disease and those who do not, and that 5 patients out of 1000 actually have the disease), then $P(A|T^+) = 0.087156$. Since $P(\overline{A}|T^+) = 0.912844$, a positive result on the test appears to indicate the absence, not the presence of the disease!

This odd result is actually due to the small incidence rate of the disease. Figure 1.10 shows $P(A|T^+)$ as a function of $r$ assuming that $p = 0.95$. We see that $P(A|T^+)$ becomes quite large ($\geq 0.8$) for $r \geq 0.21$.

It is also interesting to see how $r$ and $p$, varied together, effect $P(A|T^+)$. The surface is shown in Figure 1.11. The surface shows that $P(A|T^+)$ is large when the test is sensitive, that is, when $P(T^+|A)$ is large, or when the incidence rate $r = P(A)$ is large. But there are also combinations of these values that give large values of $P(A|T^+)$ : one of these is $r = 0.2$ and $P(T^+|A) = 0.8$ for then $P(A|T^+) = 1/2$. 

![Figure 1.9](image-url)
1.4 Conditional Probability and Independence

Figure 1.10 $P(A|T^+)$ as a function of the incidence rate, $r$, if $p = 0.95$.

Figure 1.11 $P(A|T^+)$ as a function of $r$, the incidence rate, and $P(T^+|A)$.

Example 1.4.7

A game show contestant is shown three doors, one of which conceals a valuable prize, while the other two are empty. The contestant is allowed to choose one door. Regardless of the choice made, at least one (i.e., exactly one or perhaps both) of the remaining doors is empty. The show host opens one door to show it empty. The contestant is now given the opportunity to switch doors. Should the contestant switch?

The problem is often called the Monty Hall problem because of its origin on the television show “Let’s Make A Deal.” It has been written about extensively, possibly because of its nonintuitive answer and perhaps because people unwittingly change the problem in the course of thinking about it.

The contestant clearly has a probability of $1/3$ of choosing the prize if a random choice of the door is made. So a probability of $2/3$ rests with the remaining two doors. The fact that one door is opened and revealed empty does not change these probabilities; hence the contestant should switch and will gain the prize with probability $2/3$. 
Chapter 1 Sample Spaces and Probability

Some think that showing the empty door gives information. Actually, it does not since the contestant already knows that at least one of the doors is empty.

When the empty door is opened, the problem does not suddenly become a choice between two doors (one of which conceals the prize). This change in the problem ignores the fact that the game show host sometimes has a choice of one door to open and sometimes two. Persons changing the problem in this manner may think, incorrectly, that the probability of choosing the prize is now \(\frac{1}{2}\), indicating that switching may have no effect in the long run; the strategy in reality has a great effect on the probability of choosing the prize.

To analyze the situation, suppose that the contestant chooses the first door and the host opens the second door. Other possibilities are handled here by symmetry. Let \(A_i\), \(i = 1, 2, 3\) denote the event “the prize is behind door \(i\)” and \(D\) denote the event “door 2 is opened.” The condition here is then \(D\); we now calculate the probability the prize is behind door 3, that is, the probability the contestant will win if he switches. We assume that \(P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}\).

Then \(P(D|A_1) = \frac{1}{2}\), \(P(D|A_2) = 0\), and \(P(D|A_3) = 1\).

The situation is shown in Figure 1.12.

It is clear from the shaded area in Figure 1.12 that the probability the contestant wins if the first choice is switched to door 3 is

\[
P(A_3|D) = \frac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 1} = \frac{2}{3},
\]

which verifies our previous analysis.

This example illustrates that some events are highly dependent on others. We now turn our attention to events for which this is not so.

![Figure 1.12 Diagram for the Monty Hall problem.](image)
Independence

We have found that \( P(A \cap B) = P(A) \cdot P(B|A) \). Occasionally, the occurrence of \( A \) has no effect on the occurrence of \( B \) so that \( P(B|A) = P(B) \). If this is the case, we call \( A \) and \( B \) as independent events. When \( A \) and \( B \) are independent, we have \( P(A \cap B) = P(A) \cdot P(B) \).

**Definition** Events \( A \) and \( B \) are called independent events if \( P(A \cap B) = P(A) \cdot P(B) \).

If we draw cards from a deck, replacing each drawn card before the next card is drawn, then the events denoting the cards drawn are clearly independent since the deck is full before each drawing and each drawing occurs under exactly the same conditions. If the cards are not replaced, however, then the events are not independent.

For three events, say \( A, B, \) and \( C \), we define the events as independent if

\[
P(A \cap B) = P(A) \cdot P(B),
\]
\[
P(A \cap C) = P(A) \cdot P(C),
\]
\[
P(B \cap C) = P(B) \cdot P(C), \quad \text{and}
\]
\[
P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C). \tag{1.3}
\]

The first three of these conditions establishes that the events are independent in pairs, so we call events satisfying these three conditions as pairwise independent. Example 1.4.8 will show that events satisfying these three conditions may not satisfy the fourth condition so pairwise independence does not determine independence.

We also note that there is some confusion between independent events and mutually exclusive events. Often people speak of these as, “having no effect on each other,” but that is not a precise characterization in either case. Note that while mutually exclusive events cannot occur together, independent events must be able to occur together. To be specific, suppose that neither \( P(A) \) nor \( P(B) \) is 0, and that \( A \) and \( B \) are mutually exclusive. Then \( P(A \cap B) = 0 \neq P(A) \cdot P(B) \). Hence, \( A \) and \( B \) cannot be independent. So if \( A \) and \( B \) are mutually exclusive, then they cannot be independent. This is equivalent to the statement that if \( A \) and \( B \) are independent, then they cannot be mutually exclusive, but the reader may enjoy establishing this from first principles as well.

**Example 1.4.8**

This example shows that pairwise independent events are not necessarily independent.

A fair coin is tossed four times. Consider the events \( A \), the first coin shows a head; \( B \), the third coin shows a tail; and \( C \), there are equal numbers of heads and tails. Are these events independent?

Suppose the sample space consists of the 16 points showing the tosses of the coins in order. The sample space, indicating the events that occur at each point, is as follows:

<table>
<thead>
<tr>
<th>Point</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHHH</td>
<td>( A )</td>
</tr>
<tr>
<td>HHHT</td>
<td>( A )</td>
</tr>
<tr>
<td>HHTH</td>
<td>( A, B )</td>
</tr>
</tbody>
</table>
Then \( P(A) = 1/2 \) and \( P(B) = 1/2 \) while \( C \) consists of the 6 points with exactly two heads and two tails, so \( P(C) = 6/16 = 3/8 \).

Now \( P(A \cap B) = \frac{4}{16} = \frac{1}{4} = P(A) \cdot P(B) \); \( P(A \cap C) = \frac{3}{16} = P(A) \cdot P(C) \); and \( P(B \cap C) = \frac{3}{16} = P(B) \cdot P(C) \), so the events \( A, B, \) and \( C \) are pairwise independent.

Now \( A \cap B \cap C \) consists of the two points \( HTTH \) and \( HHTT \) with probability \( 2/16 = 1/8 \). Hence, \( P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C) \), so \( A, B, \) and \( C \) are not independent.

Formulas (1.3) also show that establishing only that \( P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C) \) is not sufficient to establish the independence of events \( A, B, \) and \( C \).

**EXERCISES 1.4**

1. In Example 1.4.6, verify \( P(A|T^+) \) and \( P(\overline{A}|T^+) \).

2. Example 1.4.8 defines the events \( D \): the first head occurs on an even numbered toss and \( E \): at least three heads occur. Are \( D \) and \( E \) independent?

3. Box I contains 4 green and 5 brown marbles. Box II contains 6 green and 8 brown marbles. A marble is chosen from Box I and placed in Box II, then a marble is drawn from Box II.

   (a) What is the probability the second marble chosen is green?

   (b) If the second marble chosen is green, what is the probability a brown marble was transferred?

4. A football team wins its weekly game with probability 0.7. Suppose the outcomes of games on 3 successive weekends are independent. What is the probability the number of wins exceeds the number of losses?

5. Three manufacturers of floppy disks, A, B, and C, produce 15%, 25%, and 60% of the floppy disks made, respectively. Manufacturer A produces 5% defective disks, manufacturer B produces 7% defective disks, and manufacturer C produces 4% defective disks.

   (a) What proportion of floppy disks are defective?
1.4 Conditional Probability and Independence

(b) If a floppy disk is found to be defective, what is the probability it came from manufacturer B?

6. A chest contains three drawers, each containing a coin. One coin is silver on both sides, one is gold on both sides, and the third is silver on one side and gold on the other side. A drawer is chosen at random and one face of the coin is shown to be silver. What is the probability that the other side is silver also?

7. If \( A \) and \( B \) are independent events in a sample space, show that

\[
P(A \cup B) = P(B) + P(A) - P(B) \cdot P(A) = P(A) + P(B) - P(A) \cdot P(B).
\]

8. In a sample space, events \( A \) and \( B \) are independent; events \( B \) and \( C \) are mutually exclusive, and \( A \) and \( C \) are independent. If \( P(A \cup B \cup C) = 0.9 \) while \( P(B) = 0.5 \) and \( P(C) = 0.3 \), find \( P(A) \).

9. If \( P(A \cup B) = 0.4 \) and \( P(A) = 0.3 \), find \( P(B) \) if

(a) \( A \) and \( B \) are independent.

(b) \( A \) and \( B \) are mutually exclusive.

10. A coin, loaded so that the probability it shows heads when tossed is \( \frac{3}{4} \), is tossed twice. Let the events \( A \), \( B \), and \( C \), be “first toss is heads,” “second toss is heads,” and “tosses show the same face,” respectively.

(a) Are the events \( A \) and \( B \) independent?

(b) Are the events \( A \) and \( B \cup C \) independent?

(c) Are the events \( A \), \( B \), and \( C \) independent?

11. Three missiles, whose probabilities of hitting a target are 0.7, 0.8, and 0.9, respectively, are fired at a target. Assuming independence, what is the probability the target is hit?

12. A student takes a driving test until it is passed. If the probability the test is passed on any attempt is \( \frac{4}{7} \) and if the attempts are independent, what is the probability the test is taken an even number of times?

13. (a) Let \( p \) be the probability of obtaining a 5 at least once in \( n \) independent tosses of a die. What is the least value of \( n \) so that \( p \geq 1/2 \)?

(b) Generalize the result in part (a): suppose an event has probability \( p \) of occurring at any one of \( n \) independent trials of an experiment. What is the least value of \( n \) so that the probability the event occurs at least once is \( \geq r \)?

(c) Graph the surface in part (b), showing \( n \) as a function of \( p \) and \( r \).

14. Box I contains 7 red and 3 black balls; Box II contains 4 red and 5 black balls. After a randomly selected ball is transferred from Box I to Box II, 2 balls are drawn from Box II without replacement. Given that the two balls are red, what is the probability a black ball was transferred?

15. In rolling a fair die, what is the probability of rolling 1 before rolling an even number?

16. (a) There is a fifty-fifty chance that firm A will bid for the construction of a bridge. Firm B submits a bid and the probability that it will get the job is \( \frac{2}{3} \), provided firm A does not bid; if firm A submits a bid, the probability firm B gets the job is \( \frac{1}{5} \). Firm B is awarded the job; what is the probability firm A did not bid?

(b) In part (a), suppose now that the probability firm B gets the job if firm A bids on the job is \( p \). Graph the probability that firm A did not bid given that B gets the job as a function of \( p \).
26 Chapter 1 Sample Spaces and Probability

(c) Generalize parts (a) and (b) further and suppose that the probability that B gets the job given that firm A bids on the job is \( r \). Graph the surface showing the probability firm A did not bid, given that firm B gets the job as a function of \( p \) and \( r \).

17. In a sample space, events \( A \) and \( B \) have probabilities \( P(A) = P(B) = \frac{1}{2} \), and \( P(A \cup B) = \frac{2}{3} \).

(a) Are \( A \) and \( B \) mutually exclusive?
(b) Are \( A \) and \( B \) independent?
(c) Calculate \( P(A \cap B) \).
(d) Calculate \( P(\overline{A} \cap B) \).

18. Suppose that events \( A \), \( B \), and \( C \) are independent with \( P(A) = \frac{1}{4} \), \( P(B) = \frac{1}{2} \), and \( P(A \cup B \cup C) = \frac{3}{4} \). Find \( P(C) \).

19. A fair coin is tossed until the same face occurs twice in a row, but it is tossed no more than four times. If the experiment is over no later than the third toss, what is the probability that it is over by the second toss?

20. A collection of 65 coins contains one with two heads; the remainder of the coins are fair. If a coin, selected at random from the collection, turns up heads six times in six tosses, what is the probability that it is the two-headed coin?

21. Three distinct methods, \( A \), \( B \), and \( C \), are available for teaching a certain industrial skill. The failure rates are 30\%, 20\%, and 10\%, respectively. However, due to costs, \( A \) is used twice as frequently as \( B \), which is used twice as frequently as \( C \).

(a) What is the overall failure rate in teaching the skill?
(b) A worker is taught the skill, but fails to learn it correctly. What is the probability he was taught by method \( A \)?

22. Sixty percent of new drivers have had driver education. During their first year of driving, drivers without driver education have a probability 0.08 of having an accident, but new drivers with driver education have a 0.05 probability of having an accident. What is the probability a new driver with no accidents during the first year had driver education?

23. Events \( A \), \( B \), and \( C \) have \( P(A) = 0.3 \), \( P(B) = 0.2 \), and \( P(C) = 0.4 \). Also \( A \) and \( B \) are mutually exclusive; \( A \) and \( C \) are independent and \( B \) and \( C \) are independent. Find the probability that exactly one of the events \( A \), \( B \), or \( C \) occurs.

24. A set consists of the six possible arrangements of the letters \( a, b, \) and \( c \), as well as the points \( (a, a, a), (b, b, b), \) and \( (c, c, c) \). Let \( A_k \) be the event “letter \( a \) is in position \( k \)” for \( k = 1, 2, 3 \). Show that the events \( A_k \) are pairwise independent, but that they are not independent.

25. Assume that the probability a first-born child is a boy is \( p \), and that the sex of subsequent children follows a chance mechanism so that the probability the next child is the same sex as the previous child is \( r \).

(a) Let \( P_n \) denote the probability that the \( n \)th child is a boy. Find \( P_i \), \( i = 1, 2, 3 \), in terms of \( p \) and \( r \).
(b) Are the events \( A_i \) : “the \( i \)th child is a boy”, \( i = 1, 2, 3 \) mutually independent?
(c) Find a value for \( r \) so that \( A_1 \) and \( A_2 \) are independent.

26. A message is coded into the binary symbols 0 and 1 and the message is sent over a communication channel. The probability a 0 is sent is 0.4 and the probability a 1 is
sent is 0.6. The channel, however, has a random error that changes a 1 to a 0 with probability 0.2 and changes a 0 to a 1 with probability 0.1.

(a) What is the probability a 0 is received?  
(b) If a 1 is received, what is the probability a 0 was sent?

27. (a) Hospital patients with a certain disease are known to recover with probability 1/2 if they do not receive a certain drug. The probability of recovery is 3/4 if the drug is used. Of 100 patients, 10 are selected to receive the drug. If a patient recovers, what is the probability the drug was used?  
(b) In part (a), let the probability the drug is used be $p$. Graph the probability the drug was used given the patient recovers as a function of $p$.  
(c) Find $p$ if the probability the drug was used given that the patient recovers is 1/2.

28. Two people each toss four fair coins. What is the probability they each throw the same number of heads?

29. In sample surveys, people may be asked questions which they regard as sensitive and so they may or may not answer them truthfully. An example might be, “Are you using illegal drugs?” If it is important to discover the real proportion of illegal drug users in the population, the following procedure often called a randomized response technique may be used.

The respondent is asked to flip a fair coin and not reveal the result to the questioner. If the result is heads, then the respondent answers the question, “Is your Social Security number even?” If the coin comes up tails, the respondent answers the sensitive question. Clearly the questioner cannot tell whether a response of “yes” is a consequence of illegal drug use or of an even Social Security number. Explain, however, how the results of such a survey to a large number of respondents can be used to find accurately the percentage of the respondents who are users of illegal drugs.

30. (a) The individual events in a series of independent events have probabilities

$$\frac{1}{2}, \left(\frac{1}{2}\right)^2, \left(\frac{1}{2}\right)^3, \ldots, \left(\frac{1}{2}\right)^n.$$  

Show the probability that at least one of the events occurs approaches 0.711 as $n \to \infty$.  
(b) Show, if the probabilities of the events are $1/3, (1/3)^2, (1/3)^3, \ldots, (1/3)^n$, that the probability at least one of the events occurs approaches 0.440 as $n \to \infty$.  
(c) Show, if the probabilities of the events are $p, p^2, p^3, \ldots, p^n$, that the probability at least one of the events occurs can be very well approximated by the function $1 - p - p^2 + p^5 + p^7$ for $1/11 \leq p \leq 1/2$.

31. (a) If events $A$ and $B$ are independent, show that

1. $\bar{A}$ and $B$ are independent.  
2. $A$ and $\bar{B}$ are independent.  
3. $\bar{A}$ and $\bar{B}$ are independent.  

(b) Show that events $A$ and $B$ are independent if and only if $P(A|B) = P(A|\bar{B})$.  

32. A lie detector is accurate 3/4 of the time. That is, if a person is telling the truth, the lie detector indicates he is telling the truth with probability 3/4 while if the person is lying, the lie detector indicates that he is lying with probability 3/4. Assume that a person taking the lie detector test is unable to influence its results and also assume that
95% of the people taking the test tell the truth. What is the probability that a person is lying if the lie detector indicates that he is lying?

1.5 SOME EXAMPLES

We now show two examples of probability problems that have interesting results which may counter intuition.

Example 1.5.1  \textit{(The Birthday Problem)}

This problem exists in many variations in the literature on probability and has been written about extensively. The basic problem is this: There are \( n \) people in a room; what is the probability that at least two of them have the same birthday?

Let \( A \) denote the event “at least two people have the same birthday”; we want to find \( P(A) \). It is easier in this case to calculate \( P(\overline{A}) \) (the probability the birthdays are all distinct) rather than \( P(A) \). To find \( P(\overline{A}) \), note that the first person can have any day as a birthday. The birthday of the next person cannot match that of the first person; this has probability \( \frac{364}{365} \); the birthday of the third person cannot match that of either of the first two people; this has probability \( \frac{363}{365} \), and so on. So, multiplying these conditional probabilities,

\[
P(\overline{A}) = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365 - (n - 1)}{365}.
\]

It is easy with a computer algebra system to calculate exact values for \( P(A) = 1 - P(\overline{A}) \) for various values of \( n \):

<table>
<thead>
<tr>
<th>( n )</th>
<th>( P(A) )</th>
<th>( n )</th>
<th>( P(A) )</th>
<th>( n )</th>
<th>( P(A) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.002740</td>
<td>18</td>
<td>0.346911</td>
<td>34</td>
<td>0.795317</td>
</tr>
<tr>
<td>3</td>
<td>0.008204</td>
<td>19</td>
<td>0.379119</td>
<td>35</td>
<td>0.814383</td>
</tr>
<tr>
<td>4</td>
<td>0.016356</td>
<td>20</td>
<td>0.411438</td>
<td>36</td>
<td>0.832182</td>
</tr>
<tr>
<td>5</td>
<td>0.027136</td>
<td>21</td>
<td>0.443688</td>
<td>37</td>
<td>0.848734</td>
</tr>
<tr>
<td>6</td>
<td>0.040462</td>
<td>22</td>
<td>0.475695</td>
<td>38</td>
<td>0.864068</td>
</tr>
<tr>
<td>7</td>
<td>0.056236</td>
<td>23</td>
<td>0.507297</td>
<td>39</td>
<td>0.878220</td>
</tr>
<tr>
<td>8</td>
<td>0.074335</td>
<td>24</td>
<td>0.538344</td>
<td>40</td>
<td>0.891232</td>
</tr>
<tr>
<td>9</td>
<td>0.094624</td>
<td>25</td>
<td>0.568700</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.116948</td>
<td>26</td>
<td>0.598241</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.141141</td>
<td>27</td>
<td>0.626859</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.167025</td>
<td>28</td>
<td>0.654461</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.194410</td>
<td>29</td>
<td>0.680969</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.223103</td>
<td>30</td>
<td>0.706316</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.252901</td>
<td>31</td>
<td>0.730455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.283604</td>
<td>32</td>
<td>0.753348</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.315008</td>
<td>33</td>
<td>0.774972</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We see that \( P(A) \) increases rather rapidly; it exceeds \( 1/2 \) for \( n = 23 \), a fact that surprises many, most people guessing that the value of \( n \) to make \( P(A) \geq 1/2 \) is much larger. In thinking about this, note that the problem says that \textit{any} two people in the room can share...
any birthday. If some specific date comes to mind, such as August 2, then, since the probability a particular person’s birthday is not August 2 is $\frac{364}{365}$, the probability that at least one person in a group of $n$ people has that specific birthday is

$$1 - \left(\frac{364}{365}\right)^n.$$

It is easy to solve this for some specific probability. We find, for example, that for this probability to equal $\frac{1}{2}$, $n = 253$ people are necessary.

We show a graph, in Figure 1.13, of $P(A)$ for $n = 1, 2, 3, ..., 40$. The graph indicates that $P(A)$ increases quite rapidly as $n$, the number of people, increases.

![Birthday problem](image)

Figure 1.13 The birthday problem as a function of $n$, the number of people in the group.

It would appear that $P(A)$ might be approximated by a polynomial function of $n$. To consider how such functions can be constructed would be a diversion now, so we will not discuss it. For now, we state that the least squares approximating function found by applying a principle known as least squares is

$$f(n) = -6.44778 \cdot 10^{-3} - 4.54359 \cdot 10^{-5} \cdot n + 1.51787 \cdot 10^{-3} \cdot n^2 - 2.40561 \cdot 10^{-5} \cdot n^3.$$

It can be shown that $f(n)$ fits $P(A)$ quite well in the range $2 \leq n \leq 40$. For example, if $n = 13, P(A) = 0.194410$ while $f(13) = 0.196630$; if $n = 27, P(A) = 0.626859$ while $f(27) = 0.625357$. A graph of $P(A)$ and the approximating function $f(n)$ is shown in Figure 1.14. The principle of least squares will be considered in Section 4.16.

### Example 1.5.2

How many people must be in a group so that the probability at least two of them have birthdays within at most one day of each other is at least $1/2$?

Suppose there are $n$ people in the group, and that $A$ represents the event “at least two people have birthdays within at most one day of each other.” If a person’s birthday is August
2. For example, then the second person’s birthday must not fall on August 1, 2, or 3, giving 362 choices for the second person’s birthday. The third person, however, has either 359 or 360 choices, depending on whether the second person’s birthday is August 4 or July 31 or some other day that has not previously been excluded from the possibilities. We give then an approximate solution as

\[ P(A) = \frac{365 \cdot 362 \cdot 359 \cdots (368 - 3n)}{365 \cdot 365 \cdots 365}. \]

We seek \( P(A) = 1 - P(\overline{A}) \). It is easy to make a table of values of \( n \) and \( P(A) \) with a computer algebra system.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( P(A) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.008219</td>
</tr>
<tr>
<td>3</td>
<td>0.024522</td>
</tr>
<tr>
<td>4</td>
<td>0.048575</td>
</tr>
<tr>
<td>5</td>
<td>0.079855</td>
</tr>
<tr>
<td>6</td>
<td>0.117669</td>
</tr>
<tr>
<td>7</td>
<td>0.161181</td>
</tr>
<tr>
<td>8</td>
<td>0.209442</td>
</tr>
<tr>
<td>9</td>
<td>0.261424</td>
</tr>
<tr>
<td>10</td>
<td>0.316058</td>
</tr>
<tr>
<td>11</td>
<td>0.372273</td>
</tr>
<tr>
<td>12</td>
<td>0.429026</td>
</tr>
<tr>
<td>13</td>
<td>0.485341</td>
</tr>
<tr>
<td>14</td>
<td>0.540332</td>
</tr>
<tr>
<td>15</td>
<td>0.593226</td>
</tr>
<tr>
<td>16</td>
<td>0.643376</td>
</tr>
</tbody>
</table>

So 14 people are sufficient to make the probability that at least two of the birthdays differ by at most one day exceed 1/2. In the previous example, we found that a group of 23 people was sufficient to make the probability that at least two of them shared the same
birthday to exceed 1/2. The probability is approximately 0.8915 that at least two of these people have birthdays that differ by at most one day.

**Example 1.5.3  (Mowing the Lawn)**

Jack and his daughter, Kaylyn, choose who will mow the lawn by a random process: Jack has one green and two red marbles in his pocket; two are selected at random. If the colors match, Jack mows the lawn, otherwise, Kaylyn mows the lawn. Is the game fair?

The sample space here is most easily shown by a diagram containing the colors of the marbles as vertices and the edges as the two marbles chosen. Assuming that the three possible samples are equally likely, then two of them lead to Kaylyn mowing the lawn, while Jack only mows it 1/3 of the time. If we mean by the word “fair” that each mows the lawn with probability 1/2, then the game is clearly unfair.

![Diagram of marbles](image)

Three marbles in the lawn mowing example.

If we are allowed to add marbles to Jack’s pocket, can the game be made fair? The reader might want to think about this before proceeding.

What if a green marble is added? Then the sample space becomes all the sides and diagonals of a square:

![Diagram of marbles with additional green marble](image)
Chapter 1 Sample Spaces and Probability

Four marbles in the lawn mowing example.

Although there are now six possible samples, four of them involve different colors while only two of them involve the same colors. So the probability that the colors differ is \( \frac{4}{6} = \frac{2}{3} \); the addition of the green marble has not altered the game at all! The reader will easily verify that the addition of a red marble, rather than a green marble, will produce a fair game.

The problem of course is that, while the number of red and green marbles is important, the relevant information is the number of sides and diagonals of the figure produced since these represent the samples chosen. If we wish to find other compositions of marbles in Jack’s pocket that make the game fair, we need to be able to count these sides and diagonals. We now show how to do this.

Consider a figure with \( n \) vertices, as shown in Figure 1.15.

In order to count the number of sides and diagonals, choose one of the \( n \) vertices. Now, to choose a side or diagonal, choose any of the other \( n - 1 \) vertices and join them. We have then \( n \cdot (n - 1) \) choices. Since it does not matter which vertex is chosen first, we have counted each side or diagonal twice. We conclude that there are \( \frac{n(n-1)}{2} \) sides and diagonals.

This is also called the number of combinations of \( n \) distinct objects chosen two at a time, which we denote by the symbol \( \binom{n}{2} \). So

\[
\binom{n}{2} = \frac{n \cdot (n - 1)}{2}.
\]

If the game is to be fair, and if we have \( r \) red and \( g \) green marbles, then \( \binom{r}{2} \) and \( \binom{g}{2} \) represent the number of sides and diagonals connecting two red or two green marbles, respectively. We want \( r \) and \( g \) so that the sum of these is \( \frac{1}{2} \) of the total number of sides and diagonals, that is, we want \( r \) and \( g \) so that

\[
\binom{r}{2} + \binom{g}{2} = \left( \frac{1}{2} \right) \cdot \binom{r+g}{2}.
\]

The reader can verify that \( r = 6, g = 3 \) will satisfy the above equation as will \( r = 10, g = 6 \). The reader may also enjoy trying to find a general pattern for \( r \) and \( g \) before reading problem 3 in Exercises 1.5.
1. In the birthday problem, verify the probability that, in a group of 23 people, the probability that at least two people have birthdays differing by at most 1 day is 0.8915.

2. In the birthday problem, verify that the values of \( f(n) \), the polynomial approximation to \( P(A) \), are correct for \( f(13) \) and \( f(27) \).

3. Show that the “mowing the lawn” game is fair if and only if \( r \) and \( g \), the number of red and green marbles, respectively, are consecutive triangular numbers. (The first few triangular numbers are \( 1, 1 + 2 = 3, 1 + 2 + 3 = 6, \ldots \))

4. A fair coin is tossed until a head appears or until six tails have been obtained.
   (a) What is the probability the experiment ends in an even number of tosses?
   (b) Answer part (a) if the coin has been loaded so as to show heads with probability \( p \).

5. Let \( P_r \) be the probability that among \( r \) people, at least two have the same birth month. Make a table of values of \( P_r \) for \( r = 2, 3, \ldots, 12 \). Plot a graph of \( P_r \) as a function of \( r \).

6. Two defective transistors become mixed up with two good ones. The four transistors are tested one at a time, without replacement, until all the defectives are identified. Find \( P_r \), the probability that the \( r \)th transistor tested will be the second defective, for \( r = 2, 3, 4 \).

7. A coin is tossed four times and the sequence of heads and tails is observed.
   (a) What is the probability that heads and tails occur equally often if the coin is fair and the tosses are independent?
   (b) Now suppose the coin is loaded so that \( P(H) = 1/3 \) and \( P(T) = 2/3 \) and that the tosses are independent. What is the probability that heads and tails occur equally often, given that the first toss is a head?

8. The following model is sometimes used to model the spread of a contagious disease. Suppose a box contains \( b \) black and \( r \) red marbles. A marble is drawn and \( c \) marbles of that color together with the drawn marble are replaced in the box before the next marble is drawn, so that infected persons infect others while immunity to the disease may also increase.
   (a) Find the probability that the first three marbles drawn are red.
   (b) Show that the probability of drawing a black on the second draw is the same as the probability of drawing a black on the first draw.
   (c) Show by induction that the probability the \( k \)th marble is black is the same as the probability of drawing a black on the first draw.

9. A set of 25 items contains five defective items. Items are sampled at random one at a time. What is the probability that the third and fourth defectives occur at the fifth and sixth sample draws if
   (a) the items are replaced after each is drawn?
   (b) the items are not replaced after each is drawn?

10. A biased coin has probability \( 3/8 \) of coming up heads. \( A \) and \( B \) toss this coin with \( A \) tossing first.
    (a) Show that the probability that \( A \) gets a head before \( B \) gets a tail is very close to \( 1/2 \).
    (b) How can the coin be loaded so as to make the probability in part (a) \( 1/2 \)?
1.6 RELIABILITY OF SYSTEMS

Mechanical and electrical systems are often composed of separate components which may or may not function independently. The space shuttle, for example, comprises hundreds of systems, each of which may have hundreds or thousands of components. The components are, of course, subject to possible failure and these failures in turn may cause individual systems to fail, and ultimately for the entire system to fail. We pause here to consider in some situations how the probability of failure of a component may influence the probability of failure of the system of which it is a part.

In general, we refer to the reliability, \( R(t) \), of a component as the probability the component will function properly, or survive, for a given period of time. If we denote the event “the component lasts at least \( t \) units of time” by \( T > t \), then

\[
R(t) = P(T > t),
\]

where \( t \) is fixed.

The reliability of the system depends on two factors: the reliability of its component parts as well as the manner in which they are connected. We will consider some systems in this section, which have few components and elementary patterns of connection.

We will presume that interest centers on the probability an entire system lasts a given period of time; we will calculate this as a function of the probabilities the components last for that amount of time. To do this, we repeatedly use the addition law and multiplication of probabilities.

**Series Systems**

If a system of two components functions only if both of the components function, then the components are connected in series. Such a system is shown in Figure 1.16.

Let \( p_A \) and \( p_B \) denote the reliabilities of the components \( A \) and \( B \), that is,

\[
p_A = P(A \text{ survives at least } t \text{ units of time}) \quad \text{and} \quad p_B = P(B \text{ survives at least } t \text{ units of time})
\]

for some fixed value \( t \).

If the components function independently, then the reliability of the system, say \( R \), is the product of the individual reliabilities so

\[
R = P(A \text{ survives at least } t \text{ units of time and } B \text{ survives at least } t \text{ units of time})
\]

\[
= P(A \text{ survives at least } t \text{ units of time}) \cdot P(B \text{ survives at least } t \text{ units of time}) \quad \text{so}
\]

\[
R = p_A \cdot p_B.
\]

**Figure 1.16** A series system of two components.
Parallel Systems

If a system of two components functions if either (or both) of the components function, then the components are connected in parallel. Such a system is shown in Figure 1.17.

One way to calculate the reliability of the system depends on the fact that at least one of the components must function properly for the given period of time so

\[ R = P(A \text{ or } B \text{ survives for a given period of time}) \]

so, by the addition law,

\[ R = p_A + p_B - p_A \cdot p_B. \]

It is also clear, if the system is to function, that not both of the components can fail so

\[ R = 1 - (1 - p_A) \cdot (1 - p_B). \]

These two expressions for \( R \) are equivalent.

Figure 1.18 shows the reliability of both series and parallel systems as a function of \( p_A \) and \( p_B \). The parallel system is always more reliable than the series system since, for the parallel system to function, at least one of the components must function, while the series system functions only if both components function simultaneously.

Series and parallel systems may be combined in fairly complex ways. We can calculate the reliability of the system from the formulas we have established.

Example 1.6.1

The reliability of the system shown in Figure 1.19 can be calculated by using the addition law and multiplication of probabilities.

The connection of components \( A \) and \( B \) in the top section can be replaced by a single component with reliability \( p_A \cdot p_B \). The parallel connection of switches \( C \) and \( D \) can be replaced by a single switch with reliability \( 1 - (1 - p_C) \cdot (1 - p_D) \). The reliability of the resulting parallel system is then

\[ 1 - (1 - p_A \cdot p_B) \cdot [1 - (1 - p_C) \cdot (1 - p_D)]. \]
Chapter 1 Sample Spaces and Probability

Figure 1.18 Reliability of series and parallel systems.

Figure 1.19 System for Example 1.6.1.

A graph of the surface generated, assuming $p_A = p_B$ and $p_C = p_D$, is shown in Figure 1.20.

A contour plot of a surface shows values of $p_A$ and $p_C$ for which the reliability takes on particular values. Figure 1.21 shows a contour plot of the surface for Example 1.6.1, with contours specified at levels 0.80, 0.85, 0.90, 0.95, 0.99, and 0.995 for the reliability. The contour plot shows that if either $p_A$ or $p_C$ is 1, then the reliability is 1. The next contour shows choices of $p_A$ and $p_C$ giving reliability 0.995. The surface indicates that the system is highly reliable if either of the components is highly reliable and that, otherwise, the reliability declines rapidly.
1.6 Reliability of Systems

Example 1.6.1

Figure 1.20  Reliability surface for Example 1.6.1.

Figure 1.21  Contour plot for the surface in Figure 1.20.
EXERCISES 1.6

1. In the diagram below, let $p_A, p_B,$ and $p_C$ be the reliabilities of the individual switches. Determine the reliability of the system if
   (a) at least one switch must function.
   (b) at least two switches must function.

2. Determine the reliability of the system shown below if the reliability of any of the individual components is $p$.

3. Find the reliability of the system shown below if $p_A = p_B$ and $p_C = p_D$. Then show the surface giving the reliability of the system as a function of $p_A$ and $p_C$ and draw a contour plot of the surface.
1.7 COUNTING TECHNIQUES

Occasionally, sample spaces are encountered for which the sample points are equally likely. If this is the case, and if the sample space $S$ contains $n$ points, then, since the total probability in the sample space is 1, each point has probability $1/n$. If we denote the mutually exclusive points in $A$ by $a_i$, $i = 1, 2, 3, ..., n$, then the probability of an event, $A$, is the sum of the probabilities of the sample points in $A$. That is,

$$P(A) = \sum_{a_i \in A} P(a_i) = \sum_{a_i \in A} \frac{1}{n} \quad \text{so}$$

$$P(A) = \frac{\text{Number of points in } A}{n} = \frac{\text{Number of points in } A}{\text{Number of points in } S}.$$ 

In order to consider problems leading to sample spaces with equally likely sample points, we pause to consider some techniques for counting sets of points. These techniques provide some challenging problems.

The reader is first cautioned here to beware of concluding that just because a sample space has $n$ points that each point has probability $1/n$. For example, an airplane journey is either safely completed or not. One hopes these do not each have probability 1/2!

The counting techniques considered here are based on two fundamental counting principles concerning mutually exclusive events $A$ and $B$:

Principle 1: If events $A$ and $B$ can occur in $n$ and $m$ ways, respectively, then $A$ and $B$ can occur together in $n \cdot m$ ways.

Principle 2: If events $A$ and $B$ can occur in $n$ and $m$ ways, respectively, then $A$ or $B$ (but not both) can occur in $n + m$ ways.

Principle 1 is easily established since $A$ can occur in $n$ ways and then must be followed by each way in which $B$ can occur. A tree diagram, shown in Figure 1.22, illustrates the result. Principle 2 simply uses the word “or” in an exclusive sense.
A linear arrangement of \( n \) distinct objects is called a \textit{permutation}. For example, three distinct objects, say \( A, B, \) and \( C \), can be arranged in six different ways: \( ABC, ACB, BAC, BCA, CAB, \) and \( CBA \). So there are six permutations of three distinct objects. To count these permutations for \( n \) distinct objects, we use Principle 1. We have \( n \) choices for the object in the first position; that object chosen, we have \( n-1 \) choices for the object in the second position. Principle 1 tells us that there are \( n \cdot (n-1) \) ways to fill the first two positions. Continuing, we have

\[
\prod_{k=1}^{n} k = n! = n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 3 \cdot 2 \cdot 1
\]

ways to arrange all \( n \) of the items. We call this expression \( n! \) and note, for example, that \( 3! = 3 \cdot 2 \cdot 1 = 6 \), verifying the number of permutations of \( A, B, \) and \( C \) above.

The values of \( n! \) increase very rapidly: \( 1! = 1, \ 2! = 2, \ 3! = 6, \ 4! = 24, \ 5! = 120 \) and \( 10! \) is over 3 million. If we are interested in the number of permutations of even a small set, we must be prepared to deal with immense quantities. For example, the cards in a deck of 52 cards can be arranged in \( 52! \approx 8.06 	imes 10^{67} \) different ways. The reader may be surprised to find out how long it would take to enumerate these, even at a rate of 10,000 different permutations per second. This consideration may also persuade us that shuffling a deck so that each of these orders is equally likely is extremely unlikely.

A fact that is useful is that

\[
\frac{n!}{n-r!} = n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-(r-1))
\]

If we wish to permute only \( r \), say, of the \( n \) distinct objects, this can be done in

\[
\frac{n!}{(n-r)!} = n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-(r-1)) \]

ways.

Multiplying and dividing by \( (n-r)! \) shows that we can permute \( r \) of the \( n \) distinct objects in

\[
\frac{n!}{(n-r)!} \]

So that this formula will work when \( r = n \), we define \( 0! = 1 \). If we wish to permute 5 cards chosen from a deck of 52, this can be done in

\[
52 \cdot 51 \cdot 50 \cdot 49 \cdot 48 = 311,875,200 \text{ ways.}
\]
We note, multiplying and dividing by $47!$, that this can also be written as $\frac{52!}{47!}$, a fact that will be useful later.

If a list of permutations is desired, then the reader is advised to do this using a computer algebra system. The $4! = 24$ permutations of the set \{a, b, c, d\} is shown by a computer algebra system to be:

\[
\begin{align*}
&abcd \quad cabd \quad abdc \quad cadb \\
&acbd \quad cbad \quad acdb \quad cbda \\
&adbc \quad cdab \quad adcb \quad cdba \\
&bacd \quad dabc \quad badc \quad dacb \\
&bcad \quad dbac \quad bcda \quad dbca \\
&bdac \quad dcab \quad bdca \quad dcba
\end{align*}
\]

If we regard these permutations as being equally likely and if we want to find the probability that a particular letter, say $b$, occupies its normal place, we can count the points for which that is true and find that there are six of them. So

\[
P(b \text{ is in second place}) = \frac{6}{24}.
\]

What if the number of letters is large? An easy way to think about the problem is as follows: $b$ is in its place, and if the set contains $n$ distinct letters, we can arrange the remaining $(n - 1)$ letters in $(n - 1)!$ ways. Since the entire set can be permuted in $n!$ ways, the probability that $b$, or any of the other particular letters, occupies its own place is $\frac{(n-1)!}{n!} = \frac{1}{n}$.

This raises the question of other letters also occupying their own position. If we arrange the letters entirely at random, what is the probability that at least one of the letters is in its own place? The problem has been posed in the literature in many different ways one of which is this: $n$ men enter a restaurant and each checks his hat; the hats become mixed up during the evening and are passed out at the end of the evening in an entirely random way. What is the probability that at least one man gets his own hat? Equivalently, if we assume some natural order for the cards in a deck, what, after thorough shuffling, is the probability that at least one of the cards is in its own position?

One way to solve the problem is to determine the number of derangements (where no object occupies its own place) of a set of objects. For the permutations of the set \{1, 2, 3, 4\}, we find the derangements are as follows:

\[
\begin{align*}
&2, 1, 4, 3 \\
&2, 3, 4, 1 \\
&2, 4, 1, 3 \\
&3, 1, 4, 2 \\
&3, 4, 1, 2 \\
&3, 4, 2, 1 \\
&4, 1, 2, 3 \\
&4, 3, 1, 2 \\
&4, 3, 2, 1
\end{align*}
\]
a total of 9 derangements in this case. It follows that the probability that at least one object occupies its own place is \(1 - \frac{9}{24} = \frac{15}{24} = 0.625\). Surely this is an awkward way to handle larger sets, such as the deck of 52 cards. Surprisingly, we will find that the probability that at least one card occupies its own place after a thorough shuffling of the deck is very close to the probability above for four objects! We will explain this when we return to this problem later in this section.

For now, consider arranging a set of objects when the objects are not all distinct. If we permute the elements in the set \(\{a, a, b, c\}\), we find there are 12 permutations:

\[
\begin{align*}
&\ a, a, b, c \\
&\ a, a, c, b \\
&\ a, b, a, c \\
&\ a, b, c, a \\
&\ a, c, a, b \\
&\ a, c, b, a \\
&\ b, a, a, c \\
&\ b, a, c, a \\
&\ b, c, a, a \\
&\ c, a, a, b \\
&\ c, a, b, a \\
&\ c, b, a, a
\end{align*}
\]

so the set of 24 permutations (where the objects were distinct) has been cut in half. We can arrive at this result by starting with the set \(\{a, a, b, c\}\). Let \(R\) denote the number of distinct permutations. If we then tag the \(a\)'s with subscripts, say as \(a_1\) and \(a_2\), then each of the permutations in the above list yields \(2!\) permutations with the subscripted \(a\)'s. Hence, \(2! \cdot R = 4!\) so \(R = \frac{4!}{2!} = 12\). We could do exactly the same procedure with any set. Consider, for example, the set \(\{a, a, b, b, b, c, c, c\}\). By subscripting the \(a\)'s, \(b\)'s, and \(c\)'s, respectively, and again letting \(R\) denote the number of distinct permutations, we conclude that

\[
2! \cdot 3! \cdot 4! \cdot R = 9!\ 
\]

so

\[
R = \frac{9!}{2! \cdot 3! \cdot 4!} = 1260.
\]

The example is perfectly typical of the general situation: if the set has \(n_1\) objects of one kind, \(n_2\) of another, and so on until we have, say \(n_k\) of the \(k\)th kind where \(\sum_{i=1}^{k} n_i = n\), then there are

\[
\frac{n!}{n_1! \cdot n_2! \cdots n_k!}
\]

distinct permutations of the \(n\) objects.
1.7 Counting Techniques

Example 1.7.1

In how many distinct ways can 10 A’s, 5 B’s, and 2 C’s be awarded to a class of 17 students?

Put the students in some order. Then each distinct permutation of the letters leads to a different assignment of the grades. So there are

\[
\frac{17!}{10! \cdot 5! \cdot 2!} = 408,408
\]

different ways to assign the grades.

We turn now to combinations, that is, the distinguishable sets or samples of objects that can be chosen from a set of \(n\) distinct objects, without regard for order. We denote these combinations of \(r\) objects chosen from \(n\) distinct objects by \(\binom{n}{r}\), which we read “\(n\) choose \(r\).” We have already seen that \(\binom{5}{2} = \frac{n(n-1)}{2}\) in Example 1.5.3. Now suppose we have a set of objects and we want to choose a subset or sample of size 3. To be specific, suppose there are four items: \(a, b, c, d\). It is easy to write down the four combinations of size 3: \(a, b, c\); \(a, b, d\); \(a, c, d\); and \(b, c, d\). However, if we were dealing with larger set, it might be very difficult to write down a complete list without a procedure in mind. As a suggestion, to create the samples of size 3, we could choose each of the samples of size 2 and then attach a third item. The resulting list is as follows:

\[
\begin{align*}
&\{a, b, c\} \\
&a, b, d \\
&b, c, a \\
&b, c, d \\
&a, b, d \\
&b, c, a \\
&b, c, d \\
&b, a, d \\
&b, c, a
\end{align*}
\]

Since we have 2 choices for the third item, the resulting list contains \(2 \cdot \binom{4}{3}\) items. But each of the combinations has occurred three times. Therefore,

\[
2 \cdot \binom{4}{2} = 3 \cdot \binom{4}{3}, \quad \text{so}
\]

\[
\binom{4}{3} = \frac{2 \cdot \frac{4}{2}}{3} = \frac{2 \cdot 6}{3} = 4.
\]

This would appear to be a difficult way to arrive at \(\binom{4}{3}\). The reasoning here, however, can easily be extended and therein lies its advantage. Suppose we have a set of \(n\) distinct items and we wish to choose a sample of size \(r\). If we choose all the possible samples of size \(r - 1\) and then attach one of the \(n - r + 1\) remaining items to each, the resulting list
Chapter 1 Sample Spaces and Probability

has \((n - r + 1) \cdot \binom{n}{r-1}\) items. But this counts each of the \(\binom{n}{r}\) combinations \(r\) times. So,

\[
(n - r + 1) \cdot \binom{n}{r-1} = r \cdot \binom{n}{r}
\]

or

\[
\binom{n}{r} = \frac{(n - r + 1) \cdot \binom{n}{r-1}}{r}.
\] (1.4)

This is a recurrence formula since it expresses some values of a function, here \(\binom{n}{r}\), in terms of other values of the same function. If we have a starting place, we can calculate any value of the function we want. Here, since \(\binom{n}{1} = n\), formula (1.4) shows that

\[
\binom{n}{2} = \frac{n - 2 + 1}{2} \cdot \binom{n}{1} = \frac{n \cdot (n - 1)}{2} = \frac{n!}{2! \cdot (n - 2)!}
\]

verifying our previous result. We continue to apply formula (1.4) to find

\[
\binom{n}{3} = \frac{n - 3 + 1}{3} \cdot \binom{n}{2} = \frac{n \cdot (n - 1) \cdot (n - 2)}{3 \cdot 2} = \frac{n \cdot (n - 1) \cdot (n - 2) \cdot (n - 3)!}{3! \cdot (n - 3)!}
\]

\[
\binom{n}{3} = \frac{n!}{3! \cdot (n - 3)!}.
\]

It can be concluded by an inductive proof that

\[
\binom{n}{r} = \frac{n!}{r! \cdot (n - r)!}, \quad r = 0, 1, ..., n
\]

using the recurrence formula above.

If we have a set of \(n\) distinct objects and \(r\) are chosen, then \(n - r\) objects must remain unchosen. Since each time the chosen set is altered, so is the unchosen set, it follows that

\[
\binom{n}{r} = \binom{n}{n-r}.
\]

The quantities \(\binom{n}{r}\) are often called binomial coefficients since they occur in the binomial expansion:

Binomial Theorem:

\[
(a + b)^n = \sum_{r=0}^{n} \binom{n}{r} \cdot a^{n-r} \cdot b^r = \sum_{r=0}^{n} \binom{n}{r} \cdot a^r \cdot b^{n-r}.
\] (1.5)

For example,

\[
(a + b)^5 = \binom{5}{0} a^5 + \binom{5}{1} a^4 b + \binom{5}{2} a^3 b^2 + \binom{5}{3} a^2 b^3 + \binom{5}{4} ab^4 + \binom{5}{5} b^5
\]

\[
= a^5 + 5a^4 b + 10a^3 b^2 + 10a^2 b^3 + 5ab^4 + b^5.
\]
Many interesting identities are known concerning the binomial coefficients. If \( a = 1 \) and \( b = 1 \) are substituted in formula (1.5), the result is

\[
(1 + 1)^n = 2^n = \sum_{r=0}^{n} \binom{n}{r} = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}.
\]

Each side of this result may be recognized as the number of possible subsets (including the null set) that can be chosen from a set of \( n \) distinct items.

If we differentiate (1.5) with respect to \( a \) and then let \( a = 1 \) and \( b = 1 \), the result is

\[
n \cdot 2^n = \sum_{r=0}^{n} \binom{n}{r} \cdot r = 0 \cdot \binom{n}{0} + 1 \cdot \binom{n}{1} + 2 \cdot \binom{n}{2} + \cdots + n \cdot \binom{n}{n}.
\]

We show one more fact concerning the binomial coefficients. Suppose we want to choose a committee of size \( r \) chosen from a group of \( n \) people, one of whom is Sam. Sam is a member of \( \binom{n-1}{r-1} \) committees and he is not a member of \( \binom{n-1}{r} \) committees, so, since we have exhausted the possibilities,

\[
\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}.
\]

This is also often known as Pascal’s identity since it occurs in Pascal’s triangle of binomial coefficients.

It is also necessary for us, although this may seem unnatural to the reader, to ascribe some meaning to a symbol such as \( \binom{-7}{3} \). Clearly, we cannot interpret this as the choice of 3 objects from −7 objects! The following definition, while including our previous interpretation of \( \binom{n}{r} \), allows us to extend its meaning as well.

**Definition:** \( \binom{n}{r} = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!} \) provided that \( r \) is a nonnegative integer.

Using the above definition, we have that \( \binom{-7}{3} = \frac{(-7)(-8)(-9)}{3!} = -84 \). We will need facts such as this in subsequent chapters.

Using this definition, the binomial theorem can also be used with negative exponents. For example,

\[
(a + b)^{-5} = a^{-5} + \binom{-5}{1} a^{-6} b + \binom{-5}{2} a^{-7} b^2 + \binom{-5}{3} a^{-8} b^3 + \cdots \quad \text{or}
\]

\[
(a + b)^{-5} = a^{-5} - \binom{5}{1} a^{-6} b + \binom{6}{2} a^{-7} b^2 - \binom{7}{3} a^{-8} b^3 + \cdots
\]

We now use some of the results found here in some examples.

**Example 1.7.2**

A box of manufactured items contains 8 items that are good and 3 that are not usable. What is the probability that a sample of 5 items contains exactly 1 unusable item?
Suppose that each of the samples has probability \( \frac{1}{\binom{11}{5}} \). There are \( \binom{8}{4} \) ways to choose the 4 good items and \( \binom{3}{1} \) ways to choose the unusable item. The multiplication principle then gives \( \binom{3}{1} \cdot \binom{8}{4} \) ways to choose exactly 1 unusable item. So the probability we seek is

\[
\binom{3}{1} \cdot \binom{8}{4} \cdot \frac{1}{\binom{11}{5}} = \frac{5}{11}.
\]

Finally, in this chapter, we consider the general addition law for \( n \) events, having established the addition law for two and for three events. So we seek to prove

**Theorem 4:** \( P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum P(A_i) - \sum P(A_i \cap A_j) + \sum P(A_i \cap A_j \cap A_k) - \cdots + \left( -1 \right)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n) \), where the sums are over all the distinct items in the summand, that is, where \( i > j > k > \cdots \).

**Proof** We again use the principle of inclusion and exclusion. Consider a point in \( A_1 \cup A_2 \cup \cdots \cup A_n \) which is in exactly \( k \) of the events \( A_i \). It will be convenient to renumber the \( A_i \)'s if necessary so that the point is in the first \( k \) of these events. We will now show that the right-hand side of Theorem 4 counts this point exactly once, showing the theorem to be correct.

The point is counted on the right-hand side of Theorem 4

\[
\binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \cdots \pm \binom{k}{k} \times \text{times.}
\]

But the binomial expansion of \( 0 = [1 + (-1)]^k = \sum_{i=0}^{k} \binom{k}{i} \cdot (-1)^i \) shows that

\[
\binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \cdots \pm \binom{k}{k} = \binom{k}{0} = 1,
\]

establishing the result.

**Example 1.7.3**

We return to the matching problem stated earlier in this section: If \( n \) integers are randomly arranged in a row, what is the probability that at least one of them occupies its own place? The general addition law can be used to provide the solution.

Let \( A_i \) denote the event, "number \( i \) is in the \( i \)th place." We seek \( P(A_1 \cup A_2 \cup \cdots \cup A_n) \).

Here \( P(A_i) = \frac{(n-1)!}{n!} \), since, after \( i \) is put in its own place, there are \( (n-1)! \) ways to arrange the remaining numbers; \( P(A_i \cap A_j) = \frac{(n-2)!}{n!} \), since if \( i \) and \( j \) occupy their own places we can permute the remaining \( n-2 \) objects in \( (n-2)! \) ways; and, in general, \( P(A_1 \cap A_2 \cap \cdots \cap A_k) = \frac{(n-k)!}{n!} \). Now we note that there are \( \binom{n}{1} \) choices for an individual number \( i \); there
are \( \binom{n}{2} \) choices for pairs of numbers \( i \) and \( j \); and, in general, there are \( \binom{n}{k} \) choices for \( k \) of the numbers. So, applying Theorem 4,

\[
P(A_1 \cup A_2 \cup \cdots \cup A_n) = \binom{n}{1} \cdot \frac{(n - 1)!}{n!} - \binom{n}{2} \cdot \frac{(n - 2)!}{n!} + \binom{n}{3} \cdot \frac{(n - 3)!}{n!} - \cdots \pm \binom{n}{n} \cdot \frac{(n - n)!}{n!}.
\]

This simplifies to

\[
P(A_1 \cup A_2 \cup \cdots \cup A_n) = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots \pm \frac{1}{n!}.
\]

A table of values of this expression is shown below.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000000</td>
</tr>
<tr>
<td>2</td>
<td>0.500000</td>
</tr>
<tr>
<td>3</td>
<td>0.666667</td>
</tr>
<tr>
<td>4</td>
<td>0.625000</td>
</tr>
<tr>
<td>5</td>
<td>0.633333</td>
</tr>
<tr>
<td>6</td>
<td>0.631944</td>
</tr>
<tr>
<td>7</td>
<td>0.632118</td>
</tr>
<tr>
<td>8</td>
<td>0.632121</td>
</tr>
</tbody>
</table>

To six decimal places, the probability that at least one number is in its natural position remains at 0.632121 for \( n \geq 9 \). An explanation for this comes from a series expansion for \( e^x \):

\[
e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots.
\]

So

\[
e^{-1} = 1 - 1 + \frac{(-1)^2}{2!} + \frac{(-1)^3}{3!} + \frac{(-1)^4}{4!} + \cdots \text{ or } e^{-1} = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} + \cdots.
\]

So we see that \( \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots \pm \frac{1}{n!} \) approaches \( 1 - \frac{1}{e} = 0.632120559 \ldots \) This is our first, but certainly not our last, encounter with \( e \) in a probability problem. This also explains why we remarked that the probability at least one card in a shuffled deck of 52 cards was in its natural position differed little from that for a deck consisting of only 9 cards.

We turn now to some examples using the results established in this section.

**Example 1.7.4**

Five red and four blue marbles are arranged in a row. What is the probability that both the end marbles are blue?

A basic decision in the solution of the problem concerns the type of sample space to be used. Clearly, the problem involves order, but should we consider the marbles to be distinct or not?
Initially, consider the marbles to be alike, except for color of course. There are \( \frac{9!}{5! \cdot 4!} = 126 \) possible orderings of the marbles and we consider each of these to be equally likely. Since the blue marbles are indistinct from each other, and since our only choice here is the arrangement of the 7 marbles in the middle, it follows that there are \( \frac{7!}{5! \cdot 2!} = 21 \) arrangements with blue marbles at the ends. The probability we seek is then \( \frac{21}{126} = \frac{1}{6} \).

Now if we consider each of the marbles to be distinct, there are \( 9! \) possible arrangements. Of these, we have \( \binom{4}{2} \cdot 2! = 12 \) ways to arrange the blue marbles at the ends and \( 7! \) ways to arrange the marbles in the middle. This produces a probability of \( \frac{7! \cdot 12}{9!} = \frac{1}{6} \).

The two methods must produce the same result, but the reader may find one method easier to use than another. In any event, it is crucial that the sample space be established as a first step in the solution of the problem and that the events of interest be dealt with consistently for this sample space.

The reader may enjoy showing that, if we have \( n \) marbles, \( r \) of which are red and \( b \) of which are blue, then the probability both ends are blue in a random arrangement of the marbles is given by the product

\[
\left(1 - \frac{r}{n}\right) \cdot \left(1 - \frac{r}{n-1}\right).
\]

This answer may indicate yet another way to solve the problem, namely this: the probability the first marble is blue is \( \frac{n-r}{n} \). Given that the first end is blue, the conditional probability the other end is also blue is \( \frac{n-r-1}{n-1} \). Often probability problems involving counting techniques can be solved in a variety of ways.

**Example 1.7.5**

Ten race cars, numbered from 1 to 10, are running around a track. An observer sees three cars go by. If the cars appear in random order, what is the probability that the largest number seen is 6?

The choice of the sample space here is natural: consider all the \( \binom{10}{3} \) samples of three cars that could be observed. If the largest is to be 6, then 6 must be in the sample, together with two cars chosen from the first 5, so the probability of the event “Maximum = 6” is

\[
P(\text{Maximum} = 6) = \frac{\binom{1}{1} \cdot \binom{5}{2}}{\binom{10}{3}} = \frac{1}{12}.
\]

It is also interesting now to look at the median or the number in the middle when the three observed numbers are arranged in order. What is the probability that the median of the group of three is 6?

For the median to be 6, 6 must be chosen and we must choose exactly one number from the set \( \{1, 2, 3, 4, 5\} \) and exactly one number from \( \{7, 8, 9, 10\} \). Then

\[
P(\text{Median} = 6) = \frac{\binom{1}{1} \cdot \binom{5}{1} \cdot \binom{4}{1}}{\binom{10}{3}} = \frac{1}{6}.
\]
This can be generalized to

\[ P(\text{Median} = k) = \frac{\binom{1}{1} \cdot \binom{k - 1}{1} \cdot \binom{10 - k}{1}}{\binom{10}{3}} \]

\[ = \frac{(k - 1)(10 - k)}{120}, \quad k = 2, 3, \ldots, 9. \]

Figure 1.23 shows a graph of \( P(\text{Median} = k) \) for \( k = 2, 3, \ldots, 9 \). It reveals a symmetry in the function around \( k = 5.5 \).

The problem is easily generalized with a result that may be surprising. Suppose there are 100 cars and we observe a sample of 9 of them. The median of the sample must be at least 5 and can be at most 96. The probability the median is \( k \) then is

\[ P(\text{Median} = k) = \frac{\binom{1}{1} \cdot \binom{k - 1}{4} \cdot \binom{100 - k}{4}}{\binom{100}{9}}, \quad k = 5, 6, \ldots, 96. \]

A graph of this function (an eighth degree polynomial in \( k \)) is shown in Figure 1.24.

The graph here shows a “bell shape” that, as we will see, is very common in probability problems. The curve is very close to what we will call a normal curve. Larger values for the number of cars involved will, surprisingly, not change the approximate normal shape of the curve! An approximation for the actual curve involved here can be found when we study the normal curve thoroughly in Chapter 3.

Example 1.7.6

We can use the result of Example 1.7.3 to count the number of derangements of a set of \( n \) objects. That is, we want to count the number of permutations in which no object occupies
its own place. Example 1.7.3 shows that the number of permutations of \( n \) distinct objects in which at least one object occupies its own place is

\[
n! \left( \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots \pm \frac{1}{n!} \right).
\]

It follows that the number of derangements of \( n \) distinct objects is

\[
n! - n! \left( \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots \pm \frac{1}{n!} \right) = n! \left( \frac{1}{2!} - \frac{1}{3!} + \cdots \pm \frac{1}{n!} \right).
\]

Formula (1.6) also suggests that the number of derangements of \( n \) distinct objects is approximately \( n! \cdot e^{-1} \) (see the series expansion for \( e^{-1} \) in Example 1.7.5). The following table compares the results of formula (1.6) and the approximation:

<table>
<thead>
<tr>
<th>( n )</th>
<th>Number of derangements</th>
<th>( n! \cdot e^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0.7358</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2.207</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>8.829</td>
</tr>
<tr>
<td>5</td>
<td>44</td>
<td>44.146</td>
</tr>
<tr>
<td>6</td>
<td>265</td>
<td>264.83</td>
</tr>
<tr>
<td>7</td>
<td>1854</td>
<td>1854.11</td>
</tr>
</tbody>
</table>

We see that in every case, the number of derangements is given by \( \lfloor n! \cdot e^{-1} + 0.5 \rfloor \) where the symbols indicate the greatest integer function.

**Example 1.7.7**  
*Ken–Ken Puzzles*

The *New York Times* as well as many other newspapers publish a Ken–Ken puzzle daily. The problem consists of a square with 4 rows and 4 columns. The problem is to insert
1.7 Counting Techniques

each of the digits 1, 2, 3, and 4 into each row and each column so that each digit appears exactly once in each row and each column. The reader is given arithmetic clues for some squares. For example, 5+ may indicate the proper entries are 4 1 (but there are many other possibilities. Here is an example of a solved puzzle (without the clues):

```
3 4 2 1
2 1 3 4
1 3 4 2
4 2 1 3
```

Clearly, each row (and hence each column) is a derangement of the integers 1 through 4, but each row (and column) must be a derangement of each of the previous rows (or columns). How many Ken–Ken puzzles are there?

Since we will be permuting the rows and columns later, we might just as well start with the row 1 2 3 4. For the second row, we must select one of the 9 derangements of the integers 1, 2, 3, 4, as shown in Example 1.7.1. We will choose 2 4 1 3, so now we have

```
1 2 3 4
2 4 1 3
```

By examining the 9 derangements again, we find only two choices for the third row: 3 1 4 2 or 4 3 2 1. When one of these is chosen, there is only one choice for the fourth row—the derangement that was not selected for the third row. Selecting the first choice for the third row, we have

```
1 2 3 4
2 4 1 3
3 1 4 2
4 3 2 1
```

Now the rows and the columns may be permuted in 4! * 4! ways, so the total number of Ken–Ken puzzles with 4 rows and 4 columns is 9 * 2 * 4! * 4! = 10368.

**EXERCISES 1.7**

1. The integers 1, 2, 3, …, 9 are arranged in a row, resulting in a nine-digit integer. What is the probability that
   (a) the integer resulting is even?
   (b) the integer resulting is divisible by 5?
   (c) the digits 6 and 4 are next to each other?

2. License plates in Indiana consist of a number from 1 to 99 (indicating the county of registration), a letter of the alphabet, and finally an integer from 1 to 9999. How many cars may be licensed in Indiana?

3. Prove that at least two people in Colordao Springs, Colorado, have the same three initials.

4. In a small school, 5 centers, 8 guards, and 6 forwards try out for the basketball team.
   (a) How many five-member teams can be formed from these players? (Assume a team has two guards, two forwards, and one center.)
   (b) Intercollegiate regulations require that no more than 8 players can be listed for the team roster. How many rosters can be formed consisting of exactly 8 players?
52 Chapter 1 Sample Spaces and Probability

5. A restaurant offers 5 appetizers, 7 main courses, and 8 desserts. How many meals can be ordered
   (a) assuming all three courses are ordered?
   (b) not assuming all three courses are necessarily ordered?

6. A club of 56 people has 40 men and 16 women. What is the probability the board of directors, consisting of 8 members, contains no women?

7. In a controlled experiment, 12 patients are to be randomly assigned to each of three different drug regimens. In how many ways can this be done if each drug is to be tested on 4 patients?

8. In the game Keno, the casino draws 20 balls from a set of balls numbered from 1 to 80. A player must choose 10 numbers in advance of this drawing. What is the probability the player has exactly five of the 20 numbers drawn?

9. A lot of 10 refrigerators contains 3 which are defective. The refrigerators are randomly chosen and shipped to customers. What is the probability that by the seventh shipment, none of the defective refrigerators remain?

10. In how many different ways can the letters in the word “repetition” be arranged?

11. In a famous correspondence in the very early history of probability, the Chevalier de Méré wrote to the mathematician Blaise Pascal and asked the following question, “Which is more likely—at least one six in four rolls of a fair die or at least one sum of 12 in 24 rolls of a pair of dice?”
   (a) Show that the two questions above have nearly equal answers. Which is more likely?
   (b) A generalization of the Pascal–de Méré problem is: what is the probability that the sum \( 6n \) occurs at least once in \( 4 \cdot 6^{n-1} \) rolls of \( n \) fair dice? Show that the answer is very nearly 1/2 for \( n \leq 5 \).
   (c) Show that in part (b) the probability approaches \( 1 - e^{-2/3} \) as \( n \to \infty \).

12. A box contains 8 red and 5 yellow marbles from which a sample of 3 is drawn.
   (a) Find the probability that the sample contains no yellow marbles if
      (1) the sampling is done without replacement; and,
      (2) if the sampling is done with replacement.
   (b) Now suppose the box contains 24 red and 15 yellow marbles (so that the ratio of reds to yellows is the same as in part (a)). Calculate the answers to part (a). What do you expect to happen as the number of marbles in the box increases but the ratio of reds to yellows remains the same?

13. (a) From a group of 20 people, two samples of size 3 are chosen, the first sample being replaced before the second sample is chosen. What is the probability the samples have at least one person in common?
   (b) Show that two bridge hands, the first being replaced before the second is drawn, are virtually certain to contain at least one card in common.

14. A shipment of 20 components will be accepted by a buyer if a random sample of 3 (chosen without replacement) contains no defectives. What is the probability the shipment will be rejected if actually 2 of the components are defective?

15. A deck of cards is shuffled and the cards turned up one at a time. What is the probability that all the aces will appear before any of the 10’s?
1.7 Counting Techniques

16. In how many distinguishable ways can 6 A’s, 4 B’s, and 8 C’s be assigned as grades to 18 students?

17. What is the probability a poker hand (5 cards drawn from a deck of 52 cards) has exactly 2 aces?

18. In how many ways can 6 students be seated in 10 chairs?

19. Ten children are to be grouped into two clubs, say the Lions and the Tigers, with five children in each club. Each club is then to elect a president and a secretary. In how many ways can this be done?

20. A small pond contains 50 fish, 10 of which have been tagged. If a catch of 7 fish is made, in how many ways can the catch contain exactly 2 tagged fish?

21. From a fleet of 12 limousines, 6 are to go to hotel I, 4 to hotel II, and the remainder to hotel III. In how many different ways can this be done?

22. The grid shows a region of city blocks defined by 7 streets running North–South and 8 streets running East–West. Joe will walk from corner A to corner B. At each corner between A and B, Joe will choose to walk either North or East.

(a) How many possible routes are there?

(b) Assuming that each route is equally likely, find the probability that Joe will pass through intersection C.

23. Suppose that N people are arranged in a line. What is the probability that two particular people, say A and B, are not next to each other?

24. The Hawaiian language has only 12 letters: the vowels a, e, i, o, and u and the consonants h, k, l, m, n, p, and w.

(a) How many possible three-letter Hawaiian “words” are there? (Some of these may be nonsense words.)

(b) How many three-letter “words” have no repeated letter?

(c) What is the probability a randomly selected three-letter “word” begins with a consonant and ends with 2 different vowels?
(d) What is the probability that a randomly selected three-letter “word” contains all vowels?

25. How many partial derivatives of order 4 are there for a function of 4 variables?

26. A set of 15 marbles contains 4 red and 11 green marbles. They are selected, one at a time, without replacement. In how many ways can the last red marble be drawn on the seventh selection?

27. A true–false test has four questions. A student is not prepared for the test and so must guess the answer to each question.

(a) What is the probability the student answers at least half of the questions correctly?

(b) Now suppose, in a sudden flash of insight, he knows the answer to question 2 is “false.” What is the probability he answers at least half of the questions correctly?

28. What is the probability of being dealt a bridge hand (13 cards selected from a deck of 52 cards) that does not contain a heart?

29. Explain why the number of derangements of \( n \) distinct objects is given by \( [n! \cdot e^{-1} + 0.5] \). Explain why \( n! \cdot e^{-1} \) sometimes underestimates the number of derangements and sometimes overestimates the number of derangements. \([x]\) denotes the greatest integer in \( x \).

30. Find the number of Ken–Ken puzzles if the grid is 5 \( \times \) 5 for the integers 1, 2, 3, 4, 5.

**CHAPTER REVIEW**

In dealing with an experiment or situation involving random or chance elements, it is reasonable to begin an analysis of the situation by asking the question, “What can happen?” An enumeration of all the possibilities is called a sample space. Generally, situations admit of more than one sample space; the appropriate one chosen is usually governed by the probabilities that one wants to compute. Several examples of sample spaces are given in this chapter, each of them discrete, that is, either the sample space has a finite number of points or a countably infinite number of points.

Tossing two dice yields a sample space with a finite number of points; observing births until a girl is born gives a sample space with an infinite (but countable) number of points. In the next chapter, we will encounter continuous sample spaces that are characterized by a noncountably infinite number of points.

Assessing the long-range relative frequency, or probability, of any of the points or sets of points (which we refer to as events) is the primary goal of this chapter. We use the set symbols \( \cup \) for the union of two events and \( \cap \) for the intersection of two events. We begin with three assumptions or axioms concerning sample spaces:

1. \( P(A) \geq 0 \), where \( A \) is an event;
2. \( P(S) = 1 \), where \( S \) is the entire sample space; and,
3. \( P(A \cup B) = P(A) + P(B) \) if \( A \) and \( B \) are disjoint, or mutually exclusive, they have no sample points in common.

From these assumptions, we derived several theorems concerning probability, among them:

1. \( P(A) = \sum_{a_i \in A} P(a_i) \), where the \( a_i \) are distinct point in \( S \).
1.7 Counting Techniques

(2) \( P(A \cup B) = P(A) + P(B) - P(A \cap B) \) (the addition law for two events).

(3) \( P(\overline{A}) = 1 - P(A) \)

We showed the Law of Total Probability.

**Theorem (Law of Total Probability)**: If the sample space \( S = A_1 \cup A_2 \cup \ldots \cup A_n \) where \( A_i \) and \( A_j \) have no sample points in common if \( i \neq j \), then, if \( B \) is an event,

\[
P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_n) \cdot P(B|A_n)
\]

We then turned our attention to problems of conditional probability where we sought the probability of some event, say \( A \), on the condition that some other event, say \( B \), has occurred. We showed that

\[
P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B) + P(\overline{A}) \cdot P(B|\overline{A})}
\]

This can be generalized using the Law of Total Probability as follows:

**Theorem (Bayes’ Theorem)**: If \( S = A_1 \cup A_2 \cup \ldots \cup A_n \) where \( A_i \) and \( A_j \) have no sample points in common for \( i \neq j \), then, if \( B \) is an event,

\[
P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i) \cdot P(B|A_i)}{P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_n) \cdot P(B|A_n)}
\]

\[
P(A_i|B) = \frac{P(A_i) \cdot P(A_i|B)}{\sum_{i=1}^{n} P(A_i) \cdot P(A_i|B)}
\]

Bayes’ theorem has a simple geometric interpretation. The chapter contains many examples of this.

We defined the independence of two events, \( A \) and \( B \) as follows:

\( A \) and \( B \) are independent if \( P(A \cap B) = P(A) \cdot P(B) \).

We then applied the results of this chapter to some specific probability problems, such as the well-known birthday problem and a geometric problem involving the sides and diagonals of a polygonal figure.

Finally, we considered some very special counting techniques which are useful, it is to be emphasized, only if the points in the sample space are equally likely. If that is so, then the probability of an event, say \( A \), is

\[
P(A) = \frac{\text{Number of points in } A}{\text{Number of points in } S}
\]

If order is important, then all the permutations of objects may well comprise the sample space. We showed that there are \( n! = n \cdot (n - 1) \cdot (n - 2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1 \) permutations of \( n \) distinct objects.
If order is not important, then the sample space may well comprise various combinations of items. We showed that there are
\[ \binom{n}{r} = \frac{n!}{r!(n-r)!} \]
samples of size \( r \) that can be selected from \( n \) distinct objects and applied this formula to several examples. A large number of identities are known concerning these combinations, or binomial coefficients, among them:
\[ \sum_{r=0}^{n} \binom{n}{r} = 2^n. \]
\[ \binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}. \]

One very important result from this section is the general addition law:

**Theorem:**

\[ P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum P(A_i) - \sum P(A_i \cap A_j) + \sum P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n), \]

where the summations are over \( i > j > k > \cdots \)

**PROBLEMS FOR REVIEW**

Exercises 1.1 #1, 2, 5, 7, 9, 11
Exercises 1.3 #1, 2, 6, 7, 9, 13
Exercises 1.4 #1, 2, 3, 6, 10, 15, 16, 18, 19, 21, 24
Exercises 1.5 #2, 3, 6, 7
Exercises 1.6 #1, 3
Exercises 1.7 #1, 6, 8, 10, 12, 13, 16, 17, 20, 23, 28

**SUPPLEMENTARY EXERCISES FOR CHAPTER 1**

1. A hat contains slips of paper on which each of the integers 1, 2, \ldots, 20 is written. A sample of size 6 is drawn (without replacement) and the sample values, \( x_i \), put in order so that \( x_1 < x_2 < \cdots < x_6 \). Find the probability that \( x_3 = 12 \).
2. Show that \((n-k) \binom{n}{n-k} = (k+1) \binom{n}{k+1} \).
3. Suppose that events \( A, B, \) and \( C \) are independent with \( P(A) = 1/3, P(B) = 1/4, \) and \( P(A \cup B \cup C) = 3/4 \). Find \( P(C) \).
4. Events \( A \) and \( B \) are such that \( P(A \cup B) = 0.8 \) and \( P(A) = 0.2 \). For what value of \( P(B) \) are
   (a) \( A \) and \( B \) independent?
   (b) \( A \) and \( B \) mutually exclusive?
5. Events \( A, B, \) and \( C \) in a sample space have \( P(A) = 0.2, P(B) = 0.4, \) and \( P(A \cup B \cup C) = 0.9 \). Find \( P(C) \) if \( A \) and \( B \) are mutually exclusive, \( A \) and \( C \) are independent, and \( B \) and \( C \) are independent.
6. How many distinguishable arrangements of the letters in PROBABILITY are there?

7. How many people must be in a group so that the probability that at least two were born on the same day of the week is at least 1/2?

8. A and B are special dice. The faces on die A are 2, 2, 5, 5, 5, 5 and the faces on die B are 3, 3, 3, 6, 6, 6. The two dice are rolled. What is the probability that the number showing on die B is greater than the number showing on die A?

9. A committee of 5 is chosen from a group of 8 men and 4 women. What is the probability the group contains a majority of women?

10. A college senior finds he needs one more course for graduation and finds only courses in Mathematics, Chemistry, and Computer Science available. On the basis of interest, he assigns probabilities of 0.1, 0.6 and 0.3, respectively, to the events of choosing each of these. After considering his past performance, his advisor estimates his probabilities of passing these courses as 0.8, 0.7, and 0.6, respectively, regarding the passing of courses as independent events.
   (a) What is the probability he passes the course if he chooses a course at random?
   (b) Later we find that the student graduated. What is the probability he took Chemistry?

11. A number, \( X \), is chosen at random from the set \{10, 11, 12, ..., 99\}.
   (a) Find the probability that the 10’s digit in \( X \) is less than the units digit.
   (b) Find the probability that \( X \) is at least 50.
   (c) Find the probability that the 10’s digit in \( X \) is the square of the units digit.

12. If the integers 1, 2, 3, and 4 are randomly permuted, what is the probability that 4 is to the left of 2?

13. In a sample space, events \( A \) and \( B \) are such that \( P(A) = P(B) \), \( P(\overline{A} \cap \overline{B}) = P(A \cap B) = 1/6 \). Find
   (a) \( P(A) \).
   (b) \( P(\overline{A} \cup \overline{B}) \).
   (c) \( P(\text{Exactly one of the events } A \text{ or } B) \).

14. A fair coin is tossed four times. Let \( A \) be the event “2nd toss is heads,” \( B \) be the event “Exactly 3 heads,” and \( C \) be the event “4th toss is tails if the 2nd toss is heads.” Are \( A \), \( B \), and \( C \) independent?

15. An instructor has decided to grade each of his students \( A \), \( B \), or \( C \). He wants the probability a student receives a grade of \( B \) or better to be 0.7 and the probability a student receives at most a grade of \( B \) to be 0.8. Is this possible? If so, what proportions of each letter grade must be assigned?

16. How many bridge hands are there containing 3 hearts, 4 clubs, and 6 spades?

17. A day’s production of 100 fuses is inspected by a quality control inspector who tests 10 fuses at random, sampling without replacement. If he finds 2 or fewer defective fuses, he accepts the entire lot of 100 fuses. What is the probability the lot is accepted if it actually contains 20 defective fuses?

18. Suppose that \( A \) and \( B \) are events for which \( P(A) = a \), \( P(B) = b \) and \( P(A \cap B) = c \). Express each of the following in terms of \( a \), \( b \), and \( c \).
   (a) \( P(\overline{A} \cup \overline{B}) \)
   (b) \( P(\overline{A} \cap B) \)
58 Chapter 1 Sample Spaces and Probability

(c) \( P(\overline{A} \cup B) \)

(d) \( P(\overline{A} \cap \overline{B}) \)

(e) \( P(\text{exactly one of } A \text{ or } B \text{ occurs}) \).

19. An elevator starts with 10 people on the first floor of an eight-story building and stops at each floor.
   (a) In how many ways can all the people get off the elevator?
   (b) How many ways are there for everyone to get off if no one gets off on some two specific floors?
   (c) In how many ways are there for everyone to get off if at least one person gets off at each floor?

20. A manufacturer of calculators buys integrated circuits from suppliers \( A \), \( B \), and \( C \). Fifty percent of the circuits come from \( A \), 30% from \( B \), and 20% from \( C \). One percent of the circuits supplied by \( A \) have been defective in the past, 3% of \( B \)'s have been defective, and 4% of \( C \)'s have been defective. A circuit is selected at random and found to be defective. What is the probability it was manufactured by \( B \)?

21. Suppose that \( E \) and \( T \) are independent events with \( P(E) = P(T) \) and \( P(E \cup T) = \frac{1}{2} \). What is \( P(E) \)?

22. A quality control inspector draws parts one at a time and without replacement from a set containing 5 defective and 10 good parts. What is the probability the third defective is found on the eighth drawing?

23. If \( A \), \( B \), and \( C \) are independent events, show that the events \( A \) and \( B \cup C \) are independent.

24. Bean seeds from supplier \( A \) have an 85% germination rate and those from supplier \( B \) have a 75% germination rate. A seed company purchases 40% of their bean seeds from supplier \( A \) and the remaining 60% from supplier \( B \) and mixes these together. If a seed germinates, what is the probability it came from supplier \( A \)?

25. An experiment consists of choosing two numbers without replacement from the set \( \{1, 2, 3, 4, 5, 6\} \) with the restriction that the second number chosen must be greater than the first.
   (a) Describe the sample space.
   (b) What is the probability the second number is even?
   (c) What is the probability the sum of the two numbers is at least 5?

26. What is the probability a poker hand contains exactly one pair?

27. A box contains 6 good and 8 defective light bulbs. The bulbs are drawn out one at a time, without replacement, and tested. What is the probability that the fifth good item is found on the ninth test?

28. An individual tried by a three-judge panel is declared guilty if at least two judges cast votes of guilty. Suppose that when the defendant is, in fact, guilty, each judge will independently vote guilty with probability 0.7 but, if the defendant is, in fact, innocent, each judge will independently vote guilty with probability 0.2. Assume that 70% of the defendants are actually guilty. If a defendant is judged guilty by the panel of judges, what is the probability he is actually innocent?

29. What is the probability a bridge hand is missing cards in at least one suit?

30. Suppose 0.1% of the population is infected with a certain disease. On a medical test for the disease, 98% of those infected give a positive result while 1% of those not infected
1.7 Counting Techniques

give a positive result. If a randomly chosen person is tested and gives a positive result, what is the probability the person has the disease?

31. A committee of 50 politicians is to be chosen from the 100 US Senators (2 are from each state). If the selection is done at random, what is the probability that each state will be represented?

32. In a roll of a pair of dice (one red and one green), let $A$ be the event “red die shows 3, 4, or 5,” $B$ the event “green die shows a 1 or a 2,” and $C$ the event “dice total 7.” Show that $A$, $B$, and $C$ are independent.

33. An oil wildcatter thinks there is a 50-50 chance that oil is on the property he purchased. He has a test for oil that is 80% reliable: that is, if there is oil, it indicates this with probability 0.80 and if there is no oil, it indicates that with probability 0.80. The test indicates oil on the property. What is the probability there really is oil on the property?

34. Given: $A$ and $B$ are events with $P(A) = 0.3$, $P(B) = 0.7$ and $P(A \cup B) = 0.9$. Find
   (a) $P(A \cap B)$
   (b) $P(B|\bar{A})$.

35. Two good transistors become mixed up with three defective transistors. A person is assigned to sampling the mixture by drawing out three items without replacement. However, the instructions are not followed and the first item is replaced, but the second and third items are not replaced.
   (a) What is the probability the sample contains exactly two items that test as good?
   (b) What is the probability the two items finally drawn are both good transistors?

36. How many lines are determined by 8 points, no three of which are collinear?

37. Show that if $A$ and $B$ are independent, then $\bar{A}$ and $\bar{B}$ are independent.

38. How many tosses of a fair coin are needed so that the probability of at least one head is at least 0.99?

39. A lot of 24 tubes contains 13 defective ones. The lot is randomly divided into two equal groups, and each group is placed in a box.
   (a) What is the probability that one box contains only defective tubes?
   (b) Suppose the tubes were divided so that one box contains only defective tubes. A box is chosen at random and one tube is chosen from the chosen box and is found to be defective. What is the probability a second tube chosen from the same box is also defective?

40. A machine is composed of two components, $A$ and $B$, which function (or fail) independently. The machine works only if both components work. It is known that component $A$ is 98% reliable and the machine is 95% reliable. How reliable is component $B$?

41. Suppose $A$ and $B$ are events. Explain why $P(\text{exactly one of events } A,B \text{ occurs}) = P(A) + P(B) - 2P(A \cap B)$.

42. A box contains 8 red, 3 white, and 9 blue balls. Three balls are to be drawn, without replacement. What is the probability that more blues than whites are drawn?

43. A marksman, whose probability of hitting a moving target is 0.6, fires three shots. Suppose the shots are independent.
   (a) What is the probability the target is hit?
   (b) How many shots must be fired to make the probability at least 0.99 that the target will be hit?
44. A box contains 6 green and 11 yellow balls. Three are chosen at random. The first and third balls are yellow. Which method of sampling—with replacement or without replacement—gives the higher probability of this event?

45. A box contains slips of paper numbered from 1 to \( m \). One slip is drawn from the box; if it is 1, it is kept; otherwise, it is returned to the box. A second slip is drawn from the box. What is the probability the second slip is numbered 2?

46. Three integers are selected at random from the set \{1, 2, ..., 10\}. What is the probability the largest of these is 5?

47. A pair of dice is rolled until a 5 or a 7 appears. What is the probability a 5 occurs first?

48. The probability is 1 that a fisherman will say he had a good day when, in fact, he did, but the probability is only 0.6 that he will say he had a good day when, in fact, he did not. Only 1/4 of his fishing days are actually good days. What is the probability he had a good day if he says he had a good day?

49. An inexperienced employee mistakenly samples \( n \) items from a lot of \( N \) items, with replacement. What is the probability the sample contains at least one duplicate?

50. A roulette wheel has 38 slots—18 red, 18 black, and 2 green (the house wins on green). Suppose the spins of the wheel are independent and that the wheel is fair. The wheel is spun twice and we know that at least one spin is green. What is the probability that both spins are green?

51. A “rook” deck of cards consists of four suits of cards: red, green, black, and yellow, each suit having 14 cards. In addition, the deck has an uncolored “rook” card. A hand contains 14 cards.

(a) How many different hands are possible?
(b) How many hands have the rook card?
(c) How many hands contain only two colors with equal numbers of cards of each color?
(d) How many hands have at most three colors and no rook card?

52. Find the probability a poker hand contains 3 of a kind (exactly 3 cards of one face value and 2 cards of different face values).

53. A box contains tags numbered 1, 2, ..., \( n \). Two tags are chosen without replacement. What is the probability they are consecutive integers?

54. In how many different ways can \( n \) people be seated around a circular table?

55. A production lot has 100 units of which 25 are known to be defective. A random sample of 4 units is chosen without replacement. What is the probability that the sample will contain no more than 2 defective units?

56. A recent issue of a newspaper said that given a 5% probability of an unusual event in a 1-year study, one should expect a 35% probability in a 7-year study. This is obviously faulty. What is the correct probability?

57. Independent events \( A \) and \( B \) have probabilities \( p_A \) and \( p_B \), respectively. Show that the probability of either two successes or two failures in two trials has probability 1/2 if and only if at least one of \( p_A \) and \( p_B \) is 1/2.