Chapter 5

Probability: An Introduction to Modeling Uncertainty

CONTENTS

ANALYTICS IN ACTION

National Aeronautics and Space Administration*

WASHINGTON, D.C.

The National Aeronautics and Space Administration (NASA) is the U.S. government agency that is responsible for the U.S. civilian space program and for aeronautics and aerospace research. NASA is best known for its manned space exploration; its mission statement is to "pioneer the future in space exploration, scientific discovery and aeronautics research." With 18,800 employees, NASA is currently working on the design of a new Space Launch System that will take the astronauts farther into space than ever before and provide the cornerstone for future space exploration.

Although NASA's primary mission is space exploration, its expertise has been called on in assisting countries and organizations throughout the world in nonspace endeavors. In one such situation, the San José copper and gold mine in Copiapó, Chile, caved in, trapping 33 men more than 2,000 feet underground. It was important to bring the men safely to the surface as quickly as possible, but it was also imperative that the rescue effort be carefully designed and implemented to save as many miners as possible. The Chilean government asked NASA to provide assistance in developing a rescue method. NASA sent a four-person team consisting of an engineer with expertise in vehicle design, two physicians, and a psychologist with knowledge about issues of long-term confinement.

The probability of success and the failure of various other rescue methods was prominent in the thoughts of everyone involved. Since no historical data were available to apply to this unique rescue situation, NASA scientists developed subjective probability estimates for the success and failure of various rescue methods based on similar circumstances experienced by astronauts returning from short- and long-term space missions. The probability estimates provided by NASA guided officials in the selection of a rescue method and provided insight as to how the miners would survive the ascent in a rescue cage. The rescue method designed by the Chilean officials in consultation with the NASA team resulted in the construction of 13-foot-long, 924-pound steel rescue capsule that would be used to bring up the miners one at a time. All miners were rescued, with the last emerging 68 days after the cave-in occurred.

In this chapter, you will learn about probability as well as how to compute and interpret probabilities for a variety of situations. The basic relationships of probability, conditional probability, and Bayes' theorem will be covered. We will also discuss the concepts of random variables and probability distributions and illustrate the use of some of the more common discrete and continuous probability distributions.

*The authors are indebted to Dr. Michael Duncan and Clinton Cragg at NASA for providing this Analytics in Action.

Identifying uncertainty in data was introduced in Chapters 2 and 3 through descriptive statistics and data-visualization techniques, respectively. In this chapter, we expand on our discussion of modeling uncertainty by formalizing the concept of probability and introducing the concept of probability distributions.

Uncertainty is an ever-present fact of life for decision makers, and much time and effort are spent trying to plan for, and respond to, uncertainty. Consider the CEO who has to make decisions about marketing budgets and production amounts using forecasted demands. Or consider the financial analyst who must determine how to build a client's portfolio of stocks and bonds when the rates of return for these investments are not known with certainty. In many business scenarios, data are available to provide information on possible outcomes for some decisions, but the exact outcome from a given decision is almost never known with certainty because many factors are outside the control of the decision maker (e.g., actions taken by competitors, the weather, etc.).

Probability is the numerical measure of the likelihood that an event will occur.[1] Therefore, it can be used as a measure of the uncertainty associated with an event. This measure of uncertainty is often communicated through a probability distribution. Probability distributions are extremely helpful in providing additional information about an

[1]Note that there are several different possible definitions of probability, depending on the method used to assign probabilities. This includes the classical definition, the relative frequency definition, and the subjective definition of probability. In this text, we most often use the relative frequency definition of probability, which assumes that probabilities are based on empirical data. For a more thorough discussion of the different possible definitions of probability see Chapter 4 of Anderson, Sweeney, Williams, Camm, and Cochran, *An Introduction to Statistics for Business and Economics*, 13e Revised (2018).

event, and as we will see in later chapters in this textbook, they can be used to help a decision maker evaluate possible actions and determine the best course of action.

5.1 Events and Probabilities

In discussing probabilities, we often start by defining a **random experiment** as a process that generates well-defined outcomes. Several examples of random experiments and their associated outcomes are shown in Table 5.1.

By specifying all possible outcomes, we identify the **sample space** for a random experiment. Consider the first random experiment in Table 5.1—a coin toss. The possible outcomes are head and tail. If we let S denote the sample space, we can use the following notation to describe the sample space.

$$S = \{\text{Head, Tail}\}$$

Suppose we consider the second random experiment in Table 5.1—rolling a die. The possible experimental outcomes, defined as the number of dots appearing on the upward face of the die, are the six points in the sample space for this random experiment.

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

Outcomes and events form the foundation of the study of probability. Formally, an **event** is defined as a collection of outcomes. For example, consider the case of an expansion project being undertaken by California Power & Light Company (CP&L). CP&L is starting a project designed to increase the generating capacity of one of its plants in Southern California. An analysis of similar construction projects indicates that the possible completion times for the project are 8, 9, 10, 11, and 12 months. Each of these possible completion times represents a possible outcome for this project. Table 5.2 shows the number of past construction projects that required 8, 9, 10, 11, and 12 months.

Let us assume that the CP&L project manager is interested in completing the project in 10 months or less. Referring to Table 5.2, we see that three possible outcomes (8 months, 9 months, and 10 months) provide completion times of 10 months or less. Letting C denote the event that the project is completed in 10 months or less, we write:

$$C = \{8, 9, 10\}$$

Event C is said to occur if *any one* of these outcomes occur.

A variety of additional events can be defined for the CP&L project:

L = The event that the project is completed in *less than* 10 months = $\{8, 9\}$

M = The event that the project is completed in *more than* 10 months = $\{11, 12\}$

In each case, the event must be identified as a collection of outcomes for the random experiment.

TABLE 5.1 Random Experiments and Experimental Outcomes

Random Experiment	Experimental Outcomes
Toss a coin	Head, tail
Roll a die	1, 2, 3, 4, 5, 6
Conduct a sales call	Purchase, no purchase
Hold a particular share of stock for one year	Price of stock goes up, price of stock goes down, no change in stock price
Reduce price of product	Demand goes up, demand goes down, no change in demand

TABLE 5.2 Completion Times for 40 CP&L Projects

Completion Time (months)	No. of Past Projects Having This Completion Time	Probability of Outcome
8	6	6/40 = 0.15
9	10	10/40 = 0.25
10	12	12/40 = 0.30
11	6	6/40 = 0.15
12	6	6/40 = 0.15
Total	40	1.00

The **probability of an event** is equal to the sum of the probabilities of outcomes for the event. Using this definition and given the probabilities of outcomes shown in Table 5.2, we can now calculate the probability of the event $C = \{8, 9, 10\}$. The probability of event C, denoted $P(C)$, is given by

$$P(C) = P(8) + P(9) + P(10) = 0.15 + 0.25 + 0.30 = 0.70$$

Similarly, because the event that the project is completed in less than 10 months is given by $L = \{8, 9\}$, the probability of this event is given by

$$P(L) = P(8) + P(9) = 0.15 + 0.25 = 0.40$$

Finally, for the event that the project is completed in more than 10 months, we have $M = \{11, 12\}$ and thus

$$P(M) = P(11) + P(12) = 0.15 + 0.15 = 0.30$$

Using these probability results, we can now tell CP&L management that there is a 0.70 probability that the project will be completed in 10 months or less, a 0.40 probability that it will be completed in less than 10 months, and a 0.30 probability that it will be completed in more than 10 months.

5.2 Some Basic Relationships of Probability

Complement of an Event

The complement of event A is sometimes written as $\bar{A}$ or A′ in other textbooks.

Given an event A, the **complement of A** is defined to be the event consisting of all outcomes that are *not* in A. The complement of A is denoted by A^C. Figure 5.1 shows what is known as a **Venn diagram**, which illustrates the concept of a complement. The rectangular area represents the sample space for the random experiment and, as such, contains all possible outcomes. The circle represents event A and contains only the outcomes that belong to A. The shaded region of the rectangle contains all outcomes not in event A and is by definition the complement of A.

In any probability application, either event A or its complement A^C must occur. Therefore, we have

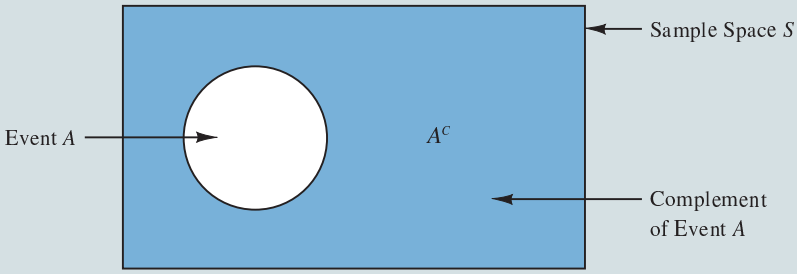
$$P(A) + P(A^C) = 1$$

Solving for $P(A)$, we obtain the following result:

COMPUTING PROBABILITY USING THE COMPLEMENT

$$P(A) = 1 - P(A^C) \tag{5.1}$$

FIGURE 5.1 Venn Diagram for Event A



Equation (5.1) shows that the probability of an event A can be computed easily if the probability of its complement, $P(A^C)$, is known.

As an example, consider the case of a sales manager who, after reviewing sales reports, states that 80% of new customer contacts result in no sale. By allowing A to denote the event of a sale and A^C to denote the event of no sale, the manager is stating that $P(A^C) = 0.80$. Using equation (5.1), we see that

$$P(A) = 1 - P(A^C) = 1 - 0.80 = 0.20$$

We can conclude that a new customer contact has a 0.20 probability of resulting in a sale.

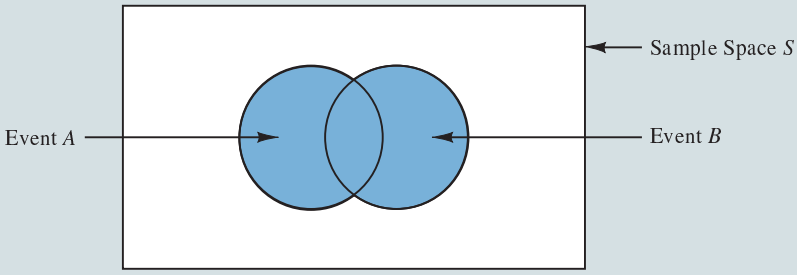
Addition Law

The addition law is helpful when we are interested in knowing the probability that at least one of two events will occur. That is, with events A and B we are interested in knowing the probability that event A or event B occurs or both events occur.

Before we present the addition law, we need to discuss two concepts related to the combination of events: the *union* of events and the *intersection* of events. Given two events A and B, the **union of A and B** is defined as the event containing all outcomes belonging to A or B or both. The union of A and B is denoted by $A \cup B$.

The Venn diagram in Figure 5.2 depicts the union of A and B. Note that one circle contains all the outcomes in A and the other all the outcomes in B. The fact that the circles overlap indicates that some outcomes are contained in both A and B.

FIGURE 5.2 Venn Diagram for the Union of Events A and B



The definition of the **intersection of A and B** is the event containing the outcomes that belong to both A and B. The intersection of A and B is denoted by $A \cap B$. The Venn diagram depicting the intersection of A and B is shown in Figure 5.3. The area in which the two circles overlap is the intersection; it contains outcomes that are in both A and B.

The **addition law** provides a way to compute the probability that event A or event B occurs or both events occur. In other words, the addition law is used to compute the probability of the union of two events. The addition law is written as follows:

ADDITION LAW

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{5.2}$$

To understand the addition law intuitively, note that the first two terms in the addition law, $P(A) + P(B)$, account for all the sample points in $A \cup B$. However, because the sample points in the intersection $A \cap B$ are in both A and B, when we compute $P(A) + P(B)$, we are in effect counting each of the sample points in $A \cap B$ twice. We correct for this double counting by subtracting $P(A \cap B)$.

As an example of the addition law, consider a study conducted by the human resources manager of a major computer software company. The study showed that 30% of the employees who left the firm within two years did so primarily because they were dissatisfied with their salary, 20% left because they were dissatisfied with their work assignments, and 12% of the former employees indicated dissatisfaction with *both* their salary and their work assignments. What is the probability that an employee who leaves within two years does so because of dissatisfaction with salary, dissatisfaction with the work assignment, or both?

Let

S = the event that the employee leaves because of salary
W = the event that the employee leaves because of work assignment

We can also think of this probability in the following manner: What proportion of employees either left because of salary or left because of work assignment?

From the survey results, we have $P(S) = 0.30$, $P(W) = 0.20$, and $P(S \cap W) = 0.12$. Using the addition law from equation (5.2), we have

$$P(S \cup W) = P(S) + P(W) - P(S \cap W) = 0.30 + 0.20 - 0.12 = 0.38$$

This calculation tells us that there is a 0.38 probability that an employee will leave for salary or work assignment reasons.

Before we conclude our discussion of the addition law, let us consider a special case that arises for **mutually exclusive events**. Events A and B are mutually exclusive if the occurrence of one event precludes the occurrence of the other. Thus, a requirement for A and B

FIGURE 5.3 Venn Diagram for the Intersection of Events A and B

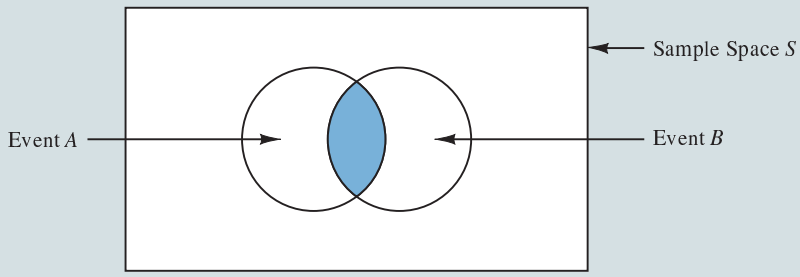
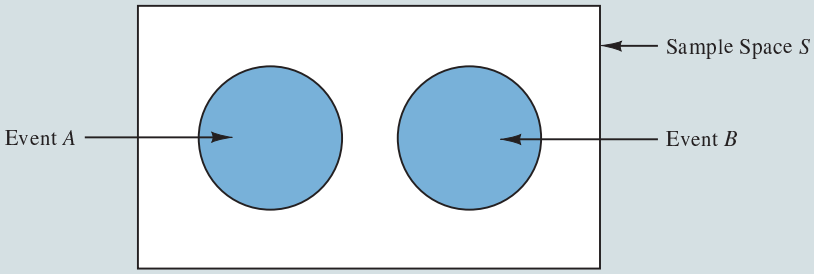


FIGURE 5.4 Venn Diagram for Mutually Exclusive Events



to be mutually exclusive is that their intersection must contain no sample points. The Venn diagram depicting two mutually exclusive events A and B is shown in Figure 5.4. In this case $P(A \cap B) = 0$ and the addition law can be written as follows:

ADDITION LAW FOR MUTUALLY EXCLUSIVE EVENTS

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

More generally, two events are said to be mutually exclusive if the events have no outcomes in common.

NOTES + COMMENTS

The addition law can be extended beyond two events. For example, the addition law for three events A, B, and C is $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$. Similar logic can be used to derive the expressions for the addition law for more than three events.

5.3 Conditional Probability

Often, the probability of one event is dependent on whether some related event has already occurred. Suppose we have an event A with probability $P(A)$. If we learn that a related event, denoted by B, has already occurred, we take advantage of this information by calculating a new probability for event A. This new probability of event A is called a **conditional probability** and is written $P(A \mid B)$. The notation $\mid$ indicates that we are considering the probability of event A *given* the condition that event B has occurred. Hence, the notation $P(A \mid B)$ reads "the probability of A given B."

To illustrate the idea of conditional probability, consider a bank that is interested in the mortgage default risk for its home mortgage customers. Table 5.3 shows the first 25 records of the 300 home mortgage customers at Lancaster Savings and Loan, a company that specializes in high-risk subprime lending. Some of these home mortgage customers have defaulted on their mortgages and others have continued to make on-time payments. These data include the age of the customer at the time of mortgage origination, the marital status of the customer (single or married), the annual income of the customer, the mortgage amount, the number of payments made by the customer per year on the mortgage, the total amount paid by the customer over the lifetime of the mortgage, and whether or not the customer defaulted on her or his mortgage.

TABLE 5.3 Subset of Data from 300 Home Mortgages of Customers at Lancaster Savings and Loan

Customer No.	Age	Marital Status	Annual Income	Mortgage Amount	Payments per Year	Total Amount Paid	Default on Mortgage?
1	37	Single	$ 172,125.70	$ 473,402.96	24	$ 581,885.13	Yes
2	31	Single	$ 108,571.04	$ 300,468.60	12	$ 489,320.38	No
3	37	Married	$ 124,136.41	$ 330,664.24	24	$ 493,541.93	Yes
4	24	Married	$ 79,614.04	$ 230,222.94	24	$ 449,682.09	Yes
5	27	Single	$ 68,087.33	$ 282,203.53	12	$ 520,581.82	No
6	30	Married	$ 59,959.80	$ 251,242.70	24	$ 356,711.58	Yes
7	41	Single	$ 99,394.05	$ 282,737.29	12	$ 524,053.46	No
8	29	Single	$ 38,527.35	$ 238,125.19	12	$ 468,595.99	No
9	31	Married	$ 112,078.62	$ 297,133.24	24	$ 399,617.40	Yes
10	36	Single	$ 224,899.71	$ 622,578.74	12	$1,233,002.14	No
11	31	Married	$ 27,945.36	$ 215,440.31	24	$ 285,900.10	Yes
12	40	Single	$ 48,929.74	$ 252,885.10	12	$ 336,574.63	No
13	39	Married	$ 82,810.92	$ 183,045.16	12	$ 262,537.23	No
14	31	Single	$ 68,216.88	$ 165,309.34	12	$ 253,633.17	No
15	40	Single	$ 59,141.13	$ 220,176.18	12	$ 424,749.80	No
16	45	Married	$ 72,568.89	$ 233,146.91	12	$ 356,363.93	No
17	32	Married	$ 101,140.43	$ 245,360.02	24	$ 388,429.41	Yes
18	37	Married	$ 124,876.53	$ 320,401.04	4	$ 360,783.45	Yes
19	32	Married	$ 133,093.15	$ 494,395.63	12	$ 861,874.67	No
20	32	Single	$ 85,268.67	$ 159,010.33	12	$ 308,656.11	No
21	37	Single	$ 92,314.96	$ 249,547.14	24	$ 342,339.27	Yes
22	29	Married	$ 120,876.13	$ 308,618.37	12	$ 472,668.98	No
23	24	Single	$ 86,294.13	$ 258,321.78	24	$ 380,347.56	Yes
24	32	Married	$ 216,748.68	$ 634,609.61	24	$ 915,640.13	Yes
25	44	Single	$ 46,389.75	$ 194,770.91	12	$ 385,288.86	No

Lancaster Savings and Loan is interested in whether the probability of a customer defaulting on a mortgage differs by marital status. Let

S = event that a customer is single
M = event that a customer is married
D = event that a customer defaulted on his or her mortgage
D^C = event that a customer did not default on his or her mortgage

Table 5.4 shows a crosstabulation for two events that can be derived from the Lancaster Savings and Loan mortgage data.

Chapter 3 discusses PivotTables in more detail.

Note that we can easily create Table 5.4 in Excel using a PivotTable by using the following steps:

Step 1. In the *Values* worksheet of *MortgageDefaultData* file
Click the **Insert** tab on the Ribbon
Step 2. Click **PivotTable** in the **Tables** group
Step 3. When the **Create PivotTable** dialog box appears:
Choose **Select a Table or Range**
Enter *A1:H301* in the **Table/Range:** box

DATA file
MortgageDefaultData

Select **New Worksheet** as the location for the PivotTable Report
Click **OK**

Step 4. In the **PivotTable Fields** area go to **Drag fields between areas below:**
Drag the **Marital Status** field to the **ROWS** area
Drag the **Default on Mortgage?** field to the **COLUMNS** area
Drag the **Customer Number** field to the **VALUES** area

Step 5. Click on **Sum of Customer Number** in the **VALUES** area and select **Value Field Settings**

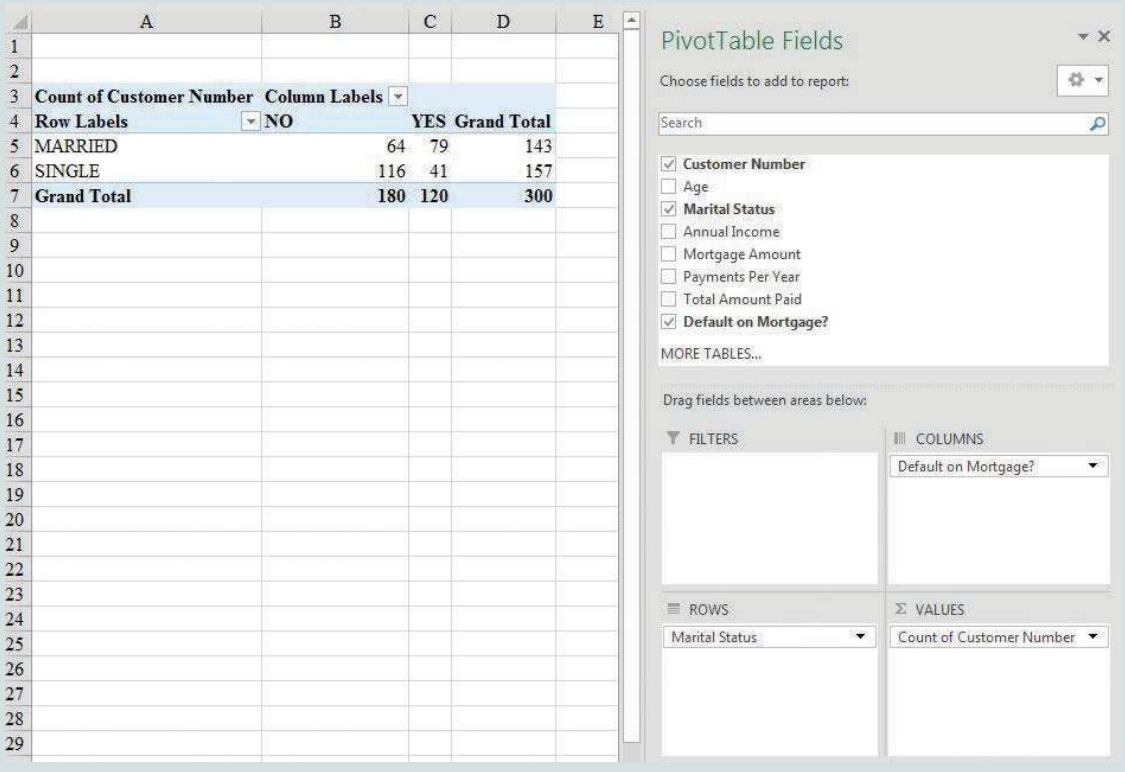
Step 6. When the **Value Field Settings** dialog box appears:
Under **Summarize value field by**, select **Count**

These steps produce the PivotTable shown in Figure 5.5.

TABLE 5.4 Crosstabulation of Marital Status and if Customer Defaults on Mortgage

Marital Status	No Default	Default	Total
Married	64	79	143
Single	116	41	157
Total	*180*	*120*	*300*

FIGURE 5.5 PivotTable for Marital Status and Whether Customer Defaults on Mortgage



From Table 5.4 or Figure 5.5, the probability that a customer defaults on his or her mortgage is 120/300 = 0.4. The probability that a customer does not default on his or her mortgage is 1 − 0.4 = 0.6(or180/300 = 0.6). But is this probability different for married customers as compared with single customers? Conditional probability allows us to answer this question.

We can also think of this joint probability in the following manner: What proportion of all customers are both married and defaulted on their loans?

But first, let us answer a related question: What is the probability that a randomly selected customer does not default on his or her mortgage and the customer is married? The probability that a randomly selected customer is married and the customer defaults on his or her mortgage is written as $P(M \cap D)$. This probability is calculated as $P(M \cap D) = \frac{79}{300} = 0.2633$.

Similarly,

$P(M \cap D^C) = \frac{64}{300} = 0.2133$ is the probability that a randomly selected customer is married and that the customer does not default on his or her mortgage.
$P(S \cap D) = \frac{41}{300} = 0.1367$ is the probability that a randomly selected customer is single and that the customer defaults on his or her mortgage.
$P(S \cap D^C) = \frac{116}{300} = 0.3867$ is the probability that a randomly selected customer is single and that the customer does not default on his or her mortgage.

Because each of these values gives the probability of the intersection of two events, the probabilities are called **joint probabilities**. Table 5.5, which provides a summary of the probability information for customer defaults on mortgages, is referred to as a joint probability table.

The values in the Total column and Total row (the margins) of Table 5.5 provide the probabilities of each event separately. That is, $P(M) = 0.4766$, $P(S) = 0.5234$, $P(D^C) = 0.6000$, and $P(D) = 0.4000$. These probabilities are referred to as **marginal probabilities** because of their location in the margins of the joint probability table. The marginal probabilities are found by summing the joint probabilities in the corresponding row or column of the joint probability table. From the marginal probabilities, we see that 60% of customers do not default on their mortgage, 40% of customers default on their mortgage, 47.66% of customers are married, and 52.34% of customers are single.

Let us begin the conditional probability analysis by computing the probability that a customer defaults on his or her mortgage given that the customer is married. In conditional probability notation, we are attempting to determine $P(D \mid M)$, which is read as "the probability that the customer defaults on the mortgage given that the customer is married." To calculate $P(D \mid M)$, first we note that we are concerned only with the 143 customers who are married (M). Because 79 of the 143 married customers defaulted on their mortgages, the probability of a customer defaulting given that the customer is married is 79/143 = 0.5524. In other words, given that a customer is married, there is a 55.24% chance that he or she will default. Note also that the conditional probability $P(D \mid M)$ can be computed as the ratio of the joint probability $P(D \cap M)$ to the marginal probability $P(M)$.

$$P(D \mid M) = \frac{P(D \cap M)}{P(M)} = \frac{0.2633}{0.4766} = 0.5524$$

We can use the PivotTable from Figure 5.5 to easily create the joint probability table in Excel. To do so, right-click on any of the numerical values in the PivotTable, select ***Show Values As****, and choose* ***% of Grand Total****. The resulting values, which are percentages of the total, can then be divided by 100 to create the probabilities in the joint probability table.*

TABLE 5.5 Joint Probability Table for Customer Mortgage Prepayments

Joint Probabilities

	No Default (D^C)	Default (D)	Total
Married (M)	0.2133	0.2633	0.4766
Single (S)	0.3867	0.1367	0.5234
Total	*0.6000*	*0.4000*	*1.0000*

Marginal Probabilities

The fact that conditional probabilities can be computed as the ratio of a joint probability to a marginal probability provides the following general formula for conditional probability calculations for two events A and B.

CONDITIONAL PROBABILITY

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{5.3}$$

or

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{5.4}$$

We have already determined the probability that a customer who is married will default is 0.5524. How does this compare to a customer who is single? In other words, we want to find $P(D \mid S)$. From equation (5.3), we can compute $P(D \mid S)$ as

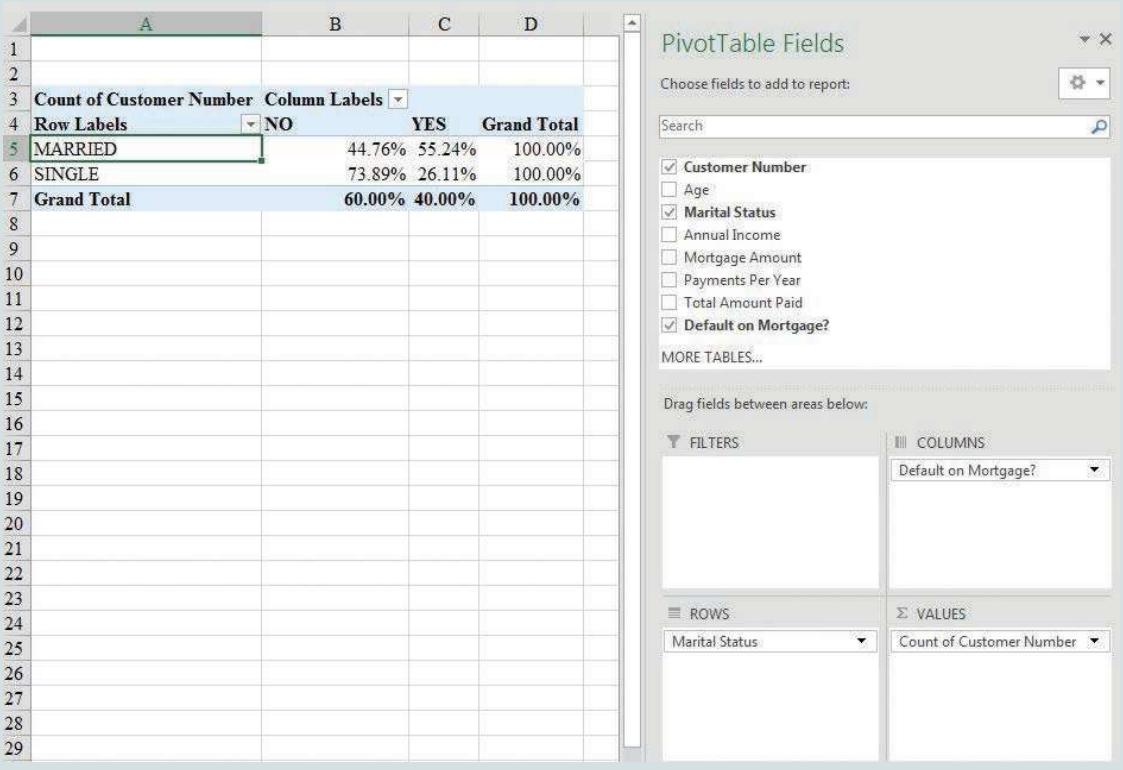
$$P(D \mid S) = \frac{P(D \cap S)}{P(S)} = \frac{0.1367}{0.5234} = 0.2611$$

In other words, the chance that a customer will default if the customer is single is 26.11%. This is substantially less than the chance of default if the customer is married.

Note that we could also answer this question using the Excel PivotTable in Figure 5.5. We can calculate these conditional probabilities by right-clicking on any numerical value in the body of the PivotTable and then selecting **Show Values As** and choosing **% of Row Total**. The modified Excel PivotTable is shown in Figure 5.6.

FIGURE 5.6 Using Excel PivotTable to Calculate Conditional Probabilities

Count of Customer Number	Column Labels		
Row Labels	NO	YES	Grand Total
MARRIED	44.76%	55.24%	100.00%
SINGLE	73.89%	26.11%	100.00%
Grand Total	60.00%	40.00%	100.00%



By calculating the **% of Row Total**, the Excel PivotTable in Figure 5.6 shows that 55.24% of married customers defaulted on mortgages, but only 26.11% of single customers defaulted.

Independent Events

Note that in our example, $P(D) = 0.4000$, $P(D \mid M) = 0.5524$, and $P(D \mid S) = 0.2611$. So the probability that a customer defaults is influenced by whether the customer is married or single. Because $P(D \mid M) \neq P(D)$, we say that events D and M are dependent. However, if the probability of event D is not changed by the existence of event M—that is, if $P(D \mid M) = P(D)$—then we would say that events D and M are **independent events**. This is summarized for two events A and B as follows:

INDEPENDENT EVENTS

Two events A and B are independent if

$$P(A \mid B) = P(A) \tag{5.5}$$

or

$$P(B \mid A) = P(B) \tag{5.6}$$

Otherwise, the events are dependent.

Multiplication Law

The multiplication law can be used to calculate the probability of the intersection of two events. The multiplication law is based on the definition of conditional probability. Solving equations (5.3) and (5.4) for $P(A \cap B)$, we obtain the **multiplication law**.

MULTIPLICATION LAW

$$P(A \cap B) = P(B)P(A \mid B) \tag{5.7}$$

or

$$P(A \cap B) = P(A)P(B \mid A) \tag{5.8}$$

To illustrate the use of the multiplication law, we will calculate the probability that a customer defaults on his or her mortgage and the customer is married, $P(D \cap M)$. From equation (5.7), this is calculated as $P(D \cap M) = P(M)P(D \mid M)$.

From Table 5.5 we know that $P(M) = 0.4766$, and from our previous calculations we know that the conditional probability $P(D \mid M) = 0.5524$. Therefore,

$$P(D \cap M) = P(M)P(D \mid M) = (0.4766)(0.5524) = 0.2633$$

This value matches the value shown for $P(D \cap M)$ in Table 5.5. The multiplication law is useful when we know conditional probabilities but do not know the joint probabilities.

Consider the special case in which events A and B are independent. From equations (5.5) and (5.6), $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$. Using these equations to simplify equations (5.7) and (5.8) for this special case, we obtain the following multiplication law for independent events.

MULTIPLICATION LAW FOR INDEPENDENT EVENTS

$$P(A \cap B) = P(A)P(B) \tag{5.9}$$

To compute the probability of the intersection of two independent events, we simply multiply the probabilities of each event.

Bayes' Theorem

Revising probabilities when new information is obtained is an important aspect of probability analysis. Often, we begin the analysis with initial or **prior probability** estimates for specific events of interest. Then, from sources such as a sample survey or a product test, we obtain additional information about the events. Given this new information, we update the prior probability values by calculating revised probabilities, referred to as **posterior probabilities**. **Bayes' theorem** provides a means for making these probability calculations.

Bayes' theorem is also discussed in Chapter 15 in the context of decision analysis.

As an application of Bayes' theorem, consider a manufacturing firm that receives shipments of parts from two different suppliers. Let A_1 denote the event that the part is from supplier 1 and let A_2 denote the event that a part is from supplier 2. Currently, 65% of the parts purchased by the company are from supplier 1 and the remaining 35% are from supplier 2. Hence, if a part is selected at random, we would assign the prior probabilities $P(A_1) = 0.65$ and $P(A_2) = 0.35$.

The quality of the purchased parts varies according to their source. Historical data suggest that the quality ratings of the two suppliers are as shown in Table 5.6.

If we let G be the event that a part is good and we let B be the event that a part is bad, the information in Table 5.6 enables us to calculate the following conditional probability values:

$$P(G \mid A_1) = 0.98 \quad P(B \mid A_1) = 0.02$$
$$P(G \mid A_2) = 0.95 \quad P(B \mid A_2) = 0.05$$

Figure 5.7 shows a diagram that depicts the process of the firm receiving a part from one of the two suppliers and then discovering that the part is good or bad as a two-step random experiment. We see that four outcomes are possible; two correspond to the part being good and two correspond to the part being bad.

Each of the outcomes is the intersection of two events, so we can use the multiplication rule to compute the probabilities. For instance,

$$P(A_1, G) = P(A_1 \cap G) = P(A_1)P(G \mid A_1)$$

The process of computing these joint probabilities can be depicted in what is called a probability tree (see Figure 5.8). From left to right through the tree, the probabilities for each branch at step 1 are prior probabilities and the probabilities for each branch at step 2 are conditional probabilities. To find the probability of each experimental outcome, simply multiply the probabilities on the branches leading to the outcome. Each of these joint probabilities is shown in Figure 5.8 along with the known probabilities for each branch.

Now suppose that the parts from the two suppliers are used in the firm's manufacturing process and that a machine breaks down while attempting the process using a bad part. Given the information that the part is bad, what is the probability that it came from supplier 1 and what is the probability that it came from supplier 2? With the information in the probability tree (Figure 5.8), Bayes' theorem can be used to answer these questions.

For the case in which there are only two events (A_1 and A_2), Bayes' theorem can be written as follows:

BAYES' THEOREM (TWO-EVENT CASE)

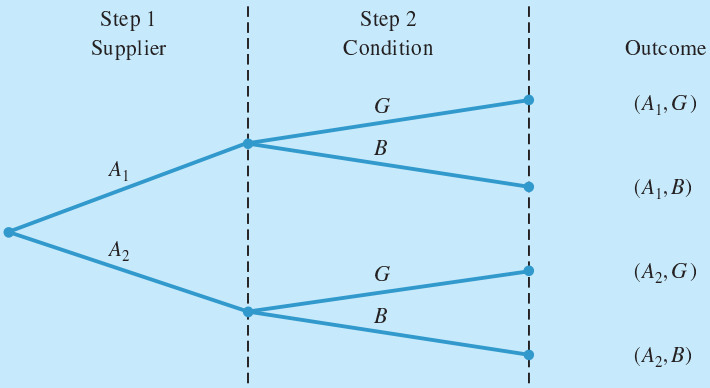
$$P(A_1 \mid B) = \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \tag{5.10}$$

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_2)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \tag{5.11}$$

TABLE 5.6 Historical Quality Levels for Two Suppliers

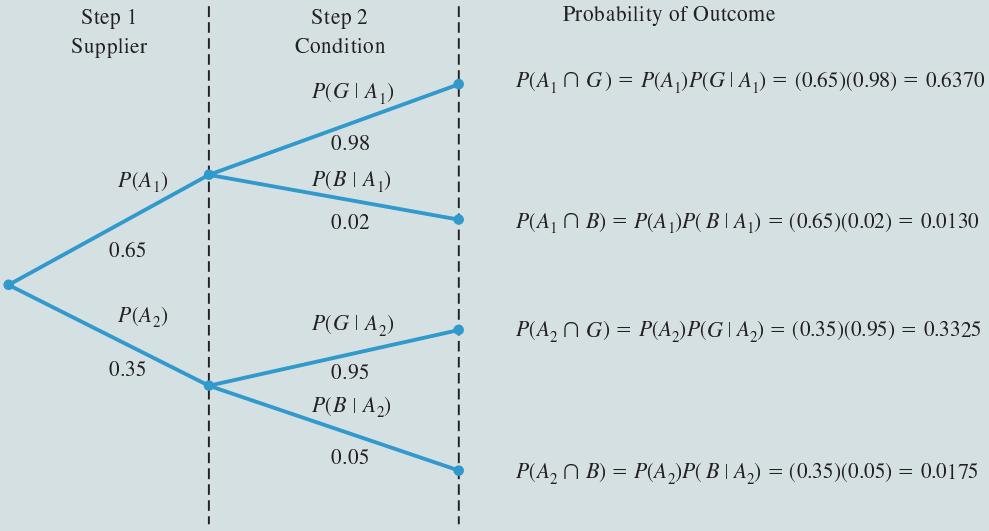
	% Good Parts	% Bad Parts
Supplier 1	98	2
Supplier 2	95	5

FIGURE 5.7 Diagram for Two-Supplier Example: Step 1 shows that the part comes from one of two suppliers and step 2 shows whether the part is good or bad



Note: Step 1 shows that the part comes from one of two suppliers and Step 2 shows whether the part is good or bad.

FIGURE 5.8 Probability Tree for Two-Supplier Example



Using equation (5.10) and the probability values provided in Figure 5.8, we have

$$P(A_1 \mid B) = \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)}$$
$$= \frac{(0.65)(0.02)}{(0.65)(0.02) + (0.35)(0.05)} = \frac{0.0130}{0.0130 + 0.0175}$$
$$= \frac{0.0130}{0.0305} = 0.4262$$

Using equation (5.11), we find $P(A_2 \mid B)$ as

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_2)}{P(A_2)P(B \mid A_1) + P(A_2)P(B \mid A_2)}$$
$$= \frac{(0.35)(0.05)}{(0.65)(0.02) + (0.35)(0.05)} = \frac{0.0175}{0.0130 + 0.0175}$$
$$= \frac{0.0175}{0.0305} = 0.5738$$

Note that in this application we started with a probability of 0.65 that a part selected at random was from supplier 1. However, given information that the part is bad, the probability that the part is from supplier 1 drops to 0.4262. In fact, if the part is bad, the chance is better than 50–50 that it came from supplier 2; that is, $P(A_2 \mid B) = 0.5738$.

If the union of events is the entire sample space, the events are said to be collectively exhaustive.

Bayes' theorem is applicable when events for which we want to compute posterior probabilities are mutually exclusive and their union is the entire sample space. For the case of n mutually exclusive events $A_1, A_2, \ldots, A_n$, whose union is the entire sample space, Bayes' theorem can be used to compute any posterior probability $P(A_i \mid B)$ as shown in equation 5.12.

BAYES' THEOREM

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \cdots + P(A_n)P(B \mid A_n)} \quad \textbf{(5.12)}$$

NOTES + COMMENTS

By applying basic algebra we can derive the multiplication law from the definition of conditional probability. For two events A and B, the probability of A given B is $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$. If we multiply both sides of this expression by $P(B)$, the $P(B)$ in the numerator and denominator on the right side of the expression will cancel and we are left with $P(A \mid B)P(B) = P(A \cap B)$, which is the multiplication law.

5.4 Random Variables

Chapter 2 introduces the concept of random variables and the use of data to describe them.

In probability terms, a **random variable** is a numerical description of the outcome of a random experiment. Because the outcome of a random experiment is not known with certainty, a random variable can be thought of as a quantity whose value is not known with certainty. A random variable can be classified as being either discrete or continuous depending on the numerical values it can assume.

Discrete Random Variables

A random variable that can take on only specified discrete values is referred to as a **discrete random variable**. Table 5.7 provides examples of discrete random variables.

Returning to our example of Lancaster Savings and Loan, we can define a random variable x to indicate whether or not a customer defaults on his or her mortgage. As previously

TABLE 5.7 Examples of Discrete Random Variables

Random Experiment	Random Variable (x)	Possible Values for the Random Variable
Flip a coin	Face of coin showing	1 if heads; 0 if tails
Roll a die	Number of dots showing on top of die	1, 2, 3, 4, 5, 6
Contact five customers	Number of customers who place an order	0, 1, 2, 3, 4, 5
Operate a health care clinic for one day	Number of patients who arrive	0, 1, 2, 3, ...
Offer a customer the choice of two products	Product chosen by customer	0 if none; 1 if choose product A; 2 if choose product B

stated, the values of a random variable must be numerical, so we can define random variable x such that $x = 1$ if the customer defaults on his or her mortgage and $x = 0$ if the customer does not default on his or her mortgage. An additional random variable, y, could indicate whether the customer is married or single. For instance, we can define random variable y such that $y = 1$ if the customer is married and $y = 0$ if the customer is single. Yet another random variable, z, could be defined as the number of mortgage payments per year made by the customer. For instance, a customer who makes monthly payments would make $z = 12$ payments per year, a customer who makes payments quarterly would make $z = 4$ payments per year.

Table 5.8 repeats the joint probability table for the Lancaster Savings and Loan data, but this time with the values labeled as random variables.

Continuous Random Variables

A random variable that may assume any numerical value in an interval or collection of intervals is called a **continuous random variable**. Technically, relatively few random variables are truly continuous; these include values related to time, weight, distance, and temperature. An example of a continuous random variable is x = the time between consecutive incoming calls to a call center. This random variable can take on any value $x > 0$ such as $x = 1.26$ minutes, $x = 2.571$ minutes, $x = 4.3333$ minutes, etc. Table 5.9 provides examples of continuous random variables.

As illustrated by the final example in Table 5.9, many discrete random variables have a large number of potential outcomes and so can be effectively modeled as continuous random variables. Consider our Lancaster Savings and Loan example. We can define a random variable x = total amount paid by customer over the lifetime of the mortgage. Because we typically measure financial values only to two decimal places, one could consider this a discrete random variable. However, because in any practical interval there are many possible values for this random variable, then it is usually appropriate to model the amount as a continuous random variable.

TABLE 5.8 Joint Probability Table for Customer Mortgage Prepayments

	No Default ($x = 0$)	Default ($x = 1$)	$f(y)$
Married ($y = 1$)	0.2133	0.2633	*0.4766*
Single ($y = 0$)	0.3867	0.1367	*0.5234*
$f(x)$	*0.6000*	*0.4000*	*1.0000*

TABLE 5.9 Examples of Continuous Random Variables

Random Experiment	Random Variable (x)	Possible Values for the Random Variable
Customer visits a web page	Time customer spends on web page in minutes	$x \geq 0$
Fill a soft drink can (max capacity = 12.1 ounces)	Number of ounces	$0 \leq x \leq 12.1$
Test a new chemical process	Temperature when the desired reaction takes place (min temperature = 150°F; max temperature = 212°F)	$150 \leq x \leq 212$
Invest $10,000 in the stock market	Value of investment after one year	$x \geq 0$

NOTES + COMMENTS

1. In this section we again use the relative frequency method to assign probabilities for the Lancaster Savings and Loan example. Technically, the concept of random variables applies only to populations; probabilities that are found using sample data are only estimates of the true probabilities. However, larger samples generate more reliable estimated probabilities, so if we have a large enough data set (as we are assuming here for the Lancaster Savings and Loan data), then we can treat the data as if they are from a population and the relative frequency method is appropriate to assign probabilities to the outcomes.
2. Random variables can be used to represent uncertain future values. Chapter 11 explains how random variables can be used in simulation models to evaluate business decisions in the presence of uncertainty.

5.5 Discrete Probability Distributions

The **probability distribution** for a random variable describes the range and relative likelihood of possible values for a random variable. For a discrete random variable x, the probability distribution is defined by a **probability mass function**, denoted by $f(x)$. The probability mass function provides the probability for each value of the random variable.

Returning to our example of mortgage defaults, consider the data shown in Table 5.3 for Lancaster Savings and Loan and the associated joint probability table in Table 5.8. From Table 5.8, we see that $f(0) = 0.6$ and $f(1) = 0.4$. Note that these values satisfy the required conditions of a discrete probability distribution that (1) $f(x) \geq 0$ and (2) $\Sigma f(x) = 1$.

We can also present probability distributions graphically. In Figure 5.9, the values of the random variable x are shown on the horizontal axis and the probability associated with these values is shown on the vertical axis.

Custom Discrete Probability Distribution

A probability distribution that is generated from observations such as that shown in Figure 5.9 is called an **empirical probability distribution**. This particular empirical probability distribution is considered a custom discrete distribution because it is discrete and the possible values of the random variable have different values.

A **custom discrete probability distribution** is very useful for describing different possible scenarios that have different probabilities of occurring. The probabilities associated with each scenario can be generated using either the subjective method or the relative frequency method. Using a subjective method, probabilities are based on experience or intuition when little relevant data are available. If sufficient data exist, the relative frequency method can be used to determine probabilities. Consider the random variable describing the number of payments made per year by a randomly chosen customer. Table 5.10 presents a summary of the number of payments made per year by the 300 home mortgage

FIGURE 5.9 Graphical Representation of the Probability Distribution for Whether a Customer Defaults on a Mortgage

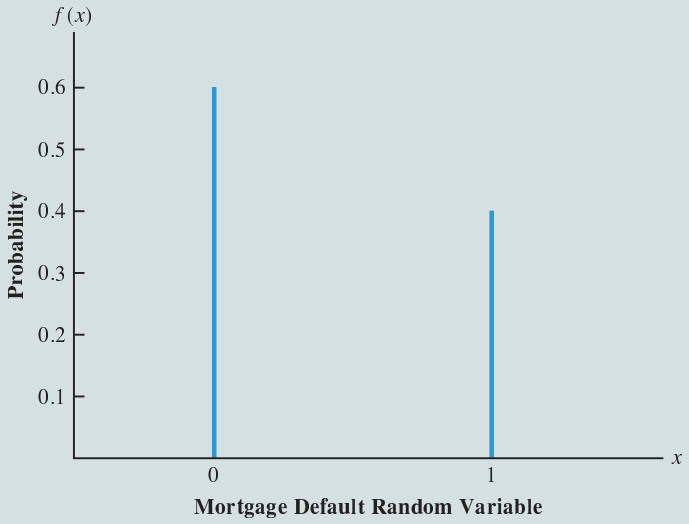


TABLE 5.10 Summary Table of Number of Payments Made per Year

	Number of Payments Made per Year			
	$x = 4$	$x = 12$	$x = 24$	Total
Number of observations	45	180	75	300
$f(x)$	0.15	0.60	0.25	

customers. This table shows us that 45 customers made quarterly payments ($x = 4$), 180 customers made monthly payments ($x = 12$), and 75 customers made two payments each month ($x = 24$). We can then calculate $f(4) = 45/300 = 0.15$, $f(12) = 180/300 = 0.60$, and $f(24) = 75/300 = 0.25$. In other words, the probability that a randomly selected customer makes 4 payments per year is 0.15, the probability that a randomly selected customer makes 12 payments per year is 0.60, and the probability that a randomly selected customer makes 24 payments per year is 0.25.

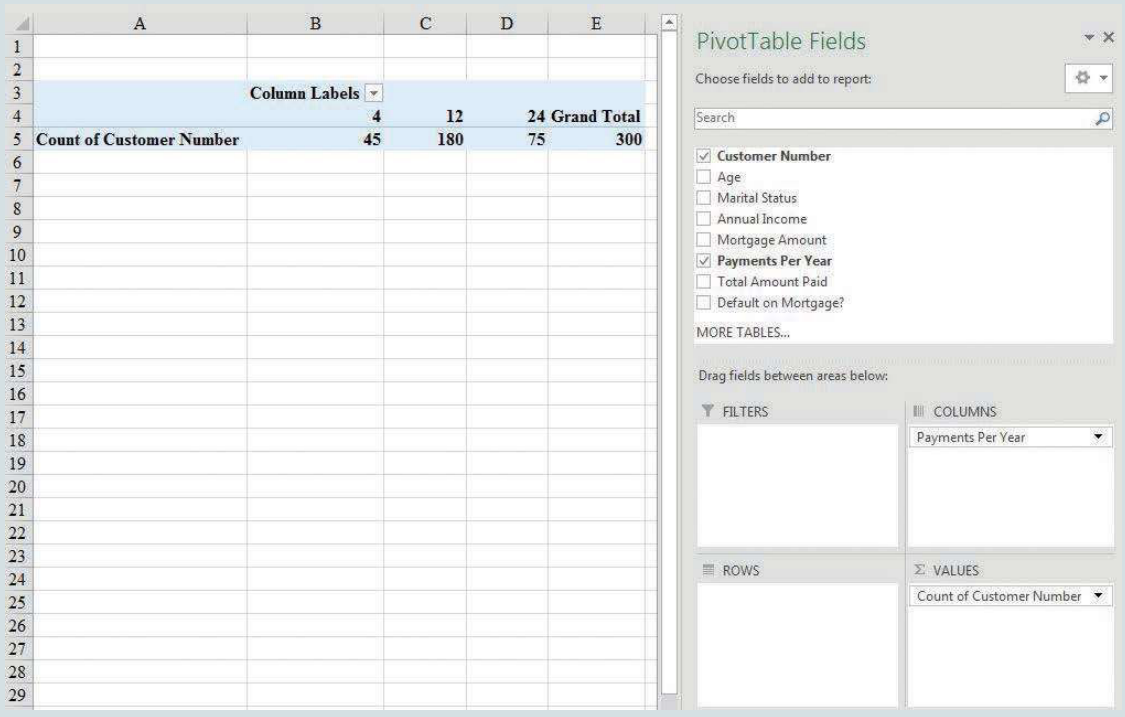
We can write this probability distribution as a function in the following manner:

$$f(x) = \begin{cases} 0.15 \text{ if } x = 4 \\ 0.60 \text{ if } x = 12 \\ 0.25 \text{ if } x = 24 \\ 0 \quad \text{otherwise} \end{cases}$$

This probability mass function tells us in a convenient way that $f(x) = 0.15$ when $x = 4$ (the probability that the random variable $x = 4$ is 0.15); $f(x) = 0.60$ when $x = 12$ (the probability that the random variable $x = 12$ is 0.60); $f(x) = 0.25$ when $x = 24$ (the probability that the random variable $x = 24$ is 0.25); and $f(x) = 0$ when x is any other value (there is zero probability that the random variable x is some value other than 4, 12, or 24).

Note that we can also create Table 5.10 in Excel using a PivotTable as shown in Figure 5.10.

FIGURE 5.10 Excel PivotTable for Number of Payments Made per Year



Expected Value and Variance

Chapter 2 discusses the computation of the mean of a random variable based on data.

The **expected value**, or mean, of a random variable is a measure of the central location for the random variable. It is the weighted average of the values of the random variable, where the weights are the probabilities. The formula for the expected value of a discrete random variable x follows:

EXPECTED VALUE OF A DISCRETE RANDOM VARIABLE

$$E(x) = \mu = \Sigma x f(x) \tag{5.13}$$

Both the notations $E(x)$ and μ are used to denote the expected value of a random variable. Equation (5.13) shows that to compute the expected value of a discrete random variable, we must multiply each value of the random variable by the corresponding probability $f(x)$ and then add the resulting products. Table 5.11 calculates the expected value of the number of payments made by a mortgage customer in a year. The sum of the entries in the $xf(x)$ column shows that the expected value is 13.8 payments per year. Therefore, if Lancaster Savings and Loan signs up a new mortgage customer, the expected number of payments per year made by this new customer is 13.8. Obviously, no customer will make exactly 13.8 payments per year, but this value represents our expectation for the number of payments per year made by a new customer absent any other information about the new customer. Some customers will make fewer payments (4 or 12 per year), some customers will make more payments (24 per year), but 13.8 represents the expected number of payments per year based on the probabilities calculated in Table 5.10.

The SUMPRODUCT function in Excel can easily be used to calculate the expected value for a discrete random variable. This is illustrated in Figure 5.11. We can also

TABLE 5.11 Calculation of the Expected Value for Number of Payments Made per Year by a Lancaster Savings and Loan Mortgage Customer

x	$f(x)$	$xf(x)$
4	0.15	$(4)(0.15) = 0.6$
12	0.60	$(12)(0.60) = 7.2$
24	0.25	$(24)(0.25) = 6.0$
		13.8 ← $E(x) = \mu = \sum xf(x)$

calculate the expected value of the random variable directly from the Lancaster Savings and Loan data using the Excel function AVERAGE, as shown in Figure 5.12. Column F contains the data on the number of payments made per year by each mortgage customer in the data set. Using the Excel formula =AVERAGE(F2:F301) gives us a value of 13.8 for the expected value, which is the same as the value we calculated in Table 5.11.

Note that we cannot simply use the AVERAGE function on the x values for a custom discrete random variable. If we did, this would give us a calculated value of $(4 + 12 + 24)/3 = 13.333$, which is not the correct expected value in this scenario. This is because using the AVERAGE function in this way assumes that each value of the random variable x is equally likely. But in this case, we know that $x = 12$ is much more likely than $x = 4$ or $x = 24$. Therefore, we must use equation (5.13) to calculate the expected value of a custom discrete random variable, or we can use the Excel function AVERAGE on the entire data set, as shown in Figure 5.12.

FIGURE 5.11 Using Excel SUMPRODUCT Function to Calculate the Expected Value for Number of Payments Made per Year by a Lancaster Savings and Loan Mortgage Customer

	A	B	C	D
1	*x*	*f(x)*		
2	4	0.15		
3	12	0.6		
4	24	0.25		
5				
6	**Expected Value:**	=SUMPRODUCT(A2:A4,B2:B4)		
7				
8				
9				
10				

	A	B	C	D
1	*x*	*f(x)*		
2	4	0.15		
3	12	0.60		
4	24	0.25		
5				
6	**Expected Value:**	13.8		
7				
8				
9				
10				

FIGURE 5.12 Excel Calculation of the Expected Value for Number of Payments Made per Year by a Lancaster Savings and Loan Mortgage Customer

	A	B	C	D	E	F	G	H
1	Customer Number	Age	Marital Status	Annual Income	Mortgage Amount	Payments Per Year	Total Amount Paid	Prepay Mortgage?
2	1	37	SINGLE	172125.7	473402.96	24	581885.13	YES
3	2	31	SINGLE	108571.04	300468.6	12	489320.38	NO
4	3	37	MARRIED	124136.41	330664.24	24	493541.93	YES
5	4	24	MARRIED	79614.04	230222.94	24	449682.09	YES
6	5	27	SINGLE	68087.33	282203.53	12	520581.82	NO
296	295	37	MARRIED	84791.08	179676.63	24	256361.65	YES
297	296	33	MARRIED	83498.89	235907.5	12	437145.85	NO
298	297	41	SINGLE	16597.53	151972.2	4	171289.87	YES
299	298	30	SINGLE	49293.95	186043.13	12	376694.27	NO
300	299	35	SINGLE	84241.8	194417.84	12	352597.79	NO
301	300	31	MARRIED	94428.15	264175.55	24	434102.49	YES
302								
304					Expected Value:	=AVERAGE(F2:F301)		
305					Variance:	=VAR.P(F2:F301)		
306					Standard Deviation:	=STDEV.P(F2:F301)		

	A	B	C	D	E	F	G	H
1	Customer Number	Age	Marital Status	Annual Income	Mortgage Amount	Payments Per Year	Total Amount Paid	Prepay Mortgage?
2	1	37	SINGLE	$ 172,125.70	$ 473,402.96	24	$ 581,885.13	YES
3	2	31	SINGLE	$ 108,571.04	$ 300,468.60	12	$ 489,320.38	NO
4	3	37	MARRIED	$ 124,136.41	$ 330,664.24	24	$ 493,541.93	YES
5	4	24	MARRIED	$ 79,614.04	$ 230,222.94	24	$ 449,682.09	YES
6	5	27	SINGLE	$ 68,087.33	$ 282,203.53	12	$ 520,581.82	NO
296	295	37	MARRIED	$ 84,791.08	$ 179,676.63	24	$ 256,361.65	YES
297	296	33	MARRIED	$ 83,498.89	$ 235,907.50	12	$ 437,145.85	NO
298	297	41	SINGLE	$ 16,597.53	$ 151,972.20	4	$ 171,289.87	YES
299	298	30	SINGLE	$ 49,293.95	$ 186,043.13	12	$ 376,694.27	NO
300	299	35	SINGLE	$ 84,241.80	$ 194,417.84	12	$ 352,597.79	NO
301	300	31	MARRIED	$ 94,428.15	$ 264,175.55	24	$ 434,102.49	YES
302								
304					Expected Value:	13.8		
305					Variance:	42.360		
306					Standard Deviation:	6.508		

Chapter 2 discusses the computation of the variance of a random variable based on data.

Variance is a measure of variability in the values of a random variable. It is a weighted average of the squared deviations of a random variable from its mean where the weights are the probabilities. Below we define the formula for calculating the variance of a discrete random variable.

VARIANCE OF A DISCRETE RANDOM VARIABLE

$$\text{Var}(x) = \sigma^2 = \Sigma(x - \mu)^2 f(x) \qquad \textbf{(5.14)}$$

As equation (5.14) shows, an essential part of the variance formula is the deviation, $x - \mu$, which measures how far a particular value of the random variable is from the expected value, or mean, μ. In computing the variance of a random variable, the deviations are squared and then weighted by the corresponding value of the probability mass function. The sum of these weighted squared deviations for all values of the random variable is referred to as the *variance*. The notations Var(x) and σ^2 are both used to denote the variance of a random variable.

The calculation of the variance of the number of payments made per year by a mortgage customer is summarized in Table 5.12. We see that the variance is 42.360. The **standard deviation**, σ, is defined as the positive square root of the variance. Thus, the standard deviation for the number of payments made per year by a mortgage customer is $\sqrt{42.360} = 6.508$.

Chapter 2 discusses the computation of the standard deviation of a random variable based on data.

The Excel function SUMPRODUCT can be used to easily calculate equation (5.14) for a custom discrete random variable. We illustrate the use of the SUMPRODUCT function to calculate variance in Figure 5.13.

We can also use Excel to find the variance directly from the data when the values in the data occur with relative frequencies that correspond to the probability distribution of the random variable. Cell F305 in Figure 5.12 shows that we use the Excel formula =VAR.P(F2:F301)

TABLE 5.12 Calculation of the Variance for Number of Payments Made per Year by a Lancaster Savings and Loan Mortgage Customer

x	$x - \mu$	$f(x)$	$(x - \mu)^2 f(x)$
4	$4 - 13.8 = -9.8$	0.15	$(-9.8)^{2*}\ 0.15 = 15.606$
12	$12 - 13.8 = -1.8$	0.60	$(-1.8)^{2*}\ 0.60 = 2.904$
21	$21 - 13.8 = 10.2$	0.25	$(10.2)^{2*}\ 0.25 = 24.010$
			42.360 ← $\sigma^2 = \sum (x - \mu)^2 f(x)$

FIGURE 5.13 Excel Calculation of the Variance for Number of Payments Made per Year by a Lancaster Savings and Loan Mortgage Customer

	A	B	C	D
1	x	$f(x)$	$(x - \mu)^2$	
2	4	0.15	=(A2-$B$6)^2	
3	12	0.6	=(A3-$B$6)^2	
4	24	0.25	=(A4-$B$6)^2	
5				
6	**Expected Value:**	=SUMPRODUCT(A2:A4,B2:B4)		
7				
8	**Variance:**	=SUMPRODUCT(B2:B4,C2:C4)		
9				
10	**Stadard Deviation:**	=SQRT(B8)		

	A	B	C	D
1	x	$f(x)$	$(x - \mu)^2$	
2	4	0.15	96.04	
3	12	0.60	3.24	
4	24	0.25	104.04	
5				
6	**Expected Value:**	13.8		
7				
8	**Variance:**	42.360		
9				
10	**Stadard Deviation:**	6.508		

Note that here we are using the Excel functions VAR.P and STDEV.P rather than VAR.S and STDEV.S. This is because we are assuming that the sample of 300 Lancaster Savings and Loan mortgage customers is a perfect representation of the population.

to calculate the variance from the complete data. This formula gives us a value of 42.360, which is the same as that calculated in Table 5.12 and Figure 5.13. Similarly, we can use the formula =STDEV.P(F2:F301) to calculate the standard deviation of 6.508.

As with the AVERAGE function and expected value, we cannot use the Excel functions VAR.P and STDEV.P directly on the x values to calculate the variance and standard deviation of a custom discrete random variable if the x values are not equally likely to occur. Instead we must either use the formula from equation (5.14) or use the Excel functions on the entire data set as shown in Figure 5.12.

Discrete Uniform Probability Distribution

When the possible values of the probability mass function, $f(x)$, are all equal, then the probability distribution is a **discrete uniform probability distribution**. For instance, the values that result from rolling a single fair die is an example of a discrete uniform distribution

because the possible outcomes $y = 1$, $y = 2$, $y = 3$, $y = 4$, $y = 5$, and $y = 6$ all have the same values $f(1) = f(2) = f(3) = f(4) = f(5) = f(6) = 1/6$. The general form of the probability mass function for a discrete uniform probability distribution is given below as follows:

DISCRETE UNIFORM PROBABILITY MASS FUNCTION

$$f(x) = 1/n \tag{5.15}$$

where n = the number of unique values that may be assumed by the random variable.

Binomial Probability Distribution

As an example of the use of the binomial probability distribution, consider an online specialty clothing company called Martin's. Martin's commonly sends out targeted e-mails to its best customers notifying them about special discounts that are available only to the recipients of the e-mail. The e-mail contains a link that takes the customer directly to a web page for the discounted item. The exact number of customers who will click on the link is obviously unknown, but from previous data, Martin's estimates that the probability that a customer clicks on the link in the e-mail is 0.30. Martin's is interested in knowing more about the probabilities associated with one, two, three, etc. customers clicking on the link in the targeted e-mail.

The probability distribution related to the number of customers who click on the targeted e-mail link can be described using a **binomial probability distribution**. A binomial probability distribution is a discrete probability distribution that can be used to describe many situations in which a fixed number (n) of repeated identical and independent trials has two, and only two, possible outcomes. In general terms, we refer to these two possible outcomes as either a success or a failure. A success occurs with probability p in each trial and a failure occurs with probability $1 - p$ in each trial. In the Martin's example, the "trial" refers to a customer receiving the targeted e-mail. We will define a success as a customer clicking on the e-mail link ($p = 0.30$) and a failure as a customer not clicking on the link ($1 - p = 0.70$). The binomial probability distribution can then be used to calculate the probability of a given number of successes (customers who click on the e-mail link) out of a given number of independent trials (number of e-mails sent to customers). Other examples that can often be described by a binomial probability distribution include counting the number of heads resulting from flipping a coin 20 times, the number of customers who click on a particular advertisement link on web site in a day, the number of days on which a particular financial stock increases in value over a month, and the number of nondefective parts produced in a batch.

Whether or not a customer clicks on the link is an example of what is known as a Bernoulli trial—a trial in which: (1) there are two possible outcomes, success or failure, and (2) the probability of success is the same every time the trial is executed. The probability distribution related to the number of successes in a set of n independent Bernoulli trials can be described by a binomial probability distribution.

Equation (5.16) provides the probability mass function for a binomial random variable that calculates the probability of x successes in n independent events.

BINOMIAL PROBABILITY MASS FUNCTION

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)}$$

where

$$\begin{aligned} x &= \text{the number of successes} \\ p &= \text{the probability of a success on one trial} \\ n &= \text{the number of trials} \\ f(x) &= \text{the probability of } x \text{ successes in } n \text{ trials} \end{aligned} \tag{5.16}$$

and

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

n! is read as "n factorial," and n! = n × n − 1 × n − 2 × ⋯ × 2 × 1. For example, 4! = 4 × 3 × 2 × 1 = 24. The Excel formula = FACT(n) can be used to calculate n factorial.

In the Martin's example, use equation (5.16) to compute the probability that out of three customers who receive the e-mail: (1) no customer clicks on the link; (2) exactly one customer clicks on the link; (3) exactly two customers click on the link; and (4) all three customers click on the link. The calculations are summarized in Table 5.13, which gives the probability distribution of the number of customers who click on the targeted e-mail link. Figure 5.14 is a graph of this probability distribution. Table 5.13 and Figure 5.14 show that the highest probability is associated with exactly one customer clicking on the Martin's targeted e-mail link and the lowest probability is associated with all three customers clicking on the link.

Because the outcomes in the Martin's example are mutually exclusive, we can easily use these results to answer interesting questions about various events. For example, using the information in Table 5.13, the probability that no more than one customer clicks on the link is $P(x \leq 1) = P(x = 0) + P(x = 1) = 0.343 + 0.441 = 0.784$.

TABLE 5.13 Probability Distribution for the Number of Customers Who Click on the Link in the Martin's Targeted E-Mail

x	$f(x)$
0	$\frac{3!}{0!3!}(0.30)^0(0.70)^3 = 0.343$
1	$\frac{3!}{1!2!}(0.30)^1(0.70)^2 = 0.441$
2	$\frac{3!}{2!1!}(0.30)^2(0.70)^1 = 0.189$
3	$\frac{3!}{3!0!}(0.30)^3(0.70)^0 = \frac{0.027}{1.000}$

FIGURE 5.14 Graphical Representation of the Probability Distribution for the Number of Customers Who Click on the Link in the Martin's Targeted E-Mail

If we consider a scenario in which 10 customers receive the targeted e-mail, the binomial probability mass function given by equation (5.16) is still applicable. If we want to find the probability that exactly 4 of the 10 customers click on the link and $p = 0.30$, then we calculate:

$$f(4) = \frac{10!}{4!6!}(0.30)^4(0.70)^6 = 0.2001$$

In Excel we can use the BINOM.DIST function to compute binomial probabilities. Figure 5.15 reproduces the Excel calculations from Table 5.13 for the Martin's problem with three customers.

The BINOM.DIST function in Excel has four input values: the first is the value of x, the second is the value of n, the third is the value of p, and the fourth is FALSE or TRUE. We choose FALSE for the fourth input if a probability mass function value $f(x)$ is desired, and TRUE if a cumulative probability is desired. The formula =BINOM.DIST(A5,$D$1:$D$2,FALSE) has been entered into cell B5 to compute the probability of 0 successes in three trials, $f(0)$. Figure 5.15 shows that this value is 0.343, the same as in Table 5.13.

Cells C5:C8 show the cumulative probability distribution values for this example. Note that these values are computed in Excel by entering TRUE as the fourth input in the BINOM.DIST. The cumulative probability for x using a binomial distribution is the probability of x or fewer successes out of n trials. Cell C5 computes the cumulative probability for $x = 0$, which is the same as the probability for $x = 0$ because the probability of 0 successes is the same as the probability of 0 or fewer successes. Cell C7 computes the cumulative probability for $x = 2$ using the formula =BINOM.DIST(A7,$D$1,$D$2,TRUE). This value is 0.973, meaning that the probability that two or fewer customers click on the targeted e-mail link is 0.973. Note that the value 0.973 simply corresponds to $f(0) + f(1) + f(2) = 0.343 + 0.441 + 0.189 = 0.973$ because it is the probability of two or fewer customers clicking on the link, which could be zero customers, one customer, or two customers.

FIGURE 5.15 Excel Worksheet for Computing Binomial Probabilities of the Number of Customers Who Make a Purchase at Martin's

	A	B	C	D
1			**Number of Trials (n):**	3
2			**Probability of Success (p):**	0.3
3				
4	***x***	***f(x)***	**Cumulative Probability**	
5	0	=BINOM.DIST(A5,$D$1,$D$2,FALSE)	=BINOM.DIST(A5,$D$1,$D$2,TRUE)	
6	1	=BINOM.DIST(A6,$D$1,$D$2,FALSE)	=BINOM.DIST(A6,$D$1,$D$2,TRUE)	
7	2	=BINOM.DIST(A7,$D$1,$D$2,FALSE)	=BINOM.DIST(A7,$D$1,$D$2,TRUE)	
8	3	=BINOM.DIST(A8,$D$1,$D$2,FALSE)	=BINOM.DIST(A8,$D$1,$D$2,TRUE)	

	A	B	C	D
1			**Number of Trials (n):**	3
2			**Probability of Success (p):**	0.3
3				
4	***x***	***f(x)***	**Cumulative Probability**	
5	0	0.343	0.343	
6	1	0.441	0.784	
7	2	0.189	0.973	
8	3	0.027	1.000	

Poisson Probability Distribution

In this section, we consider a discrete random variable that is often useful in estimating the number of occurrences of an event over a specified interval of time or space. For example, the random variable of interest might be the number of patients who arrive at a health care clinic in 1 hour, the number of computer-server failures in a month, the number of repairs needed in 10 miles of highway, or the number of leaks in 100 miles of pipeline. If the following two properties are satisfied, the number of occurrences is a random variable that is described by the **Poisson probability distribution**: (1) the probability of an occurrence is the same for any two intervals (of time or space) of equal length; and (2) the occurrence or nonoccurrence in any interval (of time or space) is independent of the occurrence or nonoccurrence in any other interval.

The Poisson probability mass function is defined by equation (5.17).

POISSON PROBABILITY MASS FUNCTION

$$f(x) = \frac{\mu^x e^{-\mu}}{x!} \tag{5.17}$$

where

$f(x)$ = the probability of x occurrences in an interval
μ = expected value or mean number of occurrences in an interval
$e \approx 2.71828$

The number e is a mathematical constant that is the base of the natural logarithm. Although it is an irrational number, 2.71828 is a sufficient approximation for our purposes.

For the Poisson probability distribution, x is a discrete random variable that indicates the number of occurrences in the interval. Since there is no stated upper limit for the number of occurrences, the probability mass function $f(x)$ is applicable for values $x = 0,1,2,\ldots$ without limit. In practical applications, x will eventually become large enough so that $f(x)$ is approximately zero and the probability of any larger values of x becomes negligible.

Suppose that we are interested in the number of patients who arrive at the emergency room of a large hospital during a 15-minute period on weekday mornings. Obviously, we do not know exactly how many patients will arrive at the emergency room in any defined interval of time, so the value of this variable is uncertain. It is important for administrators at the hospital to understand the probabilities associated with the number of arriving patients, as this information will have an impact on staffing decisions such as how many nurses and doctors to hire. It will also provide insight into possible wait times for patients to be seen once they arrive at the emergency room. If we can assume that the probability of a patient arriving is the same for any two periods of equal length during this 15-minute period and that the arrival or nonarrival of a patient in any period is independent of the arrival or nonarrival in any other period during the 15-minute period, the Poisson probability mass function is applicable. Suppose these assumptions are satisfied and an analysis of historical data shows that the average number of patients arriving during a 15-minute period of time is 10; in this case, the following probability mass function applies:

$$f(x) = \frac{10^x e^{-10}}{x!}$$

The random variable here is x = number of patients arriving at the emergency room during any 15-minute period.

If the hospital's management team wants to know the probability of exactly five arrivals during 15 minutes, we would set $x = 5$ and obtain:

$$\text{Probability of exactly 5 arrivals in 15 minutes} = f(5) = \frac{10^5 e^{-10}}{5!} = 0.0378$$

In the preceding example, the mean of the Poisson distribution is $\mu = 10$ arrivals per 15-minute period. A property of the Poisson distribution is that the mean of the distribution

and the variance of the distribution are *always equal*. Thus, the variance for the number of arrivals during all 15-minute periods is $\sigma^2 = 10$, and so the standard deviation is $\sigma = \sqrt{10} = 3.16$. Our illustration involves a 15-minute period, but other amounts of time can be used. Suppose we want to compute the probability of one arrival during a 3-minute period. Because 10 is the expected number of arrivals during a 15-minute period, we see that $10/15 = 2/3$ is the expected number of arrivals during a 1-minute period and that $(2/3)(3\text{minutes}) = 2$ is the expected number of arrivals during a 3-minute period. Thus, the probability of x arrivals during a 3-minute period with $\mu = 2$ is given by the following Poisson probability mass function:

$$f(x) = \frac{2^x e^{-2}}{x!}$$

The probability of one arrival during a 3-minute period is calculated as follows:

$$\text{Probability of exactly 1 arrival in 3 minutes} = f(1) = \frac{2^1 e^{-2}}{1!} = 0.2707$$

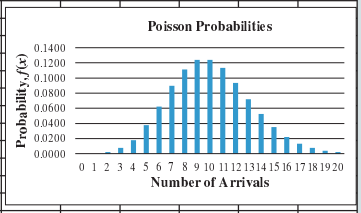
One might expect that because $(5\text{arrivals})/5 = 1\text{arrival}$ and $(15\text{minutes})/5 = 3\text{minutes}$, we would get the same probability for one arrival during a 3-minute period as we do for five arrivals during a 15-minute period. Earlier we computed the probability of five arrivals during a 15-minute period as 0.0378. However, note that the probability of one arrival during a 3-minute period is 0.2707, which is not the same. When computing a Poisson probability for a different time interval, we must first convert the mean arrival rate to the period of interest and then compute the probability.

In Excel we can use the POISSON.DIST function to compute Poisson probabilities. Figure 5.16 shows how to calculate the probabilities of patient arrivals at the emergency room if patients arrive at a mean rate of 10 per 15-minute interval.

FIGURE 5.16 Excel Worksheet for Computing Poisson Probabilities of the Number of Patients Arriving at the Emergency Room

	A	B	C	D
1			Mean Number of Occurrences:	10
2				
3	Number of Arrivals	Probability, $f(x)$	Cumulative Probability	
4	0	=POISSON.DIST(A4,$D$1,FALSE)	=POISSON.DIST(A4,$D$1,TRUE)	
5	1	=POISSON.DIST(A5,$D$1,FALSE)	=POISSON.DIST(A5,$D$1,TRUE)	
6	2	=POISSON.DIST(A6,$D$1,FALSE)	=POISSON.DIST(A6,$D$1,TRUE)	
7	3	=POISSON.DIST(A7,$D$1,FALSE)	=POISSON.DIST(A7,$D$1,TRUE)	
8	4	=POISSON.DIST(A8,$D$1,FALSE)	=POISSON.DIST(A8,$D$1,TRUE)	
9	5	=POISSON.DIST(A9,$D$1,FALSE)	=POISSON.DIST(A9,$D$1,TRUE)	
10	6	=POISSON.DIST(A10,$D$1,FALSE)	=POISSON.DIST(A10,$D$1,TRUE)	
11	7	=POISSON.DIST(A11,$D$1,FALSE)	=POISSON.DIST(A11,$D$1,TRUE)	
12	8	=POISSON.DIST(A12,$D$1,FALSE)	=POISSON.DIST(A12,$D$1,TRUE)	
13	9	=POISSON.DIST(A13,$D$1,FALSE)	=POISSON.DIST(A13,$D$1,TRUE)	
14	10	=POISSON.DIST(A14,$D$1,FALSE)	=POISSON.DIST(A14,$D$1,TRUE)	
15	11	=POISSON.DIST(A15,$D$1,FALSE)	=POISSON.DIST(A15,$D$1,TRUE)	
16	12	=POISSON.DIST(A16,$D$1,FALSE)	=POISSON.DIST(A16,$D$1,TRUE)	
17	13	=POISSON.DIST(A17,$D$1,FALSE)	=POISSON.DIST(A17,$D$1,TRUE)	
18	14	=POISSON.DIST(A18,$D$1,FALSE)	=POISSON.DIST(A18,$D$1,TRUE)	
19	15	=POISSON.DIST(A19,$D$1,FALSE)	=POISSON.DIST(A19,$D$1,TRUE)	
20	16	=POISSON.DIST(A20,$D$1,FALSE)	=POISSON.DIST(A20,$D$1,TRUE)	
21	17	=POISSON.DIST(A21,$D$1,FALSE)	=POISSON.DIST(A21,$D$1,TRUE)	
22	18	=POISSON.DIST(A22,$D$1,FALSE)	=POISSON.DIST(A22,$D$1,TRUE)	
23	19	=POISSON.DIST(A23,$D$1,FALSE)	=POISSON.DIST(A23,$D$1,TRUE)	
24	20	=POISSON.DIST(A24,$D$1,FALSE)	=POISSON.DIST(A24,$D$1,TRUE)	

	A	B	C	D	E	F	G	H	I
1			Mean Number of Occurrences:	10					
2									
3	Number of Arrivals	Probability, $f(x)$	Cumulative Probability						
4	0	0.0000	0.0000						
5	1	0.0005	0.0005						
6	2	0.0023	0.0028						
7	3	0.0076	0.0103						
8	4	0.0189	0.0293						
9	5	0.0378	0.0671						
10	6	0.0631	0.1301						
11	7	0.0901	0.2202						
12	8	0.1126	0.3328						
13	9	0.1251	0.4579						
14	10	0.1251	0.5830						
15	11	0.1137	0.6968						
16	12	0.0948	0.7916						
17	13	0.0729	0.8645						
18	14	0.0521	0.9165						
19	15	0.0347	0.9513						
20	16	0.0217	0.9730						
21	17	0.0128	0.9857						
22	18	0.0071	0.9928						
23	19	0.0037	0.9965						
24	20	0.0019	0.9984						



The POISSON.DIST function in Excel has three input values: the first is the value of x, the second is the mean of the Poisson distribution, and the third is FALSE or TRUE. We choose FALSE for the third input if a probability mass function value $f(x)$ is desired, and TRUE if a cumulative probability is desired. The formula =POISSON.DIST(A4,$D$1,FALSE) has been entered into cell B4 to compute the probability of 0 occurrences, $f(0)$. Figure 5.16 shows that this value (to four decimal places) is 0.0000, which means that it is highly unlikely (probability near 0) that we will have 0 patient arrivals during a 15-minute interval. The value in cell B12 shows that the probability that there will be exactly eight arrivals during a 15-minute interval is 0.1126.

The cumulative probability for x using a Poisson distribution is the probability of x or fewer occurrences during the interval. Cell C4 computes the cumulative probability for $x = 0$, which is the same as the probability for $x = 0$ because the probability of 0 occurrences is the same as the probability of 0 or fewer occurrences. Cell C12 computes the cumulative probability for $x = 8$ using the formula =POISSON.DIST(A12,$D$1,TRUE). This value is 0.3328, meaning that the probability that eight or fewer patients arrive during a 15-minute interval is 0.3328. This value corresponds to

$$f(0) + f(1) + f(2) + \cdots + f(7) + f(8) = 0.0000 + 0.0005 + 0.0023 + \cdots + 0.0901 + 0.1126 = 0.3328$$

Let us illustrate an application not involving time intervals in which the Poisson distribution is useful. Suppose we want to determine the occurrence of major defects in a highway one month after it has been resurfaced. We assume that the probability of a defect is the same for any two highway intervals of equal length and that the occurrence or nonoccurrence of a defect in any one interval is independent of the occurrence or nonoccurrence of a defect in any other interval. Hence, the Poisson distribution can be applied.

Suppose we learn that major defects one month after resurfacing occur at the average rate of two per mile. Let us find the probability of no major defects in a particular 3-mile section of the highway. Because we are interested in an interval with a length of 3 miles, $\mu = (2\text{defects/mile})(3\text{miles}) = 6$ represents the expected number of major defects over the 3-mile section of highway. Using equation (5.17), the probability of no major defects is $f(0) = \dfrac{6^0 e^{-6}}{0!} = 0.0025$. Thus, it is unlikely that no major defects will occur in the 3-mile section. In fact, this example indicates a $1 - 0.0025 = 0.9975$ probability of at least one major defect in the 3-mile highway section.

NOTES + COMMENTS

1. If sample data are used to estimate the probabilities of a custom discrete distribution, equation (5.13) yields the sample mean $\bar{x}$ rather than the population mean μ. However, as the sample size increases, the sample generally becomes more representative of the population and the sample mean $\bar{x}$ converges to the population mean μ. In this chapter we have assumed that the sample of 300 Lancaster Savings and Loan mortgage customers is sufficiently large to be representative of the population of mortgage customers at Lancaster Savings and Loan.
2. We can use the Excel function AVERAGE only to compute the expected value of a custom discrete random variable when the values in the data occur with relative frequencies that correspond to the probability distribution of the random variable. If this assumption is not satisfied, then the estimate of the expected value with the AVERAGE function will be inaccurate. In practice, this assumption is satisfied with an increasing degree of accuracy as the size of the sample is increased. Otherwise, we must use equation (5.13) to calculate the expected value for a custom discrete random variable.
3. If sample data are used to estimate the probabilities for a custom discrete distribution, equation (5.14) yields the sample variance s^2 rather than the population variance σ^2. However, as the sample size increases the sample generally becomes more representative of the population and the sample variance s^2 converges to the population variance σ^2.

5.6 Continuous Probability Distributions

In the preceding section we discussed discrete random variables and their probability distributions. In this section we consider continuous random variables. Specifically, we discuss some of the more useful continuous probability distributions for analytics models: the uniform, the triangular, the normal, and the exponential.

A fundamental difference separates discrete and continuous random variables in terms of how probabilities are computed. For a discrete random variable, the probability mass function $f(x)$ provides the probability that the random variable assumes a particular value. With continuous random variables, the counterpart of the probability mass function is the **probability density function**, also denoted by $f(x)$. The difference is that the probability density function does not directly provide probabilities. However, the area under the graph of $f(x)$ corresponding to a given interval does provide the probability that the continuous random variable x assumes a value in that interval. So when we compute probabilities for continuous random variables, we are computing the probability that the random variable assumes any value in an interval. Because the area under the graph of $f(x)$ at any particular point is zero, one of the implications of the definition of probability for continuous random variables is that the probability of any particular value of the random variable is zero.

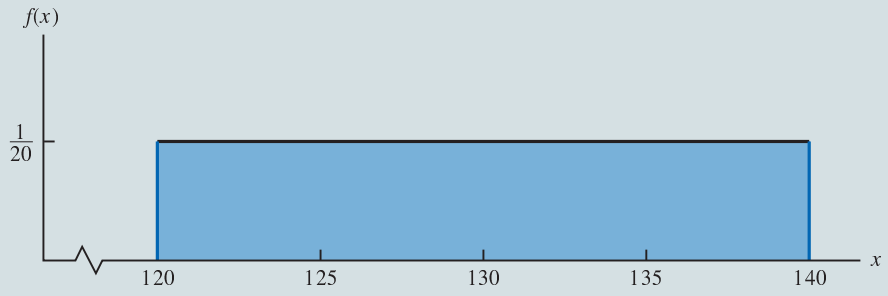
Uniform Probability Distribution

Consider the random variable x representing the flight time of an airplane traveling from Chicago to New York. The exact flight time from Chicago to New York is uncertain because it can be affected by weather (headwinds or storms), flight traffic patterns, and other factors that cannot be known with certainty. It is important to characterize the uncertainty associated with the flight time because this can have an impact on connecting flights and how we construct our overall flight schedule. Suppose the flight time can be any value in the interval from 120 minutes to 140 minutes. Because the random variable x can assume any value in that interval, x is a continuous rather than a discrete random variable. Let us assume that sufficient actual flight data are available to conclude that the probability of a flight time within any interval of a given length is the same as the probability of a flight time within any other interval of the same length that is contained in the larger interval from 120 to 140 minutes. With every interval of a given length being equally likely, the random variable x is said to have a **uniform probability distribution**. The probability density function, which defines the uniform distribution for the flight-time random variable, is:

$$f(x) = \begin{cases} 1/20 & \text{for } 120 \le x \le 140 \\ 0 & \text{elsewhere} \end{cases}$$

Figure 5.17 shows a graph of this probability density function.

FIGURE 5.17 Uniform Probability Distribution for Flight Time



In general, the uniform probability density function for a random variable x is defined by the following formula:

UNIFORM PROBABILITY DENSITY FUNCTION

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{elsewhere} \end{cases} \tag{5.18}$$

For the flight-time random variable, $a = 120$ and $b = 140$.

For a continuous random variable, we consider probability only in terms of the likelihood that a random variable assumes a value within a specified interval. In the flight time example, an acceptable probability question is: What is the probability that the flight time is between 120 and 130 minutes? That is, what is $P(120 \le x \le 130)$?

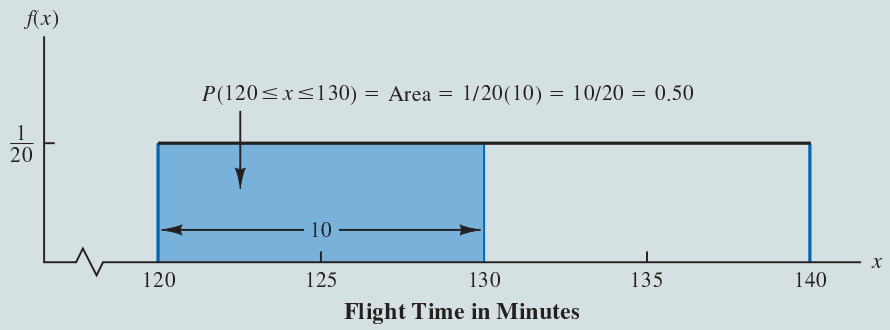
To answer this question, consider the area under the graph of $f(x)$ in the interval from 120 to 130 (see Figure 5.18). The area is rectangular, and the area of a rectangle is simply the width multiplied by the height. With the width of the interval equal to $130 - 120 = 10$ and the height equal to the value of the probability density function $f(x) = 1/20$, we have area = width × height = $10(1/20) = 10/20 = 0.50$.

The area under the graph of $f(x)$ and probability are identical for all continuous random variables. Once a probability density function $f(x)$ is identified, the probability that x takes a value between some lower value x_1 and some higher value x_2 can be found by computing the area under the graph of $f(x)$ over the interval from x_1 to x_2.

Given the uniform distribution for flight time and using the interpretation of area as probability, we can answer any number of probability questions about flight times. For example:

- What is the probability of a flight time between 128 and 136 minutes? The width of the interval is $136 - 128 = 8$. With the uniform height of $f(x) = 1/20$, we see that $P(128 \le x \le 136) = 8(1/20) = 0.40$.
- What is the probability of a flight time between 118 and 123 minutes? The width of the interval is $123 - 118 = 5$, but the height is $f(x) = 0$ for $118 \le x < 120$ and $f(x) = 1/20$ for $120 \le x \le 123$, so we have that $P(118 \le x \le 123) = P(118 \le x < 120) + P(120 \le x \le 123) = 2(0) + 3(1/20) = 0.15$.

FIGURE 5.18 The Area Under the Graph Provides the Probability of a Flight Time Between 120 and 130 Minutes



Note that $P(120 \le x \le 140) = 20(1/20) = 1$; that is, the total area under the graph of $f(x)$ is equal to 1. This property holds for all continuous probability distributions and is the analog of the condition that the sum of the probabilities must equal 1 for a discrete probability mass function.

Note also that because we know that the height of the graph of $f(x)$ for a uniform distribution is $\frac{1}{b-a}$ for $a \le x \le b$, then the area under the graph of $f(x)$ for a uniform distribution evaluated from a to a point x_0 when $a \le x_0 \le b$ is width $\times$ height $= (x_0 - a) \times (b - a)$. This value provides the cumulative probability of obtaining a value for a uniform random variable of less than or equal to some specific value denoted by x_0 and the formula is given in equation (5.19).

UNIFORM DISTRIBUTION: CUMULATIVE PROBABILITIES

$$P(x \le x_0) = \frac{x_0 - a}{b - a} \text{ for } a \le x_0 \le b \qquad \textbf{(5.19)}$$

The calculation of the expected value and variance for a continuous random variable is analogous to that for a discrete random variable. However, because the computational procedure involves integral calculus, we do not show the formulas here.

For the uniform continuous probability distribution introduced in this section, the formulas for the expected value and variance are as follows:

$$E(x) = \frac{a + b}{2}$$

$$\text{Var}(x) = \frac{(b - a)^2}{12}$$

In these formulas, a is the minimum value and b is the maximum value that the random variable may assume.

Applying these formulas to the uniform distribution for flight times from Chicago to New York, we obtain

$$E(x) = \frac{(120 + 140)}{2} = 130$$

$$\text{Var}(x) = \frac{(140 - 120)^2}{12} = 33.33$$

The standard deviation of flight times can be found by taking the square root of the variance. Thus, for flight times from Chicago to New York, $\sigma = \sqrt{33.33} = 5.77$ minutes.

Triangular Probability Distribution

The triangular probability distribution is useful when only subjective probability estimates are available. There are many situations for which we do not have sufficient data and only subjective estimates of possible values are available. In the **triangular probability distribution**, we need only to specify the minimum possible value a, the maximum possible value b, and the most likely value (or mode) of the distribution m. If these values can be knowledgeably estimated for a continuous random variable by a subject-matter expert, then as an approximation of the actual probability density function, we can assume that the triangular distribution applies.

Consider a situation in which a project manager is attempting to estimate the time that will be required to complete an initial assessment of the capital project of constructing a new corporate headquarters. The assessment process includes completing environmental-impact studies, procuring the required permits, and lining up all the contractors and

subcontractors needed to complete the project. There is considerable uncertainty regarding the duration of these tasks, and generally little or no historical data are available to help estimate the probability distribution for the time required for this assessment process.

Suppose that we are able to discuss this project with several subject-matter experts who have worked on similar projects. From these expert opinions and our own experience, we estimate that the minimum required time for the initial assessment phase is six months and that the worst-case estimate is that this phase could require 24 months if we are delayed in the permit process or if the results from the environmental-impact studies require additional action. While a time of six months represents a best case and 24 months a worst case, the consensus is that the most likely amount of time required for the initial assessment phase of the project is 12 months. From these estimates, we can use a triangular distribution as an approximation for the probability density function for the time required for the initial assessment phase of constructing a new corporate headquarters.

Figure 5.19 shows the probability density function for this triangular distribution. Note that the probability density function is a triangular shape.

The general form of the triangular probability density function is as follows:

TRIANGULAR PROBABILITY DENSITY FUNCTION

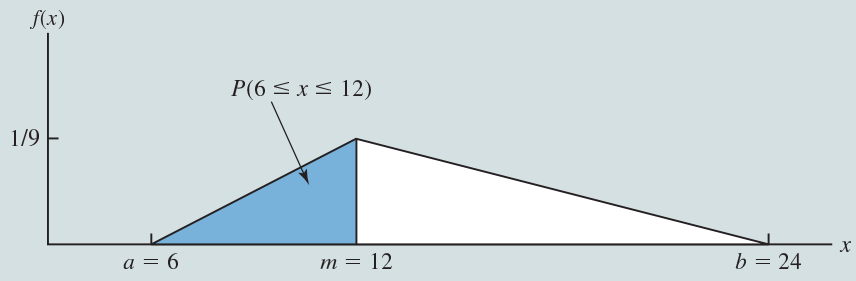
$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(m-a)} & \text{for } a \le x \le m \\ \dfrac{2(b-x)}{(b-a)(b-m)} & \text{for } m < x \le b \end{cases} \tag{5.20}$$

where

a = minimum value
b = maximum value
m = mode

In the example of the time required to complete the initial assessment phase of constructing a new corporate headquarters, the minimum value a is six months, the maximum value b is 24 months, and the mode m is 12 months. As with the explanation given for the uniform distribution above, we can calculate probabilities by using the area under the graph of $f(x)$. We can calculate the probability that the time required is less than 12 months by finding the area under the graph of $f(x)$ from $x = 6$ to $x = 12$ as shown in Figure 5.19.

FIGURE 5.19 Triangular Probability Distribution for Time Required for Initial Assessment of Corporate Headquarters Construction



The geometry required to find this area for any given value is slightly more complex than that required to find the area for a uniform distribution, but the resulting formula for a triangular distribution is relatively simple:

TRIANGULAR DISTRIBUTION: CUMULATIVE PROBABILITIES

$$P(x \le x_0) = \begin{cases} \dfrac{(x_0 - a)^2}{(b - a)(m - a)} & \text{for } a \le x_0 \le m \\ 1 - \dfrac{(b - x_0)^2}{(b - a)(b - m)} & \text{for } m < x_0 \le b \end{cases} \tag{5.21}$$

Equation (5.21) provides the cumulative probability of obtaining a value for a triangular random variable of less than or equal to some specific value denoted by x_0.

To calculate $P(x \le 12)$ we use equation (5.20) with $a = 6$, $b = 24$, $m = 12$, and $x_0 = 12$.

$$P(x \le 12) = \frac{(12 - 6)^2}{(24 - 6)(12 - 6)} = 0.3333$$

Thus, the probability that the assessment phase of the project requires less than 12 months is 0.3333. We can also calculate the probability that the project requires more than 10 months, but less than or equal to 18 months by subtracting $P(x \le 10)$ from $P(x \le 18)$. This is shown graphically in Figure 5.20. The calculations are as follows:

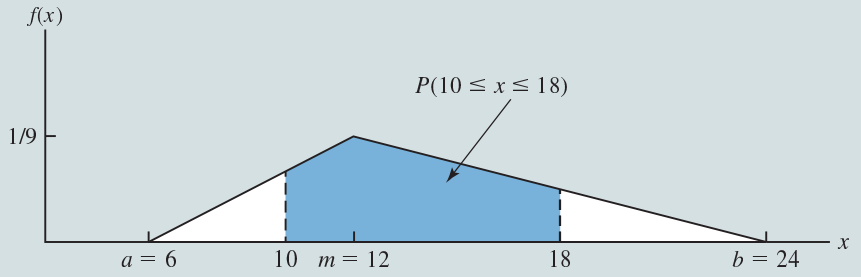
$$P(x \le 18) - P(x \le 10) = \left[1 - \frac{(24 - 18)^2}{(24 - 6)(24 - 12)}\right] - \left[\frac{(10 - 6)^2}{(24 - 6)(10 - 6)}\right] = 0.6111$$

Thus, the probability that the assessment phase of the project requires at least 10 months but less than 18 months is 0.6111.

Normal Probability Distribution

One of the most useful probability distributions for describing a continuous random variable is the **normal probability distribution**. The normal distribution has been used in a wide variety of practical applications in which the random variables are heights and weights of people, test scores, scientific measurements, amounts of rainfall, and other similar values. It is also widely used in business applications to describe uncertain quantities such as demand for products, the rate of return for stocks and bonds, and the time it takes to manufacture a part or complete many types of service-oriented activities such as medical surgeries and consulting engagements.

FIGURE 5.20 Triangular Distribution to Determine $P(10 \le x \le 18) = P(x \le 18) - P(x \le 10)$



The form, or shape, of the normal distribution is illustrated by the bell-shaped normal curve in Figure 5.21.

The probability density function that defines the bell-shaped curve of the normal distribution follows.

NORMAL PROBABILITY DENSITY FUNCTION

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{5.22}$$

where

μ = mean
σ = standard deviation
$\pi \approx 3.14159$
$e \approx 2.71828$

Although π and e are irrational numbers, 3.14159 and 2.71828, respectively, are sufficient approximations for our purposes.

We make several observations about the characteristics of the normal distribution.

1. The entire family of normal distributions is differentiated by two parameters: the mean μ and the standard deviation σ. The mean and standard deviation are often referred to as the location and shape parameters of the normal distribution, respectively.
2. The highest point on the normal curve is at the mean, which is also the median and mode of the distribution.
3. The mean of the distribution can be any numerical value: negative, zero, or positive. Three normal distributions with the same standard deviation but three different means (−10, 0, and 20) are shown in Figure 5.22.

FIGURE 5.21 Bell-Shaped Curve for the Normal Distribution

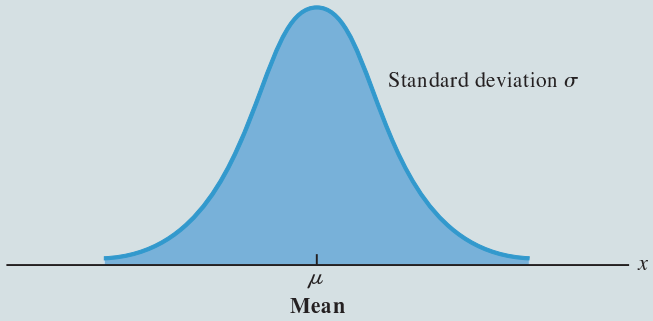
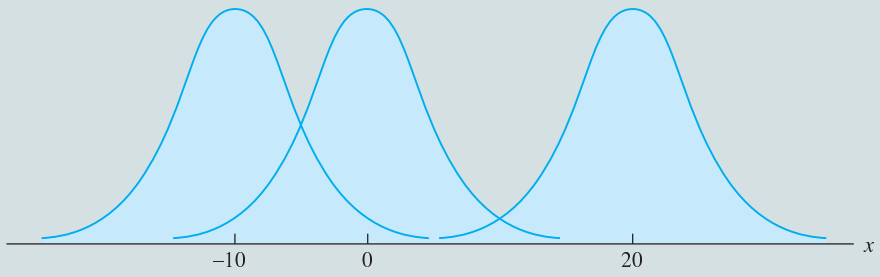


FIGURE 5.22 Three Normal Distributions with the Same Standard Deviation but Different Means ($\mu = -10$, $\mu = 0$, $\mu = 20$)



4. The normal distribution is symmetric, with the shape of the normal curve to the left of the mean a mirror image of the shape of the normal curve to the right of the mean.
5. The tails of the normal curve extend to infinity in both directions and theoretically never touch the horizontal axis. Because it is symmetric, the normal distribution is not skewed; its skewness measure is zero.
6. The standard deviation determines how flat and wide the normal curve is. Larger values of the standard deviation result in wider, flatter curves, showing more variability in the data. More variability corresponds to greater uncertainty. Two normal distributions with the same mean but with different standard deviations are shown in Figure 5.23.
7. Probabilities for the normal random variable are given by areas under the normal curve. The total area under the curve for the normal distribution is 1. Because the distribution is symmetric, the area under the curve to the left of the mean is 0.50 and the area under the curve to the right of the mean is 0.50.
8. The percentages of values in some commonly used intervals are as follows:
   a. 68.3% of the values of a normal random variable are within plus or minus one standard deviation of its mean.
   b. 95.4% of the values of a normal random variable are within plus or minus two standard deviations of its mean.
   c. 99.7% of the values of a normal random variable are within plus or minus three standard deviations of its mean.

These percentages are the basis for the empirical rule discussed in Section 2.7.

Figure 5.24 shows properties (a), (b), and (c) graphically.

We turn now to an application of the normal probability distribution. Suppose Grear Aircraft Engines sells aircraft engines to commercial airlines. Grear is offering a new performance-based sales contract in which Grear will guarantee that its engines will provide a certain amount of lifetime flight hours subject to the airline purchasing a preventive-maintenance service plan that is also provided by Grear. Grear believes that this performance-based contract will lead to additional sales as well as additional income from providing the associated preventive maintenance and servicing.

From extensive flight testing and computer simulations, Grear's engineering group has estimated that if their engines receive proper parts replacement and preventive maintenance, the mean lifetime flight hours achieved is normally distributed with a mean $\mu = 36{,}500$ hours and standard deviation $\sigma = 5{,}000$ hours. Grear would like to know what percentage of its aircraft engines will be expected to last more than 40,000 hours. In other words, what is the probability that the aircraft lifetime flight hours x will exceed 40,000? This question can be answered by finding the area of the darkly shaded region in Figure 5.25.

FIGURE 5.23 Two Normal Distributions with the Same Mean but Different Standard Deviations ($\sigma = 5$, $\sigma = 10$)

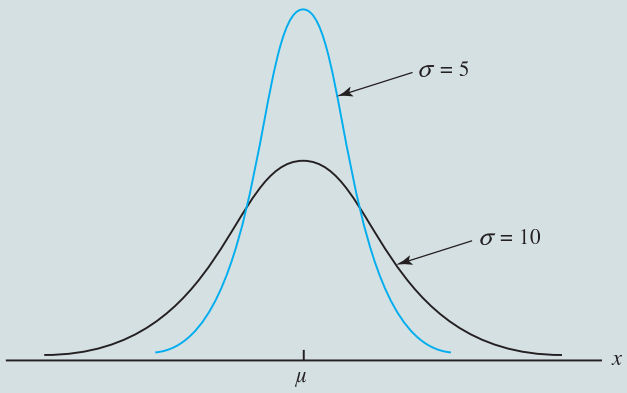


FIGURE 5.24 Areas Under the Curve for Any Normal Distribution

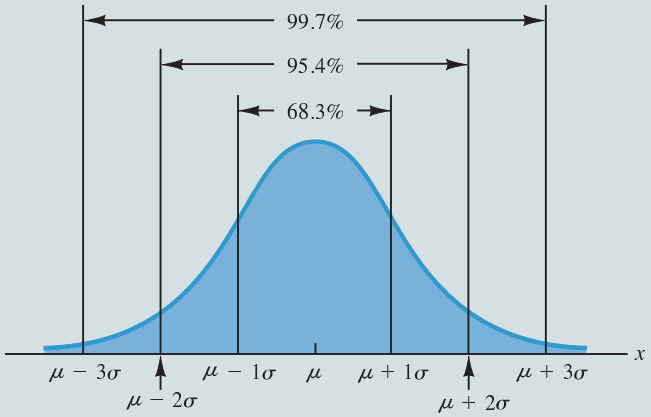
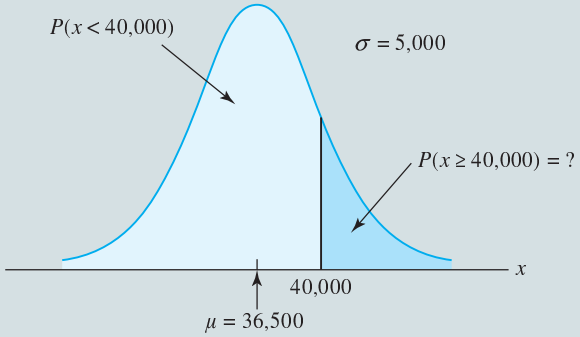


FIGURE 5.25 Grear Aircraft Engines Lifetime Flight Hours Distribution



The Excel function NORM.DIST can be used to compute the area under the curve for a normal probability distribution. The NORM.DIST function has four input values. The first is the value of interest corresponding to the probability you want to calculate, the second is the mean of the normal distribution, the third is the standard deviation of the normal distribution, and the fourth is TRUE or FALSE. We enter TRUE for the fourth input if we want the cumulative distribution function and FALSE if we want the probability density function.

Figure 5.26 shows how we can answer the question of interest for Grear using Excel—in cell B5, we use the formula =NORM.DIST(40,000, $B$1, $B$2, TRUE). Cell B1 contains the mean of the normal distribution and cell B2 contains the standard deviation. Because we want to know the area under the curve, we want the cumulative distribution function, so we use TRUE as the fourth input value in the formula. This formula provides a value of 0.7580 in cell B5. But note that this corresponds to $P(x \leq 40{,}000) = 0.7580$. In other words, this gives us the area under the curve to the left of $x = 40{,}000$ in Figure 5.25, and we are interested in the area under the curve to the right of $x = 40{,}000$. To find this value, we simply use $1 - 0.7580 = 0.2420$ (cell B6). Thus, 0.2420 is the probability that x will exceed 40,000 hours. We can conclude that about 24.2% of aircraft engines will exceed 40,000 lifetime flight hours.

FIGURE 5.26 Excel Calculations for Grear Aircraft Engines Example

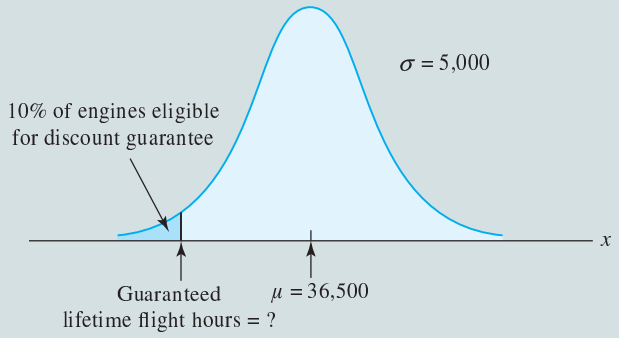
	A	B	C
1	**Mean:**	36500	
2	**Standard Deviation:**	5000	
3			
4			
5	$P(x \leq 40{,}000) =$	=NORM.DIST(40000, $B$1, $B$2,TRUE)	
6	$P(x > 40{,}000) = 1 - P(x \leq 40{,}000) =$	=1-B5	
7			
8	**Guarantee on Lifetime Flight Hours for 10% of engines eligible for discount guarantee:**	=NORM.INV(0.1, $B$1, $B$2)	

	A	B	C
1	**Mean:**	36500	
2	**Standard Deviation:**	5000	
3			
4			
5	$P(x \leq 40{,}000) =$	0.7580	
6	$P(x > 40{,}000) = 1 - P(x \leq 40{,}000) =$	0.2420	
7			
8	**Guarantee on Lifetime Flight Hours for 10% of engines eligible for discount guarantee:**	30092.24	

Let us now assume that Grear is considering a guarantee that will provide a discount on a replacement aircraft engine if the original engine does not meet the lifetime-flight-hour guarantee. How many lifetime flight hours should Grear guarantee if Grear wants no more than 10% of aircraft engines to be eligible for the discount guarantee? This question is interpreted graphically in Figure 5.27.

According to Figure 5.27, the area under the curve to the left of the unknown guarantee on lifetime flight hours must be 0.10. To find the appropriate value using Excel, we use the function NORM.INV. The NORM.INV function has three input values. The first is the probability of interest, the second is mean of the normal distribution, and the third is the standard deviation of the normal distribution. Figure 5.26 shows how we can use Excel to answer the Grear's question about a guarantee on lifetime flight hours. In cell B8 we use

FIGURE 5.27 Grear's Discount Guarantee



With the guarantee set at 30,000 hours, the actual percentage eligible for the guarantee will be =NORM.DIST(30000,36500, 5000,TRUE) = 0.0968, or 9.68%

the formula =NORM.INV(0.10, $B$1, $B$2), where the mean of the normal distribution is contained in cell B1 and the standard deviation in cell B2. This provides a value of 30,092.24. Thus, a guarantee of 30,092 hours will meet the requirement that approximately 10% of the aircraft engines will be eligible for the guarantee. This information could be used by Grear's analytics team to suggest a lifetime flight hours guarantee of 30,000 hours.

Perhaps Grear is also interested in knowing the probability that an engine will have a lifetime of flight hours greater than 30,000 hours but less than 40,000 hours. How do we calculate this probability? First, we can restate this question as follows. What is $P(30,000 \le x \le 40,000)$? Figure 5.28 shows the area under the curve needed to answer this question. The area that corresponds to $P(30,000 \le x \le 40.000)$ can be found by subtracting the area corresponding to $P(x \le 30,000)$ from the area corresponding to $P(x \le 40,000)$. In other words, $P(30,000 \le x \le 40,000) = P(x \le 40,000) - P(x \le 30,000)$. Figure 5.29 shows how we can find the value for $P(30,000 \le x \le 40,000)$ using Excel. We calculate $P(x \le 40,000)$ in cell B5 and $P(x \le 30,000)$ in cell B6 using the NORM.DIST function. We then calculate $P(30,000 \le x \le 40,000)$ in cell B8 by subtracting the value in cell B6 from the value in cell B5. This tells us that $P(30,000 \le x \le 40,000) = 0.7580 - 0.0968 = 0.6612$. In other words, the probability that the lifetime flight hours for an aircraft engine will be between 30,000 hours and 40,000 hours is 0.6612.

Note that we can calculate P(30,000 ≤ x ≤ 40,000) in a single cell using the formula =NORM.DIST(40000, $B$1, $B$2, TRUE) – NORM.DIST(30000, $B$1, $B$2, TRUE).

Exponential Probability Distribution

The **exponential probability distribution** may be used for random variables such as the time between patient arrivals at an emergency room, the distance between major defects in a highway, and the time until default in certain credit-risk models. The exponential probability density function is as follows:

FIGURE 5.28 Graph Showing the Area Under the Curve Corresponding to $P(30,000 \le x \le 40,000)$ in the Grear Aircraft Engines Example

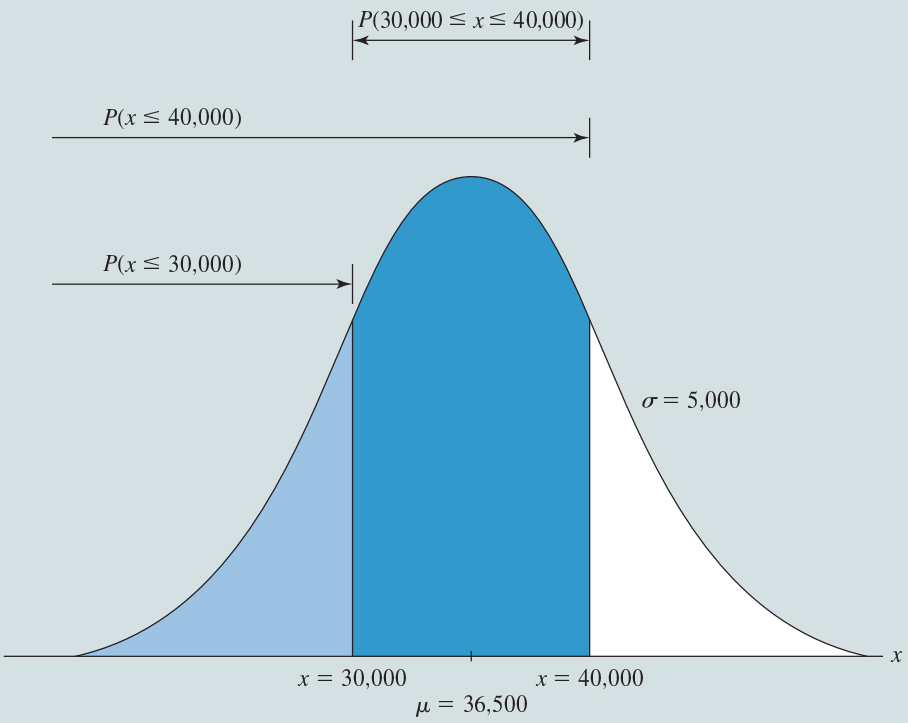


FIGURE 5.29 Using Excel to Find $P(30{,}000 \le x \le 40.000)$ in the Grear Aircraft Engines Example

	A	B	C
1	**Mean:**	36500	
2	**Standard Deviation:**	5000	
3			
4			
5	$P(x \le 40{,}000) =$	=NORM.DIST(40000, $B$1, $B$2,TRUE)	
6	$P(x \le 30{,}000) =$	=NORM.DIST(30000, $B$1, $B$2,TRUE)	
7			
8	$P(30{,}000 \le x \le 40{,}000) = P(x \le 40{,}000) - P(x \le 30{,}000) =$	=B5-B6	

	A	B	C
1	**Mean:**	36500	
2	**Standard Deviation:**	5000	
3			
4			
5	$P(x \le 40{,}000) =$	0.7580	
6	$P(x \le 30{,}000) =$	0.0968	
7			
8	$P(30{,}000 \le x \le 40{,}000) = P(x \le 40{,}000) - P(x \le 30{,}000) =$	0.6612	

EXPONENTIAL PROBABILITY DENSITY FUNCTION

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \quad \text{for } x \ge 0 \tag{5.23}$$

where

μ = expected value or mean
e = 2.71828

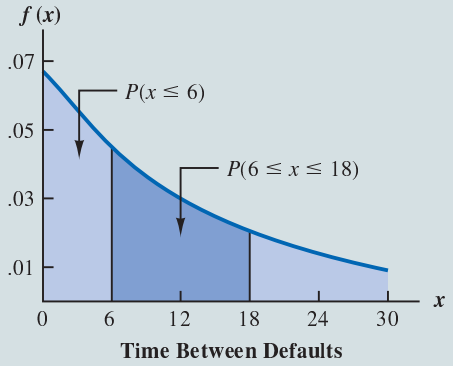
As an example, suppose that x represents the time between business loan defaults for a particular lending agency. If the mean, or average, time between loan defaults is 15 months ($\mu = 15$), the appropriate density function for x is

$$f(x) = \frac{1}{15} e^{-x/15}$$

Figure 5.30 is the graph of this probability density function.

As with any continuous probability distribution, the area under the curve corresponding to an interval provides the probability that the random variable assumes a value in that interval. In the time between loan defaults example, the probability that the time between defaults is six months or less, $P(x \le 6)$, is defined to be the area under the curve in Figure 5.30 from $x = 0$ to $x = 6$. Similarly, the probability that the time between defaults will be 18 months or less, $P(x \le 18)$, is the area under the curve from $x = 0$ to $x = 18$. Note also that the probability that the time between defaults will be between 6 months and 18 months, $P(6 \le x \le 18)$, is given by the area under the curve from $x = 6$ to $x = 18$.

FIGURE 5.30 Exponential Distribution for the Time Between Business Loan Defaults Example



To compute exponential probabilities such as those just described, we use the following formula, which provides the cumulative probability of obtaining a value for the exponential random variable of less than or equal to some specific value denoted by x_0.

EXPONENTIAL DISTRIBUTION: CUMULATIVE PROBABILITIES

$$P(x \leq x_0) = 1 - e^{-x_0/\mu} \tag{5.24}$$

For the time between defaults example, x = time between business loan defaults in months and $\mu = 15$ months. Using equation (5.24),

$$P(x \leq x_0) = 1 - e^{-x_0/15}$$

Hence, the probability that the time between defaults is six months or less is:

$$P(x \leq 6) = 1 - e^{-6/15} = 0.3297$$

Using equation (5.24), we calculate the probability that the time between defaults is 18 months or less:

$$P(x \leq 18) = 1 - e^{-18/15} = 0.6988$$

Thus, the probability that the time between business loan defaults is 6 months and 18 months is equal to $0.6988 - 0.3297 = 0.3691$. Probabilities for any other interval can be computed similarly.

Figure 5.31 shows how we can calculate these values for an exponential distribution in Excel using the function EXPON.DIST. The EXPON.DIST function has three inputs: the first input is x, the second input is $1/m$, and the third input is TRUE or FALSE. An input of TRUE for the third input provides the cumulative distribution function value and FALSE provides the probability density function value. Cell B3 calculates $P(x \leq 18)$ using the formula =EXPON.DIST(18, 1/$B$1, TRUE), where cell B1 contains the mean of the exponential distribution. Cell B4 calculates the value for $P(x \leq 6)$ and cell B5 calculates the value for $P(6 \leq x \leq 18) = P(x \leq 18) - P(x \leq 6)$ by subtracting the value in cell B4 from the value in cell B3.

We can calculate P(6 ≤ x ≤ 18) in a single cell using the formula =EXPON.DIST(18, 1/$B$1, TRUE) - EXPON.DIST(6, 1/$B$1, TRUE).

FIGURE 5.31 Using Excel to Calculate $P(6 \le x \le 18)$ for the Time Between Business Loan Defaults Example

	A	B	C
1	**Mean, μ =**	15	
2			
3	**$P(x \le 18)$ =**	=EXPON.DIST(18,1/$B$1, TRUE)	
4	**$P(x \le 6)$ =**	=EXPON.DIST(6,1/$B$1, TRUE)	
5	**$P(6 \le x \le 18) = P(x \le 18) - P(x \le 6)$ =**	=B3-B4	

	A	B	C
1	**Mean, μ =**	15	
2			
3	**$P(x \le 18)$ =**	0.6988	
4	**$P(x \le 6)$ =**	0.3297	
5	**$P(6 \le x \le 18) = P(x \le 18) - P(x \le 6)$ =**	0.3691	

NOTES + COMMENTS

1. The way we describe probabilities is different for a discrete random variable than it is for a continuous random variable. For discrete random variables, we can talk about the probability of the random variable assuming a particular value. For continuous random variables, we can only talk about the probability of the random variable assuming a value within a given interval.
2. To see more clearly why the height of a probability density function is not a probability, think about a random variable with the following uniform probability distribution:

$$f(x) = \begin{cases} 2 & \text{for } 0 \le x \le 0.5 \\ 0 & \text{elsewhere} \end{cases}$$

The height of the probability density function, $f(x)$, is 2 for values of x between 0 and 0.5. However, we know that probabilities can never be greater than 1. Thus, we see that $f(x)$ cannot be interpreted as the probability of x.

3. The standard normal distribution is the special case of the normal distribution for which the mean is 0 and the standard deviation is 1. This is useful because probabilities for all normal distributions can be computed using the standard normal distribution. We can convert any normal random variable x with mean μ and standard deviation σ to the standard normal random variable z by using the formula $z = \dfrac{x - \mu}{\sigma}$. We interpret z as the number of standard deviations that the normal random variable x is from its mean μ. Then we can use a table of standard normal probability distributions to find the area under the curve using z and the standard normal probability table. Excel contains special functions for the standard normal distribution: NORM.S.DIST and NORM.S.INV. The function NORM.S.DIST is similar to the function NORM.DIST, but it requires only two input values: the value of interest for calculating the probability and TRUE or FALSE, depending on whether you are interested in finding the probability density or the cumulative distribution function. NORM.S.INV is similar to the NORM.INV function, but it requires only the single input of the probability of interest. Both NORM.S.DIST and NORM.S.INV do not need the additional parameters because they assume a mean of 0 and standard deviation of 1 for the standard normal distribution.
4. A property of the exponential distribution is that the mean and the standard deviation are equal to each other.
5. The continuous exponential distribution is related to the discrete Poisson distribution. If the Poisson distribution provides an appropriate description of the number of occurrences per interval, the exponential distribution provides a description of the length of the interval between occurrences. This relationship often arises in queueing applications in which, if arrivals follow a Poisson distribution, the time between arrivals must follow an exponential distribution.
6. Chapter 11 explains how values for discrete and continuous random variables can be generated in Excel for use in simulation models. It also discusses how to use Analytic Solver to assess which probability distribution(s) best describe sample values of a random variable.

SUMMARY

In this chapter we introduced the concept of probability as a means of understanding and measuring uncertainty. Uncertainty is a factor in virtually all business decisions, thus an understanding of probability is essential to modeling such decisions and improving the decision-making process.

We introduced some basic relationships in probability including the concepts of outcomes, events, and calculations of related probabilities. We introduced the concept of conditional probability and discussed how to calculate posterior probabilities from prior probabilities using Bayes' theorem. We then discussed both discrete and continuous random variables as well as some of the more common probability distributions related to these types of random variables. These probability distributions included the custom discrete, discrete uniform, binomial, and Poisson probability distributions for discrete random variables, as well as the uniform, triangular, normal, and exponential probability distributions for continuous random variables. We also discussed the concepts of the expected value (mean) and variance of a random variable.

Probability is used in many chapters that follow in this textbook. The normal distribution is essential to many of the predictive modeling techniques that we introduce in later chapters. Random variables and probability distributions will be seen again in Chapter 6 when we discuss the use of statistical inference to draw conclusions about a population from sample data, Chapter 7 when we discuss regression analysis as a way of estimating relationships between variables, and Chapter 11 when we discuss simulation as a means of modeling uncertainty. Conditional probability and Bayes' theorem will be discussed in Chapter 15 in the context of decision analysis. It is very important to have a basic understanding of probability, such as is provided in this chapter, as you continue to improve your skills in business analytics.

GLOSSARY

Addition law A probability law used to compute the probability of the union of events. For two events A and B, the addition law is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. For two mutually exclusive events, $P(A \cap B) = 0$, so $P(A \cup B) = P(A) + P(B)$.

Bayes' theorem A method used to compute posterior probabilities.

Binomial probability distribution A probability distribution for a discrete random variable showing the probability of x successes in n trials.

Complement of *A* The event consisting of all outcomes that are not in A.

Conditional probability The probability of an event given that another event has already occurred. The conditional probability of A given B is $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$.

Continuous random variable A random variable that may assume any numerical value in an interval or collection of intervals. An interval can include negative and positive infinity.

Custom discrete probability distribution A probability distribution for a discrete random variable for which each value x_i that the random variable assumes is associated with a defined probability $f(x_i)$.

Discrete random variable A random variable that can take on only specified discrete values.

Discrete uniform probability distribution A probability distribution in which each possible value of the discrete random variable has the same probability.

Empirical probability distribution A probability distribution for which the relative frequency method is used to assign probabilities.

Event A collection of outcomes.

Expected value A measure of the central location, or mean, of a random variable.

Exponential probability distribution A continuous probability distribution that is useful in computing probabilities for the time it takes to complete a task or the time between arrivals. The mean and standard deviation for an exponential probability distribution are equal to each other.

Independent events Two events A and B are independent if $P(A \mid B) = P(A)$ or $P(B \mid A) = P(B)$; the events do not influence each other.

Intersection of A and B The event containing the outcomes belonging to both A and B. The intersection of A and B is denoted $A \cap B$.

Joint probabilities The probability of two events both occurring; in other words, the probability of the intersection of two events.

Marginal probabilities The values in the margins of a joint probability table that provide the probabilities of each event separately.

Multiplication law A law used to compute the probability of the intersection of events. For two events A and B, the multiplication law is $P(A \cap B) = P(B)P(A \mid B)$ or $P(A \cap B) = P(A)P(B \mid A)$. For two independent events, it reduces to $P(A \cap B) = P(A)P(B)$.

Mutually exclusive events Events that have no outcomes in common; $A \cap B$ is empty and $P(A \cap B) = 0$.

Normal probability distribution A continuous probability distribution in which the probability density function is bell shaped and determined by its mean μ and standard deviation σ.

Poisson probability distribution A probability distribution for a discrete random variable showing the probability of x occurrences of an event over a specified interval of time or space.

Posterior probabilities Revised probabilities of events based on additional information.

Prior probability Initial estimate of the probabilities of events.

Probability A numerical measure of the likelihood that an event will occur.

Probability density function A function used to compute probabilities for a continuous random variable. The area under the graph of a probability density function over an interval represents probability.

Probability distribution A description of how probabilities are distributed over the values of a random variable.

Probability mass function A function, denoted by $f(x)$, that provides the probability that x assumes a particular value for a discrete random variable.

Probability of an event Equal to the sum of the probabilities of outcomes for the event.

Random experiment A process that generates well-defined experimental outcomes. On any single repetition or trial, the outcome that occurs is determined by chance.

Random variables A numerical description of the outcome of an experiment.

Sample space The set of all outcomes.

Standard deviation Positive square root of the variance.

Triangular probability distribution A continuous probability distribution in which the probability density function is shaped like a triangle defined by the minimum possible value a, the maximum possible value b, and the most likely value m. A triangular probability distribution is often used when only subjective estimates are available for the minimum, maximum, and most likely values.

Uniform probability distribution A continuous probability distribution for which the probability that the random variable will assume a value in any interval is the same for each interval of equal length.

Union of A and B The event containing the outcomes belonging to A or B or both. The union of A and B is denoted by $A \cup B$.

Variance A measure of the variability, or dispersion, of a random variable.

Venn diagram A graphical representation of the sample space and operations involving events, in which the sample space is represented by a rectangle and events are represented as circles within the sample space.

PROBLEMS

1. On-time arrivals, lost baggage, and customer complaints are three measures that are typically used to measure the quality of service being offered by airlines. Suppose that the following values represent the on-time arrival percentage, amount of lost baggage, and customer complaints for 10 U.S. airlines.

Airline	On-Time Arrivals (%)	Mishandled Baggage per 1,000 Passengers	Customer Complaints per 1,000 Passengers
Virgin America	83.5	0.87	1.50
JetBlue	79.1	1.88	0.79
AirTran Airways	87.1	1.58	0.91
Delta Air Lines	86.5	2.10	0.73
Alaska Airlines	87.5	2.93	0.51
Frontier Airlines	77.9	2.22	1.05
Southwest Airlines	83.1	3.08	0.25
US Airways	85.9	2.14	1.74
American Airlines	76.9	2.92	1.80
United Airlines	77.4	3.87	4.24

   a. Based on the data above, if you randomly choose a Delta Air Lines flight, what is the probability that this individual flight will have an on-time arrival?
   b. If you randomly choose 1 of the 10 airlines for a follow-up study on airline quality ratings, what is the probability that you will choose an airline with less than two mishandled baggage reports per 1,000 passengers?
   c. If you randomly choose 1 of the 10 airlines for a follow-up study on airline quality ratings, what is the probability that you will choose an airline with more than one customer complaint per 1,000 passengers?
   d. What is the probability that a randomly selected AirTran Airways flight will not arrive on time?

2. Consider the random experiment of rolling a pair of dice. Suppose that we are interested in the sum of the face values showing on the dice.
   a. How many outcomes are possible?
   b. List the outcomes.
   c. What is the probability of obtaining a value of 7?
   d. What is the probability of obtaining a value of 9 or greater?

3. Suppose that for a recent admissions class, an Ivy League college received 2,851 applications for early admission. Of this group, it admitted 1,033 students early, rejected 854 outright, and deferred 964 to the regular admission pool for further consideration. In the past, this school has admitted 18% of the deferred early admission applicants during the regular admission process. Counting the students admitted early and the students admitted during the regular admission process, the total class size was 2,375. Let E, R, and D represent the events that a student who applies for early admission is admitted early, rejected outright, or deferred to the regular admissions pool.
   a. Use the data to estimate $P(E)$, $P(R)$, and $P(D)$.
   b. Are events E and D mutually exclusive? Find $P(E \cap D)$.
   c. For the 2,375 students who were admitted, what is the probability that a randomly selected student was accepted during early admission?
   d. Suppose a student applies for early admission. What is the probability that the student will be admitted for early admission or be deferred and later admitted during the regular admission process?

4. Suppose that we have two events, A and B, with $P(A) = 0.50$, $P(B) = 0.60$, and $P(A \cap B) = 0.40$.
   a. Find $P(A \mid B)$.
   b. Find $P(B \mid A)$.
   c. Are A and B independent? Why or why not?
5. Students taking the Graduate Management Admissions Test (GMAT) were asked about their undergraduate major and intent to pursue their MBA as a full-time or part-time student. A summary of their responses is as follows:

		Undergraduate Major			
		Business	Engineering	Other	Totals
Intended Enrollment Status	Full-Time	352	197	251	800
	Part-Time	150	161	194	505
	Totals	502	358	445	1,305

   a. Develop a joint probability table for these data.
   b. Use the marginal probabilities of undergraduate major (business, engineering, or other) to comment on which undergraduate major produces the most potential MBA students.
   c. If a student intends to attend classes full time in pursuit of an MBA degree, what is the probability that the student was an undergraduate engineering major?
   d. If a student was an undergraduate business major, what is the probability that the student intends to attend classes full time in pursuit of an MBA degree?
   e. Let F denote the event that the student intends to attend classes full time in pursuit of an MBA degree, and let B denote the event that the student was an undergraduate business major. Are events F and B independent? Justify your answer.
6. More than 40 million Americans are estimated to have at least one outstanding student loan to help pay college expenses ("40 Million Americans Now Have Student Loan Debt," *CNNMoney,* September 2014). Not all of these graduates pay back their debt in satisfactory fashion. Suppose that the following joint probability table shows the probabilities of student loan status and whether or not the student had received a college degree.

		College Degree		
		Yes	No	
Loan Status	Satisfactory	0.26	0.24	0.50
	Delinquent	0.16	0.34	0.50
		0.42	0.58	

   a. What is the probability that a student with a student loan had received a college degree?
   b. What is the probability that a student with a student loan had not received a college degree?
   c. Given that the student has received a college degree, what is the probability that the student has a delinquent loan?
   d. Given that the student has not received a college degree, what is the probability that the student has a delinquent loan?
   e. What is the impact of dropping out of college without a degree for students who have a student loan?

DATA file
OptilyticsLLC

7. The Human Resources Manager for Optilytics LLC is evaluating applications for the position of Senior Data Scientist. The file *OptilyticsLLC* presents summary data of the applicants for the position.
   a. Use a PivotTable in Excel to create a joint probability table showing the probabilities associated with a randomly selected applicant's sex and highest degree achieved. Use this joint probability table to answer the questions below.
   b. What are the marginal probabilities? What do they tell you about the probabilities associated with the sex of applicants and highest degree completed by applicants?
   c. If the applicant is female, what is the probability that the highest degree completed by the applicant is a PhD?
   d. If the highest degree completed by the applicant is a bachelor's degree, what is the probability that the applicant is male?
   e. What is the probability that a randomly selected applicant will be a male whose highest completed degree is a PhD?

8. The U.S. Census Bureau is a leading source of quantitative data related to the people and economy of the United States. The crosstabulation below represents the number of households (thousands) and the household income by the highest level of education for the head of household (U.S. Census Bureau web site, 2013). Use this crosstabulation to answer the following questions.

	Household Income				
Highest Level of Education	Under $25,000	$25,000 to $49,999	$50,000 to $99,999	$100,000 and Over	Total
High school graduate	9,880	9,970	9,441	3,482	32,773
Bachelor's degree	2,484	4,164	7,666	7,817	22,131
Master's degree	685	1,205	3,019	4,094	9,003
Doctoral degree	79	160	422	1,076	1,737
Total	13,128	15,499	20,548	16,469	65,644

   a. Develop a joint probability table.
   b. What is the probability the head of one of these households has a master's degree or higher education?
   c. What is the probability a household is headed by someone with a high school diploma earning $100,000 or more?
   d. What is the probability one of these households has an income below $25,000?
   e. What is the probability a household is headed by someone with a bachelor's degree earning less than $25,000?
   f. Are household income and educational level independent?

9. Cooper Realty is a small real estate company located in Albany, New York, that specializes primarily in residential listings. The company recently became interested in determining the likelihood of one of its listings being sold within a certain number of days. An analysis of company sales of 800 homes in previous years produced the following data.

		Days Listed Until Sold			
		Under 30	31–90	Over 90	Total
Initial Asking Price	Under $150,000	50	40	10	100
	$150,000–$199,999	20	150	80	250
	$200,000–$250,000	20	280	100	400
	Over $250,000	10	30	10	50
	Total	100	500	200	800

a. If A is defined as the event that a home is listed for more than 90 days before being sold, estimate the probability of A.
b. If B is defined as the event that the initial asking price is under \$150,000, estimate the probability of B.
c. What is the probability of $A \cap B$?
d. Assuming that a contract was just signed to list a home with an initial asking price of less than \$150,000, what is the probability that the home will take Cooper Realty more than 90 days to sell?
e. Are events A and B independent?

10. The prior probabilities for events A_1 and A_2 are $P(A_1) = 0.40$ and $P(A_2) = 0.60$. It is also known that $P(A_1 \cap A_2) = 0$. Suppose $P(B \mid A_1) = 0.20$ and $P(B \mid A_2) = 0.05$.
a. Are A_1 and A_2 mutually exclusive? Explain.
b. Compute $P(A_1 \cap B)$ and $P(A_2 \cap B)$.
c. Compute $P(B)$.
d. Apply Bayes' theorem to compute $P(A_1 \mid B)$ and $P(A_2 \mid B)$.

11. A local bank reviewed its credit-card policy with the intention of recalling some of its credit cards. In the past, approximately 5% of cardholders defaulted, leaving the bank unable to collect the outstanding balance. Hence, management established a prior probability of 0.05 that any particular cardholder will default. The bank also found that the probability of missing a monthly payment is 0.20 for customers who do not default. Of course, the probability of missing a monthly payment for those who default is 1.
a. Given that a customer missed a monthly payment, compute the posterior probability that the customer will default.
b. The bank would like to recall its credit card if the probability that a customer will default is greater than 0.20. Should the bank recall its credit card if the customer misses a monthly payment? Why or why not?

12. RunningWithTheDevil.com created a web site to market running shoes and other running apparel. Management would like a special pop-up offer to appear for female website visitors and a different special pop-up offer to appear for male web site visitors. From a sample of past web site visitors, RunningWithTheDevil's management learns that 60% of the visitors are male and 40% are female.
a. What is the probability that a current visitor to the web site is female?
b. Suppose that 30% of RunningWithTheDevil's female visitors previously visited LetsRun.com and 10% of male customers previously visited LetsRun.com. If the current visitor to RunningWithTheDevil's web site previously visited LetsRun.com, what is the revised probability that the current visitor is female? Should the RunningWithTheDevil's web site display the special offer that appeals to female visitors or the special offer that appeals to male visitors?

13. An oil company purchased an option on land in Alaska. Preliminary geologic studies assigned the following prior probabilities.

$$P(\text{high-quality oil}) = 0.50$$
$$P(\text{medium-quality oil}) = 0.20$$
$$P(\text{no oil}) = 0.30$$

a. What is the probability of finding oil?
b. After 200 feet of drilling on the first well, a soil test is taken. The probabilities of finding the particular type of soil identified by the test are as follows.

$$P(\text{soil} \mid \text{high-quality oil}) = 0.20$$
$$P(\text{soil} \mid \text{medium-quality oil}) = 0.80$$
$$P(\text{soil} \mid \text{no oil}) = 0.20$$

c. How should the firm interpret the soil test? What are the revised probabilities, and what is the new probability of finding oil?

14. Suppose the following data represent the number of persons unemployed for a given number of months in Killeen, Texas. The values in the first column show the number of months unemployed and the values in the second column show the corresponding number of unemployed persons.

Months Unemployed	Number Unemployed
1	1,029
2	1,686
3	2,269
4	2,675
5	3,487
6	4,652
7	4,145
8	3,587
9	2,325
10	1,120

Let x be a random variable indicating the number of months a randomly selected person is unemployed.

a. Use the data to develop an empirical discrete probability distribution for x.
b. Show that your probability distribution satisfies the conditions for a valid discrete probability distribution.
c. What is the probability that a person is unemployed for two months or less? Unemployed for more than two months?
d. What is the probability that a person is unemployed for more than six months?

15. The percent frequency distributions of job satisfaction scores for a sample of information systems (IS) senior executives and middle managers are as follows. The scores range from a low of 1 (very dissatisfied) to a high of 5 (very satisfied).

Job Satisfaction Score	IS Senior Executives (%)	IS Middle Managers (%)
1	5	4
2	9	10
3	3	12
4	42	46
5	41	28

a. Develop a probability distribution for the job satisfaction score of a randomly selected senior executive.
b. Develop a probability distribution for the job satisfaction score of a randomly selected middle manager.
c. What is the probability that a randomly selected senior executive will report a job satisfaction score of 4 or 5?
d. What is the probability that a randomly selected middle manager is very satisfied?
e. Compare the overall job satisfaction of senior executives and middle managers.

16. The following table provides a probability distribution for the random variable y.

y	$f(y)$
2	0.20
4	0.30
7	0.40
8	0.10

a. Compute $E(y)$.
b. Compute Var(y) and σ.

17. The probability distribution for damage claims paid by the Newton Automobile Insurance Company on collision insurance is as follows.

Payment ($)	Probability
0	0.85
500	0.04
1,000	0.04
3,000	0.03
5,000	0.02
8,000	0.01
10,000	0.01

a. Use the expected collision payment to determine the collision insurance premium that would enable the company to break even.
b. The insurance company charges an annual rate of $520 for the collision coverage. What is the expected value of the collision policy for a policyholder? (*Hint:* It is the expected payments from the company minus the cost of coverage.) Why does the policyholder purchase a collision policy with this expected value?

18. The J.R. Ryland Computer Company is considering a plant expansion to enable the company to begin production of a new computer product. The company's president must determine whether to make the expansion a medium- or large-scale project. Demand for the new product is uncertain, which for planning purposes may be low demand, medium demand, or high demand. The probability estimates for demand are 0.20, 0.50, and 0.30, respectively. Letting x and y indicate the annual profit in thousands of dollars, the firm's planners developed the following profit forecasts for the medium- and large-scale expansion projects.

		Medium-Scale Expansion Profit		Large-Scale Expansion Profit	
		x	$f(x)$	y	$f(y)$
Demand	Low	50	0.20	0	0.20
	Medium	150	0.50	100	0.50
	High	200	0.30	300	0.30

a. Compute the expected value for the profit associated with the two expansion alternatives. Which decision is preferred for the objective of maximizing the expected profit?
b. Compute the variance for the profit associated with the two expansion alternatives. Which decision is preferred for the objective of minimizing the risk or uncertainty?

19. Consider a binomial experiment with $n = 10$ and $p = 0.10$.
    a. Compute $f(0)$.
    b. Compute $f(2)$.
    c. Compute $P(x \leq 2)$.
    d. Compute $P(x \geq 1)$.
    e. Compute $E(x)$.
    f. Compute $\text{Var}(x)$ and σ.
20. Many companies use a quality control technique called acceptance sampling to monitor incoming shipments of parts, raw materials, and so on. In the electronics industry, component parts are commonly shipped from suppliers in large lots. Inspection of a sample of n components can be viewed as the n trials of a binomial experiment. The outcome for each component tested (trial) will be that the component is classified as good or defective. Reynolds Electronics accepts a lot from a particular supplier if the defective components in the lot do not exceed 1%. Suppose a random sample of five items from a recent shipment is tested.
    a. Assume that 1% of the shipment is defective. Compute the probability that no items in the sample are defective.
    b. Assume that 1% of the shipment is defective. Compute the probability that exactly one item in the sample is defective.
    c. What is the probability of observing one or more defective items in the sample if 1% of the shipment is defective?
    d. Would you feel comfortable accepting the shipment if one item was found to be defective? Why or why not?
21. A university found that 20% of its students withdraw without completing the introductory statistics course. Assume that 20 students registered for the course.
    a. Compute the probability that two or fewer will withdraw.
    b. Compute the probability that exactly four will withdraw.
    c. Compute the probability that more than three will withdraw.
    d. Compute the expected number of withdrawals.
22. Consider a Poisson distribution with $\mu = 3$.
    a. Write the appropriate Poisson probability mass function.
    b. Compute $f(2)$.
    c. Compute $f(1)$.
    d. Compute $P(x \geq 2)$.
23. Emergency 911 calls to a small municipality in Idaho come in at the rate of one every 2 minutes. Assume that the number of 911 calls is a random variable that can be described by the Poisson distribution.
    a. What is the expected number of 911 calls in 1 hour?
    b. What is the probability of three 911 calls in 5 minutes?
    c. What is the probability of no 911 calls during a 5-minute period?
24. A regional director responsible for business development in the state of Pennsylvania is concerned about the number of small business failures. If the mean number of small business failures per month is 10, what is the probability that exactly 4 small businesses will fail during a given month? Assume that the probability of a failure is the same for any two months and that the occurrence or nonoccurrence of a failure in any month is independent of failures in any other month.
25. The random variable x is known to be uniformly distributed between 10 and 20.
    a. Show the graph of the probability density function.
    b. Compute $P(x < 15)$.
    c. Compute $P(12 \leq x \leq 18)$.
    d. Compute $E(x)$.
    e. Compute $\text{Var}(x)$.

26. Most computer languages include a function that can be used to generate random numbers. In Excel, the RAND function can be used to generate random numbers between 0 and 1. If we let x denote a random number generated using RAND, then x is a continuous random variable with the following probability density function:

$$f(x) = \begin{cases} 1 & \text{for } 0 \le x \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

    a. Graph the probability density function.
    b. What is the probability of generating a random number between 0.25 and 0.75?
    c. What is the probability of generating a random number with a value less than or equal to 0.30?
    d. What is the probability of generating a random number with a value greater than 0.60?
    e. Generate 50 random numbers by entering =RAND() into 50 cells of an Excel worksheet.
    f. Compute the mean and standard deviation for the random numbers in part (e).

27. Suppose we are interested in bidding on a piece of land and we know one other bidder is interested. The seller announced that the highest bid in excess of $10,000 will be accepted. Assume that the competitor's bid x is a random variable that is uniformly distributed between $10,000 and $15,000.
    a. Suppose you bid $12,000. What is the probability that your bid will be accepted?
    b. Suppose you bid $14,000. What is the probability that your bid will be accepted?
    c. What amount should you bid to maximize the probability that you get the property?
    d. Suppose you know someone who is willing to pay you $16,000 for the property. Would you consider bidding less than the amount in part (c)? Why or why not?

28. A random variable has a triangular probability density function with $a = 50$, $b = 375$, and $m = 250$.
    a. Sketch the probability distribution function for this random variable. Label the points $a = 50$, $b = 375$, and $m = 250$ on the x-axis.
    b. What is the probability that the random variable will assume a value between 50 and 250?
    c. What is the probability that the random variable will assume a value greater than 300?

29. The Siler Construction Company is about to bid on a new industrial construction project. To formulate their bid, the company needs to estimate the time required for the project. Based on past experience, management expects that the project will require at least 24 months, and could take as long as 48 months if there are complications. The most likely scenario is that the project will require 30 months.
    a. Assume that the actual time for the project can be approximated using a triangular probability distribution. What is the probability that the project will take less than 30 months?
    b. What is the probability that the project will take between 28 and 32 months?
    c. To submit a competitive bid, the company believes that if the project takes more than 36 months, then the company will lose money on the project. Management does not want to bid on the project if there is greater than a 25% chance that they will lose money on this project. Should the company bid on this project?

30. Suppose that the return for a particular large-cap stock fund is normally distributed with a mean of 14.4% and standard deviation of 4.4%.
    a. What is the probability that the large-cap stock fund has a return of at least 20%?
    b. What is the probability that the large-cap stock fund has a return of 10% or less?

31. A person must score in the upper 2% of the population on an IQ test to qualify for membership in Mensa, the international high IQ society. If IQ scores are normally distributed with a mean of 100 and a standard deviation of 15, what score must a person have to qualify for Mensa?
32. Assume that the traffic to the web site of Smiley's People, Inc., which sells customized T-shirts, follows a normal distribution, with a mean of 4.5 million visitors per day and a standard deviation of 820,000 visitors per day.
    a. What is the probability that the web site has fewer than 5 million visitors in a single day?
    b. What is the probability that the web site has 3 million or more visitors in a single day?
    c. What is the probability that the web site has between 3 million and 4 million visitors in a single day?
    d. Assume that 85% of the time, the Smiley's People web servers can handle the daily web traffic volume without purchasing additional server capacity. What is the amount of web traffic that will require Smiley's People to purchase additional server capacity?
33. Suppose that Motorola uses the normal distribution to determine the probability of defects and the number of defects in a particular production process. Assume that the production process manufactures items with a mean weight of 10 ounces. Calculate the probability of a defect and the suspected number of defects for a 1,000-unit production run in the following situations.
    a. The process standard deviation is 0.15, and the process control is set at plus or minus one standard deviation. Units with weights less than 9.85 or greater than 10.15 ounces will be classified as defects.
    b. Through process design improvements, the process standard deviation can be reduced to 0.05. Assume that the process control remains the same, with weights less than 9.85 or greater than 10.15 ounces being classified as defects.
    c. What is the advantage of reducing process variation, thereby causing process control limits to be at a greater number of standard deviations from the mean?
34. Consider the following exponential probability density function:

$$f(x) = \frac{1}{3}e^{-x/3} \quad \text{for } x \geq 0$$

    a. Write the formula for $P(x \leq x_0)$.
    b. Find $P(x \leq 2)$.
    c. Find $P(x \geq 3)$.
    d. Find $P(x \leq 5)$.
    e. Find $P(2 \leq x \leq 5)$.
35. The time between arrivals of vehicles at a particular intersection follows an exponential probability distribution with a mean of 12 seconds.
    a. Sketch this exponential probability distribution.
    b. What is the probability that the arrival time between vehicles is 12 seconds or less?
    c. What is the probability that the arrival time between vehicles is 6 seconds or less?
    d. What is the probability of 30 or more seconds between vehicle arrivals?
36. Suppose that the time spent by players in a single session on the *World of Warcraft* multiplayer online role-playing game follows an exponential distribution with a mean of 38.3 minutes.
    a. Write the exponential probability distribution function for the time spent by players on a single session of *World of Warcraft*.
    b. What is the probability that a player will spend between 20 and 40 minutes on a single session of *World of Warcraft*?
    c. What is the probability that a player will spend more than 1 hour on a single session of *World of Warcraft*?

CASE PROBLEM: HAMILTON COUNTY JUDGES

Hamilton County judges try thousands of cases per year. In an overwhelming majority of the cases disposed, the verdict stands as rendered. However, some cases are appealed, and of those appealed, some of the cases are reversed. Kristen DelGuzzi of the *Cincinnati Enquirer* newspaper conducted a study of cases handled by Hamilton County judges over a three-year period. Shown in the table below are the results for 182,908 cases handled (disposed) by 38 judges in Common Pleas Court, Domestic Relations Court, and Municipal Court. Two of the judges (Dinkelacker and Hogan) did not serve in the same court for the entire three-year period.

The purpose of the newspaper's study was to evaluate the performance of the judges. Appeals are often the result of mistakes made by judges, and the newspaper wanted to know which judges were doing a good job and which were making too many mistakes. You are called in to assist in the data analysis. Use your knowledge of probability and conditional probability to help with the ranking of the judges. You also may be able to analyze the likelihood of appeal and reversal for cases handled by different courts.

Total Cases Disposed, Appealed, and Reversed in Hamilton County Courts

Common Pleas Court

Judge	Total Cases Disposed	Appealed Cases	Reversed Cases
Fred Cartolano	3,037	137	12
Thomas Crush	3,372	119	10
Patrick Dinkelacker	1,258	44	8
Timothy Hogan	1,954	60	7
Robert Kraft	3,138	127	7
William Mathews	2,264	91	18
William Morrissey	3,032	121	22
Norbert Nadel	2,959	131	20
Arthur Ney, Jr.	3,219	125	14
Richard Niehaus	3,353	137	16
Thomas Nurre	3,000	121	6
John O'Connor	2,969	129	12
Robert Ruehlman	3,205	145	18
J. Howard Sundermann	955	60	10
Ann Marie Tracey	3,141	127	13
Ralph Winkler	3,089	88	6
Total	43,945	1,762	199

Domestic Relations Court

Judge	Total Cases Disposed	Appealed Cases	Reversed Cases
Penelope Cunningham	2,729	7	1
Patrick Dinkelacker	6,001	19	4
Deborah Gaines	8,799	48	9
Ronald Panioto	12,970	32	3
Total	30,499	106	17

Municipal Court

Judge	Total Cases Disposed	Appealed Cases	Reversed Cases
Mike Allen	6,149	43	4
Nadine Allen	7,812	34	6
Timothy Black	7,954	41	6
David Davis	7,736	43	5
Leslie Isaiah Gaines	5,282	35	13
Karla Grady	5,253	6	0
Deidra Hair	2,532	5	0
Dennis Helmick	7,900	29	5
Timothy Hogan	2,308	13	2
James Patrick Kenney	2,798	6	1
Joseph Luebbers	4,698	25	8
William Mallory	8,277	38	9
Melba Marsh	8,219	34	7
Beth Mattingly	2,971	13	1
Albert Mestemaker	4,975	28	9
Mark Painter	2,239	7	3
Jack Rosen	7,790	41	13
Mark Schweikert	5,403	33	6
David Stockdale	5,371	22	4
John A. West	2,797	4	2
Total	108,464	500	104

Managerial Report

Prepare a report with your rankings of the judges. Also, include an analysis of the likelihood of appeal and case reversal in the three courts. At a minimum, your report should include the following:

1. The probability of cases being appealed and reversed in the three different courts.
2. The probability of a case being appealed for each judge.
3. The probability of a case being reversed for each judge.
4. The probability of reversal given an appeal for each judge.
5. Rank the judges within each court. State the criteria you used and provide a rationale for your choice.