52. How many committees of 4 people are possible from a group of 9 people if
(A) There are no restrictions?
(B) Both Juan and Mary must be on the committee?
(C) Either Juan or Mary, but not both, must be on the committee?

53. A 5-card hand is dealt from a standard 52-card deck. Which is more likely: the hand contains exactly 1 king or the hand contains no hearts?

54. A 10-card hand is dealt from a standard 52-card deck. Which is more likely: all cards in the hand are red or the hand contains all four aces?

Section 6-5 Sample Spaces and Probability

Experiments
Sample Spaces and Events
Probability of an Event
Equally Likely Assumption
Empirical Probability

This section provides an introduction to probability, a topic that has whole books and courses devoted to it. Probability studies involve many subtle notions, and care must be taken at the beginning to understand the fundamental concepts on which the studies are based. First, we develop a mathematical model for probability studies. Our development, because of space, must be somewhat informal. More formal and precise treatments can be found in books on probability.

Experiments
Our first step in constructing a mathematical model for probability studies is to describe the type of experiments on which probability studies are based. Some types of experiments do not yield the same results, no matter how carefully they are repeated under the same conditions. These experiments are called random experiments. Familiar examples of random experiments are flipping coins, rolling dice, observing the frequency of defective items from an assembly line, or observing the frequency of deaths in a certain age group.

Probability theory is a branch of mathematics that has been developed to deal with outcomes of random experiments, both real and conceptual. In the work that follows, the word experiment will be used to mean a random experiment.

Sample Spaces and Events
Associated with outcomes of experiments are sample spaces and events. Our second step in constructing a mathematical model for probability studies is to define these two terms. Set concepts will be useful in this regard.

Consider the experiment, “A single six-sided die is rolled.” What outcomes might we observe? We might be interested in the number of dots facing up, or whether the number of dots facing up is an even number, or whether the number of dots facing up is divisible by 3, and so on. The list of possible outcomes appears endless. In general, there is no unique method of analyzing all possible outcomes of an experiment. Therefore, before conducting an experiment, it is important to decide just what outcomes are of interest.

In the die experiment, suppose we limit our interest to the number of dots facing up when the die comes to rest. Having decided what to observe, we make a
list of outcomes of the experiment, called *simple events*, such that in each trial of the experiment, one and only one of the results on the list will occur. The set of simple events for the experiment is called a *sample space* for the experiment. The sample space \( S \) we have chosen for the die-rolling experiment is

\[
S = \{1, 2, 3, 4, 5, 6\}
\]

Now consider the outcome, “The number of dots facing up is an even number.” This outcome is not a simple event, since it will occur whenever 2, 4, or 6 dots appear, that is, whenever an element in the subset

\[
E = \{2, 4, 6\}
\]

occurs. Subset \( E \) is called a *compound event*. In general, we have the following definition:

---

**DEFINITION**

Given a sample space \( S \) for an experiment, we define an *event* \( E \) to be any subset of \( S \). If an event \( E \) has only one element in it, it is called a *simple event*. If event \( E \) has more than one element, it is called a *compound event*. We say that an event \( E \) occurs if any of the simple events in \( E \) occurs.

---

**EXAMPLE 1**

Choosing a Sample Space

A nickel and a dime are tossed. How will we identify a sample space for this experiment?

There are a number of possibilities, depending on our interest. We will consider three.

(A) If we are interested in whether each coin falls heads (H) or tails (T), then, using a tree diagram, we can easily determine an appropriate sample space for the experiment:

<table>
<thead>
<tr>
<th>Nickel Outcomes</th>
<th>Dime Outcomes</th>
<th>Combined Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>H</td>
<td>HH</td>
</tr>
<tr>
<td>H</td>
<td>T</td>
<td>HT</td>
</tr>
<tr>
<td>T</td>
<td>H</td>
<td>TH</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>TT</td>
</tr>
</tbody>
</table>

Thus,

\[
S_1 = \{HH, HT, TH, TT\}
\]

and there are four simple events in the sample space.
(B) If we are interested only in the number of heads that appear on a single toss of the two coins, then we can let

\[ S_2 = \{0, 1, 2\} \]

and there are three simple events in the sample space.

(C) If we are interested in whether the coins match \((M)\) or don’t match \((D)\), then we can let

\[ S_3 = \{M, D\} \]

and there are only two simple events in the sample space.

An experiment consists of recording the boy–girl composition of families with 2 children.

(A) What is an appropriate sample space if we are interested in the sex of each child in the order of their births? Draw a tree diagram.

(B) What is an appropriate sample space if we are interested only in the number of girls in a family?

(C) What is an appropriate sample space if we are interested only in whether the sexes are alike \((A)\) or different \((D)\)?

(D) What is an appropriate sample space for all three interests expressed above?

In Example 1, sample space \(S_1\) contains more information than either \(S_2\) or \(S_3\). If we know which outcome has occurred in \(S_1\), then we know which outcome has occurred in \(S_2\) and \(S_3\). However, the reverse is not true. In this sense, we say that \(S_1\) is a more fundamental sample space than either \(S_2\) or \(S_3\).

**Important Remark:** There is no one correct sample space for a given experiment. When specifying a sample space for an experiment, we include as much detail as necessary to answer all questions of interest regarding the outcomes of the experiment. If in doubt, include more elements in the sample space rather than fewer.

Now let’s return to the 2-coin problem in Example 1 and the sample space

\[ S_1 = \{HH, HT, TH, TT\} \]

Suppose we are interested in the outcome, “Exactly 1 head is up.” Looking at \(S_1\), we find that it occurs if either of the two simple events HT or TH occurs.* Thus, to say that the event, “Exactly 1 head is up” occurs is the same as saying the experiment has an outcome in the set

\[ E = \{HT, TH\} \]

This is a subset of the sample space \(S_1\). The event \(E\) is a compound event.

*Technically, we should write \(\{HT\}\) and \(\{TH\}\), since there is a logical distinction between an element of a set and a subset consisting of only that element. But we will just keep this in mind and drop the braces for simple events to simplify the notation.
Consider an experiment of rolling 2 dice. A convenient sample space that will enable us to answer many questions about interesting events is shown in Figure 1. Let $S$ be the set of all ordered pairs listed in the figure. Note that the simple event $(3, 2)$ is to be distinguished from the simple event $(2, 3)$. The former indicates a 3 turned up on the first die and a 2 on the second, while the latter indicates a 2 turned up on the first die and a 3 on the second. What is the event that corresponds to each of the following outcomes?

(A) A sum of 7 turns up. 
(B) A sum of 11 turns up. 
(C) A sum less than 4 turns up. 
(D) A sum of 12 turns up.

**Solutions**

(A) By “A sum of 7 turns up,” we mean that the sum of all dots on both turned-up faces is 7. This outcome corresponds to the event

$$\{(6, 1), (5, 2), (4, 3), (3, 4), (2, 5), (1, 6)\}$$

(B) “A sum of 11 turns up” corresponds to the event

$$\{(6, 5), (5, 6)\}$$

(C) “A sum less than 4 turns up” corresponds to the event

$$\{(1, 1), (2, 1), (1, 2)\}$$

(D) “A sum of 12 turns up” corresponds to the event

$$\{(6, 6)\}$$

**Matched Problem**

Refer to the sample space in Example 2 (Fig. 1). What is the event that corresponds to each of the following outcomes?

(A) A sum of 5 turns up.

(B) A sum that is a prime number greater than 7 turns up.
Informally, to facilitate discussion, we often use the terms *event* and *outcome of an experiment* interchangeably. Thus, in Example 2 we might say “the event ‘A sum of 11 turns up’” in place of “the outcome ‘A sum of 11 turns up,’” or even write

\[ E = \text{A sum of 11 turns up} = \{(6, 5), (5, 6)\} \]

Technically speaking, an event is the mathematical counterpart of an outcome of an experiment.

**Probability of an Event**

The next step in developing our mathematical model for probability studies is the introduction of a *probability function*. This is a function that assigns to an arbitrary event associated with a sample space a real number between 0 and 1, inclusive. We start by discussing ways in which probabilities are assigned to simple events in \( S \).

**DEFINITION**

**Probabilities for Simple Events**

Given a sample space

\[ S = \{ e_1, e_2, \ldots, e_n \} \]

with \( n \) simple events, to each simple event \( e_i \) we assign a real number, denoted by \( P(e_i) \), that is called the *probability of the event \( e_i \)*. These numbers may be assigned in an arbitrary manner as long as the following two conditions are satisfied:

1. \( 0 \leq P(e_i) \leq 1 \)
2. \( P(e_1) + P(e_2) + \cdots + P(e_n) = 1 \)

The sum of the probabilities of all simple events in the sample space is 1.

Any probability assignment that meets conditions 1 and 2 is said to be an [acceptable probability assignment](#).

Our mathematical theory does not explain how acceptable probabilities are assigned to simple events. These assignments are generally based on the expected or actual percentage of times a simple event occurs when an experiment is repeated a large number of times. Assignments based on this principle are called *reasonable*.

Let an experiment be the flipping of a single coin, and let us choose a sample space \( S \) to be

\[ S = \{ H, T \} \]

If a coin appears to be fair, we are inclined to assign probabilities to the simple events in \( S \) as follows:

\[ P(H) = \frac{1}{2} \quad \text{and} \quad P(T) = \frac{1}{2} \]
These assignments are based on reasoning that, since there are two ways a coin can land, in the long run, a head will turn up half the time and a tail will turn up half the time. These probability assignments are acceptable, since both of the conditions for acceptable probability assignments in Definition 2 are satisfied:

1. $0 \leq P(H) \leq 1$, $0 \leq P(T) \leq 1$
2. $P(H) + P(T) = \frac{1}{2} + \frac{1}{2} = 1$

But there are other acceptable assignments. Maybe after flipping a coin 1,000 times we find that the head turns up 376 times and the tail turns up 624 times. With this result, we might suspect that the coin is not fair and assign the simple events in the sample space $S$ the probabilities

$$P(H) = .376 \quad \text{and} \quad P(T) = .624$$

This is also an acceptable assignment. But the probability assignment

$$P(H) = 1 \quad \text{and} \quad P(T) = 0$$

though acceptable, is not reasonable, unless the coin has 2 heads. The assignment

$$P(H) = .6 \quad \text{and} \quad P(T) = .8$$

is not acceptable, since $.6 + .8 = 1.4$, which violates condition 2 in Definition 2.

In probability studies, the 0 to the left of the decimal is usually omitted. Thus, we write .8 and not 0.8.

It is important to keep in mind that out of the infinitely many possible acceptable probability assignments to simple events in a sample space, we are generally inclined to choose one assignment over another based on reasoning or experimental results.

Given an acceptable probability assignment for simple events in a sample space $S$, how do we define the probability of an arbitrary event $E$ associated with $S$?

### PROBABILITY OF AN EVENT $E$

Given an acceptable probability assignment for the simple events in a sample space $S$, we define the **probability of an arbitrary event $E$**, denoted by $P(E)$, as follows:

1. If $E$ is the empty set, then $P(E) = 0$.
2. If $E$ is a simple event, then $P(E)$ has already been assigned.
3. If $E$ is a compound event, then $P(E)$ is the sum of the probabilities of all the simple events in $E$.
4. If $E$ is the sample space $S$, then $P(E) = P(S) = 1$. This is a special case of 3.

### Example 3

**Finding Probabilities of Events**

Let’s return to Example 1, the tossing of a nickel and dime, and the sample space

$$S = \{HH, HT, TH, TT\}$$
Since there are four simple outcomes and the coins are assumed to be fair, it appears that each outcome should occur in the long run 25% of the time. Let’s assign the same probability of \( \frac{1}{4} \) to each simple event in \( S \):

<table>
<thead>
<tr>
<th>Simple event, ( e_i )</th>
<th>HH</th>
<th>HT</th>
<th>TH</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(e_i) )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{4} )</td>
</tr>
</tbody>
</table>

This is an acceptable assignment according to Definition 2 and a reasonable assignment for ideal coins that are perfectly balanced or coins close to ideal.

(A) What is the probability of getting exactly 1 head?
(B) What is the probability of getting at least 1 head?
(C) What is the probability of getting a head or a tail?
(D) What is the probability of getting 3 heads?

**Solutions**

(A) \( E_1 = \text{Getting 1 head} = \{HT, TH\} \)

Since \( E_1 \) is a compound event, we use item 3 in Definition 3 and find \( P(E_1) \) by adding the probabilities of the simple events in \( E_1 \). Thus,

\[
P(E_1) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}
\]

(B) \( E_2 = \text{Getting at least 1 head} = \{HH, HT, TH\} \)

\[
P(E_2) = P(HH) + P(HT) + P(TH)
\]

\[
= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}
\]

(C) \( E_3 = \{HH, HT, TH, TT\} = S \)

\[
P(E_3) = P(S) = 1 = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1
\]

(D) \( E_4 = \text{Getting 3 heads} = \emptyset \quad \text{Empty set} \)

\[
P(\emptyset) = 0
\]

**Steps for Finding Probabilities of Events**

- **Step 1.** Set up an appropriate sample space \( S \) for the experiment.
- **Step 2.** Assign acceptable probabilities to the simple events in \( S \).
- **Step 3.** To obtain the probability of an arbitrary event \( E \), add the probabilities of the simple events in \( E \).

The function \( P \) defined in steps 2 and 3 is called a **probability function**. The domain of this function is all possible events in the sample space \( S \), and the range is a set of real numbers between 0 and 1, inclusive.
Return to Matched Problem 1, recording the boy–girl composition of families with 2 children and the sample space

\[ S = \{BB, BG, GB, GG\} \]

Statistics from the U.S. Census Bureau indicate that an acceptable and reasonable probability for this sample space is

<table>
<thead>
<tr>
<th>Simple event, ( e_i )</th>
<th>BB</th>
<th>BG</th>
<th>GB</th>
<th>GG</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(e_i) )</td>
<td>.26</td>
<td>.25</td>
<td>.25</td>
<td>.24</td>
</tr>
</tbody>
</table>

Find the probabilities of the following events:
(A) \( E_1 = \text{Having at least one girl in the family} \)
(B) \( E_2 = \text{Having at most one girl in the family} \)
(C) \( E_3 = \text{Having two children of the same sex in the family} \)

**Equally Likely Assumption**

In tossing a nickel and dime (Example 3), we assigned the same probability, \( \frac{1}{4} \), to each simple event in the sample space \( S = \{HH, HT, TH, TT\} \). By assigning the same probability to each simple event in \( S \), we are actually making the assumption that each simple event is as likely to occur as any other. We refer to this as an **equally likely assumption**. In general, we have Definition 4.

**Definition 4**

**Probability of a Simple Event Under an Equally Likely Assumption**

If, in a sample space

\[ S = \{e_1, e_2, \ldots, e_n\} \]

with \( n \) elements, we assume each simple event \( e_i \) is as likely to occur as any other, then we assign the probability \( \frac{1}{n} \) to each. That is,

\[ P(e_i) = \frac{1}{n} \]

Under an equally likely assumption, we can develop a very useful formula for finding probabilities of arbitrary events associated with a sample space \( S \). Consider the following example.

If a single die is rolled and we assume each face is as likely to come up as any other, then for the sample space

\[ S = \{1, 2, 3, 4, 5, 6\} \]
we assign a probability of \( \frac{1}{6} \) to each simple event, since there are 6 simple events. Then the probability of

\[ E = \text{Rolling a prime number} = \{2, 3, 5\} \]

is

\[ P(E) = P(2) + P(3) + P(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} \]

Thus, under the assumption that each simple event is as likely to occur as any other, the computation of the probability of the occurrence of any event \( E \) in a sample space \( S \) is the number of elements in \( E \) divided by the number of elements in \( S \).

**THEOREM 1**

**PROBABILITY OF AN ARBITRARY EVENT UNDER AN EQUALLY LIKELY ASSUMPTION**

If we assume each simple event in sample space \( S \) is as likely to occur as any other, then the probability of an arbitrary event \( E \) in \( S \) is given by

\[ P(E) = \frac{\text{Number of elements in } E}{\text{Number of elements in } S} = \frac{n(E)}{n(S)} \]

**EXAMPLE 4**

**Finding Probabilities of Events**

If in rolling 2 dice we assume each simple event in the sample space shown in Figure 1 (page 495) is as likely as any other, find the probabilities of the following events:

(A) \( E_1 \) = A sum of 7 turns up
(B) \( E_2 \) = A sum of 11 turns up
(C) \( E_3 \) = A sum less than 4 turns up
(D) \( E_4 \) = A sum of 12 turns up

**Solutions**

Referring to Figure 1, we see that:

(A) \( P(E_1) = \frac{n(E_1)}{n(S)} = \frac{6}{36} = \frac{1}{6} \)
(B) \( P(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{36} = \frac{1}{18} \)
(C) \( P(E_3) = \frac{n(E_3)}{n(S)} = \frac{3}{36} = \frac{1}{12} \)
(D) \( P(E_4) = \frac{n(E_4)}{n(S)} = \frac{1}{36} \)

**MATCHED PROBLEM 4**

Under the conditions in Example 4, find the probabilities of the following events:

(A) \( E_5 \) = A sum of 5 turns up
(B) \( E_6 \) = A sum that is a prime number greater than 7 turns up
A box contains 4 red balls and 7 green balls. A ball is drawn at random and then, without replacing the first ball, a second ball is drawn. Discuss whether or not the equally likely assumption would be appropriate for the sample space \( S = \{RR, RG, GR, GG\} \).

We now turn to some examples that make use of the counting techniques developed in Section 6-4.

**Example 5**

**Drawing Cards**

In drawing 5 cards from a 52-card deck without replacement, what is the probability of getting 5 spades?

**Solution**

Let the sample space \( S \) be the set of all 5-card hands from a 52-card deck. Since the order in a hand does not matter, \( n(S) = C_{52,5} \). The event we seek is

\[ E = \text{Set of all 5-card hands from 13 spades} \]

Again, the order does not matter and \( n(E) = C_{13,5} \). Thus, assuming each 5-card hand is as likely as any other,

\[ P(E) = \frac{n(E)}{n(S)} = \frac{C_{13,5}}{C_{52,5}} = \frac{13!}{5!8!} \cdot \frac{52!}{5!47!} \approx .0005 \]

**Matched Problem 5**

In drawing 7 cards from a 52-card deck without replacement, what is the probability of getting 7 hearts?

**Example 6**

**Selecting Committees**

The board of regents of a university is made up of 12 men and 16 women. If a committee of 6 is chosen at random, what is the probability that it will contain 3 men and 3 women?

**Solution**

Let \( S = \text{Set of all 6-person committees out of 28 people:} \)

\[ n(S) = C_{28,6} \]

Let \( E = \text{Set of all 6-person committees with 3 men and 3 women.} \) To find \( n(E) \), we use the multiplication principle and the following two operations:

\[ O_1: \text{Select 3 men out of the 12 available} \quad N_1: \quad C_{12,3} \]

\[ O_2: \text{Select 3 women out of the 16 available} \quad N_2: \quad C_{16,3} \]
Thus,

\[ n(E) = C_{12,3} \cdot C_{16,3} \]

and

\[ P(E) = \frac{n(E)}{n(S)} = \frac{C_{12,3} \cdot C_{16,3}}{C_{28,6}} = .327 \]

**Empirical Probability**

In the earlier examples in this section, we made a reasonable assumption about an experiment and used deductive reasoning to assign probabilities. For example, it is reasonable to assume that an ordinary coin will come up heads about as often as it will come up tails. Probabilities determined in this manner are called theoretical probabilities. No experiments are ever conducted. But what if the theoretical probabilities are not obvious? Then we assign probabilities to simple events based on the results of actual experiments. Probabilities determined from the results of actually performing an experiment are called empirical probabilities. As an experiment is repeated over and over, the percentage of times an event occurs may get closer and closer to a single fixed number. If so, this single fixed number is generally called the actual probability of the event.

Like a coin, a thumbtack tossed into the air will land in one of two positions, point up or point down [see Fig. 2(a)]. Unlike a coin, we would not expect both events to occur with the same frequency. Indeed, the frequencies of landing point up and point down may well vary from one thumbtack to another [see Fig. 2(b)]. Find two thumbtacks of different sizes and guess which one is likely to land point up more frequently. Then toss each tack 100 times and record the number of times each lands point up. Did the experiment confirm your initial guess?

Suppose when tossing one of the thumbtacks in Explore/Discuss 2, we observe that the tack lands point up 43 times and point down 57 times. Based on this experiment, it seems reasonable to say that for this particular thumbtack
Probability assignments based on the results of repeated trials of an experiment are called **approximate empirical probabilities**.

In general, if we conduct an experiment \( n \) times and an event \( E \) occurs with frequency \( f(E) \), then the ratio \( f(E)/n \) is called the **relative frequency** of the occurrence of event \( E \) in \( n \) trials. We define the **empirical probability** of \( E \), denoted by \( P(E) \), by the number, if it exists, that the relative frequency \( f(E)/n \) approaches as \( n \) gets larger and larger. Of course, for any particular \( n \), the relative frequency \( f(E)/n \) is generally only approximately equal to \( P(E) \). However, as \( n \) increases in size, we expect the approximation to improve.

**Empirical Probability**

If \( f(E) \) is the frequency of event \( E \) in \( n \) trials, then

\[
P(E) = \frac{\text{Frequency of occurrence of } E}{\text{Total number of trials}} = \frac{f(E)}{n}
\]

If we can also deduce theoretical probabilities for an experiment, then we expect the approximate empirical probabilities to approach the theoretical probabilities. If this does not happen, then we should begin to suspect the manner in which the theoretical probabilities were computed. If \( P(E) \) is the theoretical probability of an event \( E \) and the experiment is performed \( n \) times, then the **expected frequency** of the occurrence of \( E \) is \( n \cdot P(E) \).

**Example 7**

Finding Approximate Empirical and Theoretical Probabilities

Two coins are tossed 500 times with the following frequencies of outcomes:

- 2 heads: 121
- 1 head: 262
- 0 heads: 117

(A) Compute the approximate empirical probability for each outcome.

(B) Compute the theoretical probability for each outcome.

(C) Compute the expected frequency for each outcome.

**Solutions**

(A) \[ P(2 \text{ heads}) = \frac{121}{500} = .242 \]

\[ P(1 \text{ head}) = \frac{262}{500} = .524 \]

\[ P(0 \text{ heads}) = \frac{117}{500} = .234 \]
(B) A sample space of equally likely simple events is \( S = \{HH, HT, TH, TT\} \). Let

\[
E_1 = \text{2 heads} = \{HH\} \\
E_2 = \text{1 head} = \{HT, TH\} \\
E_3 = \text{0 heads} = \{TT\}
\]

Then

\[
P(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{4} = .25 \\
P(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{4} = .50 \\
P(E_3) = \frac{n(E_3)}{n(S)} = \frac{1}{4} = .25
\]

(C) The expected frequencies are

\[
E_1: \ 500(.25) = 125 \\
E_2: \ 500(.5) = 250 \\
E_3: \ 500(.25) = 125
\]

The actual frequencies obtained from performing the experiment are reasonably close to the expected frequencies. Increasing the number of trials of the experiment would produce even better approximations.

**Matched Problem 7**

One die is rolled 500 times with the following frequencies of outcomes:

<table>
<thead>
<tr>
<th>Outcome</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>89</td>
<td>83</td>
<td>77</td>
<td>91</td>
<td>72</td>
<td>88</td>
</tr>
</tbody>
</table>

(A) Compute the approximate empirical probability for each outcome.

(B) Compute the theoretical probability for each outcome.

(C) Compute the expected frequency for each outcome.

**Figure 3**

Using a random number generator.

Tossing two coins 500 times is certainly a tedious task and we did not do this to generate the data in Example 7. Instead, we used a random number generator on a graphing utility to simulate this experiment. Specifically, we used the command `randInt(i,k,n)` on a Texas Instruments TI-83, which generates a random sequence of \( n \) integers between \( i \) and \( k \), inclusively. If we let 0 represent tails and 1 represent heads, then a random sequence of 0s and 1s can be used to represent the outcomes of repeated tosses of one coin (see the first two lines of Fig. 3). Thus, in six tosses, we obtained 2 heads and 4 tails. To simulate tossing two coins, we simply add together two similar statements, as shown in lines three through
five of Figure 3. We see that in these six tosses 0 heads occurred once, 1 head occurred four times, and 2 heads occurred once. Of course, to obtain meaningful results, we need to toss the coins many more times. Figure 4(a) shows a command that will simulate 500 tosses of two coins. To determine the frequency of each outcome, we construct a histogram [see Figs. 4(b) and 4(c)] and use the TRACE command to determine the following frequencies [see Figs. 5(a), 5(b), and 5(c)].

If you perform the same simulation on your graphing utility, you are not likely to get exactly the same results. But the approximate empirical probabilities you obtain will be close to the theoretical probabilities.

This discussion assumes that your graphing utility has the ability to generate and manipulate sequences of random integers.

(A) As an alternative to using the histogram in Figure 5 to count the outcomes of the sequence of random integers in Figure 4(a), enter the following function and evaluate it for \( x = 0, 1, \) and 2:

\[
y_i = \text{sum(seq(L1(I)=X,I,1,dim(L1))))}
\]

(B) Simulate the experiment of rolling a single die and compare your empirical results with the results in Matched Problem 7.

**Empirical Probabilities for an Insurance Company**

An insurance company selected 1,000 drivers at random in a particular city to determine a relationship between age and accidents. The data obtained are listed in Table 1. Compute the approximate empirical probabilities of the following events for a driver chosen at random in the city:

(A) \( E_1 \): being under 20 years old and having exactly 3 accidents in 1 year

(B) \( E_2 \): being 30–39 years old and having 1 or more accidents in 1 year
(C) \( E_3 \): having no accidents in 1 year
(D) \( E_4 \): being under 20 years old or having exactly 3 accidents in 1 year

<table>
<thead>
<tr>
<th>Age</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Over 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 20</td>
<td>50</td>
<td>62</td>
<td>53</td>
<td>35</td>
<td>20</td>
</tr>
<tr>
<td>20–29</td>
<td>64</td>
<td>93</td>
<td>67</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td>30–39</td>
<td>82</td>
<td>68</td>
<td>32</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>40–49</td>
<td>38</td>
<td>32</td>
<td>20</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Over 49</td>
<td>43</td>
<td>50</td>
<td>35</td>
<td>28</td>
<td>24</td>
</tr>
</tbody>
</table>

### Solutions

(A) \( P(E_1) \approx \frac{35}{1,000} = .035 \)

(B) \( P(E_2) \approx \frac{68 + 32 + 14 + 4}{1,000} = .118 \)

(C) \( P(E_3) \approx \frac{50 + 64 + 82 + 38 + 43}{1,000} = .277 \)

(D) \( P(E_4) \approx \frac{50 + 62 + 53 + 35 + 20 + 40 + 14 + 7 + 28}{1,000} = .309 \)

Notice that in this type of problem, which is typical of many realistic problems, approximate empirical probabilities are the only type we can compute.

### Matched Problem

Referring to Table 1 in Example 8, compute the approximate empirical probabilities of the following events for a driver chosen at random in the city:

(A) \( E_1 \): being under 20 years old with no accidents in 1 year
(B) \( E_2 \): being 20–29 years old and having fewer than 2 accidents in 1 year
(C) \( E_3 \): not being over 49 years old

Approximate empirical probabilities are often used to test theoretical probabilities. Equally likely assumptions may not be justified in reality. In addition to this use, there are many situations in which it is either very difficult or impossible to compute the theoretical probabilities for given events. For example, insurance companies use past experience to establish approximate empirical probabilities to predict future accident rates, baseball teams use batting averages, which are approximate empirical probabilities based on past experience, to predict the future performance of a player, and pollsters use approximate empirical probabilities to predict outcomes of elections.
Answers to Matched Problems

1. (A) \( S_A = \{BB, BG, GB, GG\} \):
   
   \[
   \begin{array}{ccc}
   & \text{Sex of First Child} & \text{Sex of Second Child} \\
   B & B & B \\
   G & B & G \\
   B & G & B \\
   G & G & G \\
   \end{array}
   \]
   
   (B) \( S_B = \{0, 1, 2\} \)  (C) \( S_C = \{A, D\} \)  (D) The sample space in part A.

2. (A) \( \{(4, 1), (3, 2), (2, 3), (1, 4)\} \)  (B) \( \{(6, 5), (5, 6)\} \)

3. (A) .74  (B) .76  (C) .5

4. (A) \( P(E_1) = \frac{1}{5} \)  (B) \( P(E_2) = \frac{1}{18} \)  
5. \( C_{13,2}/C_{52,7} = .000013 \)
6. \( C_{12,4} \cdot C_{16,2}/C_{28,6} = .158 \)
7. (A) \( P(E_1) \approx .178, P(E_2) = .166, P(E_3) = .154, P(E_4) = .182, P(E_5) = .144, P(E_6) = .176 \)  
   (B) \( \frac{5}{6} \approx .167 \) for each  
   (C) 83.3 for each
8. (A) \( P(E_1) \approx .05 \)  (B) \( P(E_2) = .157 \)  (C) \( P(E_3) = .82 \)

---

EXERCISE 6-5

A

1. How would you interpret \( P(E) = 1 \)?
2. How would you interpret \( P(E) = 0 \)?
3. A spinner can land on 4 different colors: red (R), green (G), yellow (Y), and blue (B). If we do not assume each color is as likely to turn up as any other, which of the following probability assignments have to be rejected, and why?
   (A) \( P(R) = .15, P(G) = -.35, P(Y) = .50, P(B) = .70 \)
   (B) \( P(R) = .32, P(G) = .28, P(Y) = .24, P(B) = .30 \)
   (C) \( P(R) = .26, P(G) = .14, P(Y) = .30, P(B) = .30 \)
4. Under the probability assignments in Problem 3, part C, what is the probability that the spinner will not land on blue?
5. Under the probability assignments in Problem 3, part C, what is the probability that the spinner will land on red or yellow?
6. Under the probability assignments in Problem 3, part C, what is the probability that the spinner will not land on red or yellow?
7. A ski jumper has jumped over 300 feet in 25 out of 250 jumps. What is the approximate empirical probability of the next jump being over 300 feet?
8. In a certain city there are 4,000 youths between 16 and 20 years old who drive cars. If 560 of them were involved in accidents last year, what is the approximate empirical probability of a youth in this age group being involved in an accident this year?
9. Out of 420 times at bat, a baseball player gets 189 hits. What is the approximate empirical probability that the player will get a hit next time at bat?
10. In a medical experiment, a new drug is found to help 2,400 out of 3,000 people. If a doctor prescribes the drug for a particular patient, what is the approximate empirical probability that the patient will be helped?

B

11. A small combination lock on a suitcase has 3 wheels, each labeled with the 10 digits from 0 to 9. If an opening combination is a particular sequence of 3 digits with no repeats, what is the probability of a person guessing the right combination?
12. A combination lock has 5 wheels, each labeled with the 10 digits from 0 to 9. If an opening combination is a particular sequence of 5 digits with no repeats, what is the probability of a person guessing the right combination?

An experiment consists of dealing 5 cards from a standard 52-card deck. In Problems 13–16, what is the probability of being dealt each of the following hands?

13. 5 black cards
14. 5 hearts
15. 5 face cards if aces are considered to be face cards
16. 5 nonface cards if an ace is considered to be 1 and not a face card
17. If 4-digit numbers less than 5,000 are randomly formed from the digits 1, 3, 5, 7, and 9, what is the probability of forming a number divisible by 5? Digits may be repeated; for example, 1,355 is acceptable.
18. If code words of four letters are generated at random using the letters A, B, C, D, E, and F, what is the probability of forming a word without a vowel in it? Letters may be repeated.

19. Suppose 5 thank-you notes are written and 5 envelopes are addressed. Accidentally, the notes are randomly inserted into the envelopes and mailed without checking the addresses. What is the probability that all 5 notes will be inserted into the correct envelopes?

20. Suppose 6 people check their coats in a checkroom. If all claim checks are lost and the 6 coats are randomly returned, what is the probability that all 6 people will get their own coats back?

An experiment consists of rolling 2 fair dice and adding the dots on the 2 sides facing up. Using the sample space shown in Figure 1 (page 495) and assuming each simple event is as likely as any other, find the probabilities of the sums of dots indicated in Problems 21—36.

21. Sum is 2.  
22. Sum is 10.

23. Sum is 6.  
24. Sum is 8.

25. Sum is less than 5.  
26. Sum is greater than 8.

27. Sum is not 7 or 11.  
28. Sum is not 2, 4, or 6.

29. Sum is 1.  
30. Sum is not 13.

31. Sum is divisible by 3.  
32. Sum is divisible by 4.

33. Sum is 7 or 11 (a “natural”).  
34. Sum is 2, 3, or 12 (“craps”).

35. Sum is divisible by 2 or 3.  
36. Sum is divisible by 2 and 3.

37. Five thousand people work in a large auto plant. An individual is selected at random and his or her birthday (month and day, not year) is recorded. Set up an appropriate sample space for this experiment and assign acceptable probabilities to the simple events.

38. In a hotly contested three-way race for governor of Minnesota, the leading candidates are running neck-and-neck while the third candidate is receiving half the support of either of the others. Registered voters are chosen at random and are asked of which of the three they are most likely to vote. Set up an appropriate sample space for the random survey experiment and assign acceptable probabilities to the simple events.

39. A pair of dice is rolled 500 times with the following frequencies:

<table>
<thead>
<tr>
<th>Sum</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>71</td>
</tr>
<tr>
<td>7</td>
<td>89</td>
</tr>
<tr>
<td>8</td>
<td>72</td>
</tr>
<tr>
<td>9</td>
<td>52</td>
</tr>
<tr>
<td>10</td>
<td>36</td>
</tr>
<tr>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
</tr>
</tbody>
</table>

40. Three coins are flipped 500 times with the following frequencies of outcomes:

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 heads</td>
<td>58</td>
</tr>
<tr>
<td>2 heads</td>
<td>198</td>
</tr>
<tr>
<td>1 head</td>
<td>190</td>
</tr>
<tr>
<td>0 heads</td>
<td>54</td>
</tr>
</tbody>
</table>

41. (A) Is it possible to get 29 heads in 30 flips of a fair coin? Explain.

(B) If you flip a coin 50 times and get 42 heads, would you suspect that the coin was unfair? Why or why not? If you suspect an unfair coin, what empirical probabilities would you assign to the simple events of the sample space?

42. (A) Is it possible to get 9 double sixes in 12 rolls of a pair of fair dice? Explain.

(B) If you roll a pair of dice 40 times and get 14 double sixes, would you suspect that the dice were unfair? Why or why not? If you suspect loaded dice, what empirical probability would you assign to the event of rolling a double six?

An experiment consists of tossing 3 fair coins, but 1 of the 3 coins has a head on both sides. Compute the probabilities of obtaining the indicated results in Problems 43—48.

43. 1 head  
44. 2 heads

45. 3 heads  
46. 0 heads

47. More than 1 head  
48. More than 1 tail
An experiment consists of rolling 2 fair dice and adding the dots on the 2 sides facing up. Each die has 1 dot on two opposite faces, 2 dots on two opposite faces, and 3 dots on two opposite faces. Compute the probabilities of obtaining the indicated sums in Problems 49–56.

49. 2
50. 3
51. 4
52. 5
53. 6
54. 7
55. An odd sum
56. An even sum

An experiment consists of dealing 5 cards from a standard 52-card deck. In Problems 57–64, what is the probability of being dealt the following cards?

57. 5 cards, jacks through aces
58. 5 cards, 2 through 10
59. 4 aces
60. 4 of a kind
61. Straight flush, ace high; that is, 10, jack, queen, king, ace in one suit
62. Straight flush, starting with 2; that is, 2, 3, 4, 5, 6 in one suit
63. 2 aces and 3 queens
64. 2 kings and 3 aces

**APPLICATIONS**

65. Market Analysis. A company selected 1,000 households at random and surveyed them to determine a relationship between income level and the number of television sets in a home. The information gathered is listed in the table:

<table>
<thead>
<tr>
<th>Yearly Income</th>
<th>Televisions per Household</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Less than $12,000</td>
<td>0</td>
</tr>
<tr>
<td>$12,000–19,999</td>
<td>0</td>
</tr>
<tr>
<td>$20,000–39,999</td>
<td>2</td>
</tr>
<tr>
<td>$40,000–59,999</td>
<td>10</td>
</tr>
<tr>
<td>$60,000 or more</td>
<td>30</td>
</tr>
</tbody>
</table>

Compute the approximate empirical probabilities:

(A) Of a household earning $12,000–$19,999 per year and owning exactly 3 television sets
(B) Of a household earning $20,000–$39,999 per year and owning more than 1 television set
(C) Of a household earning $60,000 or more per year or owning more than 3 television sets
(D) Of a household not owning 0 television sets

66. Market Analysis. Use the sample results in Problem 65 to compute the approximate empirical probabilities:

(A) Of a household earning $40,000–$59,999 per year and owning 0 television sets
(B) Of a household earning $12,000–$39,999 per year and owning more than 2 television sets
(C) Of a household earning less than $20,000 per year or owning exactly 2 television sets
(D) Of a household not owning more than 3 television sets

---

**Section 6-6 Binomial Formula**

- Pascal’s Triangle
- The Binomial Formula
- Proof of the Binomial Formula

The binomial form

\[(a + b)^n\]

where \(n\) is a natural number, appears more frequently than you might expect. It turns out that the coefficients in the expansion are related to probability concepts that we have already discussed.