



This is “Basic Concepts of Probability”, chapter 3 from the book [Beginning Statistics \(index.html\)](#) (v. 1.0).

This book is licensed under a [Creative Commons by-nc-sa 3.0](http://creativecommons.org/licenses/by-nc-sa/3.0/) license. See the license for more details, but that basically means you can share this book as long as you credit the author (but see below), don't make money from it, and do make it available to everyone else under the same terms.

This content was accessible as of December 29, 2012, and it was downloaded then by [Andy Schmitz](#) (<http://lardbucket.org>) in an effort to preserve the availability of this book.

Normally, the author and publisher would be credited here. However, the publisher has asked for the customary Creative Commons attribution to the original publisher, authors, title, and book URI to be removed. Additionally, per the publisher's request, their name has been removed in some passages. More information is available on this project's [attribution page](http://2012books.lardbucket.org/attribution.html?utm_source=header).

For more information on the source of this book, or why it is available for free, please see [the project's home page](#) (<http://2012books.lardbucket.org/>). You can browse or download additional books there.

## Chapter 3

---

### Basic Concepts of Probability

Suppose a polling organization questions 1,200 voters in order to estimate the proportion of all voters who favor a particular bond issue. We would expect the proportion of the 1,200 voters in the survey who are in favor to be close to the proportion of all voters who are in favor, but this need not be true. There is a degree of randomness associated with the survey result. If the survey result is highly likely to be close to the true proportion, then we have confidence in the survey result. If it is not particularly likely to be close to the population proportion, then we would perhaps not take the survey result too seriously. The likelihood that the survey proportion is close to the population proportion determines our confidence in the survey result. For that reason, we would like to be able to compute that likelihood. The task of computing it belongs to the realm of probability, which we study in this chapter.

### 3.1 Sample Spaces, Events, and Their Probabilities

#### LEARNING OBJECTIVES

1. To learn the concept of the sample space associated with a random experiment.
2. To learn the concept of an event associated with a random experiment.
3. To learn the concept of the probability of an event.

#### Sample Spaces and Events

Rolling an ordinary six-sided die is a familiar example of a *random experiment*, an action for which all possible outcomes can be listed, but for which the actual outcome on any given trial of the experiment cannot be predicted with certainty. In such a situation we wish to assign to each outcome, such as rolling a two, a number, called the *probability* of the outcome, that indicates how likely it is that the outcome will occur. Similarly, we would like to assign a probability to any *event*, or collection of outcomes, such as rolling an even number, which indicates how likely it is that the event will occur if the experiment is performed. This section provides a framework for discussing probability problems, using the terms just mentioned.

#### Definition

A **random experiment** is a mechanism that produces a definite outcome that cannot be predicted with certainty. The **sample space**<sup>1</sup> associated with a random experiment is the set of all possible outcomes. An **event**<sup>2</sup> is a subset of the sample space.

#### Definition

An event  $E$  is said to **occur** on a particular trial of the experiment if the outcome observed is an element of the set  $E$ .

1. The set of all possible outcomes of a random experiment.
2. Any set of outcomes.

## EXAMPLE 1

Construct a sample space for the experiment that consists of tossing a single coin.

Solution:

The outcomes could be labeled  $h$  for heads and  $t$  for tails. Then the sample space is the set  $S = \{h, t\}$ .

## EXAMPLE 2

Construct a sample space for the experiment that consists of rolling a single die. Find the events that correspond to the phrases “an even number is rolled” and “a number greater than two is rolled.”

Solution:

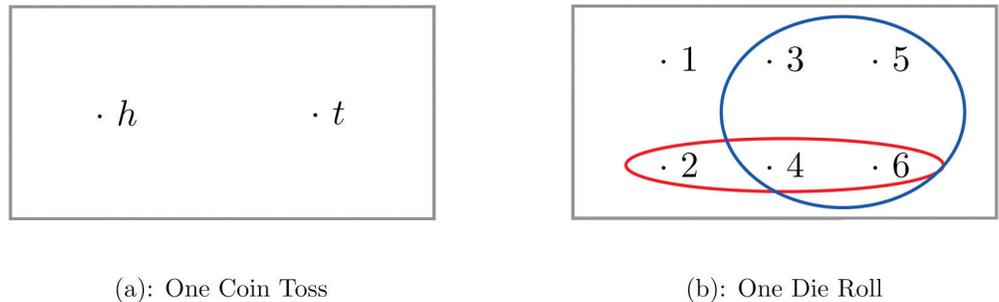
The outcomes could be labeled according to the number of dots on the top face of the die. Then the sample space is the set  $S = \{1, 2, 3, 4, 5, 6\}$ .

The outcomes that are even are 2, 4, and 6, so the event that corresponds to the phrase “an even number is rolled” is the set  $\{2, 4, 6\}$ , which it is natural to denote by the letter  $E$ . We write  $E = \{2, 4, 6\}$ .

Similarly the event that corresponds to the phrase “a number greater than two is rolled” is the set  $T = \{3, 4, 5, 6\}$ , which we have denoted  $T$ .

A graphical representation of a sample space and events is a **Venn diagram**, as shown in [Figure 3.1 "Venn Diagrams for Two Sample Spaces"](#) for [Note 3.6 "Example 1"](#) and [Note 3.7 "Example 2"](#). In general the sample space  $S$  is represented by a rectangle, outcomes by points within the rectangle, and events by ovals that enclose the outcomes that compose them.

Figure 3.1 Venn Diagrams for Two Sample Spaces



### EXAMPLE 3

A random experiment consists of tossing two coins.

- Construct a sample space for the situation that the coins are indistinguishable, such as two brand new pennies.
- Construct a sample space for the situation that the coins are distinguishable, such as one a penny and the other a nickel.

Solution:

- After the coins are tossed one sees either two heads, which could be labeled  $2h$ , two tails, which could be labeled  $2t$ , or coins that differ, which could be labeled  $d$ . Thus a sample space is  $S = \{2h, 2t, d\}$ .
- Since we can tell the coins apart, there are now two ways for the coins to differ: the penny heads and the nickel tails, or the penny tails and the nickel heads. We can label each outcome as a pair of letters, the first of which indicates how the penny landed and the second of which indicates how the nickel landed. A sample space is then  $S' = \{hh, ht, th, tt\}$ .

A device that can be helpful in identifying all possible outcomes of a random experiment, particularly one that can be viewed as proceeding in stages, is what is called a **tree diagram**. It is described in the following example.

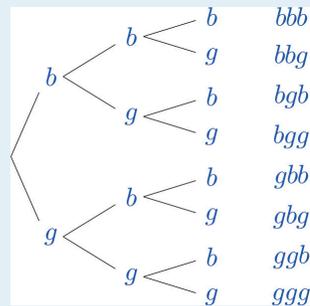
## EXAMPLE 4

Construct a sample space that describes all three-child families according to the genders of the children with respect to birth order.

Solution:

Two of the outcomes are “two boys then a girl,” which we might denote  $bbg$ , and “a girl then two boys,” which we would denote  $gbb$ . Clearly there are many outcomes, and when we try to list all of them it could be difficult to be sure that we have found them all unless we proceed systematically. The tree diagram shown in [Figure 3.2 "Tree Diagram For Three-Child Families"](#), gives a systematic approach.

Figure 3.2  
Tree Diagram For  
Three-Child Families



The diagram was constructed as follows. There are two possibilities for the first child, boy or girl, so we draw two line segments coming out of a starting point, one ending in a  $b$  for “boy” and the other ending in a  $g$  for “girl.” For each of these two possibilities for the first child there are two possibilities for the second child, “boy” or “girl,” so from each of the  $b$  and  $g$  we draw two line segments, one segment ending in a  $b$  and one in a  $g$ . For each of the four ending points now in the diagram there are two possibilities for the third child, so we repeat the process once more.

The line segments are called **branches** of the tree. The right ending point of each branch is called a **node**. The nodes on the extreme right are the **final nodes**; to each one there corresponds an outcome, as shown in the figure.

From the tree it is easy to read off the eight outcomes of the experiment, so the sample space is, reading from the top to the bottom of the final nodes in the tree,

$$S = \{bbb, bbg, bgb, bgg, gbb, gbg, ggb, ggg\}$$

## Probability

### Definition

The **probability of an outcome**<sup>3</sup>  $e$  in a sample space  $S$  is a number  $p$  between 0 and 1 that measures the likelihood that  $e$  will occur on a single trial of the corresponding random experiment. The value  $p = 0$  corresponds to the outcome  $e$  being impossible and the value  $p = 1$  corresponds to the outcome  $e$  being certain.

### Definition

The **probability of an event**<sup>4</sup>  $A$  is the sum of the probabilities of the individual outcomes of which it is composed. It is denoted  $P(A)$ .

The following formula expresses the content of the definition of the probability of an event:

If an event  $E$  is  $E = \{e_1, e_2, \dots, e_k\}$ , then

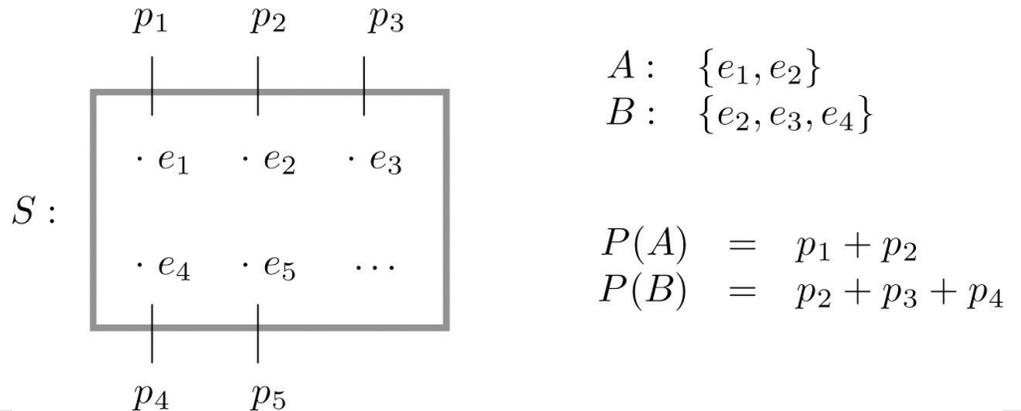
$$P(E) = P(e_1) + P(e_2) + \dots + P(e_k)$$

3. A number that measures the likelihood of the outcome.

4. A number that measures the likelihood of the event.

**Figure 3.3 "Sample Spaces and Probability"** graphically illustrates the definitions.

Figure 3.3 Sample Spaces and Probability



Since the whole sample space  $S$  is an event that is certain to occur, the sum of the probabilities of all the outcomes must be the number 1.

In ordinary language probabilities are frequently expressed as percentages. For example, we would say that there is a 70% chance of rain tomorrow, meaning that the probability of rain is 0.70. We will use this practice here, but in all the computational formulas that follow we will use the form 0.70 and not 70%.

### EXAMPLE 5

A coin is called “balanced” or “fair” if each side is equally likely to land up. Assign a probability to each outcome in the sample space for the experiment that consists of tossing a single fair coin.

**Solution:**

With the outcomes labeled  $h$  for heads and  $t$  for tails, the sample space is the set  $S = \{h, t\}$ . Since the outcomes have the same probabilities, which must add up to 1, each outcome is assigned probability  $1/2$ .

## EXAMPLE 6

A die is called “balanced” or “fair” if each side is equally likely to land on top. Assign a probability to each outcome in the sample space for the experiment that consists of tossing a single fair die. Find the probabilities of the events  $E$ : “an even number is rolled” and  $T$ : “a number greater than two is rolled.”

Solution:

With outcomes labeled according to the number of dots on the top face of the die, the sample space is the set  $S = \{1,2,3,4,5,6\}$ . Since there are six equally likely outcomes, which must add up to 1, each is assigned probability  $1/6$ .

$$\begin{aligned} \text{Since } E &= \{2,4,6\}, \\ P(E) &= 1/6 + 1/6 + 1/6 = 3/6 = 1/2. \end{aligned}$$

$$\text{Since } T = \{3,4,5,6\}, P(T) = 4/6 = 2/3.$$

## EXAMPLE 7

Two fair coins are tossed. Find the probability that the coins match, i.e., either both land heads or both land tails.

Solution:

In [Note 3.8 "Example 3"](#) we constructed the sample space  $S = \{2h, 2t, d\}$  for the situation in which the coins are identical and the sample space  $S' = \{hh, ht, th, tt\}$  for the situation in which the two coins can be told apart.

The theory of probability does not tell us *how* to assign probabilities to the outcomes, only what to do with them once they are assigned. Specifically, using sample space  $S$ , matching coins is the event  $M = \{2h, 2t\}$ , which has probability  $P(2h) + P(2t)$ . Using sample space  $S'$ , matching coins is the event  $M' = \{hh, tt\}$ , which has probability  $P(hh) + P(tt)$ . In the physical world it should make no difference whether the coins are identical or not, and so we would like to assign probabilities to the outcomes so that the numbers  $P(M)$  and  $P(M')$  are the same and best match what we observe when actual physical experiments are performed with coins that seem to be fair. Actual experience suggests that the outcomes in  $S'$  are equally likely, so we assign to each probability  $1/4$ , and then

$$P(M') = P(hh) + P(tt) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Similarly, from experience appropriate choices for the outcomes in  $S$  are:

$$P(2h) = \frac{1}{4} \quad P(2t) = \frac{1}{4} \quad P(d) = \frac{1}{2}$$

which give the same final answer

$$P(M) = P(2h) + P(2t) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

The previous three examples illustrate how probabilities can be computed simply by counting when the sample space consists of a finite number of equally likely

outcomes. In some situations the individual outcomes of any sample space that represents the experiment are unavoidably unequally likely, in which case probabilities cannot be computed merely by counting, but the computational formula given in the definition of the probability of an event must be used.

### EXAMPLE 8

The breakdown of the student body in a local high school according to race and ethnicity is 51% white, 27% black, 11% Hispanic, 6% Asian, and 5% for all others. A student is randomly selected from this high school. (To select “randomly” means that every student has the same chance of being selected.) Find the probabilities of the following events:

- $B$ : the student is black,
- $M$ : the student is minority (that is, not white),
- $N$ : the student is not black.

Solution:

The experiment is the action of randomly selecting a student from the student population of the high school. An obvious sample space is  $S = \{w, b, h, a, o\}$ . Since 51% of the students are white and all students have the same chance of being selected,  $P(w) = 0.51$ , and similarly for the other outcomes. This information is summarized in the following table:

| Outcome     | $w$  | $b$  | $h$  | $a$  | $o$  |
|-------------|------|------|------|------|------|
| Probability | 0.51 | 0.27 | 0.11 | 0.06 | 0.05 |

- Since  $B = \{b\}$ ,  $P(B) = P(b) = 0.27$ .
- Since  $M = \{b, h, a, o\}$ ,  
 $P(M) = P(b) + P(h) + P(a) + P(o) = 0.27 + 0.11 + 0.06 + 0.05 = 0.49$
- Since  $N = \{w, h, a, o\}$ ,  
 $P(N) = P(w) + P(h) + P(a) + P(o) = 0.51 + 0.11 + 0.06 + 0.05 = 0.73$

## EXAMPLE 9

The student body in the high school considered in [Note 3.18 "Example 8"](#) may be broken down into ten categories as follows: 25% white male, 26% white female, 12% black male, 15% black female, 6% Hispanic male, 5% Hispanic female, 3% Asian male, 3% Asian female, 1% male of other minorities combined, and 4% female of other minorities combined. A student is randomly selected from this high school. Find the probabilities of the following events:

- $B$ : the student is black,
- $MF$ : the student is minority female,
- $FN$ : the student is female and is not black.

Solution:

Now the sample space is

$S = \{wm, bm, hm, am, om, wf, bf, hf, af, of\}$ . The information given in the example can be summarized in the following table, called a *two-way contingency table*:

| Gender | Race / Ethnicity |       |          |       |        |
|--------|------------------|-------|----------|-------|--------|
|        | White            | Black | Hispanic | Asian | Others |
| Male   | 0.25             | 0.12  | 0.06     | 0.03  | 0.01   |
| Female | 0.26             | 0.15  | 0.05     | 0.03  | 0.04   |

- Since  $B = \{bm, bf\}$ ,  

$$P(B) = P(bm) + P(bf) = 0.12 + 0.15 = 0.27.$$
- Since  $MF = \{bf, hf, af, of\}$ ,  

$$P(M) = P(bf) + P(hf) + P(af) + P(of) = 0.15 + 0.05 + 0.03 + 0.04 = 0.27.$$
- Since  $FN = \{wf, hf, af, of\}$ ,  

$$P(FN) = P(wf) + P(hf) + P(af) + P(of) = 0.26 + 0.05 + 0.03 + 0.04 = 0.38.$$

### KEY TAKEAWAYS

- The sample space of a random experiment is the collection of all possible outcomes.
- An event associated with a random experiment is a subset of the sample space.
- The probability of any outcome is a number between 0 and 1. The probabilities of all the outcomes add up to 1.
- The probability of any event  $A$  is the sum of the probabilities of the outcomes in  $A$ .

## EXERCISES

## BASIC

1. A box contains 10 white and 10 black marbles. Construct a sample space for the experiment of randomly drawing out, with replacement, two marbles in succession and noting the color each time. (To draw “with replacement” means that the first marble is put back before the second marble is drawn.)
2. A box contains 16 white and 16 black marbles. Construct a sample space for the experiment of randomly drawing out, with replacement, three marbles in succession and noting the color each time. (To draw “with replacement” means that each marble is put back before the next marble is drawn.)
3. A box contains 8 red, 8 yellow, and 8 green marbles. Construct a sample space for the experiment of randomly drawing out, with replacement, two marbles in succession and noting the color each time.
4. A box contains 6 red, 6 yellow, and 6 green marbles. Construct a sample space for the experiment of randomly drawing out, with replacement, three marbles in succession and noting the color each time.
5. In the situation of Exercise 1, list the outcomes that comprise each of the following events.
  - a. At least one marble of each color is drawn.
  - b. No white marble is drawn.
6. In the situation of Exercise 2, list the outcomes that comprise each of the following events.
  - a. At least one marble of each color is drawn.
  - b. No white marble is drawn.
  - c. More black than white marbles are drawn.
7. In the situation of Exercise 3, list the outcomes that comprise each of the following events.
  - a. No yellow marble is drawn.
  - b. The two marbles drawn have the same color.
  - c. At least one marble of each color is drawn.
8. In the situation of Exercise 4, list the outcomes that comprise each of the following events.
  - a. No yellow marble is drawn.

- b. The three marbles drawn have the same color.  
 c. At least one marble of each color is drawn.
9. Assuming that each outcome is equally likely, find the probability of each event in Exercise 5.
10. Assuming that each outcome is equally likely, find the probability of each event in Exercise 6.
11. Assuming that each outcome is equally likely, find the probability of each event in Exercise 7.
12. Assuming that each outcome is equally likely, find the probability of each event in Exercise 8.
13. A sample space is  $S = \{a, b, c, d, e\}$ . Identify two events as  $U = \{a, b, d\}$  and  $V = \{b, c, d\}$ . Suppose  $P(a)$  and  $P(b)$  are each 0.2 and  $P(c)$  and  $P(d)$  are each 0.1.
- a. Determine what  $P(e)$  must be.  
 b. Find  $P(U)$ .  
 c. Find  $P(V)$ .
14. A sample space is  $S = \{u, v, w, x\}$ . Identify two events as  $A = \{v, w\}$  and  $B = \{u, w, x\}$ . Suppose  $P(u) = 0.22$ ,  $P(w) = 0.36$ , and  $P(x) = 0.27$ .
- a. Determine what  $P(v)$  must be.  
 b. Find  $P(A)$ .  
 c. Find  $P(B)$ .
15. A sample space is  $S = \{m, n, q, r, s\}$ . Identify two events as  $U = \{m, q, s\}$  and  $V = \{n, q, r\}$ . The probabilities of some of the outcomes are given by the following table:
- |             |      |      |      |      |     |
|-------------|------|------|------|------|-----|
| Outcome     | $m$  | $n$  | $q$  | $r$  | $s$ |
| Probability | 0.18 | 0.16 | 0.24 | 0.21 |     |
- a. Determine what  $P(q)$  must be.  
 b. Find  $P(U)$ .  
 c. Find  $P(V)$ .
16. A sample space is  $S = \{d, e, f, g, h\}$ . Identify two events as  $M = \{e, f, g, h\}$  and  $N = \{d, g\}$ . The probabilities of some of the outcomes are given by the following table:

| Outcome     | $d$  | $e$  | $f$  | $g$ | $h$  |
|-------------|------|------|------|-----|------|
| Probability | 0.22 | 0.13 | 0.27 |     | 0.19 |

- Determine what  $P(g)$  must be.
- Find  $P(M)$ .
- Find  $P(N)$ .

### APPLICATIONS

- The sample space that describes all three-child families according to the genders of the children with respect to birth order was constructed in [Note 3.9 "Example 4"](#). Identify the outcomes that comprise each of the following events in the experiment of selecting a three-child family at random.
  - At least one child is a girl.
  - At most one child is a girl.
  - All of the children are girls.
  - Exactly two of the children are girls.
  - The first born is a girl.
- The sample space that describes three tosses of a coin is the same as the one constructed in [Note 3.9 "Example 4"](#) with “boy” replaced by “heads” and “girl” replaced by “tails.” Identify the outcomes that comprise each of the following events in the experiment of tossing a coin three times.
  - The coin lands heads more often than tails.
  - The coin lands heads the same number of times as it lands tails.
  - The coin lands heads at least twice.
  - The coin lands heads on the last toss.
- Assuming that the outcomes are equally likely, find the probability of each event in Exercise 17.
- Assuming that the outcomes are equally likely, find the probability of each event in Exercise 18.

### ADDITIONAL EXERCISES

- The following two-way contingency table gives the breakdown of the population in a particular locale according to age and tobacco usage:

| Age      | Tobacco Use |            |
|----------|-------------|------------|
|          | Smoker      | Non-smoker |
| Under 30 | 0.05        | 0.20       |
| Over 30  | 0.20        | 0.55       |

A person is selected at random. Find the probability of each of the following events.

- The person is a smoker.
  - The person is under 30.
  - The person is a smoker who is under 30.
22. The following two-way contingency table gives the breakdown of the population in a particular locale according to party affiliation (*A*, *B*, *C*, or *None*) and opinion on a bond issue:

| Affiliation | Opinion |         |           |
|-------------|---------|---------|-----------|
|             | Favors  | Opposes | Undecided |
| <i>A</i>    | 0.12    | 0.09    | 0.07      |
| <i>B</i>    | 0.16    | 0.12    | 0.14      |
| <i>C</i>    | 0.04    | 0.03    | 0.06      |
| <i>None</i> | 0.08    | 0.06    | 0.03      |

A person is selected at random. Find the probability of each of the following events.

- The person is affiliated with party *B*.
  - The person is affiliated with some party.
  - The person is in favor of the bond issue.
  - The person has no party affiliation and is undecided about the bond issue.
23. The following two-way contingency table gives the breakdown of the population of married or previously married women beyond child-bearing age in a particular locale according to age at first marriage and number of children:

| Age      | Number of Children |        |           |
|----------|--------------------|--------|-----------|
|          | 0                  | 1 or 2 | 3 or More |
| Under 20 | 0.02               | 0.14   | 0.08      |
| 20–29    | 0.07               | 0.37   | 0.11      |

| Age          | Number of Children |        |           |
|--------------|--------------------|--------|-----------|
|              | 0                  | 1 or 2 | 3 or More |
| 30 and above | 0.10               | 0.10   | 0.01      |

A woman is selected at random. Find the probability of each of the following events.

- The woman was in her twenties at her first marriage.
  - The woman was 20 or older at her first marriage.
  - The woman had no children.
  - The woman was in her twenties at her first marriage and had at least three children.
24. The following two-way contingency table gives the breakdown of the population of adults in a particular locale according to highest level of education and whether or not the individual regularly takes dietary supplements:

| Education              | Use of Supplements |               |
|------------------------|--------------------|---------------|
|                        | Takes              | Does Not Take |
| No High School Diploma | 0.04               | 0.06          |
| High School Diploma    | 0.06               | 0.44          |
| Undergraduate Degree   | 0.09               | 0.28          |
| Graduate Degree        | 0.01               | 0.02          |

An adult is selected at random. Find the probability of each of the following events.

- The person has a high school diploma and takes dietary supplements regularly.
- The person has an undergraduate degree and takes dietary supplements regularly.
- The person takes dietary supplements regularly.
- The person does not take dietary supplements regularly.

### LARGE DATA SET EXERCISES

25. Large Data Sets 4 and 4A record the results of 500 tosses of a coin. Find the relative frequency of each outcome 1, 2, 3, 4, 5, and 6. Does the coin appear to be “balanced” or “fair”?

<http://www.gone.2012books.lardbucket.org/sites/all/files/data4.xls>

<http://www.gone.2012books.lardbucket.org/sites/all/files/data4A.xls>

26. Large Data Sets 6, 6A, and 6B record results of a random survey of 200 voters in each of two regions, in which they were asked to express whether they prefer Candidate A for a U.S. Senate seat or prefer some other candidate.
- Find the probability that a randomly selected voter among these 400 prefers Candidate A.
  - Find the probability that a randomly selected voter among the 200 who live in Region 1 prefers Candidate A (separately recorded in Large Data Set 6A).
  - Find the probability that a randomly selected voter among the 200 who live in Region 2 prefers Candidate A (separately recorded in Large Data Set 6B).

<http://www.gone.2012books.lardbucket.org/sites/all/files/data6.xls>

<http://www.gone.2012books.lardbucket.org/sites/all/files/data6A.xls>

<http://www.gone.2012books.lardbucket.org/sites/all/files/data6B.xls>

## ANSWERS

1.  $S = \{bb, bw, wb, ww\}$
3.  $S = \{rr, ry, rg, yr, yy, yg, gr, gy, gg\}$
5.
  - a.  $\{bw, wb\}$
  - b.  $\{bb\}$
7.
  - a.  $\{rr, rg, gr, gg\}$
  - b.  $\{rr, yy, gg\}$
  - c.  $\emptyset$
9.
  - a.  $2/4$
  - b.  $1/4$
11.
  - a.  $4/9$
  - b.  $3/9$
  - c.  $0$
13.
  - a.  $0.4$
  - b.  $0.5$
  - c.  $0.4$
15.
  - a.  $0.21$
  - b.  $0.6$
  - c.  $0.61$
17.
  - a.  $\{bbg, bgb, bgg, gbb, gbg, ggb, ggg\}$
  - b.  $\{bbb, bbg, bgb, gbb\}$
  - c.  $\{ggg\}$
  - d.  $\{bgg, gbg, ggb\}$
  - e.  $\{gbb, gbg, ggb, ggg\}$
19.
  - a.  $7/8$
  - b.  $4/8$
  - c.  $1/8$
  - d.  $3/8$
  - e.  $4/8$
21.
  - a.  $0.25$
  - b.  $0.25$
  - c.  $0.05$

23.     a. 0.55  
          b. 0.76  
          c. 0.19  
          d. 0.11
25. The relative frequencies for 1 through 6 are 0.16, 0.194, 0.162, 0.164, 0.154 and 0.166. It would appear that the die is not balanced.

## 3.2 Complements, Intersections, and Unions

### LEARNING OBJECTIVES

1. To learn how some events are naturally expressible in terms of other events.
2. To learn how to use special formulas for the probability of an event that is expressed in terms of one or more other events.

Some events can be naturally expressed in terms of other, sometimes simpler, events.

### Complements

#### Definition

The **complement of an event**<sup>5</sup>  $A$  in a sample space  $S$ , denoted  $A^c$ , is the collection of all outcomes in  $S$  that are not elements of the set  $A$ . It corresponds to negating any description in words of the event  $A$ .

5. The event does not occur.

## EXAMPLE 10

Two events connected with the experiment of rolling a single die are  $E$ : “the number rolled is even” and  $T$ : “the number rolled is greater than two.” Find the complement of each.

Solution:

In the sample space  $S = \{1, 2, 3, 4, 5, 6\}$  the corresponding sets of outcomes are  $E = \{2, 4, 6\}$  and  $T = \{3, 4, 5, 6\}$ . The complements are  $E^c = \{1, 3, 5\}$  and  $T^c = \{1, 2\}$ .

In words the complements are described by “the number rolled is not even” and “the number rolled is not greater than two.” Of course easier descriptions would be “the number rolled is odd” and “the number rolled is less than three.”

If there is a 60% chance of rain tomorrow, what is the probability of fair weather? The obvious answer, 40%, is an instance of the following general rule.

### Probability Rule for Complements

$$P(A^c) = 1 - P(A)$$

This formula is particularly useful when finding the probability of an event directly is difficult.

## EXAMPLE 11

Find the probability that at least one heads will appear in five tosses of a fair coin.

Solution:

Identify outcomes by lists of five *hs* and *ts*, such as *tthtt* and *hhttt*. Although it is tedious to list them all, it is not difficult to count them. Think of using a tree diagram to do so. There are two choices for the first toss. For each of these there are two choices for the second toss, hence  $2 \times 2 = 4$  outcomes for two tosses. For each of these four outcomes, there are two possibilities for the third toss, hence  $4 \times 2 = 8$  outcomes for three tosses. Similarly, there are  $8 \times 2 = 16$  outcomes for four tosses and finally  $16 \times 2 = 32$  outcomes for five tosses.

Let  $O$  denote the event “at least one heads.” There are many ways to obtain at least one heads, but only one way to fail to do so: all tails. Thus although it is difficult to list all the outcomes that form  $O$ , it is easy to write  $O^c = \{ttttt\}$ . Since there are 32 equally likely outcomes, each has probability  $1/32$ , so  $P(O^c) = 1/32$ , hence  $P(O) = 1 - 1/32 \approx 0.97$  or about a 97% chance.

## Intersection of Events

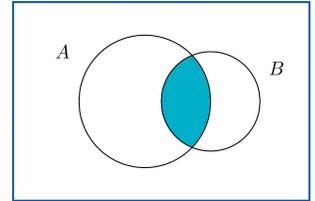
## Definition

The **intersection of events**<sup>6</sup>  $A$  and  $B$ , denoted  $A \cap B$ , is the collection of all outcomes that are elements of both of the sets  $A$  and  $B$ . It corresponds to combining descriptions of the two events using the word “and.”

To say that the event  $A \cap B$  occurred means that on a particular trial of the experiment both  $A$  and  $B$  occurred. A visual representation of the intersection of events  $A$  and  $B$  in a sample space  $S$  is given in [Figure 3.4 "The Intersection of Events"](#). The intersection corresponds to the shaded lens-shaped region that lies within both ovals.

6. Both events occur.

Figure 3.4 The Intersection of Events A and B



### EXAMPLE 12

In the experiment of rolling a single die, find the intersection  $E \cap T$  of the events  $E$ : “the number rolled is even” and  $T$ : “the number rolled is greater than two.”

Solution:

The sample space is  $S = \{1, 2, 3, 4, 5, 6\}$ . Since the outcomes that are common to  $E = \{2, 4, 6\}$  and  $T = \{3, 4, 5, 6\}$  are 4 and 6,  $E \cap T = \{4, 6\}$ .

In words the intersection is described by “the number rolled is even and is greater than two.” The only numbers between one and six that are both even and greater than two are four and six, corresponding to  $E \cap T$  given above.

## EXAMPLE 13

A single die is rolled.

- Suppose the die is fair. Find the probability that the number rolled is both even and greater than two.
- Suppose the die has been “loaded” so that  $P(1) = 1/12$ ,  $P(6) = 3/12$ , and the remaining four outcomes are equally likely with one another. Now find the probability that the number rolled is both even and greater than two.

Solution:

In both cases the sample space is  $S = \{1, 2, 3, 4, 5, 6\}$  and the event in question is the intersection  $E \cap T = \{4, 6\}$  of the previous example.

- Since the die is fair, all outcomes are equally likely, so by counting we have  $P(E \cap T) = 2/6$ .
- The information on the probabilities of the six outcomes that we have so far is

|             |                |     |     |     |     |                |
|-------------|----------------|-----|-----|-----|-----|----------------|
| Outcome     | 1              | 2   | 3   | 4   | 5   | 6              |
| Probability | $\frac{1}{12}$ | $p$ | $p$ | $p$ | $p$ | $\frac{3}{12}$ |

Since  $P(1) + P(6) = 4/12 = 1/3$  and the probabilities of all six outcomes add up to 1,

$$P(2) + P(3) + P(4) + P(5) = 1 - \frac{1}{3} = \frac{2}{3}$$

Thus  $4p = 2/3$ , so  $p = 1/6$ . In particular  $P(4) = 1/6$ .  
Therefore

$$P(E \cap T) = P(4) + P(6) = \frac{1}{6} + \frac{3}{12} = \frac{5}{12}$$

### Definition

Events  $A$  and  $B$  are **mutually exclusive**<sup>7</sup> if they have no elements in common.

For  $A$  and  $B$  to have no outcomes in common means precisely that it is impossible for both  $A$  and  $B$  to occur on a single trial of the random experiment. This gives the following rule.

### Probability Rule for Mutually Exclusive Events

Events  $A$  and  $B$  are mutually exclusive if and only if

$$P(A \cap B) = 0$$

Any event  $A$  and its complement  $A^c$  are mutually exclusive, but  $A$  and  $B$  can be mutually exclusive without being complements.

### EXAMPLE 14

In the experiment of rolling a single die, find three choices for an event  $A$  so that the events  $A$  and  $E$ : “the number rolled is even” are mutually exclusive.

Solution:

Since  $E = \{2,4,6\}$  and we want  $A$  to have no elements in common with  $E$ , any event that does not contain any even number will do. Three choices are  $\{1,3,5\}$  (the complement  $E^c$ , the odds),  $\{1,3\}$ , and  $\{5\}$ .

7. Events that cannot both occur at once.

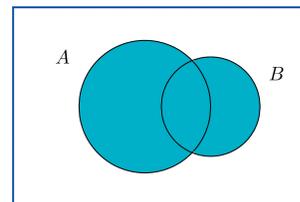
## Union of Events

### Definition

The **union of events**<sup>8</sup>  $A$  and  $B$ , denoted  $A \cup B$ , is the collection of all outcomes that are elements of one or the other of the sets  $A$  and  $B$ , or of both of them. It corresponds to combining descriptions of the two events using the word “or.”

To say that the event  $A \cup B$  occurred means that on a particular trial of the experiment either  $A$  or  $B$  occurred (or both did). A visual representation of the union of events  $A$  and  $B$  in a sample space  $S$  is given in [Figure 3.5 "The Union of Events"](#). The union corresponds to the shaded region.

Figure 3.5 The Union of Events  $A$  and  $B$



8. One or the other event occurs.

## EXAMPLE 15

In the experiment of rolling a single die, find the union of the events  $E$ : “the number rolled is even” and  $T$ : “the number rolled is greater than two.”

Solution:

Since the outcomes that are in either  $E = \{2,4,6\}$  or

$T = \{3,4,5,6\}$  (or both) are 2, 3, 4, 5, and 6,

$E \cup T = \{2,3,4,5,6\}$ . Note that an outcome such as 4 that is in both sets is still listed only once (although strictly speaking it is not incorrect to list it twice).

In words the union is described by “the number rolled is even or is greater than two.” Every number between one and six except the number one is either even or is greater than two, corresponding to  $E \cup T$  given above.

## EXAMPLE 16

A two-child family is selected at random. Let  $B$  denote the event that at least one child is a boy, let  $D$  denote the event that the genders of the two children differ, and let  $M$  denote the event that the genders of the two children match. Find  $B \cup D$  and  $B \cup M$ .

Solution:

A sample space for this experiment is  $S = \{bb, bg, gb, gg\}$ , where the first letter denotes the gender of the firstborn child and the second letter denotes the gender of the second child. The events  $B, D,$  and  $M$  are

$$B = \{bb, bg, gb\} \quad D = \{bg, gb\} \quad M = \{bb, gg\}$$

Each outcome in  $D$  is already in  $B$ , so the outcomes that are in at least one or the other of the sets  $B$  and  $D$  is just the set  $B$  itself:

$$B \cup D = \{bb, bg, gb\} = B.$$

Every outcome in the whole sample space  $S$  is in at least one or the other of the sets  $B$  and  $M$ , so  $B \cup M = \{bb, bg, gb, gg\} = S$ .

The following **Additive Rule of Probability** is a useful formula for calculating the probability of  $A \cup B$ .

### Additive Rule of Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

The next example, in which we compute the probability of a union both by counting and by using the formula, shows why the last term in the formula is needed.

## EXAMPLE 17

Two fair dice are thrown. Find the probabilities of the following events:

- a. both dice show a four
- b. at least one die shows a four

Solution:

As was the case with tossing two identical coins, actual experience dictates that for the sample space to have equally likely outcomes we should list outcomes as if we could distinguish the two dice. We could imagine that one of them is red and the other is green. Then any outcome can be labeled as a pair of numbers as in the following display, where the first number in the pair is the number of dots on the top face of the green die and the second number in the pair is the number of dots on the top face of the red die.

|    |    |    |    |    |    |
|----|----|----|----|----|----|
| 11 | 12 | 13 | 14 | 15 | 16 |
| 21 | 22 | 23 | 24 | 25 | 26 |
| 31 | 32 | 33 | 34 | 35 | 36 |
| 41 | 42 | 43 | 44 | 45 | 46 |
| 51 | 52 | 53 | 54 | 55 | 56 |
| 61 | 62 | 63 | 64 | 65 | 66 |

- a. There are 36 equally likely outcomes, of which exactly one corresponds to two fours, so the probability of a pair of fours is  $1/36$ .
- b. From the table we can see that there are 11 pairs that correspond to the event in question: the six pairs in the fourth row (the green die shows a four) plus the additional five pairs other than the pair 44, already counted, in the fourth column (the red die is four), so the answer is  $11/36$ . To see how the formula gives the same number, let  $A_G$  denote the event that the green die is a four and let  $A_R$  denote the event that the red die is a four. Then clearly by counting we get  $P(A_G) = 6/36$  and  $P(A_R) = 6/36$ . Since  $A_G \cap A_R = \{44\}$ ,  $P(A_G \cap A_R) = 1/36$ , this is the computation in part (a), of course. Thus by the Additive Rule of Probability,

$$P(A_G \cup A_R) = P(A_G) + P(A_R) - P(A_G \cap A_R) = \frac{6}{36} + \frac{6}{36} - \frac{1}{36}$$

**EXAMPLE 18**

A tutoring service specializes in preparing adults for high school equivalence tests. Among all the students seeking help from the service, 63% need help in mathematics, 34% need help in English, and 27% need help in both mathematics and English. What is the percentage of students who need help in either mathematics or English?

Solution:

Imagine selecting a student at random, that is, in such a way that every student has the same chance of being selected. Let  $M$  denote the event “the student needs help in mathematics” and let  $E$  denote the event “the student needs help in English.” The information given is that  $P(M) = 0.63$ ,  $P(E) = 0.34$ , and  $P(M \cap E) = 0.27$ . The Additive Rule of Probability gives

$$P(M \cup E) = P(M) + P(E) - P(M \cap E) = 0.63 + 0.34 - 0.27 = 0.70$$

Note how the naïve reasoning that if 63% need help in mathematics and 34% need help in English then 63 plus 34 or 97% need help in one or the other gives a number that is too large. The percentage that need help in both subjects must be subtracted off, else the people needing help in both are counted twice, once for needing help in mathematics and once again for needing help in English. The simple sum of the probabilities would work if the events in question were mutually exclusive, for then  $P(A \cap B)$  is zero, and makes no difference.

## EXAMPLE 19

Volunteers for a disaster relief effort were classified according to both specialty ( $C$ : construction,  $E$ : education,  $M$ : medicine) and language ability ( $S$ : speaks a single language fluently,  $T$ : speaks two or more languages fluently). The results are shown in the following *two-way classification table*:

| Specialty | Language Ability |     |
|-----------|------------------|-----|
|           | $S$              | $T$ |
| $C$       | 12               | 1   |
| $E$       | 4                | 3   |
| $M$       | 6                | 2   |

The first row of numbers means that 12 volunteers whose specialty is construction speak a single language fluently, and 1 volunteer whose specialty is construction speaks at least two languages fluently. Similarly for the other two rows.

A volunteer is selected at random, meaning that each one has an equal chance of being chosen. Find the probability that:

- his specialty is medicine and he speaks two or more languages;
- either his specialty is medicine or he speaks two or more languages;
- his specialty is something other than medicine.

Solution:

When information is presented in a two-way classification table it is typically convenient to adjoin to the table the row and column totals, to produce a new table like this:

| Specialty | Language Ability |     | Total |
|-----------|------------------|-----|-------|
|           | $S$              | $T$ |       |
| $C$       | 12               | 1   | 13    |
| $E$       | 4                | 3   | 7     |
| $M$       | 6                | 2   | 8     |

| Specialty | Language Ability |   | Total |
|-----------|------------------|---|-------|
|           | S                | T |       |
| Total     | 22               | 6 | 28    |

- a. The probability sought is  $P(M \cap T)$ . The table shows that there are 2 such people, out of 28 in all, hence  $P(M \cap T) = 2 / 28 \approx 0.07$  or about a 7% chance.
- b. The probability sought is  $P(M \cup T)$ . The third row total and the grand total in the sample give  $P(M) = 8 / 28$ . The second column total and the grand total give  $P(T) = 6 / 28$ . Thus using the result from part (a),

$$P(M \cup T) = P(M) + P(T) - P(M \cap T) = \frac{8}{28} + \frac{6}{28} - \frac{2}{28} = \frac{12}{28}$$

or about a 43% chance.

- c. This probability can be computed in two ways. Since the event of interest can be viewed as the event  $C \cup E$  and the events  $C$  and  $E$  are mutually exclusive, the answer is, using the first two row totals,

$$P(C \cup E) = P(C) + P(E) - P(C \cap E) = \frac{13}{28} + \frac{7}{28} - \frac{0}{28} = \frac{20}{28} \approx$$

On the other hand, the event of interest can be thought of as the complement  $M^c$  of  $M$ , hence using the value of  $P(M)$  computed in part (b),

$$P(M^c) = 1 - P(M) = 1 - \frac{8}{28} = \frac{20}{28} \approx 0.71$$

as before.

KEY TAKEAWAY

- The probability of an event that is a complement or union of events of known probability can be computed using formulas.

## EXERCISES

## BASIC

1. For the sample space  $S = \{a, b, c, d, e\}$  identify the complement of each event given.

- $A = \{a, d, e\}$
- $B = \{b, c, d, e\}$
- $S$

2. For the sample space  $S = \{r, s, t, u, v\}$  identify the complement of each event given.

- $R = \{t, u\}$
- $T = \{r\}$
- $\emptyset$  (the “empty” set that has no elements)

3. The sample space for three tosses of a coin is

$$S = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$$

Define events

$H$  : at least one head is observed

$M$  : more heads than tails are observed

- List the outcomes that comprise  $H$  and  $M$ .
  - List the outcomes that comprise  $H \cap M$ ,  $H \cup M$ , and  $H^c$ .
  - Assuming all outcomes are equally likely, find  $P(H \cap M)$ ,  $P(H \cup M)$ , and  $P(H^c)$ .
  - Determine whether or not  $H^c$  and  $M$  are mutually exclusive. Explain why or why not.
4. For the experiment of rolling a single six-sided die once, define events

$T$  : the number rolled is three

$G$  : the number rolled is four or greater

- List the outcomes that comprise  $T$  and  $G$ .
- List the outcomes that comprise  $T \cap G$ ,  $T \cup G$ ,  $T^c$ , and  $(T \cup G)^c$ .
- Assuming all outcomes are equally likely, find  $P(T \cap G)$ ,  $P(T \cup G)$ , and  $P(T^c)$ .
- Determine whether or not  $T$  and  $G$  are mutually exclusive. Explain why or why not.

5. A special deck of 16 cards has 4 that are blue, 4 yellow, 4 green, and 4 red. The four cards of each color are numbered from one to four. A single card is drawn at random. Define events

$B$  : the card is blue

$R$  : the card is red

$N$  : the number on the card is at most two

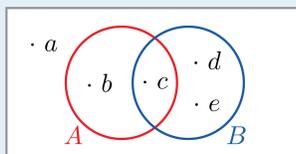
- List the outcomes that comprise  $B$ ,  $R$ , and  $N$ .
  - List the outcomes that comprise  $B \cap R$ ,  $B \cup R$ ,  $B \cap N$ ,  $R \cup N$ ,  $B^c$ , and  $(B \cup R)^c$ .
  - Assuming all outcomes are equally likely, find the probabilities of the events in the previous part.
  - Determine whether or not  $B$  and  $N$  are mutually exclusive. Explain why or why not.
6. In the context of the previous problem, define events

$Y$  : the card is yellow

$I$  : the number on the card is not a one

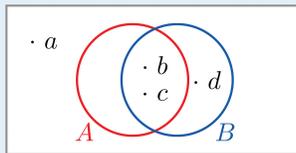
$J$  : the number on the card is a two or a four

- List the outcomes that comprise  $Y$ ,  $I$ , and  $J$ .
  - List the outcomes that comprise  $Y \cap I$ ,  $Y \cup J$ ,  $I \cap J$ ,  $I^c$ , and  $(Y \cup J)^c$ .
  - Assuming all outcomes are equally likely, find the probabilities of the events in the previous part.
  - Determine whether or not  $I^c$  and  $J$  are mutually exclusive. Explain why or why not.
7. The Venn diagram provided shows a sample space and two events  $A$  and  $B$ . Suppose  $P(a) = 0.13$ ,  $P(b) = 0.09$ ,  $P(c) = 0.27$ ,  $P(d) = 0.20$ , and  $P(e) = 0.31$ . Confirm that the probabilities of the outcomes add up to 1, then compute the following probabilities.



- $P(A)$ .
- $P(B)$ .

- c.  $P(A^c)$  two ways: (i) by finding the outcomes in  $A^c$  and adding their probabilities, and (ii) using the Probability Rule for Complements.
  - d.  $P(A \cap B)$ .
  - e.  $P(A \cup B)$  two ways: (i) by finding the outcomes in  $A \cup B$  and adding their probabilities, and (ii) using the Additive Rule of Probability.
8. The Venn diagram provided shows a sample space and two events  $A$  and  $B$ . Suppose  $P(a) = 0.32, P(b) = 0.17, P(c) = 0.28$ , and  $P(d) = 0.23$ . Confirm that the probabilities of the outcomes add up to 1, then compute the following probabilities.



- a.  $P(A)$ .
  - b.  $P(B)$ .
  - c.  $P(A^c)$  two ways: (i) by finding the outcomes in  $A^c$  and adding their probabilities, and (ii) using the Probability Rule for Complements.
  - d.  $P(A \cap B)$ .
  - e.  $P(A \cup B)$  two ways: (i) by finding the outcomes in  $A \cup B$  and adding their probabilities, and (ii) using the Additive Rule of Probability.
9. Confirm that the probabilities in the two-way contingency table add up to 1, then use it to find the probabilities of the events indicated.

|     | $U$  | $V$  | $W$  |
|-----|------|------|------|
| $A$ | 0.15 | 0.00 | 0.23 |
| $B$ | 0.22 | 0.30 | 0.10 |

- a.  $P(A), P(B), P(A \cap B)$ .
  - b.  $P(U), P(W), P(U \cap W)$ .
  - c.  $P(U \cup W)$ .
  - d.  $P(V^c)$ .
  - e. Determine whether or not the events  $A$  and  $U$  are mutually exclusive; the events  $A$  and  $V$ .
10. Confirm that the probabilities in the two-way contingency table add up to 1, then use it to find the probabilities of the events indicated.

|          | <b>R</b> | <b>S</b> | <b>T</b> |
|----------|----------|----------|----------|
| <b>M</b> | 0.09     | 0.25     | 0.19     |
| <b>N</b> | 0.31     | 0.16     | 0.00     |

- $P(R), P(S), P(R \cap S)$ .
- $P(M), P(N), P(M \cap N)$ .
- $P(R \cup S)$ .
- $P(R^c)$ .
- Determine whether or not the events  $N$  and  $S$  are mutually exclusive; the events  $N$  and  $T$ .

### APPLICATIONS

- Make a statement in ordinary English that describes the complement of each event (do not simply insert the word “not”).
  - In the roll of a die: “five or more.”
  - In a roll of a die: “an even number.”
  - In two tosses of a coin: “at least one heads.”
  - In the random selection of a college student: “Not a freshman.”
- Make a statement in ordinary English that describes the complement of each event (do not simply insert the word “not”).
  - In the roll of a die: “two or less.”
  - In the roll of a die: “one, three, or four.”
  - In two tosses of a coin: “at most one heads.”
  - In the random selection of a college student: “Neither a freshman nor a senior.”
- The sample space that describes all three-child families according to the genders of the children with respect to birth order is

$$S = \{bbb, bbg, bgb, bgg, gbb, gbg, ggb, ggg\}.$$

For each of the following events in the experiment of selecting a three-child family at random, state the complement of the event in the simplest possible terms, then find the outcomes that comprise the event and its complement.

- At least one child is a girl.
- At most one child is a girl.
- All of the children are girls.
- Exactly two of the children are girls.
- The first born is a girl.

14. The sample space that describes the two-way classification of citizens according to gender and opinion on a political issue is

$$S = \{mf, ma, mn, ff, fa, fn\},$$

where the first letter denotes gender ( $m$ : male,  $f$ : female) and the second opinion ( $f$ : for,  $a$ : against,  $n$ : neutral). For each of the following events in the experiment of selecting a citizen at random, state the complement of the event in the simplest possible terms, then find the outcomes that comprise the event and its complement.

- a. The person is male.
  - b. The person is not in favor.
  - c. The person is either male or in favor.
  - d. The person is female and neutral.
15. A tourist who speaks English and German but no other language visits a region of Slovenia. If 35% of the residents speak English, 15% speak German, and 3% speak both English and German, what is the probability that the tourist will be able to talk with a randomly encountered resident of the region?
16. In a certain country 43% of all automobiles have airbags, 27% have anti-lock brakes, and 13% have both. What is the probability that a randomly selected vehicle will have both airbags and anti-lock brakes?
17. A manufacturer examines its records over the last year on a component part received from outside suppliers. The breakdown on source (supplier  $A$ , supplier  $B$ ) and quality ( $H$ : high,  $U$ : usable,  $D$ : defective) is shown in the two-way contingency table.

|     | $H$    | $U$    | $D$    |
|-----|--------|--------|--------|
| $A$ | 0.6937 | 0.0049 | 0.0014 |
| $B$ | 0.2982 | 0.0009 | 0.0009 |

The record of a part is selected at random. Find the probability of each of the following events.

- a. The part was defective.
- b. The part was either of high quality or was at least usable, in two ways: (i) by adding numbers in the table, and (ii) using the answer to (a) and the Probability Rule for Complements.
- c. The part was defective and came from supplier  $B$ .
- d. The part was defective or came from supplier  $B$ , in two ways: by finding the cells in the table that correspond to this event and adding their probabilities, and (ii) using the Additive Rule of Probability.

18. Individuals with a particular medical condition were classified according to the presence ( $T$ ) or absence ( $N$ ) of a potential toxin in their blood and the onset of the condition ( $E$ : early,  $M$ : midrange,  $L$ : late). The breakdown according to this classification is shown in the two-way contingency table.

|     | $E$   | $M$   | $L$   |
|-----|-------|-------|-------|
| $T$ | 0.012 | 0.124 | 0.013 |
| $N$ | 0.170 | 0.638 | 0.043 |

One of these individuals is selected at random. Find the probability of each of the following events.

- The person experienced early onset of the condition.
  - The onset of the condition was either midrange or late, in two ways: (i) by adding numbers in the table, and (ii) using the answer to (a) and the Probability Rule for Complements.
  - The toxin is present in the person's blood.
  - The person experienced early onset of the condition and the toxin is present in the person's blood.
  - The person experienced early onset of the condition or the toxin is present in the person's blood, in two ways: (i) by finding the cells in the table that correspond to this event and adding their probabilities, and (ii) using the Additive Rule of Probability.
19. The breakdown of the students enrolled in a university course by class ( $F$ : freshman,  $So$ : sophomore,  $J$ : junior,  $Se$ : senior) and academic major ( $S$ : science, mathematics, or engineering,  $L$ : liberal arts,  $O$ : other) is shown in the two-way classification table.

| Major | Class |      |     |      |
|-------|-------|------|-----|------|
|       | $F$   | $So$ | $J$ | $Se$ |
| $S$   | 92    | 42   | 20  | 13   |
| $L$   | 368   | 167  | 80  | 53   |
| $O$   | 460   | 209  | 100 | 67   |

A student enrolled in the course is selected at random. Adjoin the row and column totals to the table and use the expanded table to find the probability of each of the following events.

- The student is a freshman.
- The student is a liberal arts major.
- The student is a freshman liberal arts major.

- d. The student is either a freshman or a liberal arts major.  
 e. The student is not a liberal arts major.
20. The table relates the response to a fund-raising appeal by a college to its alumni to the number of years since graduation.

| Response | Years Since Graduation |      |       |         |
|----------|------------------------|------|-------|---------|
|          | 0-5                    | 6-20 | 21-35 | Over 35 |
| Positive | 120                    | 440  | 210   | 90      |
| None     | 1380                   | 3560 | 3290  | 910     |

An alumnus is selected at random. Adjoin the row and column totals to the table and use the expanded table to find the probability of each of the following events.

- a. The alumnus responded.  
 b. The alumnus did not respond.  
 c. The alumnus graduated at least 21 years ago.  
 d. The alumnus graduated at least 21 years ago and responded.

### ADDITIONAL EXERCISES

21. The sample space for tossing three coins is

$$S = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$$

- a. List the outcomes that correspond to the statement “All the coins are heads.”  
 b. List the outcomes that correspond to the statement “Not all the coins are heads.”  
 c. List the outcomes that correspond to the statement “All the coins are not heads.”

## ANSWERS

1.
  - a.  $\{b, c\}$
  - b.  $\{a\}$
  - c.  $\emptyset$
  
3.
  - a.  $H = \{hhh, hht, hth, htt, thh, tht, tth\}$ ,  
 $M = \{hhh, hht, hth, thh\}$
  - b.  $H \cap M = \{hhh, hht, hth, thh\}$ ,  $H \cup M = H$ ,  $H^c = \{ttt\}$
  - c.  $P(H \cap M) = 4 / 8$ ,  $P(H \cup M) = 7 / 8$ ,  $P(H^c) = 1 / 8$
  - d. Mutually exclusive because they have no elements in common.
  
5.
  - a.  $B = \{b1, b2, b3, b4\}$ ,  $R = \{r1, r2, r3, r4\}$   
 $N = \{b1, b2, y1, y2, g1, g2, r1, r2\}$
  - b.  $B \cap R = \emptyset$ ,  $B \cup R = \{b1, b2, b3, b4, r1, r2, r3, r4\}$ ,  
 $B \cap N = \{b1, b2\}$ ,  
 $R \cup N = \{b1, b2, y1, y2, g1, g2, r1, r2, r3, r4\}$ ,  
 $B^c = \{y1, y2, y3, y4, g1, g2, g3, g4, r1, r2, r3, r4\}$ ,  
 $(B \cup R)^c = \{y1, y2, y3, y4, g1, g2, g3, g4\}$
  - c.  $P(B \cap R) = 0$ ,  $P(B \cup R) = 8 / 16$ ,  $P(B \cap N) = 2 / 16$ ,  
 $P(R \cup N) = 10 / 16$ ,  $P(B^c) = 12 / 16$ ,  
 $P((B \cup R)^c) = 8 / 16$
  - d. Not mutually exclusive because they have an element in common.
  
7.
  - a. 0.36
  - b. 0.78
  - c. 0.64
  - d. 0.27
  - e. 0.87
  
9.
  - a.  $P(A) = 0.38$ ,  $P(B) = 0.62$ ,  $P(A \cap B) = 0$
  - b.  $P(U) = 0.37$ ,  $P(W) = 0.33$ ,  $P(U \cap W) = 0$
  - c. 0.7
  - d. 0.7
  - e.  $A$  and  $U$  are not mutually exclusive because  $P(A \cap U)$  is the nonzero number 0.15.  $A$  and  $V$  are mutually exclusive because  $P(A \cap V) = 0$ .
  
11.
  - a. "four or less"
  - b. "an odd number"
  - c. "no heads" or "all tails"
  - d. "a freshman"

13. a. "All the children are boys."  
 Event:  $\{bbg, bgb, bgg, gbb, gbg, ggb, ggg\}$  ,  
 Complement:  $\{bbb\}$
- b. "At least two of the children are girls" or "There are two or three girls."  
 Event:  $\{bbb, bbg, bgb, gbb\}$  ,  
 Complement:  $\{bgg, gbg, ggb, ggg\}$
- c. "At least one child is a boy."  
 Event:  $\{ggg\}$  ,  
 Complement:  $\{bbb, bbg, bgb, bgg, gbb, gbg, ggb\}$
- d. "There are either no girls, exactly one girl, or three girls."  
 Event:  $\{bgg, gbg, ggb\}$  ,  
 Complement:  $\{bbb, bbg, bgb, gbb, ggg\}$
- e. "The first born is a boy."  
 Event:  $\{gbb, gbg, ggb, ggg\}$  ,  
 Complement:  $\{bbb, bbg, bgb, bgg\}$
15. 0.47
17. a. 0.0023  
 b. 0.9977  
 c. 0.0009  
 d. 0.3014
19. a. 920/1671  
 b. 668/1671  
 c. 368/1671  
 d. 1220/1671  
 e. 1003/1671
21. a.  $\{hhh\}$   
 b.  $\{hht, hth, htt, thh, tht, tth, ttt\}$   
 c.  $\{ttt\}$

### 3.3 Conditional Probability and Independent Events

#### LEARNING OBJECTIVES

1. To learn the concept of a conditional probability and how to compute it.
2. To learn the concept of independence of events, and how to apply it.

#### Conditional Probability

Suppose a fair die has been rolled and you are asked to give the probability that it was a five. There are six equally likely outcomes, so your answer is  $1/6$ . But suppose that before you give your answer you are given the extra information that the number rolled was odd. Since there are only three odd numbers that are possible, one of which is five, you would certainly revise your estimate of the likelihood that a five was rolled from  $1/6$  to  $1/3$ . In general, the *revised* probability that an event  $A$  has occurred, taking into account the additional information that another event  $B$  has definitely occurred on this trial of the experiment, is called the *conditional probability of  $A$  given  $B$*  and is denoted by  $P(A|B)$ . The reasoning employed in this example can be generalized to yield the computational formula in the following definition.

#### Definition

The **conditional probability**<sup>9</sup> of  $A$  given  $B$ , denoted  $P(A|B)$ , is the probability that event  $A$  has occurred in a trial of a random experiment for which it is known that event  $B$  has definitely occurred. It may be computed by means of the following formula:

Rule for Conditional Probability

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

9. The probability of the event  $A$  taking into account the fact that event  $B$  is known to have occurred.

## EXAMPLE 20

A fair die is rolled.

- Find the probability that the number rolled is a five, given that it is odd.
- Find the probability that the number rolled is odd, given that it is a five.

Solution:

The sample space for this experiment is the set  $S = \{1,2,3,4,5,6\}$  consisting of six equally likely outcomes. Let  $F$  denote the event “a five is rolled” and let  $O$  denote the event “an odd number is rolled,” so that

$$F = \{5\} \quad \text{and} \quad O = \{1,3,5\}$$

- This is the introductory example, so we already know that the answer is  $1/3$ . To use the formula in the definition to confirm this we must replace  $A$  in the formula (the event whose likelihood we seek to estimate) by  $F$  and replace  $B$  (the event we know for certain has occurred) by  $O$ :

$$P(F|O) = \frac{P(F \cap O)}{P(O)}$$

$$\text{Since } F \cap O = \{5\} \cap \{1,3,5\} = \{5\}, \\ P(F \cap O) = 1/6.$$

$$\text{Since } O = \{1,3,5\}, P(O) = 3/6.$$

Thus

$$P(F|O) = \frac{P(F \cap O)}{P(O)} = \frac{1/6}{3/6} = \frac{1}{3}$$

- This is the same problem, but with the roles of  $F$  and  $O$  reversed. Since we are given that the number that was rolled is five, which is odd, the probability in question must be 1. To apply the formula to this case we must now replace  $A$  (the event whose

likelihood we seek to estimate) by  $O$  and  $B$  (the event we know for certain has occurred) by  $F$ :

$$P(O|F) = \frac{P(O \cap F)}{P(F)}$$

Obviously  $P(F) = 1/6$ . In part (a) we found that  $P(F \cap O) = 1/6$ . Thus

$$P(O|F) = \frac{P(O \cap F)}{P(F)} = \frac{1/6}{1/6} = 1$$

Just as we did not need the computational formula in this example, we do not need it when the information is presented in a two-way classification table, as in the next example.

## EXAMPLE 21

In a sample of 902 individuals under 40 who were or had previously been married, each person was classified according to gender and age at first marriage. The results are summarized in the following two-way classification table, where the meaning of the labels is:

- $M$ : male
- $F$ : female
- $E$ : a teenager when first married
- $W$ : in one's twenties when first married
- $H$ : in one's thirties when first married

|       | $E$ | $W$ | $H$ | Total |
|-------|-----|-----|-----|-------|
| $M$   | 43  | 293 | 114 | 450   |
| $F$   | 82  | 299 | 71  | 452   |
| Total | 125 | 592 | 185 | 902   |

The numbers in the first row mean that 43 people in the sample were men who were first married in their teens, 293 were men who were first married in their twenties, 114 men who were first married in their thirties, and a total of 450 people in the sample were men. Similarly for the numbers in the second row. The numbers in the last row mean that, irrespective of gender, 125 people in the sample were married in their teens, 592 in their twenties, 185 in their thirties, and that there were 902 people in the sample in all. Suppose that the proportions in the sample accurately reflect those in the population of all individuals in the population who are under 40 and who are or have previously been married. Suppose such a person is selected at random.

- a. Find the probability that the individual selected was a teenager at first marriage.
- b. Find the probability that the individual selected was a teenager at first marriage, given that the person is male.

Solution:

It is natural to let  $E$  also denote the event that the person selected was a teenager at first marriage and to let  $M$  denote the event that the person selected is male.

- a. According to the table the proportion of individuals in the sample who were in their teens at their first marriage is  $125/902$ . This is the relative frequency of such people in the population, hence  $P(E) = 125 / 902 \approx 0.139$  or about 14%.
- b. Since it is known that the person selected is male, all the females may be removed from consideration, so that only the row in the table corresponding to men in the sample applies:

|          | <i>E</i> | <i>W</i> | <i>H</i> | Total |
|----------|----------|----------|----------|-------|
| <i>M</i> | 43       | 293      | 114      | 450   |

The proportion of males in the sample who were in their teens at their first marriage is  $43/450$ . This is the relative frequency of such people in the population of males, hence  $P(E|M) = 43 / 450 \approx 0.096$  or about 10%.

In the next example, the computational formula in the definition must be used.

## EXAMPLE 22

Suppose that in an adult population the proportion of people who are both overweight and suffer hypertension is 0.09; the proportion of people who are not overweight but suffer hypertension is 0.11; the proportion of people who are overweight but do not suffer hypertension is 0.02; and the proportion of people who are neither overweight nor suffer hypertension is 0.78. An adult is randomly selected from this population.

- Find the probability that the person selected suffers hypertension given that he is overweight.
- Find the probability that the selected person suffers hypertension given that he is not overweight.
- Compare the two probabilities just found to give an answer to the question as to whether overweight people tend to suffer from hypertension.

Solution:

Let  $H$  denote the event “the person selected suffers hypertension.” Let  $O$  denote the event “the person selected is overweight.” The probability information given in the problem may be organized into the following contingency table:

|       | $O$  | $O^c$ |
|-------|------|-------|
| $H$   | 0.09 | 0.11  |
| $H^c$ | 0.02 | 0.78  |

- Using the formula in the definition of conditional probability,

$$P(H|O) = \frac{P(H \cap O)}{P(O)} = \frac{0.09}{0.09 + 0.02} = 0.8182$$

- Using the formula in the definition of conditional probability,

$$P(H|O^c) = \frac{P(H \cap O^c)}{P(O^c)} = \frac{0.11}{0.11 + 0.78} = 0.1236$$

- c.  $P(H|O) = 0.8182$  is over six times as large as  $P(H|O^c) = 0.1236$ , which indicates a much higher rate of hypertension among people who are overweight than among people who are not overweight. It might be interesting to note that a direct comparison of  $P(H \cap O) = 0.09$  and  $P(H \cap O^c) = 0.11$  does not answer the same question.

## Independent Events

Although typically we expect the conditional probability  $P(A|B)$  to be different from the probability  $P(A)$  of  $A$ , it does not have to be different from  $P(A)$ . When  $P(A|B) = P(A)$ , the occurrence of  $B$  has no effect on the likelihood of  $A$ . Whether or not the event  $A$  has occurred is *independent* of the event  $B$ .

Using algebra it can be shown that the equality  $P(A|B) = P(A)$  holds if and only if the equality  $P(A \cap B) = P(A) \cdot P(B)$  holds, which in turn is true if and only if  $P(B|A) = P(B)$ . This is the basis for the following definition.

### Definition

Events  $A$  and  $B$  are **independent**<sup>10</sup> if

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

If  $A$  and  $B$  are not independent then they are **dependent**.

The formula in the definition has two practical but exactly opposite uses:

1. In a situation in which we can compute all three probabilities  $P(A)$ ,  $P(B)$ , and  $P(A \cap B)$ , it is used to check whether or not the events  $A$  and  $B$  are independent:
  - If  $P(A \cap B) = P(A) \cdot P(B)$ , then  $A$  and  $B$  are independent.
  - If  $P(A \cap B) \neq P(A) \cdot P(B)$ , then  $A$  and  $B$  are not independent.
2. In a situation in which each of  $P(A)$  and  $P(B)$  can be computed and it is known that  $A$  and  $B$  are independent, then we can compute

10. Events whose probability of occurring together is the product of their individual probabilities.

$P(A \cap B)$  by multiplying together  $P(A)$  and  $P(B)$ :  
 $P(A \cap B) = P(A) \cdot P(B)$ .

### EXAMPLE 23

A single fair die is rolled. Let  $A = \{3\}$  and  $B = \{1,3,5\}$ . Are  $A$  and  $B$  independent?

Solution:

In this example we can compute all three probabilities  $P(A) = 1/6$ ,  $P(B) = 1/2$ , and  $P(A \cap B) = P(\{3\}) = 1/6$ . Since the product  $P(A) \cdot P(B) = (1/6)(1/2) = 1/12$  is not the same number as  $P(A \cap B) = 1/6$ , the events  $A$  and  $B$  are not independent.

## EXAMPLE 24

The two-way classification of married or previously married adults under 40 according to gender and age at first marriage in [Note 3.48 "Example 21"](#) produced the table

|          | <i>E</i> | <i>W</i> | <i>H</i> | Total |
|----------|----------|----------|----------|-------|
| <i>M</i> | 43       | 293      | 114      | 450   |
| <i>F</i> | 82       | 299      | 71       | 452   |
| Total    | 125      | 592      | 185      | 902   |

Determine whether or not the events *F*: “female” and *E*: “was a teenager at first marriage” are independent.

Solution:

The table shows that in the sample of 902 such adults, 452 were female, 125 were teenagers at their first marriage, and 82 were females who were teenagers at their first marriage, so that

$$P(F) = \frac{452}{902} \quad P(E) = \frac{125}{902} \quad P(F \cap E) = \frac{82}{902}$$

Since

$$P(F) \cdot P(E) = \frac{452}{902} \cdot \frac{125}{902} = 0.069$$

is not the same as

$$P(F \cap E) = \frac{82}{902} = 0.091$$

we conclude that the two events are not independent.

## EXAMPLE 25

Many diagnostic tests for detecting diseases do not test for the disease directly but for a chemical or biological product of the disease, hence are not perfectly reliable. The *sensitivity* of a test is the probability that the test will be positive when administered to a person who has the disease. The higher the sensitivity, the greater the detection rate and the lower the false negative rate.

Suppose the sensitivity of a diagnostic procedure to test whether a person has a particular disease is 92%. A person who actually has the disease is tested for it using this procedure by two independent laboratories.

- What is the probability that both test results will be positive?
- What is the probability that at least one of the two test results will be positive?

Solution:

- Let  $A_1$  denote the event “the test by the first laboratory is positive” and let  $A_2$  denote the event “the test by the second laboratory is positive.” Since  $A_1$  and  $A_2$  are independent,

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2) = 0.92 \times 0.92 = 0.8464$$

- Using the Additive Rule for Probability and the probability just computed,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = 0.92 + 0.92 - 0.8464 = 0.9936$$

## EXAMPLE 26

The *specificity* of a diagnostic test for a disease is the probability that the test will be negative when administered to a person who does not have the disease. The higher the specificity, the lower the false positive rate.

Suppose the specificity of a diagnostic procedure to test whether a person has a particular disease is 89%.

- A person who does not have the disease is tested for it using this procedure. What is the probability that the test result will be positive?
- A person who does not have the disease is tested for it by two independent laboratories using this procedure. What is the probability that both test results will be positive?

Solution:

- Let  $B$  denote the event “the test result is positive.” The complement of  $B$  is that the test result is negative, and has probability the specificity of the test, 0.89. Thus

$$P(B) = 1 - P(B^c) = 1 - 0.89 = 0.11.$$

- Let  $B_1$  denote the event “the test by the first laboratory is positive” and let  $B_2$  denote the event “the test by the second laboratory is positive.” Since  $B_1$  and  $B_2$  are independent, by part (a) of the example

$$P(B_1 \cap B_2) = P(B_1) \cdot P(B_2) = 0.11 \times 0.11 = 0.0121.$$

The concept of independence applies to any number of events. For example, three events  $A$ ,  $B$ , and  $C$  are independent if  $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ . Note carefully that, as is the case with just two events, this is not a formula that is always valid, but holds precisely when the events in question are independent.

## EXAMPLE 27

The reliability of a system can be enhanced by redundancy, which means building two or more independent devices to do the same job, such as two independent braking systems in an automobile.

Suppose a particular species of trained dogs has a 90% chance of detecting contraband in airline luggage. If the luggage is checked three times by three different dogs independently of one another, what is the probability that contraband will be detected?

Solution:

Let  $D_1$  denote the event that the contraband is detected by the first dog,  $D_2$  the event that it is detected by the second dog, and  $D_3$  the event that it is detected by the third. Since each dog has a 90% of detecting the contraband, by the Probability Rule for Complements it has a 10% chance of failing. In symbols,  $P(D_1^c) = 0.10$ ,  $P(D_2^c) = 0.10$ , and  $P(D_3^c) = 0.10$ .

Let  $D$  denote the event that the contraband is detected. We seek  $P(D)$ . It is easier to find  $P(D^c)$ , because although there are several ways for the contraband to be detected, there is only one way for it to go undetected: all three dogs must fail. Thus  $D^c = D_1^c \cap D_2^c \cap D_3^c$  and

$$P(D) = 1 - P(D^c) = 1 - P(D_1^c \cap D_2^c \cap D_3^c)$$

But the events  $D_1$ ,  $D_2$ , and  $D_3$  are independent, which implies that their complements are independent, so

$$P(D_1^c \cap D_2^c \cap D_3^c) = P(D_1^c) \cdot P(D_2^c) \cdot P(D_3^c) = 0.10 \times 0.10 \times 0.10 = 0.001$$

Using this number in the previous display we obtain

$$P(D) = 1 - 0.001 = 0.999$$

That is, although any one dog has only a 90% chance of detecting the contraband, three dogs working independently have a 99.9% chance of detecting it.

### **Probabilities on Tree Diagrams**

Some probability problems are made much simpler when approached using a tree diagram. The next example illustrates how to place probabilities on a tree diagram and use it to solve a problem.

## EXAMPLE 28

A jar contains 10 marbles, 7 black and 3 white. Two marbles are drawn without replacement, which means that the first one is not put back before the second one is drawn.

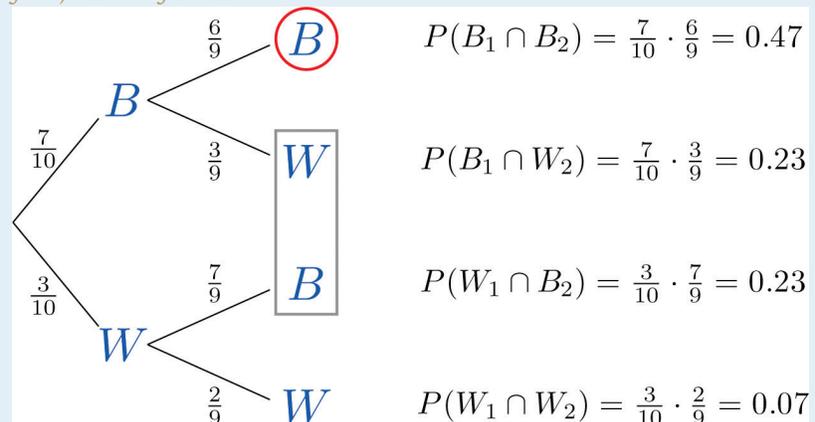
- What is the probability that both marbles are black?
- What is the probability that exactly one marble is black?
- What is the probability that at least one marble is black?

Solution:

A tree diagram for the situation of drawing one marble after the other without replacement is shown in [Figure 3.6 "Tree Diagram for Drawing Two Marbles"](#). The circle and rectangle will be explained later, and should be ignored for now.

Figure 3.6

Tree Diagram for Drawing Two Marbles



The numbers on the two leftmost branches are the probabilities of getting either a black marble, 7 out of 10, or a white marble, 3 out of 10, on the first draw. The number on each remaining branch is the probability of the event corresponding to the node on the right end of the branch occurring, given that the event corresponding to the node on the left end of the branch has occurred. Thus for the top branch, connecting the two Bs, it is  $P(B_2|B_1)$ , where  $B_1$  denotes the event “the first marble drawn is black” and  $B_2$  denotes the event “the second marble drawn is black.” Since after drawing a black marble out there are 9 marbles left, of which 6 are black, this probability is 6/9.

The number to the right of each final node is computed as shown, using the principle that if the formula in the Conditional Rule for Probability is multiplied by  $P(B)$ , then the result is

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

- a. The event “both marbles are black” is  $B_1 \cap B_2$  and corresponds to the top right node in the tree, which has been circled. Thus as indicated there, it is 0.47.
- b. The event “exactly one marble is black” corresponds to the two nodes of the tree enclosed by the rectangle. The events that correspond to these two nodes are mutually exclusive: black followed by white is incompatible with white followed by black. Thus in accordance with the Additive Rule for Probability we merely add the two probabilities next to these nodes, since what would be subtracted from the sum is zero. Thus the probability of drawing exactly one black marble in two tries is  $0.23 + 0.23 = 0.46$ .
- c. The event “at least one marble is black” corresponds to the three nodes of the tree enclosed by either the circle or the rectangle. The events that correspond to these nodes are mutually exclusive, so as in part (b) we merely add the probabilities next to these nodes. Thus the probability of drawing at least one black marble in two tries is  $0.47 + 0.23 + 0.23 = 0.93$ .

Of course, this answer could have been found more easily using the Probability Law for Complements, simply subtracting the probability of the complementary event, “two white marbles are drawn,” from 1 to obtain  $1 - 0.07 = 0.93$ .

As this example shows, finding the probability for each branch is fairly straightforward, since we compute it knowing everything that has happened in the sequence of steps so far. Two principles that are true in general emerge from this example:

### Probabilities on Tree Diagrams

1. The probability of the event corresponding to any node on a tree is the product of the numbers on the unique path of branches that leads to that node from the start.
2. If an event corresponds to several final nodes, then its probability is obtained by adding the numbers next to those nodes.

#### KEY TAKEAWAYS

- A conditional probability is the probability that an event has occurred, taking into account additional information about the result of the experiment.
- A conditional probability can always be computed using the formula in the definition. Sometimes it can be computed by discarding part of the sample space.
- Two events  $A$  and  $B$  are independent if the probability  $P(A \cap B)$  of their intersection  $A \cap B$  is equal to the product  $P(A) \cdot P(B)$  of their individual probabilities.

## EXERCISES

## BASIC

1. For two events  $A$  and  $B$ ,  $P(A) = 0.73$ ,  $P(B) = 0.48$ , and  $P(A \cap B) = 0.29$ .
  - a. Find  $P(A|B)$ .
  - b. Find  $P(B|A)$ .
  - c. Determine whether or not  $A$  and  $B$  are independent.
2. For two events  $A$  and  $B$ ,  $P(A) = 0.26$ ,  $P(B) = 0.37$ , and  $P(A \cap B) = 0.11$ .
  - a. Find  $P(A|B)$ .
  - b. Find  $P(B|A)$ .
  - c. Determine whether or not  $A$  and  $B$  are independent.
3. For independent events  $A$  and  $B$ ,  $P(A) = 0.81$  and  $P(B) = 0.27$ .
  - a. Find  $P(A \cap B)$ .
  - b. Find  $P(A|B)$ .
  - c. Find  $P(B|A)$ .
4. For independent events  $A$  and  $B$ ,  $P(A) = 0.68$  and  $P(B) = 0.37$ .
  - a. Find  $P(A \cap B)$ .
  - b. Find  $P(A|B)$ .
  - c. Find  $P(B|A)$ .
5. For mutually exclusive events  $A$  and  $B$ ,  $P(A) = 0.17$  and  $P(B) = 0.32$ .
  - a. Find  $P(A|B)$ .
  - b. Find  $P(B|A)$ .
6. For mutually exclusive events  $A$  and  $B$ ,  $P(A) = 0.45$  and  $P(B) = 0.09$ .
  - a. Find  $P(A|B)$ .
  - b. Find  $P(B|A)$ .
7. Compute the following probabilities in connection with the roll of a single fair die.
  - a. The probability that the roll is even.
  - b. The probability that the roll is even, given that it is not a two.
  - c. The probability that the roll is even, given that it is not a one.

8. Compute the following probabilities in connection with two tosses of a fair coin.
- The probability that the second toss is heads.
  - The probability that the second toss is heads, given that the first toss is heads.
  - The probability that the second toss is heads, given that at least one of the two tosses is heads.
9. A special deck of 16 cards has 4 that are blue, 4 yellow, 4 green, and 4 red. The four cards of each color are numbered from one to four. A single card is drawn at random. Find the following probabilities.
- The probability that the card drawn is red.
  - The probability that the card is red, given that it is not green.
  - The probability that the card is red, given that it is neither red nor yellow.
  - The probability that the card is red, given that it is not a four.
10. A special deck of 16 cards has 4 that are blue, 4 yellow, 4 green, and 4 red. The four cards of each color are numbered from one to four. A single card is drawn at random. Find the following probabilities.
- The probability that the card drawn is a two or a four.
  - The probability that the card is a two or a four, given that it is not a one.
  - The probability that the card is a two or a four, given that it is either a two or a three.
  - The probability that the card is a two or a four, given that it is red or green.
11. A random experiment gave rise to the two-way contingency table shown. Use it to compute the probabilities indicated.

|   | R    | S    |
|---|------|------|
| A | 0.12 | 0.18 |
| B | 0.28 | 0.42 |

- $P(A)$ ,  $P(R)$ ,  $P(A \cap R)$ .
  - Based on the answer to (a), determine whether or not the events  $A$  and  $R$  are independent.
  - Based on the answer to (b), determine whether or not  $P(A|R)$  can be predicted without any computation. If so, make the prediction. In any case, compute  $P(A|R)$  using the Rule for Conditional Probability.
12. A random experiment gave rise to the two-way contingency table shown. Use it to compute the probabilities indicated.

|   | R    | S    |
|---|------|------|
| A | 0.13 | 0.07 |
| B | 0.61 | 0.19 |

- $P(A), P(R), P(A \cap R)$ .
  - Based on the answer to (a), determine whether or not the events  $A$  and  $R$  are independent.
  - Based on the answer to (b), determine whether or not  $P(A|R)$  can be predicted without any computation. If so, make the prediction. In any case, compute  $P(A|R)$  using the Rule for Conditional Probability.
- Suppose for events  $A$  and  $B$  in a random experiment  $P(A) = 0.70$  and  $P(B) = 0.30$ . Compute the indicated probability, or explain why there is not enough information to do so.
    - $P(A \cap B)$ .
    - $P(A \cap B)$ , with the extra information that  $A$  and  $B$  are independent.
    - $P(A \cap B)$ , with the extra information that  $A$  and  $B$  are mutually exclusive.
  - Suppose for events  $A$  and  $B$  connected to some random experiment,  $P(A) = 0.50$  and  $P(B) = 0.50$ . Compute the indicated probability, or explain why there is not enough information to do so.
    - $P(A \cap B)$ .
    - $P(A \cap B)$ , with the extra information that  $A$  and  $B$  are independent.
    - $P(A \cap B)$ , with the extra information that  $A$  and  $B$  are mutually exclusive.
  - Suppose for events  $A, B,$  and  $C$  connected to some random experiment,  $A, B,$  and  $C$  are independent and  $P(A) = 0.88, P(B) = 0.65,$  and  $P(C) = 0.44$ . Compute the indicated probability, or explain why there is not enough information to do so.
    - $P(A \cap B \cap C)$
    - $P(A^c \cap B^c \cap C^c)$
  - Suppose for events  $A, B,$  and  $C$  connected to some random experiment,  $A, B,$  and  $C$  are independent and  $P(A) = 0.95, P(B) = 0.73,$  and  $P(C) = 0.62$ . Compute the indicated probability, or explain why there is not enough information to do so.
    - $P(A \cap B \cap C)$
    - $P(A^c \cap B^c \cap C^c)$

## APPLICATIONS

17. The sample space that describes all three-child families according to the genders of the children with respect to birth order is

$$S = \{bbb, bbg, bgb, bgg, gbb, gbg, ggb, ggg\}$$

In the experiment of selecting a three-child family at random, compute each of the following probabilities, assuming all outcomes are equally likely.

- The probability that the family has at least two boys.
  - The probability that the family has at least two boys, given that not all of the children are girls.
  - The probability that at least one child is a boy.
  - The probability that at least one child is a boy, given that the first born is a girl.
18. The following two-way contingency table gives the breakdown of the population in a particular locale according to age and number of vehicular moving violations in the past three years:

| Age      | Violations |      |      |
|----------|------------|------|------|
|          | 0          | 1    | 2+   |
| Under 21 | 0.04       | 0.06 | 0.02 |
| 21–40    | 0.25       | 0.16 | 0.01 |
| 41–60    | 0.23       | 0.10 | 0.02 |
| 60+      | 0.08       | 0.03 | 0.00 |

A person is selected at random. Find the following probabilities.

- The person is under 21.
  - The person has had at least two violations in the past three years.
  - The person has had at least two violations in the past three years, given that he is under 21.
  - The person is under 21, given that he has had at least two violations in the past three years.
  - Determine whether the events “the person is under 21” and “the person has had at least two violations in the past three years” are independent or not.
19. The following two-way contingency table gives the breakdown of the population in a particular locale according to party affiliation (*A*, *B*, *C*, or *None*) and opinion on a bond issue:

| Affiliation | Opinion |         |           |
|-------------|---------|---------|-----------|
|             | Favors  | Opposes | Undecided |
| A           | 0.12    | 0.09    | 0.07      |
| B           | 0.16    | 0.12    | 0.14      |
| C           | 0.04    | 0.03    | 0.06      |
| None        | 0.08    | 0.06    | 0.03      |

A person is selected at random. Find each of the following probabilities.

- The person is in favor of the bond issue.
  - The person is in favor of the bond issue, given that he is affiliated with party A.
  - The person is in favor of the bond issue, given that he is affiliated with party B.
20. The following two-way contingency table gives the breakdown of the population of patrons at a grocery store according to the number of items purchased and whether or not the patron made an impulse purchase at the checkout counter:

| Number of Items | Impulse Purchase |          |
|-----------------|------------------|----------|
|                 | Made             | Not Made |
| Few             | 0.01             | 0.19     |
| Many            | 0.04             | 0.76     |

A patron is selected at random. Find each of the following probabilities.

- The patron made an impulse purchase.
  - The patron made an impulse purchase, given that the total number of items purchased was many.
  - Determine whether or not the events “few purchases” and “made an impulse purchase at the checkout counter” are independent.
21. The following two-way contingency table gives the breakdown of the population of adults in a particular locale according to employment type and level of life insurance:

| Employment Type | Level of Insurance |        |      |
|-----------------|--------------------|--------|------|
|                 | Low                | Medium | High |
| Unskilled       | 0.07               | 0.19   | 0.00 |

| Employment Type | Level of Insurance |        |      |
|-----------------|--------------------|--------|------|
|                 | Low                | Medium | High |
| Semi-skilled    | 0.04               | 0.28   | 0.08 |
| Skilled         | 0.03               | 0.18   | 0.05 |
| Professional    | 0.01               | 0.05   | 0.02 |

An adult is selected at random. Find each of the following probabilities.

- The person has a high level of life insurance.
  - The person has a high level of life insurance, given that he does not have a professional position.
  - The person has a high level of life insurance, given that he has a professional position.
  - Determine whether or not the events “has a high level of life insurance” and “has a professional position” are independent.
22. The sample space of equally likely outcomes for the experiment of rolling two fair dice is

11 12 13 14 15 16  
 21 22 23 24 25 26  
 31 32 33 34 35 36  
 41 42 43 44 45 46  
 51 52 53 54 55 56  
 61 62 63 64 65 66

Identify the events  $N$ : the sum is at least nine,  $T$ : at least one of the dice is a two, and  $F$ : at least one of the dice is a five.

- Find  $P(N)$ .
  - Find  $P(N|F)$ .
  - Find  $P(N|T)$ .
  - Determine from the previous answers whether or not the events  $N$  and  $F$  are independent; whether or not  $N$  and  $T$  are.
23. The *sensitivity* of a drug test is the probability that the test will be positive when administered to a person who has actually taken the drug. Suppose that there are two independent tests to detect the presence of a certain type of banned drugs in athletes. One has sensitivity 0.75; the other has sensitivity 0.85. If both are applied to an athlete who has taken this type of drug, what is the chance that his usage will go undetected?

24. A man has two lights in his well house to keep the pipes from freezing in winter. He checks the lights daily. Each light has probability 0.002 of burning out before it is checked the next day (independently of the other light).
- If the lights are wired in parallel one will continue to shine even if the other burns out. In this situation, compute the probability that at least one light will continue to shine for the full 24 hours. Note the greatly increased reliability of the system of two bulbs over that of a single bulb.
  - If the lights are wired in series neither one will continue to shine even if only one of them burns out. In this situation, compute the probability that at least one light will continue to shine for the full 24 hours. Note the slightly decreased reliability of the system of two bulbs over that of a single bulb.
25. An accountant has observed that 5% of all copies of a particular two-part form have an error in Part I, and 2% have an error in Part II. If the errors occur independently, find the probability that a randomly selected form will be error-free.
26. A box contains 20 screws which are identical in size, but 12 of which are zinc coated and 8 of which are not. Two screws are selected at random, without replacement.
- Find the probability that both are zinc coated.
  - Find the probability that at least one is zinc coated.

### ADDITIONAL EXERCISES

27. Events  $A$  and  $B$  are mutually exclusive. Find  $P(A|B)$ .
28. The city council of a particular city is composed of five members of party  $A$ , four members of party  $B$ , and three independents. Two council members are randomly selected to form an investigative committee.
- Find the probability that both are from party  $A$ .
  - Find the probability that at least one is an independent.
  - Find the probability that the two have different party affiliations (that is, not both  $A$ , not both  $B$ , and not both independent).
29. A basketball player makes 60% of the free throws that he attempts, except that if he has just tried and missed a free throw then his chances of making a second one go down to only 30%. Suppose he has just been awarded two free throws.
- Find the probability that he makes both.

- b. Find the probability that he makes at least one. (A tree diagram could help.)
30. An economist wishes to ascertain the proportion  $p$  of the population of individual taxpayers who have purposely submitted fraudulent information on an income tax return. To truly guarantee anonymity of the taxpayers in a random survey, taxpayers questioned are given the following instructions.
1. Flip a coin.
  2. If the coin lands heads, answer “Yes” to the question “Have you ever submitted fraudulent information on a tax return?” even if you have not.
  3. If the coin lands tails, give a truthful “Yes” or “No” answer to the question “Have you ever submitted fraudulent information on a tax return?”

The questioner is not told how the coin landed, so he does not know if a “Yes” answer is the truth or is given only because of the coin toss.

- a. Using the Probability Rule for Complements and the independence of the coin toss and the taxpayers’ status fill in the empty cells in the two-way contingency table shown. Assume that the coin is fair. Each cell except the two in the bottom row will contain the unknown proportion (or probability)  $p$ .

| Status      | Coin |   | Probability |
|-------------|------|---|-------------|
|             | H    | T |             |
| Fraud       |      |   | $p$         |
| No fraud    |      |   |             |
| Probability |      |   | 1           |

- b. The only information that the economist sees are the entries in the following table:

| Response   | “Yes” | “No” |
|------------|-------|------|
| Proportion | $r$   | $s$  |

Equate the entry in the one cell in the table in (a) that corresponds to the answer “No” to the number  $s$  to obtain the formula  $p = 1 - 2s$  that expresses the unknown number  $p$  in terms of the known number  $s$ .

- c. Equate the sum of the entries in the three cells in the table in (a) that together correspond to the answer “Yes” to the number  $r$  to obtain the formula  $p = 2r - 1$  that expresses the unknown number  $p$  in terms of the known number  $r$ .

- d. Use the fact that  $r + s = 1$  (since they are the probabilities of complementary events) to verify that the formulas in (b) and (c) give the same value for  $p$ . (For example, insert  $s = 1 - r$  into the formula in (b) to obtain the formula in (c).)
- e. Suppose a survey of 1,200 taxpayers is conducted and 690 respond “Yes” (truthfully or not) to the question “Have you ever submitted fraudulent information on a tax return?” Use the answer to either (b) or (c) to estimate the true proportion  $p$  of all individual taxpayers who have purposely submitted fraudulent information on an income tax return.

## ANSWERS

1.
  - a. 0.6
  - b. 0.4
  - c. not independent
3.
  - a. 0.22
  - b. 0.81
  - c. 0.27
5.
  - a. 0
  - b. 0
7.
  - a. 0.5
  - b. 0.4
  - c. 0.6
9.
  - a. 0.25
  - b. 0.33
  - c. 0
  - d. 0.25
11.
  - a.  $P(A) = 0.3, P(R) = 0.4, P(A \cap R) = 0.12$
  - b. independent
  - c. without computation 0.3
13.
  - a. Insufficient information. The events  $A$  and  $B$  are not known to be either independent or mutually exclusive.
  - b. 0.21
  - c. 0
15.
  - a. 0.25
  - b. 0.02
17.
  - a. 0.5
  - b. 0.57
  - c. 0.875
  - d. 0.75
19.
  - a. 0.4
  - b. 0.43
  - c. 0.38
21.
  - a. 0.15
  - b. 0.14
  - c. 0.25

- d. not independent
- 23. 0.0375
- 25. 0.931
- 27. 0
- 29.
  - a. 0.36
  - b. 0.72