CHAPTER 3

PROBABILITY: THE BASIS OF STATISTICAL INFERENCE

CHAPTER OVERVIEW

Probability lays the foundation for statistical inference. This chapter provides a brief overview of the probability concepts necessary for the understanding of topics covered in the chapters that follow. It also provides a context for understanding the probability distributions used in statistical inference, and introduces the student to several measures commonly found in the medical literature (e.g., the sensitivity and specificity of a test).

TOPICS

LEARNING OUTCOMES

After studying this chapter, the student will

1. understand classical, relative frequency, and subjective probability.
2. understand the properties of probability and selected probability rules.
3. be able to calculate the probability of an event.
4. be able to apply Bayes' theorem when calculating screening test results.

3.1 INTRODUCTION

The theory of probability provides the foundation for statistical inference. However, this theory, which is a branch of mathematics, is not the main concern of this book, and, consequently, only its fundamental concepts are discussed here. Students who desire to pursue this subject should refer to the many books on probability available in most college and university libraries. The books by Gut (1), Isaac (2), and Larson (3) are recommended. The objectives of this chapter are to help students gain some mathematical ability in the area of probability and to assist them in developing an understanding of the more important concepts. Progress along these lines will contribute immensely to their success in understanding the statistical inference procedures presented later in this book.

The concept of probability is not foreign to health workers and is frequently encountered in everyday communication. For example, we may hear a physician say that a patient has a 50–50 chance of surviving a certain operation. Another physician may say that she is 95 percent certain that a patient has a particular disease. A public health nurse may say that nine times out of ten a certain client will break an appointment. As these examples suggest, most people express probabilities in terms of percentages. In dealing with probabilities mathematically, it is more convenient to express probabilities as fractions. (Percentages result from multiplying the fractions by 100.) Thus, we measure the probability of the occurrence of some event by a number between zero and one. The more likely the event, the closer the number is to one; and the more unlikely the event, the closer the number is to zero. An event that cannot occur has a probability of zero, and an event that is certain to occur has a probability of one.

Health sciences researchers continually ask themselves if the results of their efforts could have occurred by chance alone or if some other force was operating to produce the observed effects. For example, suppose six out of ten patients suffering from some disease are cured after receiving a certain treatment. Is such a cure rate likely to have occurred if the patients had not received the treatment, or is it evidence of a true curative effect on the part of the treatment? We shall see that questions such as these can be answered through the application of the concepts and laws of probability.

3.2 TWO VIEWS OF PROBABILITY: OBJECTIVE AND SUBJECTIVE

Until fairly recently, probability was thought of by statisticians and mathematicians only as an *objective* phenomenon derived from objective processes.

The concept of *objective probability* may be categorized further under the headings of (1) *classical*, or *a priori*, *probability*; and (2) the *relative frequency*, or *a posteriori*, concept of probability.

Classical Probability The classical treatment of probability dates back to the 17th century and the work of two mathematicians, Pascal and Fermat. Much of this theory developed out of attempts to solve problems related to games of chance, such as those involving the rolling of dice. Examples from games of chance illustrate very well the principles involved in classical probability. For example, if a fair six-sided die is rolled, the probability that a 1 will be observed is equal to 1/6 and is the same for the other five faces. If a card is picked at random from a well-shuffled deck of ordinary playing cards, the probability of picking a heart is 13/52. Probabilities such as these are calculated by the processes of abstract reasoning. It is not necessary to roll a die or draw a card to compute these probabilities. In the rolling of the die, we say that each of the six sides is *equally likely* to be observed if there is no reason to favor any one of the six sides. Similarly, if there is no reason to favor the drawing of a particular card from a deck of cards, we say that each of the 52 cards is equally likely to be drawn. We may define probability in the classical sense as follows:

DEFINITION

If an event can occur in N mutually exclusive and equally likely ways, and if m of these possess a trait E, the probability of the occurrence of E is equal to m/N.

If we read $P(E)$ as "the probability of E," we may express this definition as

$$P(E) = \frac{m}{N} \tag{3.2.1}$$

Relative Frequency Probability The relative frequency approach to probability depends on the repeatability of some process and the ability to count the number of repetitions, as well as the number of times that some event of interest occurs. In this context we may define the probability of observing some characteristic, E, of an event as follows:

DEFINITION

If some process is repeated a large number of times, n, and if some resulting event with the characteristic E occurs m times, the relative frequency of occurrence of E, m/n, will be approximately equal to the probability of E.

To express this definition in compact form, we write

$$P(E) = \frac{m}{n} \tag{3.2.2}$$

We must keep in mind, however, that, strictly speaking, m/n is only an estimate of $P(E)$.

Subjective Probability In the early 1950s, L. J. Savage (4) gave considerable impetus to what is called the "personalistic" or subjective concept of probability. This view holds that probability measures the confidence that a particular individual has in the truth of a particular proposition. This concept does not rely on the repeatability of any process. In fact, by applying this concept of probability, one may evaluate the probability of an event that can only happen once, for example, the probability that a cure for cancer will be discovered within the next 10 years.

Although the subjective view of probability has enjoyed increased attention over the years, it has not been fully accepted by statisticians who have traditional orientations.

Bayesian Methods Bayesian methods are named in honor of the Reverend Thomas Bayes (1702–1761), an English clergyman who had an interest in mathematics. Bayesian methods are an example of subjective probability, since it takes into consideration the degree of belief that one has in the chance that an event will occur. While probabilities based on classical or relative frequency concepts are designed to allow for decisions to be made solely on the basis of collected data, Bayesian methods make use of what are known as *prior probabilities* and *posterior probabilities*.

DEFINITION

The *prior probability* of an event is a probability based on prior knowledge, prior experience, or results derived from prior data collection activity.

DEFINITION

The *posterior probability* of an event is a probability obtained by using new information to update or revise a prior probability.

As more data are gathered, the more is likely to be known about the "true" probability of the event under consideration. Although the idea of updating probabilities based on new information is in direct contrast to the philosophy behind frequency-of-occurrence probability, Bayesian concepts are widely used. For example, Bayesian techniques have found recent application in the construction of e-mail spam filters. Typically, the application of Bayesian concepts makes use of a mathematical formula called *Bayes' theorem*. In Section 3.5 we employ Bayes' theorem in the evaluation of diagnostic screening test data.

3.3 ELEMENTARY PROPERTIES OF PROBABILITY

In 1933 the axiomatic approach to probability was formalized by the Russian mathematician A. N. Kolmogorov (5). The basis of this approach is embodied in three properties from which a whole system of probability theory is constructed through the use of mathematical logic. The three properties are as follows.

1. Given some process (or experiment) with n mutually exclusive outcomes (called events), $E_1, E_2, \ldots, E_n$, the probability of any event E_i is assigned a nonnegative number. That is,

$$P(E_i) \geq 0 \tag{3.3.1}$$

In other words, all events must have a probability greater than or equal to zero, a reasonable requirement in view of the difficulty of conceiving of negative probability. A key concept in the statement of this property is the concept of *mutually exclusive* outcomes. Two events are said to be mutually exclusive if they cannot occur simultaneously.

2. The sum of the probabilities of the mutually exclusive outcomes is equal to 1.

$$P(E_1) + P(E_2) + \cdots + P(E_n) = 1 \tag{3.3.2}$$

This is the property of *exhaustiveness* and refers to the fact that the observer of a probabilistic process must allow for all possible events, and when all are taken together, their total probability is 1. The requirement that the events be mutually exclusive is specifying that the events $E_1, E_2, \ldots, E_n$ do not overlap; that is, no two of them can occur at the same time.

3. Consider any two mutually exclusive events, E_i and E_j. The probability of the occurrence of either E_i or E_j is equal to the sum of their individual probabilities.

$$P(E_i + E_j) = P(E_i) + P(E_j) \tag{3.3.3}$$

Suppose the two events were not mutually exclusive; that is, suppose they could occur at the same time. In attempting to compute the probability of the occurrence of either E_i or E_j the problem of overlapping would be discovered, and the procedure could become quite complicated. This concept will be discussed further in the next section.

3.4 CALCULATING THE PROBABILITY OF AN EVENT

We now make use of the concepts and techniques of the previous sections in calculating the probabilities of specific events. Additional ideas will be introduced as needed.

EXAMPLE 3.4.1

The primary aim of a study by Carter et al. (A-1) was to investigate the effect of the age at onset of bipolar disorder on the course of the illness. One of the variables investigated was family history of mood disorders. Table 3.4.1 shows the frequency of a family history of mood disorders in the two groups of interest (Early age at onset defined to be

TABLE 3.4.1 Frequency of Family History of Mood Disorder by Age Group Among Bipolar Subjects

Family History of Mood Disorders	Early = 18(*E*)	Later > 18(*L*)	Total
Negative (*A*)	28	35	63
Bipolar disorder (*B*)	19	38	57
Unipolar (*C*)	41	44	85
Unipolar and bipolar (*D*)	53	60	113
Total	141	177	318

Source: Tasha D. Carter, Emanuela Mundo, Sagar V. Parkh, and James L. Kennedy, "Early Age at Onset as a Risk Factor for Poor Outcome of Bipolar Disorder," *Journal of Psychiatric Research, 37* (2003), 297–303.

18 years or younger and Later age at onset defined to be later than 18 years). Suppose we pick a person at random from this sample. What is the probability that this person will be 18 years old or younger?

Solution: For purposes of illustrating the calculation of probabilities, we consider this group of 318 subjects to be the largest group for which we have an interest. In other words, for this example, we consider the 318 subjects as a population. We assume that Early and Later are mutually exclusive categories and that the likelihood of selecting any one person is equal to the likelihood of selecting any other person. We define the desired probability as the number of subjects with the characteristic of interest (Early) divided by the total number of subjects. We may write the result in probability notation as follows:

$$\begin{aligned} P(E) &= \text{number of Early subjects/total number of subjects} \\ &= 141/318 = .4434 \end{aligned}$$

■

Conditional Probability On occasion, the set of "all possible outcomes" may constitute a subset of the total group. In other words, the size of the group of interest may be reduced by conditions not applicable to the total group. When probabilities are calculated with a subset of the total group as the denominator, the result is a *conditional probability*.

The probability computed in Example 3.4.1, for example, may be thought of as an unconditional probability, since the size of the total group served as the denominator. No conditions were imposed to restrict the size of the denominator. We may also think of this probability as a *marginal probability* since one of the marginal totals was used as the numerator.

We may illustrate the concept of conditional probability by referring again to Table 3.4.1.

EXAMPLE 3.4.2

Suppose we pick a subject at random from the 318 subjects and find that he is 18 years or younger (E). What is the probability that this subject will be one who has no family history of mood disorders (A)?

Solution: The total number of subjects is no longer of interest, since, with the selection of an Early subject, the Later subjects are eliminated. We may define the desired probability, then, as follows: What is the probability that a subject has no family history of mood disorders (A), given that the selected subject is Early (E)? This is a conditional probability and is written as $P(A \mid E)$ in which the vertical line is read "given." The 141 Early subjects become the denominator of this conditional probability, and 28, the number of Early subjects with no family history of mood disorders, becomes the numerator. Our desired probability, then, is

$$P(A \mid E) = 28/141 = .1986$$

■

Joint Probability Sometimes we want to find the probability that a subject picked at random from a group of subjects possesses two characteristics at the same time. Such a probability is referred to as a *joint probability*. We illustrate the calculation of a joint probability with the following example.

EXAMPLE 3.4.3

Let us refer again to Table 3.4.1. What is the probability that a person picked at random from the 318 subjects will be Early (E) *and* will be a person who has no family history of mood disorders (A)?

Solution: The probability we are seeking may be written in symbolic notation as $P(E \cap A)$ in which the symbol $\cap$ is read either as "intersection" or "and." The statement $E \cap A$ indicates the joint occurrence of conditions E and A. The number of subjects satisfying both of the desired conditions is found in Table 3.4.1 at the intersection of the column labeled E and the row labeled A and is seen to be 28. Since the selection will be made from the total set of subjects, the denominator is 318. Thus, we may write the joint probability as

$$P(E \cap A) = 28/318 = .0881$$

■

The Multiplication Rule A probability may be computed from other probabilities. For example, a joint probability may be computed as the product of an appropriate marginal probability and an appropriate conditional probability. This relationship is known as the *multiplication rule* of probability. We illustrate with the following example.

EXAMPLE 3.4.4

We wish to compute the joint probability of Early age at onset (E) and a negative family history of mood disorders (A) from knowledge of an appropriate marginal probability and an appropriate conditional probability.

Solution: The probability we seek is $P(E \cap A)$. We have already computed a marginal probability, $P(E) = 141/318 = .4434$, and a conditional probability, $P(A \mid E) = 28/141 = .1986$. It so happens that these are appropriate marginal and conditional probabilities for computing the desired joint probability. We may now compute $P(E \cap A) = P(E)P(A \mid E) = (.4434)(.1986) = .0881$. This, we note, is, as expected, the same result we obtained earlier for $P(E \cap A)$.

■

We may state the multiplication rule in general terms as follows: For any two events A and B,

$$P(A \cap B) = P(B)P(A \mid B), \qquad \text{if } P(B) \neq 0 \tag{3.4.1}$$

For the same two events A and B, the multiplication rule may also be written as $P(A \cap B) = P(A)P(B \mid A)$, if $P(A) \neq 0$.

We see that through algebraic manipulation the multiplication rule as stated in Equation 3.4.1 may be used to find any one of the three probabilities in its statement if the other two are known. We may, for example, find the conditional probability $P(A \mid B)$ by dividing $P(A \cap B)$ by $P(B)$. This relationship allows us to formally define conditional probability as follows.

DEFINITION

The *conditional probability* of A given B is equal to the probability of $A \cap B$ divided by the probability of B, provided the probability of B is not zero.

That is,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) \neq 0 \tag{3.4.2}$$

We illustrate the use of the multiplication rule to compute a conditional probability with the following example.

EXAMPLE 3.4.5

We wish to use Equation 3.4.2 and the data in Table 3.4.1 to find the conditional probability, $P(A \mid E)$.

Solution: According to Equation 3.4.2,

$$P(A \mid E) = P(A \cap E)/P(E)$$

■

Earlier we found $P(E \cap A) = P(A \cap E) = 28/318 = .0881$. We have also determined that $P(E) = 141/318 = .4434$. Using these results we are able to compute $P(A \mid E) = .0881/.4434 = .1987$, which, as expected, is the same result we obtained by using the frequencies directly from Table 3.4.1. (The slight discrepancy is due to rounding.)

The Addition Rule The third property of probability given previously states that the probability of the occurrence of either one or the other of two mutually exclusive events is equal to the sum of their individual probabilities. Suppose, for example, that we pick a person at random from the 318 represented in Table 3.4.1. What is the probability that this person will be Early age at onset (E) or Later age at onset (L)? We state this probability in symbols as $P(E \cup L)$, where the symbol $\cup$ is read either as "union" or "or." Since the two age conditions are mutually exclusive, $P(E \cap L) = (141/318) + (177/318) = .4434 + .5566 = 1$.

What if two events are not mutually exclusive? This case is covered by what is known as the *addition rule*, which may be stated as follows:

DEFINITION

Given two events A and B, the probability that event A, or event B, or both occur is equal to the probability that event A occurs, plus the probability that event B occurs, minus the probability that the events occur simultaneously.

The addition rule may be written

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{3.4.3}$$

When events A and B cannot occur simultaneously, $P(A \cap B)$ is sometimes called "exclusive or," and $P(A \cup B) = 0$. When events A and B can occur simultaneously, $P(A \cup B)$ is sometimes called "inclusive or," and we use the addition rule to calculate $P(A \cup B)$. Let us illustrate the use of the addition rule by means of an example.

EXAMPLE 3.4.6

If we select a person at random from the 318 subjects represented in Table 3.4.1, what is the probability that this person will be an Early age of onset subject (E) or will have no family history of mood disorders (A) or both?

Solution: The probability we seek is $P(E \cup A)$. By the addition rule as expressed by Equation 3.4.3, this probability may be written as $P(E \cup A) = P(E) + P(A) - P(E \cap A)$. We have already found that $P(E) = 141/318 = .4434$ and $P(E \cap A) = 28/318 = .0881$. From the information in Table 3.4.1 we calculate $P(A) = 63/318 = .1981$. Substituting these results into the equation for $P(E \cup A)$ we have $P(E \cup A) = .4434 + .1981 - .0881 = .5534$. ■

Note that the 28 subjects who are *both* Early *and* have no family history of mood disorders are included in the 141 who are Early as well as in the 63 who have no family history of mood disorders. Since, in computing the probability, these 28 have been added into the numerator twice, they have to be subtracted out once to overcome the effect of duplication, or overlapping.

Independent Events Suppose that, in Equation 3.4.2, we are told that event B has occurred, but that this fact has no effect on the probability of A. That is, suppose that the probability of event A is the same regardless of whether or not B occurs. In this situation, $P(A \mid B) = P(A)$. In such cases we say that A and B are *independent events*. The multiplication rule for two independent events, then, may be written as

$$P(A \cap B) = P(A)P(B); \quad P(A) \neq 0, \quad P(B) \neq 0 \tag{3.4.4}$$

Thus, we see that if two events are independent, the probability of their joint occurrence is equal to the product of the probabilities of their individual occurrences.

Note that when two events with nonzero probabilities are independent, each of the following statements is true:

$$P(A \mid B) = P(A), \quad P(B \mid A) = P(B), \quad P(A \cap B) = P(A)P(B)$$

Two events are not independent unless all these statements are true. It is important to be aware that the terms *independent* and *mutually exclusive* do not mean the same thing.

Let us illustrate the concept of independence by means of the following example.

EXAMPLE 3.4.7

In a certain high school class, consisting of 60 girls and 40 boys, it is observed that 24 girls and 16 boys wear eyeglasses. If a student is picked at random from this class, the probability that the student wears eyeglasses, $P(E)$, is 40/100, or .4.

(a) What is the probability that a student picked at random wears eyeglasses, given that the student is a boy?

Solution: By using the formula for computing a conditional probability, we find this to be

$$P(E \mid B) = \frac{P(E \cap B)}{P(B)} = \frac{16/100}{40/100} = .4$$

Thus the additional information that a student is a boy does not alter the probability that the student wears eyeglasses, and $P(E) = P(E \mid B)$. We say that the events being a boy and wearing eyeglasses for this group are independent. We may also show that the event of wearing eyeglasses, E, and *not* being a boy, $\overline{B}$ are also independent as follows:

$$P(E \mid \overline{B}) = \frac{P(E \cap \overline{B})}{P(\overline{B})} = \frac{24/100}{60/100} = \frac{24}{60} = .4$$

(b) What is the probability of the joint occurrence of the events of wearing eyeglasses and being a boy?

Solution: Using the rule given in Equation 3.4.1, we have

$$P(E \cap B) = P(B)P(E \mid B)$$

but, since we have shown that events E and B are independent we may replace $P(E \mid B)$ by $P(E)$ to obtain, by Equation 3.4.4,

$$\begin{aligned} P(E \cap B) &= P(B)P(E) \\ &= \left(\frac{40}{100}\right)\left(\frac{40}{100}\right) \\ &= .16 \end{aligned}$$

■

Complementary Events Earlier, using the data in Table 3.4.1, we computed the probability that a person picked at random from the 318 subjects will be an Early age of onset subject as $P(E) = 141/318 = .4434$. We found the probability of a Later age at onset to be $P(L) = 177/318 = .5566$. The sum of these two probabilities we found to be equal to 1. This is true because the events being Early age at onset and being Later age at onset are *complementary events*. In general, we may make the following statement about complementary events. The probability of an event A is equal to 1 minus the probability of its complement, which is written $\overline{A}$, and

$$P(\overline{A}) = 1 - P(A) \tag{3.4.5}$$

This follows from the third property of probability since the event, A, and its complement, $\overline{A}$, are mutually exclusive.

EXAMPLE 3.4.8

Suppose that of 1200 admissions to a general hospital during a certain period of time, 750 are private admissions. If we designate these as set A, then $\overline{A}$ is equal to 1200 minus 750, or 450. We may compute

$$P(A) = 750/1200 = .625$$

and

$$P(\overline{A}) = 450/1200 = .375$$

and see that

$$\begin{aligned} P(\overline{A}) &= 1 - P(A) \\ .375 &= 1 - .625 \\ .375 &= .375 \end{aligned}$$

■

Marginal Probability Earlier we used the term *marginal probability* to refer to a probability in which the numerator of the probability is a marginal total from a table such as Table 3.4.1. For example, when we compute the probability that a person picked at random from the 318 persons represented in Table 3.4.1 is an Early age of onset subject, the numerator of the probability is the total number of Early subjects, 141. Thus, $P(E) = 141/318 = .4434$. We may define marginal probability more generally as follows:

DEFINITION

Given some variable that can be broken down into m categories designated by $A_1, A_2, \ldots, A_i, \ldots, A_m$ and another jointly occurring variable that is broken down into n categories designated by $B_1, B_2, \ldots, B_j, \ldots, B_n$, the marginal probability of A_i, $P(A_i)$, is equal to the sum of the joint probabilities of A_i with all the categories of B. That is,

$$P(A_i) = \sum P(A_i \cap B_j), \quad \text{for all values of } j \tag{3.4.6}$$

The following example illustrates the use of Equation 3.4.6 in the calculation of a marginal probability.

EXAMPLE 3.4.9

We wish to use Equation 3.4.6 and the data in Table 3.4.1 to compute the marginal probability $P(E)$.

Solution: The variable age at onset is broken down into two categories, Early for onset 18 years or younger (E) and Later for onset occurring at an age over 18 years (L). The variable family history of mood disorders is broken down into four categories: negative family history (A), bipolar disorder only (B), unipolar disorder only (C), and subjects with a history of both unipolar and bipolar disorder (D). The category Early occurs jointly with all four categories of the variable family history of mood disorders. The four joint probabilities that may be computed are

$$P(E \cap A) = 28/318 = .0881$$
$$P(E \cap B) = 19/318 = .0597$$
$$P(E \cap C) = 41/318 = .1289$$
$$P(E \cap D) = 53/318 = .1667$$

We obtain the marginal probability $P(E)$ by adding these four joint probabilities as follows:

$$\begin{aligned} P(E) &= P(E \cap A) + P(E \cap B) + P(E \cap C) + P(E \cap D) \\ &= .0881 + .0597 + .1289 + .1667 \\ &= .4434 \end{aligned}$$

■

The result, as expected, is the same as the one obtained by using the marginal total for Early as the numerator and the total number of subjects as the denominator.

EXERCISES

3.4.1 In a study of violent victimization of women and men, Porcerelli et al. (A-2) collected information from 679 women and 345 men aged 18 to 64 years at several family practice centers in the metropolitan Detroit area. Patients filled out a health history questionnaire that included a question about victimization. The following table shows the sample subjects cross-classified by sex and the type of violent victimization reported. The victimization categories are defined as no victimization, partner victimization (and not by others), victimization by persons other than partners (friends, family members, or strangers), and those who reported multiple victimization.

	No Victimization	Partners	Nonpartners	Multiple Victimization	Total
Women	611	34	16	18	679
Men	308	10	17	10	345
Total	919	44	33	28	1024

Source: John H. Porcerelli, Ph.D., Rosemary Cogan, Ph.D. Used with permission.

(a) Suppose we pick a subject at random from this group. What is the probability that this subject will be a woman?

(b) What do we call the probability calculated in part a?

(c) Show how to calculate the probability asked for in part a by two additional methods.

(d) If we pick a subject at random, what is the probability that the subject will be a woman and have experienced partner abuse?

(e) What do we call the probability calculated in part d?

(f) Suppose we picked a man at random. Knowing this information, what is the probability that he experienced abuse from nonpartners?

(g) What do we call the probability calculated in part f?

(h) Suppose we pick a subject at random. What is the probability that it is a man or someone who experienced abuse from a partner?

(i) What do we call the method by which you obtained the probability in part h?

3.4.2 Fernando et al. (A-3) studied drug-sharing among injection drug users in the South Bronx in New York City. Drug users in New York City use the term "split a bag" or "get down on a bag" to refer to the practice of dividing a bag of heroin or other injectable substances. A common practice includes splitting drugs after they are dissolved in a common cooker, a procedure with considerable HIV risk. Although this practice is common, little is known about the prevalence of such practices. The researchers asked injection drug users in four neighborhoods in the South Bronx if they ever "got down on" drugs in bags or shots. The results classified by gender and splitting practice are given below:

Gender	Split Drugs	Never Split Drugs	Total
Male	349	324	673
Female	220	128	348
Total	569	452	1021

Source: Daniel Fernando, Robert F. Schilling, Jorge Fontdevila, and Nabila El-Bassel, "Predictors of Sharing Drugs Among Injection Drug Users in the South Bronx: Implications for HIV Transmission," *Journal of Psychoactive Drugs*, 35 (2003), 227–236.

(a) How many marginal probabilities can be calculated from these data? State each in probability notation and do the calculations.

(b) How many joint probabilities can be calculated? State each in probability notation and do the calculations.

(c) How many conditional probabilities can be calculated? State each in probability notation and do the calculations.

(d) Use the multiplication rule to find the probability that a person picked at random never split drugs and is female.

(e) What do we call the probability calculated in part d?

(f) Use the multiplication rule to find the probability that a person picked at random is male, given that he admits to splitting drugs.

(g) What do we call the probability calculated in part f?

3.4.3 Refer to the data in Exercise 3.4.2. State the following probabilities in words and calculate:

(a) $P(\text{Male} \cap \text{Split Drugs})$

(b) $P(\text{Male} \cup \text{Split Drugs})$

(c) $P(\text{Male} \mid \text{Split Drugs})$

(d) $P(\text{Male})$

3.4.4 Laveist and Nuru-Jeter (A-4) conducted a study to determine if doctor–patient race concordance was associated with greater satisfaction with care. Toward that end, they collected a national sample of African-American, Caucasian, Hispanic, and Asian-American respondents. The following table classifies the race of the subjects as well as the race of their physician:

	Patient's Race				
Physician's Race	**Caucasian**	**African-American**	**Hispanic**	**Asian-American**	**Total**
White	779	436	406	175	1796
African-American	14	162	15	5	196
Hispanic	19	17	128	2	166
Asian/Pacific-Islander	68	75	71	203	417
Other	30	55	56	4	145
Total	910	745	676	389	2720

Source: Thomas A. Laveist and Amani Nuru-Jeter, "Is Doctor-Patient Race Concordance Associated with Greater Satisfaction with Care?" *Journal of Health and Social Behavior*, *43* (2002), 296–306.

(a) What is the probability that a randomly selected subject will have an Asian/Pacific-Islander physician?

(b) What is the probability that an African-American subject will have an African-American physician?

(c) What is the probability that a randomly selected subject in the study will be Asian-American and have an Asian/Pacific-Islander physician?

(d) What is the probability that a subject chosen at random will be Hispanic or have a Hispanic physician?

(e) Use the concept of complementary events to find the probability that a subject chosen at random in the study does not have a white physician.