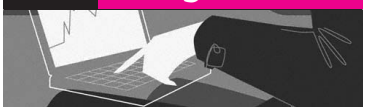
3 Random Variables

LEARNING OBJECTIVES

After studying this chapter, you should be able to:

- Distinguish between discrete and continuous random variables.
- Explain how a random variable is characterized by its probability distribution.
- Compute statistics about a random variable.
- Compute statistics about a function of a random variable.
- Compute statistics about the sum of a linear composite of random variables.
- Identify which type of distribution a given random variable is most likely to follow.
- Solve problems involving standard distributions manually using formulas.
- Solve business problems involving standard distributions using spreadsheet templates.

3–1 Using Statistics



Recent work in genetics makes assumptions about the distribution of babies of the two sexes. One such analysis concentrated on the probabilities of the number of babies of each sex in a given number of births. Consider the sample space made up of the 16 equally likely points:

BBBB	BBBG	BGGB	GBGG
GBBB	GGBB	BGBG	GGBG
BGBB	GBGB	BBGG	GGGB
BBGB	GBBG	BGGG	GGGG

All these 16 points are equally likely because when four children are born, the sex of each child is assumed to be independent of those of the other three. Hence the probability of each quadruple (e.g., GBBG) is equal to the product of the probabilities of the four separate, single outcomes—G, B, B, and G—and is thus equal to $(1/2)(1/2)\ (1/2)(1/2) = 1/16$.

Now, let's look at the variable "the number of girls out of four births." This number *varies* among points in the sample space, and it is *random*—given to chance. That's why we call such a number a **random variable.**

A **random variable** is an uncertain quantity whose value depends on chance.

A random variable has a probability law—a rule that assigns probabilities to the different values of the random variable. The probability law, the probability assignment, is called the **probability distribution** of the random variable. We usually denote the random variable by a capital letter, often X. The probability distribution will then be denoted by $P(X)$.

Look again at the sample space for the sexes of four babies, and remember that our variable is the number of girls out of four births. The first point in the sample space is BBBB; because the number of girls is zero here, $X = 0$. The next four points in the sample space all have one girl (and three boys). Hence, each one leads to the value $X = 1$. Similarly, the next six points in the sample space all lead to $X = 2$; the next four points to $X = 3$; and, finally, the last point in our sample space gives $X = 4$. The correspondence of points in the sample space with values of the random variable is as follows:

Sample Space	Random Variable
BBBB	$X = 0$
GBBB, BGBB, BBGB, BBBG	$X = 1$
GGBB, GBGB, GBBG, BGGB, BGBG, BBGG	$X = 2$
BGGG, GBGG, GGBG, GGGB	$X = 3$
GGGG	$X = 4$

This correspondence, when a sample space clearly exists, allows us to define a random variable as follows:

A **random variable** is a function of the sample space.

What is this function? The correspondence between points in the sample space and values of the random variable allows us to determine the probability distribution of X as follows: Notice that 1 of the 16 equally likely points of the sample space leads to $X = 0$. Hence, the probability that $X = 0$ is $1/16$. Because 4 of the 16 equally likely points lead to a value $X = 1$, the probability that $X = 1$ is $4/16$, and so forth. Thus, looking at the sample space and counting the number of points leading to each value of X, we find the following probabilities:

$$P(X = 0) = 1/16 = 0.0625$$
$$P(X = 1) = 4/16 = 0.2500$$
$$P(X = 2) = 6/16 = 0.3750$$
$$P(X = 3) = 4/16 = 0.2500$$
$$P(X = 4) = 1/16 = 0.0625$$

The probability statements above constitute the probability distribution of the random variable X = the number of girls in four births. Notice how this probability law was obtained simply by associating values of X with sets in the sample space. (For example, the set GBBB, BGBB, BBGB, BBBG leads to $X = 1$.) Writing the probability distribution of X in a table format is useful, but first let's make a small, simplifying notational distinction so that we do not have to write complete probability statements such as $P(X = 1)$.

As stated earlier, we use a capital letter, such as X, to denote the random variable. But we use a lowercase letter to denote a particular value that the random variable can take. For example, $x = 3$ means that some particular set of four births resulted in three girls. Think of X as random and x as known. Before a coin is tossed, the number of heads (in one toss) is an unknown, X. Once the coin lands, we have $x = 0$ or $x = 1$.

Now let's return to the number of girls in four births. We can write the probability distribution of this random variable in a table format, as shown in Table 3–1.

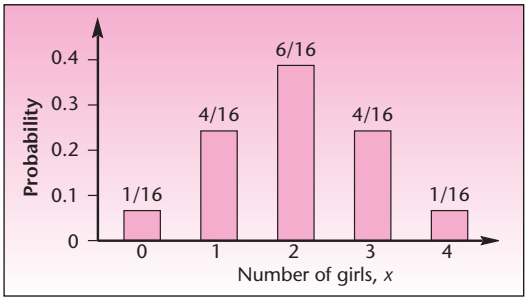
Note an important fact: The sum of the probabilities of all the values of the random variable X must be 1.00. A picture of the probability distribution of the random variable X is given in Figure 3–1. Such a picture is a **probability bar chart** for the random variable.

Marilyn is interested in the number of girls (or boys) in any fixed number of births, not necessarily four. Thus her discussion extends beyond this case. In fact, the

TABLE 3–1 Probability Distribution of the Number of Girls in Four Births

Number of Girls x	Probability $P(x)$
0	1/16
1	4/16
2	6/16
3	4/16
4	1/16
	16/16 = 1.00

FIGURE 3–1 Probability Bar Chart



random variable she describes, which in general counts the number of "successes" (here, a girl is a success) in a fixed number n of trials, is called a *binomial random variable.* We will study this particular important random variable in section 3–3.

EXAMPLE 3–1

Figure 3–2 shows the sample space for the experiment of rolling two dice. As can be seen from the sample space, the probability of every pair of outcomes is 1/36. This can be seen from the fact that, by the independence of the two dice, for example, $P(6 \text{ on red die} \cap 5 \text{ on green die}) = P(6 \text{ on red die}) \times P(5 \text{ on green die}) = (1/6)(1/6) = 1/36$, and that this holds for all 36 pairs of outcomes. Let $X =$ the sum of the dots on the two dice. What is the distribution of x?

Solution

Figure 3–3 shows the correspondence between sets in our sample space and the values of X. The probability distribution of X is given in Table 3–2. The probability distribution allows us to answer various questions about the random variable of interest. Draw a picture of this probability distribution. Such a graph need not be a histogram, used earlier, but can also be a bar graph or column chart of the probabilities of the different values of the random variable. Note from the graph you produced that the distribution of the random variable "the sum of two dice" is symmetric. The central value is $x = 7$, which has the highest probability, $P(7) = 6/36 = 1/6$. This is the *mode,*

FIGURE 3–2 Sample Space for Two Dice

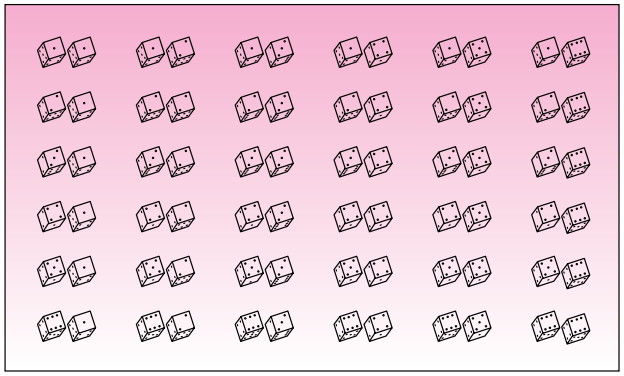


FIGURE 3–3 Correspondence between Sets and Values of *X*

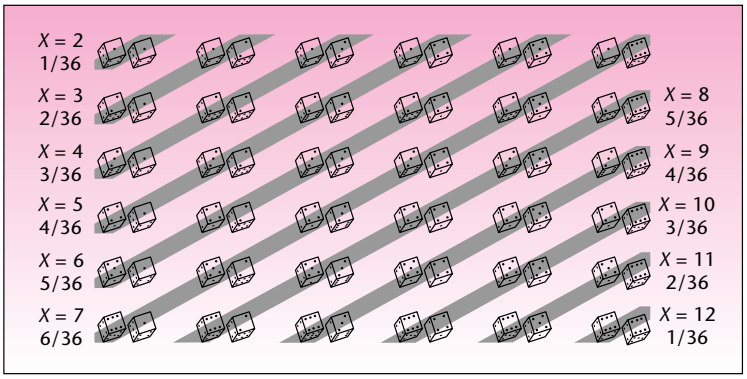


TABLE 3–2 Probability Distribution of the Sum of Two Dice

x	*P*(*x*)
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36
	36/36 = 1.00

the most likely value. Thus, if you were to bet on one sum of two dice, the best bet is that the sum will be 7.

We can answer other probability questions, such as: What is the probability that the sum will be at most 5? This is $P(X \leq 5)$. Notice that to answer this question, we require the sum of all the probabilities of the values that are less than or equal to 5:

$$P(2) + P(3) + P(4) + P(5) = 1/36 + 2/36 + 3/36 + 4/36 = 10/36$$

Similarly, we may want to know the probability that the sum is greater than 9. This is calculated as follows:

$$P(X > 9) = P(10) + P(11) + P(12) = 3/36 + 2/36 + 1/36 = 6/36 = 1/6$$

Most often, unless we are dealing with games of chance, there is no evident sample space. In such situations the probability distribution is often obtained from lists or other data that give us the relative frequency in the recorded past of the various values of the random variable. This is demonstrated in Example 3–2.

EXAMPLE 3–2

800, 900, and Now: the 500 Telephone Numbers

The new code 500 is for busy, affluent people who travel a lot: It can work with a cellular phone, your home phone, office phone, second-home phone, up to five additional phones besides your regular one. The computer technology behind this service is astounding—the new phone service can find you wherever you may be on the planet at a given moment (assuming one of the phones you specify is cellular and you keep it with you when you are not near one of your stationary telephones). What the computer does is to first ring you up at the telephone number you specify as your primary one (your office phone, for example). If there is no answer, the computer switches to search for you at your second-specified phone number (say, home); if you do not answer there, it will switch to your third phone (maybe the phone at a close companion's home, or your car phone, or a portable cellular phone); and so on up to five allowable switches. The switches are the expensive part of this service (besides arrangements to have your cellular phone reachable overseas), and the service provider wants to get information on these switches. From data available on an experimental run of the 500 program, the following probability distribution is constructed for the number of dialing switches that are necessary before a person is reached. When $X = 0$, the person was reached on her or his primary phone (no switching was necessary); when $X = 1$, a dialing switch was necessary, and the person was found at the secondary phone; and so on up to five switches. Table 3–3 gives the probability distribution for this random variable.

A plot of the probability distribution of this random variable is given in Figure 3–4. When more than two switches occur on a given call, extra costs are incurred. What is the probability that for a given call there would be extra costs?

TABLE 3–3
The Probability Distribution of the Number of Switches

x	$P(x)$
0	0.1
1	0.2
2	0.3
3	0.2
4	0.1
5	0.1
	1.00

Solution

$$P(X > 2) = P(3) + P(4) + P(5) = 0.2 + 0.1 + 0.1 = 0.4$$

What is the probability that at least one switch will occur on a given call? $1 - P(0) = 0.9$, a high probability.

Discrete and Continuous Random Variables

Refer to Example 3–2. Notice that when switches occur, the number X jumps by 1. It is impossible to have one-half a switch or 0.13278 of one. The same is true for the number of dots on two dice (you cannot see 2.3 dots or 5.87 dots) and, of course, the number of girls in four births.

A **discrete random variable** can assume at most a countable number of values.

The values of a discrete random variable do not have to be positive whole numbers; they just have to "jump" from one possible value to the next without being able to have any value in between. For example, the amount of money you make on an investment may be \$500, or it may be a loss: −\$200. At any rate, it can be measured at best to the nearest *cent,* so this variable is discrete.

What are continuous random variables, then?

A **continuous random variable** may take on any value in an interval of numbers (i.e., its possible values are uncountably infinite).

FIGURE 3–4
The Probability Distribution of the Number of Switches

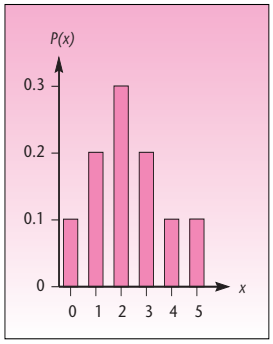
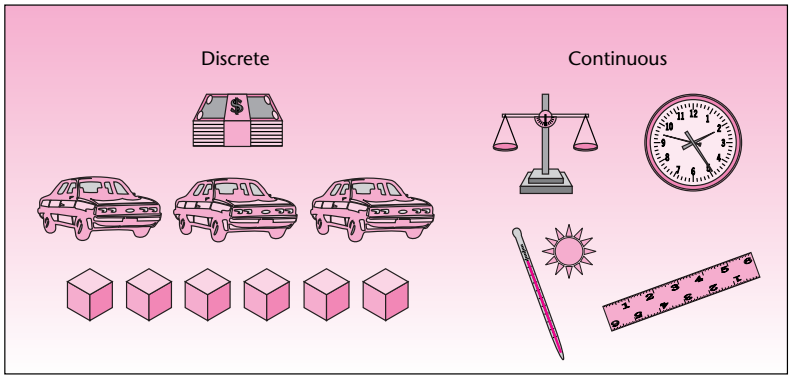


FIGURE 3–5 Discrete and Continuous Random Variables



The values of continuous random variables can be measured (at least in theory) to any degree of accuracy. They move continuously from one possible value to another, without having to jump. For example, temperature is a continuous random variable, since it can be measured as 72.00340981136 . . . °. Weight, height, and time are other examples of continuous random variables.

The difference between discrete and continuous random variables is illustrated in Figure 3–5. Is wind speed a discrete or a continuous random variable?

The probability distribution of a discrete random variable X must satisfy the following two conditions.

$$1.\ P(x) \geq 0 \text{ for all values } x \qquad (3\text{–}1)$$

$$2.\ \sum_{\text{all } x} P(x) = 1 \qquad (3\text{–}2)$$

These conditions must hold because the $P(x)$ values are probabilities. Equation 3–1 states that all probabilities must be greater than or equal to zero, as we know from Chapter 2. For the second rule, equation 3–2, note the following. For each value x, $P(x) = P(X = x)$ is the probability of the event that the random variable equals x. Since by definition *all x* means all the values the random variable X may take, and since X may take on only one value at a time, the occurrences of these values are mutually exclusive events, and one of them must take place. Therefore, the sum of all the probabilities $P(x)$ must be 1.00.

Cumulative Distribution Function

The probability distribution of a discrete random variable lists the probabilities of occurrence of different values of the random variable. We may be interested in *cumulative* probabilities of the random variable. That is, we may be interested in the probability that the value of the random variable is *at most* some value x. This is the sum of all the probabilities of the values i of X that are less than or equal to x.

TABLE 3–4 Cumulative Distribution Function of the Number of Switches (Example 3–2)

x	$P(x)$	$F(x)$
0	0.1	0.1
1	0.2	0.3
2	0.3	0.6
3	0.2	0.8
4	0.1	0.9
5	0.1	1.00
	1.00	

We define the *cumulative distribution function* (also called *cumulative probability function*) as follows.

The **cumulative distribution function,** $F(x)$, of a discrete random variable X is

$$F(x) = P(X \leq x) = \sum_{\text{all } i \leq x} P(i) \qquad (3\text{–}3)$$

Table 3–4 gives the cumulative distribution function of the random variable of Example 3–2. Note that each entry of $F(x)$ is equal to the sum of the corresponding values of $P(i)$ for all values i less than or equal to x. For example, $F(3) = P(X \leq 3) = P(0) + P(1) + P(2) + P(3) = 0.1 + 0.2 + 0.3 + 0.2 = 0.8$. Of course, $F(5) = 1.00$ because $F(5)$ is the sum of the probabilities of all values that are less than or equal to 5, and 5 is the largest value of the random variable.

Figure 3–6 shows $F(x)$ for the number of switches on a given call. All cumulative distribution functions are nondecreasing and equal 1.00 at the largest possible value of the random variable.

Let us consider a few probabilities. The probability that the number of switches will be less than or equal to 3 is given by $F(3) = 0.8$. This is illustrated, using the probability distribution, in Figure 3–7.

FIGURE 3–6 Cumulative Distribution Function of Number of Switches

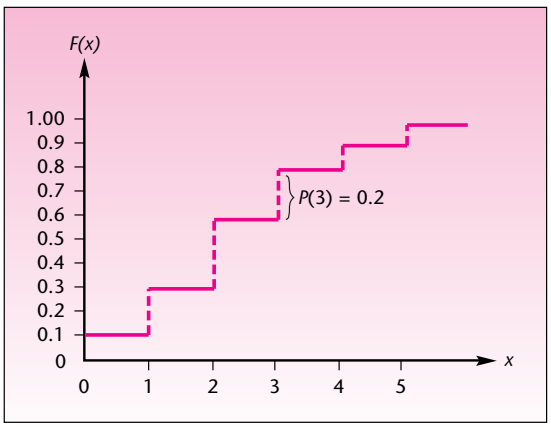


FIGURE 3–7 The Probability That at Most Three Switches Will Occur

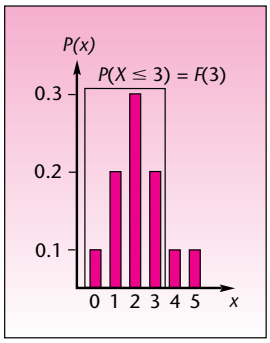
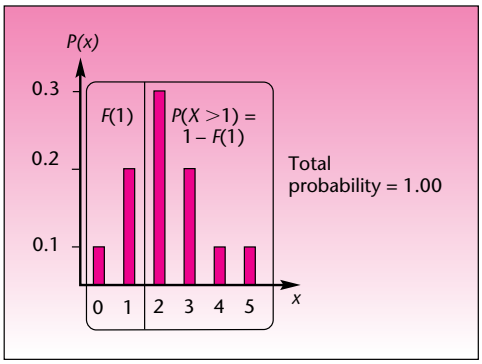


FIGURE 3–8 Probability That More than One Switch Will Occur

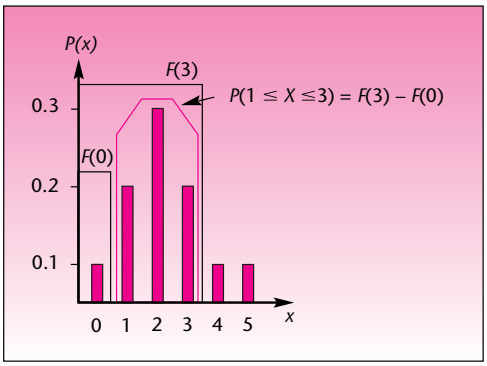


The probability that *more than* one switch will occur, $P(X > 1)$, is equal to $1 - F(1) = 1 - 0.3 = 0.7$. This is so because $F(1) = P(X \leq 1)$, and $P(X \leq 1) + P(X > 1) = 1$ (the two events are complements of each other). This is demonstrated in Figure 3–8.

The probability that anywhere from one to three switches will occur is $P(1 \leq X \leq 3)$. From Figure 3–9 we see that this is equal to $F(3) - F(0) = 0.8 - 0.1 = 0.7$. (This is the probability that the number of switches that occur will be less than or equal to 3 and greater than 0.) This, and other probability questions, could certainly be answered directly, without use of $F(x)$. We could just add the probabilities: $P(1) + P(2) + P(3) = 0.2 + 0.3 + 0.2 = 0.7$. The advantage of $F(x)$ is that probabilities may be computed by few operations [usually subtraction of two values of $F(x)$, as in this example], whereas use of $P(x)$ often requires lengthier computations.

If the probability distribution is available, use it directly. If, on the other hand, you have a cumulative distribution function for the random variable in question, you may use it as we have demonstrated. In either case, drawing a picture of the probability distribution is always helpful. You can look at the signs in the probability statement, such as $P(X \leq x)$ versus $P(X < x)$, to see which values to include and which ones to leave out of the probability computation.

FIGURE 3–9 Probability That Anywhere from One to Three Switches Will Occur



PROBLEMS

3–1. The number of telephone calls arriving at an exchange during any given minute between noon and 1:00 P.M. on a weekday is a random variable with the following probability distribution.

x	$P(x)$
0	0.3
1	0.2
2	0.2
3	0.1
4	0.1
5	0.1

a. Verify that $P(x)$ is a probability distribution.

b. Find the cumulative distribution function of the random variable.

c. Use the cumulative distribution function to find the probability that between 12:34 and 12:35 P.M. more than two calls will arrive at the exchange.

3–2. According to an article in *Travel and Leisure,* every person in a small study of sleep during vacation was found to sleep longer than average during the first vacation night.[1] Suppose that the number of additional hours slept in the first night of a vacation, over the person's average number slept per night, is given by the following probability distribution:

x	$P(x)$
0	0.01
1	0.09
2	0.30
3	0.20
4	0.20
5	0.10
6	0.10

a. Verify that $P(x)$ is a probability distribution.

b. Find the cumulative distribution function.

c. Find the probability that at most four additional hours are slept.

d. Find the probability that at least two additional hours are slept per night.

3–3. The percentage of people (to the nearest 10) responding to an advertisement is a random variable with the following probability distribution:

x(%)	$P(x)$
0	0.10
10	0.20
20	0.35
30	0.20
40	0.10
50	0.05

a. Show that $P(x)$ is a probability distribution.

b. Find the cumulative distribution function.

c. Find the probability that more than 20% will respond to the ad.

[1]Amy Farley, "Health and Fitness on the Road," *Travel and Leisure,* April 2007, p. 182.

3–4. An automobile dealership records the number of cars sold each day. The data are used in calculating the following probability distribution of daily sales:

x	$P(x)$
0	0.1
1	0.1
2	0.2
3	0.2
4	0.3
5	0.1

a. Find the probability that the number of cars sold tomorrow will be between two and four (both inclusive).

b. Find the cumulative distribution function of the number of cars sold per day.

c. Show that $P(x)$ is a probability distribution.

3–5. Consider the roll of a pair of dice, and let X denote the sum of the two numbers appearing on the dice. Find the probability distribution of X, and find the cumulative distribution function. What is the most likely sum?

3–6. The number of intercity shipment orders arriving daily at a transportation company is a random variable X with the following probability distribution:

x	$P(x)$
0	0.1
1	0.2
2	0.4
3	0.1
4	0.1
5	0.1

a. Verify that $P(x)$ is a probability distribution.

b. Find the cumulative probability function of X.

c. Use the cumulative probability function computed in (*b*) to find the probability that anywhere from one to four shipment orders will arrive on a given day.

d. When more than three orders arrive on a given day, the company incurs additional costs due to the need to hire extra drivers and loaders. What is the probability that extra costs will be incurred on a given day?

e. Assuming that the numbers of orders arriving on different days are independent of each other, what is the probability that no orders will be received over a period of five working days?

f. Again assuming independence of orders on different days, what is the probability that extra costs will be incurred two days in a row?

3–7. An article in *The New York Times* reports that several hedge fund managers now make more than a *billion dollars a year*.[2] Suppose that the annual income of a hedge

[2]Jenny Anderson and Julie Creswell, "Make Less Than $240 Million? You're Off Top Hedge Fund List," *The New York Times*, April 24, 2007, p. A1.

fund manager in the top tier, in millions of dollars a year, is given by the following probability distribution:

x ($ millions)	$P(x)$
$1,700	0.2
1,500	0.2
1,200	0.3
1,000	0.1
800	0.1
600	0.05
400	0.05

a. Find the probability that the annual income of a hedge fund manager will be between $400 million and $1 billion (both inclusive).

b. Find the cumulative distribution function of X.

c. Use $F(x)$ computed in (b) to evaluate the probability that the annual income of a hedge fund manager will be less than or equal to $1 billion.

d. Find the probability that the annual income of a hedge fund manager will be greater than $600 million and less than or equal to $1.5 billion.

3–8. The number of defects in a machine-made product is a random variable X with the following probability distribution:

x	$P(x)$
0	0.1
1	0.2
2	0.3
3	0.3
4	0.1

a. Show that $P(x)$ is a probability distribution.

b. Find the probability $P(1 < X \leq 3)$.

c. Find the probability $P(1 < X \leq 4)$.

d. Find $F(x)$.

3–9. Returns on investments overseas, especially in Europe and the Pacific Rim, are expected to be higher than those of U.S. markets in the near term, and analysts are now recommending investments in international portfolios. An investment consultant believes that the probability distribution of returns (in percent per year) on one such portfolio is as follows:

x(%)	$P(x)$
9	0.05
10	0.15
11	0.30
12	0.20
13	0.15
14	0.10
15	0.05

a. Verify that $P(x)$ is a probability distribution.

b. What is the probability that returns will be at least 12%?

c. Find the cumulative distribution of returns.

3–10. The daily exchange rate of one dollar in euros during the first three months of 2007 can be inferred to have the following distribution.[3]

x	$P(x)$
0.73	0.05
0.74	0.10
0.75	0.25
0.76	0.40
0.77	0.15
0.78	0.05

a. Show that $P(x)$ is a probability distribution.

b. What is the probability that the exchange rate on a given day during this period will be at least 0.75?

c. What is the probability that the exchange rate on a given day during this period will be less than 0.77?

d. If daily exchange rates are independent of one another, what is the probability that for two days in a row the exchange rate will be above 0.75?

3–2 Expected Values of Discrete Random Variables

In Chapter 1, we discussed summary measures of data sets. The most important summary measures discussed were the mean and the variance (also the square root of the variance, the standard deviation). We saw that the mean is a measure of *centrality,* or *location,* of the data or population, and that the variance and the standard deviation measure the *variability,* or *spread,* of our observations.

The mean of a probability distribution of a random variable is a measure of the centrality of the probability distribution. It is a measure that considers both the values of the random variable and their probabilities. The mean is a *weighted average* of the possible values of the random variable—the weights being the probabilities.

The mean of the probability distribution of a random variable is called the *expected value* of the random variable (sometimes called the *expectation* of the random variable). The reason for this name is that the mean is the (probability-weighted) average value of the random variable, and therefore it is the value we "expect" to occur. We denote the mean by two notations: μ for *mean* (as in Chapter 1 for a population) and $E(X)$ for *expected value of X.* In situations where no ambiguity is possible, we will often use μ. In cases where we want to stress the fact that we are talking about the expected value of a particular random variable (here, X), we will use the notation $E(X)$. The expected value of a discrete random variable is defined as follows.

The **expected value** of a discrete random variable X is equal to the sum of all values of the random variable, each value multiplied by its probability.

$$\mu = E(X) = \sum_{\text{all } x} xP(x) \qquad (3\text{–}4)$$

Suppose a coin is tossed. If it lands heads, you win a dollar; but if it lands tails, you lose a dollar. What is the expected value of this game? Intuitively, you know you have an even chance of winning or losing the same amount, and so the average or expected

[3]Inferred from a chart of dollars in euros published in "Business Day," *The New York Times,* April 20, 2007, p. C10.

TABLE 3–5 Computing the Expected Number of Switches for Example 3–2

x	$P(x)$	$xP(x)$
0	0.1	0
1	0.2	0.2
2	0.3	0.6
3	0.2	0.6
4	0.1	0.4
5	0.1	0.5
	1.00	2.3 ← Mean, $E(X)$

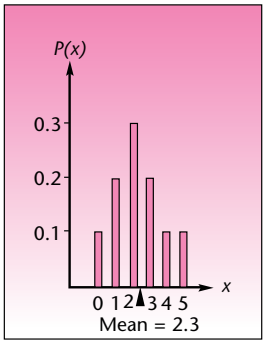
value is zero. Your payoff from this game is a random variable, and we find its expected value from equation 3–4: $E(X) = (1)(1/2) + (-1)(1/2) = 0$. The definition of an expected value, or mean, of a random variable thus conforms with our intuition. Incidentally, games of chance with an expected value of zero are called *fair games*.

Let us now return to Example 3–2 and find the expected value of the random variable involved—the expected number of switches on a given call. For convenience, we compute the mean of a discrete random variable by using a table. In the first column of the table we write the values of the random variable. In the second column we write the probabilities of the different values, and in the third column we write the products $xP(x)$ for each value x. We then add the entries in the third column, giving us $E(X) = \Sigma xP(x)$, as required by equation 3–4. This is shown for Example 3–2 in Table 3–5.

As indicated in the table, $\mu = E(X) = 2.3$. We can say that, on the average, 2.3 switches occur per call. As this example shows, the mean does not have to be one of the values of the random variable. No calls have 2.3 switches, but 2.3 is the average number of switches. It is the *expected* number of switches per call, although here the exact expectation will not be realized on any call.

As the weighted average of the values of the random variable, with probabilities as weights, the mean is the *center of mass* of the probability distribution. This is demonstrated for Example 3–2 in Figure 3–10.

FIGURE 3–10
The Mean of a Discrete Random Variable as a Center of Mass for Example 3–2



The Expected Value of a Function of a Random Variable

The expected value of a *function* of a random variable can be computed as follows. Let $h(X)$ be a function of the discrete random variable X.

The expected value of $h(X)$, a function of the discrete random variable X, is

$$E[h(X)] = \sum_{\text{all } x} h(x)P(x) \qquad (3\text{–}5)$$

The function $h(X)$ could be X^2, $3X^4$, $\log X$, or any function. As we will see shortly, equation 3–5 is most useful for computing the expected value of the special function $h(X) = X^2$. But let us first look at a simpler example, where $h(X)$ is a *linear* function of X. A linear function of X is a straight-line relation: $h(X) = a + bX$, where a and b are numbers.

EXAMPLE 3–3

Monthly sales of a certain product, recorded to the nearest thousand, are believed to follow the probability distribution given in Table 3–6. Suppose that the company has a fixed monthly production cost of $8,000 and that each item brings $2. Find the expected monthly profit from product sales.

TABLE 3–6 Probability Distribution of Monthly Product Sales for Example 3–3

Number of Items x	$P(x)$
5,000	0.2
6,000	0.3
7,000	0.2
8,000	0.2
9,000	0.1
	1.00

TABLE 3–7 Computing Expected Profit for Example 3–3

x	$h(x)$	$P(x)$	$h(x)P(x)$
5,000	2,000	0.2	400
6,000	4,000	0.3	1,200
7,000	6,000	0.2	1,200
8,000	8,000	0.2	1,600
9,000	10,000	0.1	1,000
			$E[h(X)] = 5,400$

Solution The company's profit function from sales of the product is $h(X) = 2X - 8{,}000$. Equation 3–5 tells us that the expected value of $h(X)$ is the sum of the values of $h(X)$, each value multiplied by the probability of the particular value of X. We thus add two columns to Table 3–6: a column of values of $h(x)$ for all x and a column of the products $h(x)P(x)$. At the bottom of this column we find the required sum $E[h(X)] = \Sigma_{\text{all } x}\, h(x)P(x)$. This is done in Table 3–7. As shown in the table, expected monthly profit from sales of the product is \$5,400.

In the case of a linear function of a random variable, as in Example 3–3, there is a possible simplification of our calculation of the mean of $h(X)$. The simplified formula of the expected value of a linear function of a random variable is as follows:

The expected value of a linear function of a random variable is

$$E(aX + b) = aE(X) + b \qquad (3\text{–}6)$$

where a and b are fixed numbers.

Equation 3–6 holds for *any* random variable, discrete or continuous. Once you know the expected value of X, the expected value of $aX + b$ is just $aE(X) + b$. In Example 3–3 we could have obtained the expected profit by finding the mean of X first, and then multiplying the mean of X by 2 and subtracting from this the fixed cost of \$8,000. The mean of X is 6,700 (prove this), and the expected profit is therefore $E[h(X)] = E(2X - 8{,}000) = 2E(X) - 8{,}000 = 2(6{,}700) - 8{,}000 = \$5{,}400$, as we obtained using Table 3–7.

As mentioned earlier, the most important expected value of a function of X is the expected value of $h(X) = X^2$. This is because this expected value helps us compute the *variance* of the random variable X and, through the variance, the standard deviation.

Variance and Standard Deviation of a Random Variable

The variance of a random variable is the expected squared deviation of the random variable from its mean. The idea is similar to that of the variance of a data set or a

population, defined in Chapter 1. Probabilities of the values of the random variable are used as weights in the computation of the expected squared deviation from the mean of a discrete random variable. The definition of the variance follows. As with a population, we denote the variance of a random variable by σ^2. Another notation for the variance of X is $V(X)$.

The **variance** of a discrete random variable X is given by

$$\sigma^2 = V(X) = E[(X - \mu)^2] = \sum_{\text{all } x} (x - \mu)^2 P(x) \qquad (3\text{–}7)$$

Using equation 3–7, we can compute the variance of a discrete random variable by subtracting the mean μ from each value x of the random variable, squaring the result, multiplying it by the probability $P(x)$, and finally adding the results for all x. Let us apply equation 3–7 and find the variance of the number of dialing switches in Example 3–2:

$$\begin{aligned}\sigma^2 &= \Sigma(x - \mu)^2 P(x) \\ &= (0 - 2.3)^2(0.1) + (1 - 2.3)^2(0.2) + (2 - 2.3)^2(0.3) \\ &\quad + (3 - 2.3)^2(0.2) + (4 - 2.3)^2(0.1) + (5 - 2.3)^2(0.1) \\ &= 2.01\end{aligned}$$

The variance of a discrete random variable can be computed more easily. Equation 3–7 can be shown mathematically to be equivalent to the following computational form of the variance.

Computational formula for the variance of a random variable:

$$\sigma^2 = V(X) = E(X^2) - [E(X)]^2 \qquad (3\text{–}8)$$

Equation 3–8 has the same relation to equation 3–7 as equation 1–7 has to equation 1–3 for the variance of a set of points.

Equation 3–8 states that the variance of X is equal to the expected value of X^2 minus the squared mean of X. In computing the variance using this equation, we use the definition of the expected value of a function of a discrete random variable, equation 3–5, in the special case $h(X) = X^2$. We compute x^2 for each x, multiply it by $P(x)$, and add for all x. This gives us $E(X^2)$. To get the variance, we subtract from $E(X^2)$ the mean of X, squared.

We now compute the variance of the random variable in Example 3–2, using this method. This is done in Table 3–8. The first column in the table gives the values of X, the second column gives the probabilities of these values, the third column gives the products of the values and their probabilities, and the fourth column is the product of the third column and the first [because we get $x^2P(x)$ by just multiplying each entry $xP(x)$ by x from column 1]. At the bottom of the third column we find the mean of X, and at the bottom of the fourth column we find the mean of X^2. Finally, we perform the subtraction $E(X^2) - [E(X)]^2$ to get the variance of X:

$$V(X) = E(X^2) - [E(X)]^2 = 7.3 - (2.3)^2 = 2.01$$

TABLE 3–8 Computations Leading to the Variance of the Number of Switches in Example 3–2 Using the Shortcut Formula (Equation 3–8)

x	$P(x)$	$xP(x)$	$x^2P(x)$
0	0.1	0	0
1	0.2	0.2	0.2
2	0.3	0.6	1.2
3	0.2	0.6	1.8
4	0.1	0.4	1.6
5	0.1	0.5	2.5
	1.00	2.3 ← Mean of X	7.3 ← Mean of X^2

This is the same value we found using the other formula for the variance, equation 3–7. Note that equation 3–8 holds for *all* random variables, discrete or otherwise. Once we obtain the expected value of X^2 and the expected value of X, we can compute the variance of the random variable using this equation.

For random variables, as for data sets or populations, the standard deviation is equal to the (positive) square root of the variance. We denote the standard deviation of a random variable X by σ or by $SD(X)$.

The **standard deviation** of a random variable:

$$\sigma = SD(X) = \sqrt{V(X)} \qquad (3\text{–}9)$$

In Example 3–2, the standard deviation is $\sigma = \sqrt{2.01} = 1.418$.

What are the variance and the standard deviation, and how do we interpret their meaning? By definition, the variance is the weighted average squared deviation of the values of the random variable from their mean. Thus, it is a measure of the *dispersion* of the possible values of the random variable about the mean. The variance gives us an idea of the variation or uncertainty associated with the random variable: The larger the variance, the farther away from the mean are possible values of the random variable. Since the variance is a squared quantity, it is often more useful to consider its square root—the standard deviation of the random variable. When two random variables are compared, the one with the larger variance (standard deviation) is the more variable one. The risk associated with an investment is often measured by the standard deviation of investment returns. When comparing two investments with the same average (*expected*) return, the investment with the higher standard deviation is considered riskier (although a higher standard deviation implies that returns are expected to be more variable—both below and above the mean).

Variance of a Linear Function of a Random Variable

There is a formula, analogous to equation 3–6, that gives the variance of a linear function of a random variable. For a linear function of X given by $aX + b$, we have the following:

Variance of a linear function of a random variable is

$$V(aX + b) = a^2V(X) = a^2\sigma^2 \qquad (3\text{–}10)$$

where a and b are fixed numbers.

Using equation 3–10, we will find the variance of the profit in Example 3–3. The profit is given by $2X - 8{,}000$. We need to find the variance of X in this example. We find

$$\begin{aligned} E(X^2) &= (5{,}000)^2(0.2) + (6{,}000)^2(0.3) + (7{,}000)^2(0.2) + (8{,}000)^2(0.2) \\ &\quad + (9{,}000)^2(0.1) \\ &= 46{,}500{,}000 \end{aligned}$$

The expected value of X is $E(X) = 6{,}700$. The variance of X is thus

$$V(X) = E(X^2) - [E(X)]^2 = 46{,}500{,}000 - (6{,}700)^2 = 1{,}610{,}000$$

Finally, we find the variance of the profit, using equation 3–10, as $2^2(1{,}610{,}000) = 6{,}440{,}000$. The standard deviation of the profit is $\sqrt{6{,}440{,}000} = 2{,}537.72$.

3–3 Sum and Linear Composites of Random Variables

Sometimes we are interested in the sum of several random variables. For instance, a business may make several investments, and each investment may have a random return. What finally matters to the business is the sum of all the returns. Sometimes what matters is a **linear composite** of several random variables. A linear composite of random variables $X_1, X_2, \ldots, X_k$ will be of the form

$$a_1X_1 + a_2X_2 + \cdots + a_kX_k$$

where $a_1, a_2, \ldots, a_k$ are constants. For instance, let $X_1, X_2, \ldots, X_k$ be the random quantities of k different items that you buy at a store, and let $a_1, a_2, \ldots, a_k$ be their respective prices. Then $a_1X_1 + a_2X_2 + \cdots + a_kX_k$ will be the random total amount you have to pay for the items. Note that the sum of the variables is a linear composite where all a's are 1. Also, $X_1 - X_2$ is a linear composite with $a_1 = 1$ and $a_2 = -1$.

We therefore need to know how to calculate the expected value and variance of the sum or linear composite of several random variables. The following results are useful in computing these statistics.

The expected value of the sum of several random variables is the sum of the individual expected values. That is,

$$E(X_1 + X_2 + \cdots + X_k) = E(X_1) + E(X_2) + \cdots + E(X_k)$$

Similarly, the expected value of a linear composite is given by

$$E(a_1X_1 + a_2X_2 + \cdots + a_kX_k) = a_1E(X_1) + a_2E(X_2) + \cdots + a_kE(X_k)$$

In the case of variance, we will look only at the case where $X_1, X_2, \ldots, X_k$ are **mutually independent,** because if they are not mutually independent, the computation involves covariances, which we will learn about in Chapter 10. Mutual independence means that any event $X_i = x$ and any other event $X_j = y$ are independent. We can now state the result.

If $X_1, X_2, \ldots, X_k$ are mutually independent, then the variance of their sum is the sum of their individual variances. That is,

$$V(X_1 + X_2 + \cdots + X_k) = V(X_1) + V(X_2) + \cdots + V(X_k)$$

Similarly, the variance of a linear composite is given by

$$V(a_1X_1 + a_2X_2 + \cdots + a_kX_k) = a_1^2V(X_1) + a_2^2V(X_2) + \cdots + a_k^2V(X_k)$$

We will see the application of these results through an example.

EXAMPLE 3–4

A portfolio includes stocks in three industries: financial, energy, and consumer goods (in equal proportions). Assume that these three sectors are independent of each other and that the expected annual return (in dollars) and standard deviations are as follows: financial: 1,000 and 700; energy 1,200 and 1,100; and consumer goods 600 and 300 (respectively). What are the mean and standard deviation of annual dollar-value return on this portfolio?

Solution

The mean of the sum of the three random variables is the sum of the means 1,000 + 1,200 + 600 = \$2,800. Since the three sectors are assumed independent, the variance is the sum of the three variances. It is equal to $700^2 + 1{,}100^2 + 300^2 = 1{,}790{,}000$. So the standard deviation is $\sqrt{1{,}790{,}000} = \$1{,}337.90$.

Chebyshev's Theorem

The standard deviation is useful in obtaining bounds on the possible values of the random variable with certain probability. The bounds are obtainable from a well-known theorem, *Chebyshev's theorem* (the name is sometimes spelled Tchebychev, Tchebysheff, or any of a number of variations). The theorem says that for any number k greater than 1.00, the probability that the value of a given random variable will be *within k standard deviations* of the mean is at least $1 - 1/k^2$. In Chapter 1, we listed some results for data sets that are derived from this theorem.

Chebyshev's Theorem

For a random variable X with mean μ and standard deviation σ, and for any number $k > 1$,

$$P(|X - \mu| < k\sigma) \geq 1 - 1/k^2 \qquad (3\text{–}11)$$

Let us see how the theorem is applied by selecting values of k. While k does not have to be an integer, we will use integers. When $k = 2$, we have $1 - 1/k^2 = 0.75$: The theorem says that the value of the random variable will be within a distance of 2 standard deviations away from the mean with at least a 0.75 probability. Letting $k = 3$, we find that X will be within 3 standard deviations of its mean with at least a 0.89 probability. We can similarly apply the rule for other values of k. The rule holds for data sets and populations in a similar way. When applied to a sample of observations, the rule says that at least 75% of the observations lie within 2 standard deviations of the sample mean $\bar{x}$. It says that at least 89% of the observations lie within 3 standard deviations of the mean, and so on. Applying the theorem to the random variable of Example 3–2, which has mean 2.3 and standard deviation 1.418, we find that the probability that X will be anywhere from $2.3 - 2(1.418)$ to $2.3 + 2(1.418) = -0.536$ to 5.136 is at least 0.75. From the actual probability distribution in this example, Table 3–3, we know that the probability that X will be between 0 and 5 is 1.00.

Often, we will know the distribution of the random variable in question, in which case we will be able to use the distribution for obtaining actual probabilities rather

than the bounds offered by Chebyshev's theorem. If the exact distribution of the random variable is not known, but we may assume an approximate distribution, the approximate probabilities may still be better than the general bounds offered by Chebyshev's theorem.

The Templates for Random Variables

The template shown in Figure 3–11 can be used to calculate the descriptive statistics of a random variable and also those of a function $h(x)$ of that random variable. To calculate the statistics of $h(x)$, the Excel formula for the function must be entered in cell G12. For instance, if $h(x) = 5x^2 + 8$, enter the Excel formula `=5*x^2+8` in cell G12.

The template shown in Figure 3–12 can be used to compute the statistics about the sum of mutually independent random variables. While entering the variance of the individual X's, be careful that what you enter is the variance and not the standard deviation. At times, you know only the standard deviation and not the variance. In such cases, you can make the template calculate the variance from the standard deviation. For example, if the standard deviation is 1.23, enter the formula `=1.23^2`, which will compute and use the variance.

The template shown in Figure 3–13 can be used to compute the statistics about linear composites of mutually independent random variables. You will enter the coefficients (the a_i's) in column B.

FIGURE 3–11 **Descriptive Statistics of a Random Variable X and $h(x)$ [Random Variable.xls]**

Statistics of a Random Variable

Title

x	P(x)	F(x)
0	0.1	0.1
1	0.2	0.3
2	0.3	0.6
3	0.2	0.8
4	0.1	0.9
5	0.1	1

Statistics of X	
Mean	2.3
Variance	2.01
Std. Devn.	1.41774
Skewness	0.30319
(Relative) Kurtosis	-0.63132

Definition of h(x)	x
h(x) = 8	0

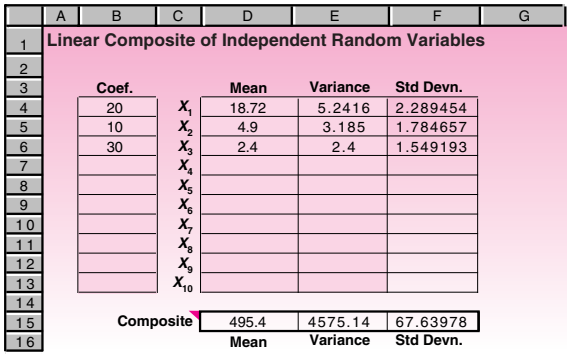
Statistics of h(x)	
Mean	44.5
Variance	1400.25
Std. Devn.	37.4199
Skewness	1.24449
(Relative) Kurtosis	0.51413

FIGURE 3–12 **Template for the Sum of Independent Variables**

Sum of Independent Random Variables

	Mean	Variance	Std Devn.
X_1	18.72	5.2416	2.289454
X_2	4.9	3.185	1.784657
X_3	2.4	2.4	1.549193
X_4			
X_5			
X_6			
X_7			
X_8			
X_9			
X_{10}			
Sum	26.02	10.8266	3.29038
	Mean	Variance	Std Devn.

FIGURE 3–13 Template for Linear Composites of Independent Variables [Random Variables.xls, Sheet: Composite]



Linear Composite of Independent Random Variables

	Coef.		Mean	Variance	Std Devn.
	20	X_1	18.72	5.2416	2.289454
	10	X_2	4.9	3.185	1.784657
	30	X_3	2.4	2.4	1.549193
		X_4			
		X_5			
		X_6			
		X_7			
		X_8			
		X_9			
		X_{10}			
	Composite		495.4	4575.14	67.63978
			Mean	Variance	Std Devn.

PROBLEMS

3–11. Find the expected value of the random variable in problem 3–1. Also find the variance of the random variable and its standard deviation.

3–12. Find the mean, variance, and standard deviation of the random variable in problem 3–2.

3–13. What is the expected percentage of people responding to an advertisement when the probability distribution is the one given in problem 3–3? What is the variance of the percentage of people who respond to the advertisement?

3–14. Find the mean, variance, and standard deviation of the number of cars sold per day, using the probability distribution in problem 3–4.

3–15. What is the expected number of dots appearing on two dice? (Use the probability distribution you computed in your answer to problem 3–5.)

3–16. Use the probability distribution in problem 3–6 to find the expected number of shipment orders per day. What is the probability that on a given day there will be more orders than the average?

3–17. Find the mean, variance, and standard deviation of the annual income of a hedge fund manager, using the probability distribution in problem 3–7.

3–18. According to Chebyshev's theorem, what is the minimum probability that a random variable will be within 4 standard deviations of its mean?

3–19. At least eight-ninths of a population lies within how many standard deviations of the population mean? Why?

3–20. The average annual return on a certain stock is 8.3%, and the variance of the returns on the stock is 2.3. Another stock has an average return of 8.4% per year and a variance of 6.4. Which stock is riskier? Why?

3–21. Returns on a certain business venture, to the nearest $1,000, are known to follow the probability distribution

x	$P(x)$
−2,000	0.1
−1,000	0.1
0	0.2
1,000	0.2
2,000	0.3
3,000	0.1

a. What is the most likely monetary outcome of the business venture?

b. Is the venture likely to be successful? Explain.

c. What is the long-term average earning of business ventures of this kind? Explain.

d. What is a good measure of the risk involved in a venture of this kind? Why? Compute this measure.

3–22. Management of an airline knows that 0.5% of the airline's passengers lose their luggage on domestic flights. Management also knows that the average value claimed for a lost piece of luggage on domestic flights is $600. The company is considering increasing fares by an appropriate amount to cover expected compensation to passengers who lose their luggage. By how much should the airline increase fares? Why? Explain, using the ideas of a random variable and its expectation.

3–23. Refer to problem 3–7. Suppose that hedge funds must withhold $300 million from the income of the manager and an additional 5% of the remaining income. Find the expected net income of a manager in this group. What property of expected values are you using?

3–24. Refer to problem 3–4. Suppose the car dealership's operation costs are well approximated by the square root of the number of cars sold, multiplied by $300. What is the expected daily cost of the operation? Explain.

3–25. In problem 3–2, suppose that a cost is imposed of an amount equal to the square of the number of additional hours of sleep. What is the expected cost? Explain.

3–26. All voters of Milwaukee County were asked a week before election day whether they would vote for a certain presidential candidate. Of these, 48% answered yes, 45% replied no, and the rest were undecided. If a yes answer is coded +1, a no answer is coded –1, and an undecided answer is coded 0, find the mean and the variance of the code.

3–27. Explain the meaning of the variance of a random variable. What are possible uses of the variance?

3–28. Why is the standard deviation of a random variable more meaningful than its variance for interpretation purposes?

3–29. Refer to problem 3–23. Find the variance and the standard deviation of hedge fund managers' income.

3–30. For problem 3–10, find the mean and the standard deviation of the dollar to euros exchange rate.

3–31. Lobsters vary in sizes. The bigger the size, the more valuable the lobster per pound (a 6-pound lobster is more valuable than two 3-pound ones). Lobster merchants will sell entire boatloads for a certain price. The boatload has a mixture of sizes. Suppose the distribution is as follows:

x(pound)	*P*(*x*)	*v*(*x*) ($)
½	0.1	2
¾	0.1	2.5
1	0.3	3.0
1¼	0.2	3.25
1½	0.2	3.40
1¾	0.05	3.60
2	0.05	5.00

What is a fair price for the shipload?

3–4 Bernoulli Random Variable

The first standard random variable we shall study is the *Bernoulli random variable,* named in honor of the mathematician Jakob Bernoulli (1654–1705). It is the building block for other random variables in this chapter. The distribution of a Bernoulli random variable X is given in Table 3–9. As seen in the table, x is 1 with probability p and 0 with probability $(1 - p)$. The case where $x = 1$ is called "success" and the case where $x = 0$ is called "failure."

TABLE 3–9
Bernoulli Distribution

x	$P(x)$
1	p
0	$1 - p$

Observe that

$$E(X) = 1 * p + 0 * (1 - p) = p$$
$$E(X^2) = 1^2 * p + 0^2 * (1 - p) = p$$
$$V(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1 - p)$$

Often the quantity $(1 - p)$, which is the probability of failure, is denoted by the symbol q, and thus $V(X) = pq$. If X is a Bernoulli random variable with probability of success p, then we write $X \sim \text{BER}(p)$, where the symbol "$\sim$" is read "is distributed as" and BER stands for Bernoulli. The characteristics of a Bernoulli random variable are summarized in the following box.

Bernoulli Distribution

If $X \sim \text{BER}(p)$, then

$$P(1) = p; \qquad P(0) = 1 - p$$
$$E[X] = p$$
$$V(X) = p(1 - p)$$

For example, if $p = 0.8$, then

$$E[X] = 0.8$$
$$V(X) = 0.8 * 0.2 = 0.16$$

Let us look at a practical instance of a Bernoulli random variable. Suppose an operator uses a lathe to produce pins, and the lathe is not perfect in that it does not always produce a good pin. Rather, it has a probability p of producing a good pin and $(1 - p)$ of producing a defective one.

Just after the operator produces one pin, let X denote the "number of good pins produced." Clearly, X is 1 if the pin is good and 0 if it is defective. Thus, X follows exactly the distribution in Table 3–9, and therefore $X \sim \text{BER}(p)$.

If the outcome of a trial can only be either a success or a failure, then the trial is a **Bernoulli trial**.

The number of successes X in one Bernoulli trial, which can be 1 or 0, is a **Bernoulli random variable**.

Another example is tossing a coin. If we take heads as 1 and tails as 0, then the outcome of a toss is a Bernoulli random variable.

A Bernoulli random variable is too simple to be of immediate practical use. But it forms the building block of the binomial random variable, which is quite useful in practice. The binomial random variable in turn is the basis for many other useful cases.

3–5 The Binomial Random Variable

In the real world we often make several trials, not just one, to achieve one or more successes. Since we now have a handle on Bernoulli-type trials, let us consider cases where there are n number of Bernoulli trials. A condition we need to impose on these trials is that the outcome of any trial be independent of the outcome of any other trial. Very often this independence condition is true. For example, when we toss a coin several times, the outcome of one toss is not affected by the outcome of any other toss.

Consider n number of *identically and independently distributed* Bernoulli random variables $X_1, X_2, \ldots, X_n$. Here, identically means that they all have the same p, and independently means that the value of one X does not in any way affect the value of another. For example, the value of X_2 does not affect the value of X_3 or X_5, and so on. Such a *sequence* of identically and independently distributed Bernoulli variables is called a **Bernoulli process.**

Suppose an operator produces n pins, one by one, on a lathe that has probability p of making a good pin at each trial. If this p remains constant throughout, then independence is guaranteed and the sequence of numbers (1 or 0) denoting the good and bad pins produced in each of the n trials is a Bernoulli process. For example, in the sequence of eight trials denoted by

$$0\,0\,1\,0\,1\,1\,0\,0$$

the third, fifth, and sixth are good pins, or successes. The rest are failures.

In practice, we are usually interested in the total number of good pins rather than the sequence of 1's and 0's. In the example above, three out of eight are good. In the general case, let X denote the total number of good pins produced in n trials. We then have

$$X = X_1 + X_2 + \cdots + X_n$$

where all $X_i \sim \text{BER}(p)$ and are independent.

An X that counts the number of successes in many independent, identical Bernoulli trials is called a **binomial random variable.**

Conditions for a Binomial Random Variable

Note the conditions that need to be satisfied for a binomial random variable:

1. The trials must be Bernoulli trials in that the outcomes can only be either success or failure.
2. The outcomes of the trials must be independent.
3. The probability of success in each trial must be constant.

The first condition is easy to understand. Coming to the second condition, we already saw that the outcomes of coin tosses will be independent. As an example of dependent outcomes, consider the following experiment. We toss a fair coin and if it is heads we record the outcome as success, or 1, and if it is tails we record it as failure, or 0. For the second outcome, we do not toss the coin but we record the opposite of the previous outcome. For the third outcome, we toss the coin again and repeat the process of writing the opposite result for every other outcome. Thus in the sequence of all

outcomes, every other outcome will be the opposite of the previous outcome. We stop after recording 20 outcomes. In this experiment, all outcomes are random and of Bernoulli type with success probability 0.5. But they are not independent in that every other outcome is the opposite of, and thus dependent on, the previous outcome. And for this reason, the number of successes in such an experiment will not be binomially distributed. (In fact, the number is not even random. Can you guess what that number will be?)

The third condition of constant probability of success is important and can be easily violated. Tossing two different coins with differing probabilities of success will violate the third condition (but not the other two). Another case that is relevant to the third condition, which we need to be aware of, is *sampling with and without replacement.* Consider an urn that contains 10 green marbles (successes) and 10 red marbles (failures). We pick a marble from the urn at random and record the outcome. The probability of success is $10/20 = 0.5$. For the second outcome, suppose *we replace the first* marble drawn and then pick one at random. In this case the probability of success remains at $10/20 = 0.5$, and the third condition is satisfied. But if *we do not replace the first* marble before picking the second, then the probability of the second outcome being a success is $9/19$ if the first was a success and $10/19$ if the first was a failure. Thus the probability of success does not remain constant (and is also dependent on the previous outcomes). Therefore, the third condition is violated (as is the second condition). This means that sampling with replacement will follow a binomial distribution, but sampling without replacement will not. Later we will see that sampling without replacement will follow a hypergeometric distribution.

Binomial Distribution Formulas

Consider the case where five trials are made, and in each trial the probability of success is 0.6. To get to the formula for calculating binomial probabilities, let us analyze the probability that the number of successes in the five trials is exactly three.

CHAPTER 4

First, we note that there are $\binom{5}{3}$ ways of getting three successes out of five trials. Next we observe that each of these $\binom{5}{3}$ possibilities has $0.6^3 * 0.4^2$ probability of occurrence corresponding to 3 successes and 2 failures. Therefore,

$$P(X = 3) = \binom{5}{3} * 0.6^3 * 0.4^2 = 0.3456$$

We can generalize this equation with n denoting the number of trials and p the probability of success:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{(n-x)} \qquad \text{for } x = 0, 1, 2, \ldots, n \qquad (3\text{–}12)$$

Equation 3–12 is the famous binomial probability formula.

To describe a binomial random variable we need two parameters, n and p. We write $X \sim \text{B}(n, p)$ to indicate that X is binomially distributed with n number of trials and p probability of success in each trial. The letter B in the expression stands for binomial.

With any random variable, we will be interested in its expected value and its variance. Let us consider the expected value of a binomial random variable X. We note that X is the sum of n number of Bernoulli random variables, each of which has an

expected value of p. Hence the expected value of X must be np, that is, $E(X) = np$. Furthermore, the variance of each Bernoulli random variable is $p(1 - p)$, and they are all independent. Therefore variance of X is $np(1 - p)$, that is, $V(X) = np(1 - p)$. The formulas for the binomial distribution are summarized in the next box, which also presents sample calculations that use these formulas.

Binomial Distribution

If $X \sim B(n, p)$, then

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{(n-x)} \qquad x = 0, 1, 2, \ldots, n$$

$$E(X) = np$$

$$V(X) = np(1 - p)$$

For example, if $n = 5$ and $p = 0.6$, then

$$P(X = 3) = 10 * 0.6^3 * 0.4^2 = 0.3456$$

$$E(X) = 5 * 0.6 = 3$$

$$V(X) = 5 * 0.6 * 0.4 = 1.2$$

The Template

The calculation of binomial probabilities, especially the cumulative probabilities, can be tedious. Hence we shall use a spreadsheet template. The template that can be used to calculate binomial probabilities is shown in Figure 3–14. When we enter the values for n and p, the template automatically tabulates the probability of "Exactly x," "At most x," and "At least x" number of successes. This tabulation can be used to solve many kinds of problems involving binomial probabilities, as explained in the next section. Besides the tabulation, a histogram is also created on the right. The histogram helps the user to visualize the shape of the distribution.

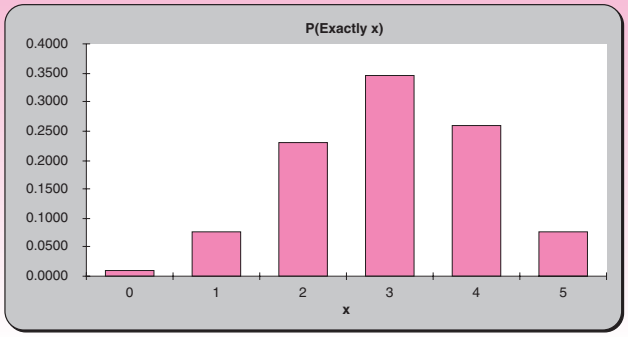
FIGURE 3–14 Binomial Distribution Template [Binomial.xls]

Binomial Distribution

n	p
5	0.6

Mean	Variance	Stdev
3	1.2	1.095445

x	P(Exactly x)	P(At most x)	P(At least x)
0	0.0102	0.0102	1.0000
1	0.0768	0.0870	0.9898
2	0.2304	0.3174	0.9130
3	0.3456	0.6630	0.6826
4	0.2592	0.9222	0.3370
5	0.0778	1.0000	0.0778



Problem Solving with the Template

Suppose an operator wants to produce *at least* two good pins. (In practice, one would want *at least* some number of *good* things, or *at most* some number of *bad* things. Rarely would one want *exactly* some number of good or bad things.) He produces the pins using a lathe that has 0.6 probability of making a good pin in each trial, and this probability stays constant throughout. Suppose he produces five pins. What is the probability that he would have made at least two good ones?

Let us see how we can answer the question using the template. After making sure that n is filled in as 5 and p as 0.6, the answer is read off as 0.9130 (in cell E9). That is, the operator can be 91.3% confident that he would have at least two good pins.

Let us go further with the problem. Suppose it is critical that the operator have at least two good pins, and therefore he wants to be at least 99% confident that he would have at least two good pins. (In this type of situation, the phrases "at least" and "at most" occur often. You should read carefully.) With five trials, we just found that he can be only 91.3% confident. To increase the confidence level, one thing he can do is increase the number of trials. How many more trials? Using the spreadsheet template, we can answer this question by progressively increasing the value of n and stopping when P(At least 2) in cell E9 just exceeds 99%. On doing this, we find that eight trials will do and seven will not. Hence the operator should make at least eight trials.

Increasing n is not the only way to increase confidence. We can increase p, if that is possible in practice. To see it, we pose another question.

Suppose the operator has enough time to produce only five pins, but he still wants to have at least 99% confidence of producing at least two good pins by improving the lathe and thus increasing p. How much should p be increased? To answer this, we can keep increasing p and stop when P(At least 2) just exceeds 99%. But this process could get tiresome if we need, say, four decimal place accuracy for p. This is where the Goal seek . . . command (see the Working with Templates file found on the student CD) in the spreadsheet comes in handy. The Goal seek command yields 0.7777. That is, p must be increased to at least 0.7777 in order to be 99% confident of getting at least two good pins in five trials.

We will complete this section by pointing out the use of the AutoCalculate command. We first note that the probability of at most x number of successes is the same as the cumulative probability $F(x)$. Certain types of probabilities are easily calculated using $F(x)$ values. For example, in our operator's problem, consider the probability that the number of successes will be between 1 and 3, both inclusive. We know that

$$P(1 \le x \le 3) = F(3) - F(0)$$

Looking at the template in Figure 3–14, we calculate this as $0.6630 - 0.0102 = 0.6528$. A quicker way is to use the AutoCalculate facility. When the range of cells containing $P(1)$ to $P(3)$ is selected, the sum of the probabilities appears in the AutoCalculate area as 0.6528.

PROBLEMS

3–32. Three of the 10 airplane tires at a hangar are faulty. Four tires are selected at random for a plane; let F be the number of faulty tires found. Is F a binomial random variable? Explain.

3–33. A salesperson finds that, in the long run, two out of three sales calls are successful. Twelve sales calls are to be made; let X be the number of concluded sales. Is X a binomial random variable? Explain.

3–34. A large shipment of computer chips is known to contain 10% defective chips. If 100 chips are randomly selected, what is the expected number of defective ones? What is the standard deviation of the number of defective chips? Use Chebyshev's theorem to give bounds such that there is at least a 0.75 chance that the number of defective chips will be within the two bounds.

3–35. A new treatment for baldness is known to be effective in 70% of the cases treated. Four bald members of the same family are treated; let X be the number of successfully treated members of the family. Is X a binomial random variable? Explain.

3–36. What are Bernoulli trials? What is the relationship between Bernoulli trials and the binomial random variable?

3–37. Look at the histogram of probabilities in the binomial distribution template [Binomial.xls] for the case $n = 5$ and $p = 0.6$.

a. Is this distribution symmetric or skewed? Now, increase the value of n to 10, 15, 20, . . . Is the distribution becoming more symmetric or more skewed? Make a formal statement about what happens to the distribution's shape when n increases.

b. With $n = 5$, change the p value to 0.1, 0.2, . . . Observe particularly the case of $p = 0.5$. Make a formal statement about how the skewness of the distribution changes with p.

3–38. A salesperson goes door-to-door in a residential area to demonstrate the use of a new household appliance to potential customers. At the end of a demonstration, the probability that the potential customer would place an order for the product is a constant 0.2107. To perform satisfactorily on the job, the salesperson needs at least four orders. Assume that each demonstration is a Bernoulli trial.

a. If the salesperson makes 15 demonstrations, what is the probability that there would be exactly 4 orders?

b. If the salesperson makes 16 demonstrations, what is the probability that there would be at most 4 orders?

c. If the salesperson makes 17 demonstrations, what is the probability that there would be at least 4 orders?

d. If the salesperson makes 18 demonstrations, what is the probability that there would be anywhere from 4 to 8 (both inclusive) orders?

e. If the salesperson wants to be at least 90% confident of getting at least 4 orders, at least how many demonstrations should she make?

f. The salesperson has time to make only 22 demonstrations, and she still wants to be at least 90% confident of getting at least 4 orders. She intends to gain this confidence by improving the quality of her demonstration and thereby improving the chances of getting an order at the end of a demonstration. At least to what value should this probability be increased in order to gain the desired confidence? Your answer should be accurate to four decimal places.

3–39. An MBA graduate is applying for nine jobs, and believes that she has in each of the nine cases a constant and independent 0.48 probability of getting an offer.

a. What is the probability that she will have at least three offers?

b. If she wants to be 95% confident of having at least three offers, how many more jobs should she apply for? (Assume each of these additional applications will also have the same probability of success.)

c. If there are no more than the original nine jobs that she can apply for, what value of probability of success would give her 95% confidence of at least three offers?

3–40. A computer laboratory in a school has 33 computers. Each of the 33 computers has 90% reliability. Allowing for 10% of the computers to be down, an instructor specifies an enrollment ceiling of 30 for his class. Assume that a class of 30 students is taken into the lab.

a. What is the probability that each of the 30 students will get a computer in working condition?

b. The instructor is surprised to see the low value of the answer to (*a*) and decides to improve it to at least 95% by doing one of the following:

i. Decreasing the enrollment ceiling.

ii. Increasing the number of computers in the lab.

iii. Increasing the reliability of all the computers.

To help the instructor, find out what the increase or decrease should be for each of the three alternatives.

3–41. A commercial jet aircraft has four engines. For an aircraft in flight to land safely, at least two engines should be in working condition. Each engine has an independent reliability of $p = 92\%$.

a. What is the probability that an aircraft in flight can land safely?

b. If the probability of landing safely must be at least 99.5%, what is the minimum value for p? Repeat the question for probability of landing safely to be 99.9%.

c. If the reliability cannot be improved beyond 92% but the number of engines in a plane can be increased, what is the minimum number of engines that would achieve at least 99.5% probability of landing safely? Repeat for 99.9% probability.

d. One would certainly desire 99.9% probability of landing safely. Looking at the answers to (*b*) and (*c*), what would you say is a better approach to safety, increasing the number of engines or increasing the reliability of each engine?

3–6 Negative Binomial Distribution

Consider again the case of the operator who wants to produce two good pins using a lathe that has 0.6 probability of making one good pin in each trial. Under binomial distribution, we assumed that he produces five pins and calculated the probability of getting at least two good ones. In practice, though, if only two pins are needed, the operator would produce the pins one by one and stop when he gets two good ones. For instance, if the first two are good, then he would stop right there; if the first and the third are good, then he would stop with the third; and so on. Notice that in this scenario, the number of successes is held constant at 2, and the number of trials is random. The number of trials could be 2, 3, 4, ... (Contrast this with the binomial distribution where the number of trials is fixed and the number of successes is random.)

The number of trials made in this scenario is said to follow a **negative binomial distribution.** Let s denote the exact number of successes desired and p the probability of success in each trial. Let X denote the number of trials made until the desired number of successes is achieved. Then X will follow a negative binomial distribution and we shall write $X \sim \text{NB}(s, p)$ where NB denotes negative binomial.

Negative Binomial Distribution Formulas

What is the formula for $P(X = x)$ when $X \sim \text{NB}(s, p)$? We know that the very last trial must be a success; otherwise, we would have already had the desired number of successes

with $x - 1$ trials, and we should have stopped right there. The last trial being a success, the first $x - 1$ trials should have had $s - 1$ successes. Thus the formula should be

$$P(X = x) = \binom{x-1}{s-1} p^s (1-p)^{(x-s)}$$

The formula for the mean can be arrived at intuitively. For instance, if $p = 0.3$, and 3 successes are desired, then the expected number of trials to achieve 3 successes is 10. Thus the mean should have the formula $\mu = s/p$. The variance is given by the formula $\sigma^2 = s(1 - p)/p^2$.

Negative Binomial Distribution

If $X \sim \text{NB}(s, p)$, then

$$P(X = x) = \binom{x-1}{s-1} p^s (1-p)^{(x-s)} \qquad x = s, s+1, s+2, \ldots$$

$$E(X) = s/p$$
$$V(X) = s(1-p)/p^2$$

For example, if $s = 2$ and $p = 0.6$, then

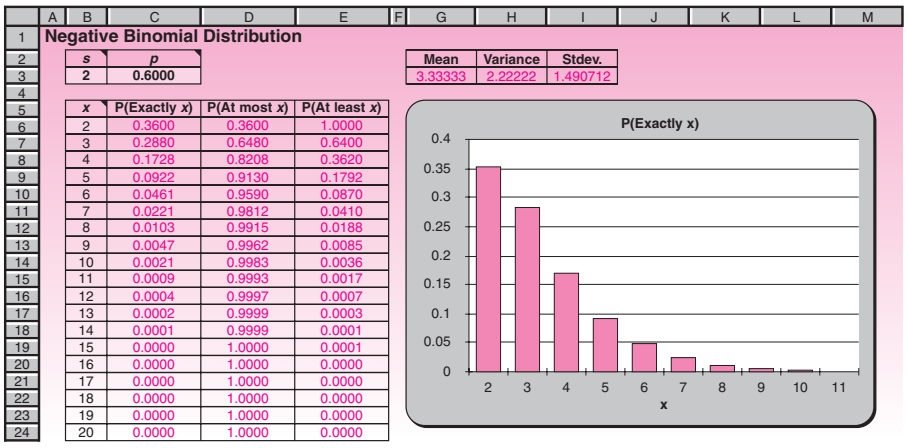
$$P(X = 5) = \binom{4}{1} * 0.6^2 * 0.4^3 = 0.0922$$

$$E(X) = 2/0.6 = 3.3333$$
$$V(X) = 2 * 0.4/0.6^2 = 2.2222$$

Problem Solving with the Template

Figure 3–15 shows the negative binomial distribution template. When we enter the s and p values, the template updates the probability tabulation and draws a histogram on the right.

FIGURE 3–15 Negative Binomial Distribution Template [Negative Binomial.xls]



Let us return to the operator who wants to keep producing pins until he has two good ones. The probability of getting a good one at any trial is 0.6. What is the probability that he would produce exactly five? Looking at the template, we see that the answer is 0.0922, which agrees with the calculation in the preceding box. We can, in addition, see that the probability of producing at most five is 0.9130 and at least five is 0.1792.

Suppose the operator has enough time to produce only four pins. How confident can he be that he would have two good ones within the available time? Looking at the template, we see that the probability of needing at most four trials is 0.8208 and hence he can be about 82% confident.

If he wants to be at least 95% confident, at least how many trials should he be prepared for? Looking at the template in the "At most" column, we infer that he should be prepared for at least six trials, since five trials yield only 91.30% confidence and six trials yield 95.90%.

Suppose the operator has enough time to produce only four pins and still wants to be at least 95% confident of getting two good pins within the available time. Suppose, further, he wants to achieve this by increasing the value of p. What is the minimum p that would achieve this? Using the Goal Seek command, this can be answered as 0.7514. Specifically, you set cell D8 to 0.95 by changing cell C3.

3–7 The Geometric Distribution

In a negative binomial distribution, the number of desired successes s can be any number. But in some practical situations, the number of successes desired is just one. For instance, if you are attempting to pass a test or get some information, it is enough to have just one success. Let X be the (random) number of Bernoulli trials, each having p probability of success, required to achieve just one success. Then X follows a **geometric distribution,** and we shall write $X \sim G(p)$. Note that the geometric distribution is a special case of the negative binomial distribution where $s = 1$. The reason for the name "geometric distribution" is that the sequence of probabilities $P(X = 1)$, $P(X = 2)$, . . . , follows a *geometric progression.*

Geometric Distribution Formulas

Because the geometric distribution is a special case of the negative binomial distribution where $s = 1$, the formulas for the negative binomial distribution with s fixed as 1 can be used for the geometric distribution.

Geometric Distribution Formulas

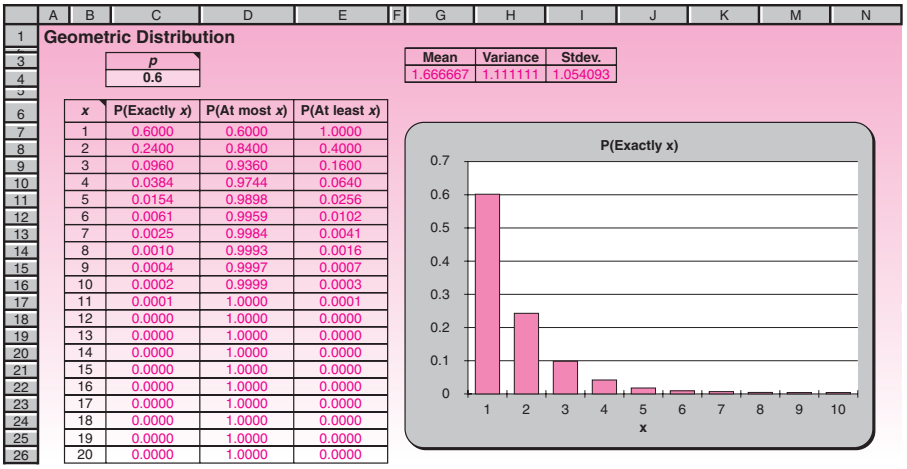
If $X \sim G(p)$, then

$$P(X = x) = p(1 - p)^{(x-1)} \qquad x = 1, 2, \ldots$$
$$E(X) = 1/p$$
$$V(X) = (1 - p)/p^2$$

For example, if $p = 0.6$, then

$$P(X = 5) = 0.6 * 0.4^4 = 0.0154$$
$$E(X) = 1/0.6 = 1.6667$$
$$V(X) = 0.4/0.6^2 = 1.1111$$

FIGURE 3–16 Geometric Distribution Template [Geometric.xls]



Geometric Distribution

p
0.6

Mean	Variance	Stdev.
1.666667	1.111111	1.054093

x	P(Exactly x)	P(At most x)	P(At least x)
1	0.6000	0.6000	1.0000
2	0.2400	0.8400	0.4000
3	0.0960	0.9360	0.1600
4	0.0384	0.9744	0.0640
5	0.0154	0.9898	0.0256
6	0.0061	0.9959	0.0102
7	0.0025	0.9984	0.0041
8	0.0010	0.9993	0.0016
9	0.0004	0.9997	0.0007
10	0.0002	0.9999	0.0003
11	0.0001	1.0000	0.0001
12	0.0000	1.0000	0.0000
13	0.0000	1.0000	0.0000
14	0.0000	1.0000	0.0000
15	0.0000	1.0000	0.0000
16	0.0000	1.0000	0.0000
17	0.0000	1.0000	0.0000
18	0.0000	1.0000	0.0000
19	0.0000	1.0000	0.0000
20	0.0000	1.0000	0.0000

Problem Solving with the Template

Consider the operator who produces pins one by one on a lathe that has 0.6 probability of producing a good pin at each trial. Suppose he wants only one good pin and stops as soon as he gets one. What is the probability that he would produce exactly five pins? The template that can be used to answer this and related questions is shown in Figure 3–16. On that template, we enter the value 0.6 for p. The answer can now be read off as 0.0154, which agrees with the example calculation in the preceding box. Further, we can read on the template that the probability of at most five is 0.9898 and at least five is 0.0256. Also note that the probability of exactly 1, 2, 3, . . . , trials follows the sequence 0.6, 0.24, 0.096, 0.0384, . . . , which is indeed a geometric progression with common ratio 0.4.

Now suppose the operator has time enough for at most two pins; how confident can he be of getting a good one within the available time? From the template, the answer is 0.8400, or 84%. What if he wants to be at least 95% confident? Again from the template, he must have enough time for four pins, because three would yield only 93.6% confidence and four yields 97.44%.

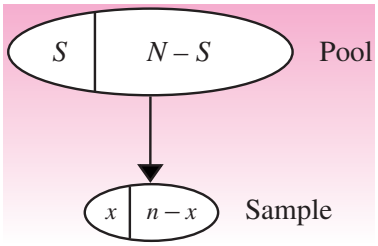
Suppose the operator wants to be 95% confident of getting a good pin by producing at most two pins. What value of p will achieve this? Using the Goal Seek command the answer is found to be 0.7761.

3–8 The Hypergeometric Distribution

CHAPTER 4

Assume that a box contains 10 pins of which 6 are good and the rest defective. An operator picks 5 pins at random from the 10, and is interested in the number of good pins picked. Let X denote the number of good pins picked. We should first note that this is a case of sampling without replacement and therefore X is *not* a binomial random variable. The probability of success p, which is the probability of picking a good pin, is neither constant nor independent from trial to trial. The first pin picked has 0.6 probability of being good; the second has either 5/9 or 6/9 probability, depending on whether or not the first was good. Therefore, X does not follow a binomial distribution, but follows what is called a **hypergeometric distribution.** In general, when a pool of size N contains S successes and $(N - S)$ failures, and a random sample of size n is drawn from the pool, the number of successes X in the sample follows

FIGURE 3–17 Schematic for Hypergeometric Distribution



a hypergeometric distribution. We shall then write $X \sim \text{HG}(n, S, N)$. The situation is depicted in Figure 3–17.

Hypergeometric Distribution Formulas

Let us derive the formula for $P(X = x)$ when X is hypergeometrically distributed. The x number of successes have to come from the S successes in the pool, which can happen in $\binom{S}{x}$ ways. The $(n - x)$ failures have to come from the $(N - S)$ failures in the pool, which can happen in $\binom{N-S}{n-x}$ ways. Together the x successes and $(n - x)$ failures can happen in $\binom{S}{x}\binom{N-S}{n-x}$ ways. Finally, there are $\binom{N}{n}$ ways of selecting a sample of size n. Putting them all together,

$$P(X = x) = \frac{\binom{S}{x}\binom{N - S}{n - x}}{\binom{N}{n}}$$

In this formula n cannot exceed N since the sample size cannot exceed the pool size. There is also a minimum possible value and a maximum possible value for x, depending on the values of n, S, and N. For instance, if $n = 9$, $S = 5$, and $N = 12$, you may verify that there would be at least two successes and at most five. In general, the minimum possible value for x is $\text{Max}(0, n - N + S)$ and the maximum possible value is $\text{Min}(n, S)$.

Hypergeometric Distribution Formulas

If $X \sim \text{HG}(n, S, N)$, then

$$P(X = x) = \frac{\binom{S}{x}\binom{N - S}{n - x}}{\binom{N}{n}} \qquad \text{Max}(0, n - N + S) \le x \le \text{Min}(n, S)$$

$$E(X) = np \qquad \text{where } p = S/N$$

$$V(X) = np(1 - p)\left[\frac{N - n}{N - 1}\right]$$

For example, if $n = 5$, $S = 6$, and $N = 10$, then

$$P(X = 2) = \frac{\binom{6}{2}\binom{10-6}{5-2}}{\binom{10}{5}} = 0.2381$$

$$E(X) = 5 * (6/10) = 3.00$$

$$V(X) = 5 * 0.6 * (1 - 0.6) * (10 - 5)/(10 - 1) = 0.6667$$

The proportion of successes in the pool, which is the ratio S/N, is the probability of the first trial being a success. This ratio is denoted by the symbol p since it resembles the p used in the binomial distribution. The expected value and variance of X are expressed using p as

$$E(X) = np$$

$$V(X) = np(1 - p)\left[\frac{N - n}{N - 1}\right]$$

Notice that the formula for $E(X)$ is the same as for the binomial case. The formula for $V(X)$ is similar to but not the same as the binomial case. The difference is the additional factor in square brackets. This additional factor approaches 1 as N becomes larger and larger compared to n and may be dropped when N is, say, 100 times as large as n. We can then approximate the hypergeometric distribution as a binomial distribution.

Problem Solving with the Template

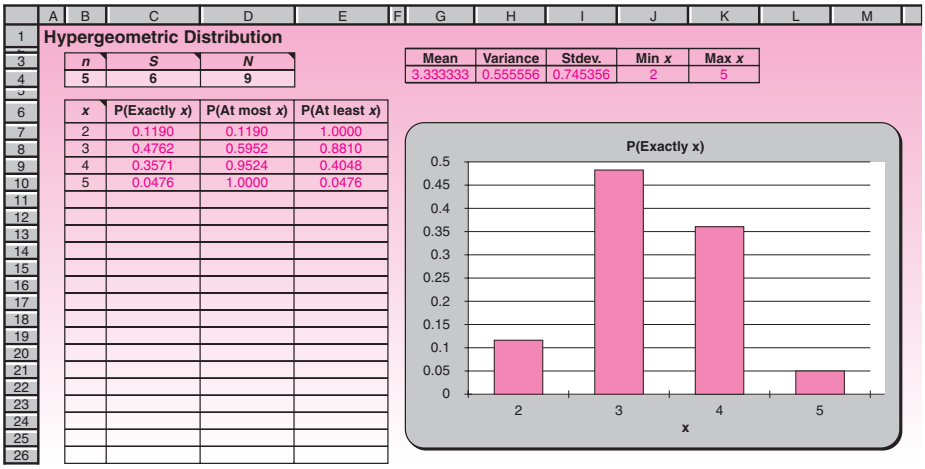
Figure 3–18 shows the template used for the hypergeometric distribution. Let us consider the case where a box contains 10 pins out of which 6 are good, and the operator picks 5 at random. What is the probability that exactly 2 good pins are picked? The answer is 0.2381 (cell C8). Additionally, the probabilities that at most two and at least two good ones are picked are, respectively, 0.2619 and 0.9762.

Suppose the operator needs at least three good pins. How confident can he be of getting at least three good pins? The answer is 0.7381 (cell E9). Suppose the operator wants to increase this confidence to 90% by adding some good pins to the pool. How many good pins should be added to the pool? This question, unfortunately, cannot be answered using the Goal Seek command for three reasons. First, the Goal Seek command works on a continuous scale, whereas S and N must be integers. Second, *when* n, S, *or* N *is changed the tabulation may shift and* P *(at least 3) may not be in cell E9!* Third, the Goal Seek command can change only one cell at a time. But in many problems, two cells (S and N) may have to change. Hence do not use the Goal Seek or the Solver on this template. Also, be careful to read the probabilities from the correct cells.

Let us solve this problem without using the Goal Seek command. If a good pin is added to the pool, what happens to S and N? They *both* increase by 1. Thus we should enter 7 for S and 11 for N. When we do, P(at least 3) = 0.8030, which is less than the desired 90% confidence. So we add one more good pin to the pool. Continuing in this fashion, we find that at least four good pins must be added to the pool.

Another way to increase P(at least 3) is to remove a bad pin from the pool. What happens to S and N when a bad pin is removed? S will remain the same and N will decrease by one. Suppose the operator wants to be 80% confident that at least three

FIGURE 3–18 The Template for the Hypergeometric Distribution [Hypergeometric.xls]



good pins will be selected. How many bad pins must be removed from the pool? Decreasing N one by one, we find that removing one bad pin is enough.

3–9 The Poisson Distribution

Imagine an automatic lathe that mass produces pins. On rare occasions, let us assume that the lathe produces a gem of a pin which is so perfect that it can be used for a very special purpose. To make the case specific, let us assume the lathe produces 20,000 pins and has 1/10,000 chance of producing a perfect one. Suppose we are interested in the number of perfect pins produced. We could try to calculate this number by using the binomial distribution with $n = 20{,}000$ and $p = 1/10{,}000$. But the calculation would be almost impossible because n is so large, p is so small, and the binomial formula calls for $n!$ and p^{n-x}, which are hard to calculate even on a computer. However, the expected number of perfect pins produced is $np = 20{,}000*(1/10{,}000) = 2$, which is neither too large nor too small. It turns out that as long as the expected value $\mu = np$ is neither too large nor too small, say, lies between 0.01 and 50, the binomial formula for $P(X = x)$ can be approximated as

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!} \qquad x = 0, 1, 2, \ldots$$

where e is the natural base of logarithms, equal to 2.71828. . . . This formula is known as the **Poisson formula,** and the distribution is called the **Poisson distribution.** In general, if we count the number of times a rare event occurs during a fixed interval, then that number would follow a Poisson distribution. We know the mean $\mu = np$.

Considering the variance of a Poisson distribution, we note that the binomial variance is $np(1 - p)$. But since p is very small, $(1 - p)$ is close to 1 and therefore can be omitted. Thus the variance of a Poisson random variable is np, which happens to be the same as its mean. The Poisson formula needs only μ, and not n or p.

We suddenly realize that we need not know n and p separately. All we need to know is their product, μ, which is the mean and the variance of the distribution. Just

one number, μ, is enough to describe the whole distribution, and in this sense, the Poisson distribution is a simple one, even simpler than the binomial. If X follows a Poisson distribution, we shall write $X \sim P(\mu)$ where μ is the expected value of the distribution. The following box summarizes the Poisson distribution.

Poisson Distribution Formulas

If $X \sim P(\mu)$, then

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!} \qquad x = 0, 1, 2, \ldots$$
$$E(X) = np = \mu$$
$$V(X) = np = \mu$$

For example, if $\mu = 2$, then

$$P(X = 3) = \frac{e^{-2}2^3}{3!} = 0.1804$$
$$E(X) = \mu = 2.00$$
$$V(X) = \mu = 2.00$$

The Poisson template is shown in Figure 3–19. The only input needed is the mean μ in cell C4. The starting value of x in cell B7 is usually zero, but it can be changed as desired.

Problem Solving with the Template

Let us return to the case of the automatic lathe that produces perfect pins on rare occasions. Assume that the lathe produces on the average two perfect pins a day, and an operator wants at least three perfect pins. What is the probability that it will produce at least three perfect pins on a given day? Looking at the template, we find the answer to be 0.3233. Suppose the operator waits for two days. In two days the lathe

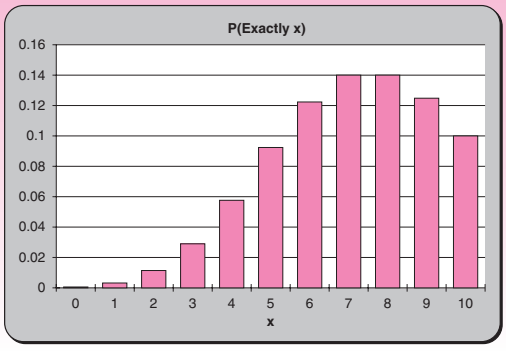
FIGURE 3–19 Poisson Distribution Template [Poisson.xls]

Poisson Distribution

Mean
8

Variance	Stdev.
8	2.8284271

x	P(Exactly x)	P(At most x)	P(At least x)
0	0.0003	0.0003	1.0000
1	0.0027	0.0030	0.9997
2	0.0107	0.0138	0.9970
3	0.0286	0.0424	0.9862
4	0.0573	0.0996	0.9576
5	0.0916	0.1912	0.9004
6	0.1221	0.3134	0.8088
7	0.1396	0.4530	0.6866
8	0.1396	0.5925	0.5470
9	0.1241	0.7166	0.4075
10	0.0993	0.8159	0.2834
11	0.0722	0.8881	0.1841
12	0.0481	0.9362	0.1119
13	0.0296	0.9658	0.0638
14	0.0169	0.9827	0.0342
15	0.0090	0.9918	0.0173
16	0.0045	0.9963	0.0082
17	0.0021	0.9984	0.0037
18	0.0009	0.9993	0.0016
19	0.0004	0.9997	0.0007



will produce on average four perfect pins. We should therefore change the mean in cell C4 to 4. What is the probability that the lathe will produce at least three perfect pins in two days? Using the template, we find the answer to be 0.7619. If the operator wants to be at least 95% confident of producing at least three perfect pins, how many days should he be prepared to wait? Again, using the template, we find that the operator should be prepared to wait at least four days.

A Poisson distribution also occurs in other types of situations leading to other forms of analysis. Consider an emergency call center. The number of distress calls received within a specific period, being a count of rare events, is usually Poisson-distributed. In this context, suppose the call center receives on average two calls per hour. In addition, suppose the crew at the center can handle up to three calls in an hour. What is the probability that the crew can handle all the calls received in a given hour? Since the crew can handle up to three calls, we look for the probability of at most three calls. From the template, the answer is 0.8571. If the crew wanted to be at least 95% confident of handling all the calls received during a given hour, how many calls should it be prepared to handle? Again, from the template, the answer is five, because the probability of at most four calls is less than 95% and of at most five calls is more than 95%.

3–10 Continuous Random Variables

Instead of depicting probability distributions by simple graphs, where the height of the line above each value represents the probability of that value of the random variable, let us use a histogram. We will associate the *area* of each rectangle of the histogram with the probability of the particular value represented. Let X be the time, measured in minutes, it takes to complete a given task. A histogram of the probability distribution of X is shown in Figure 3–20.

The probability of each value is the area of the rectangle over the value and is written on top of the rectangle. Since the rectangles all have the same base, the height of each rectangle is proportional to the probability. Note that the probabilities add to 1.00, as required. Now suppose that X can be measured more accurately. The distribution of X, with time now measured to the nearest half-minute, is shown in Figure 3–21.

Let us continue the process. Time is a continuous random variable; it can take on any value measured on an interval of numbers. We may, therefore, refine our measurement to the nearest quarter-minute, the nearest 5 seconds, or the nearest second, or we can use even more finely divided units. As we refine the measurement scale, the number of rectangles in the histogram increases and the width of each rectangle decreases. The probability of each value is still measured by the area of the rectangle above it, and the total area of all rectangles remains 1.00, as required of all probability distributions. As we keep refining our measurement scale, the discrete distribution of

FIGURE 3–20 Histogram of the Probability Distribution of Time to Complete a Task, with Time Measured to the Nearest Minute

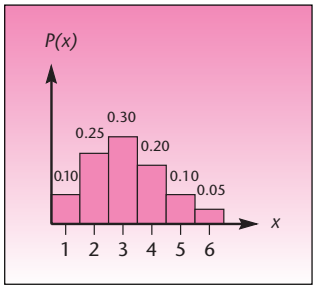


FIGURE 3–21 Histogram of the Probability Distribution of Time to Complete a Task, with Time Measured to the Nearest Half-Minute

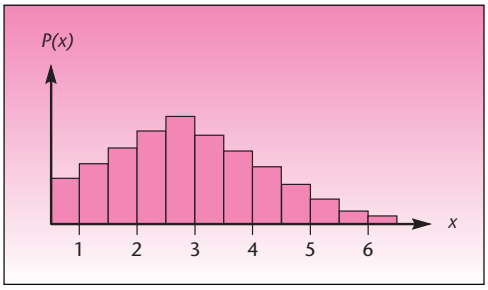
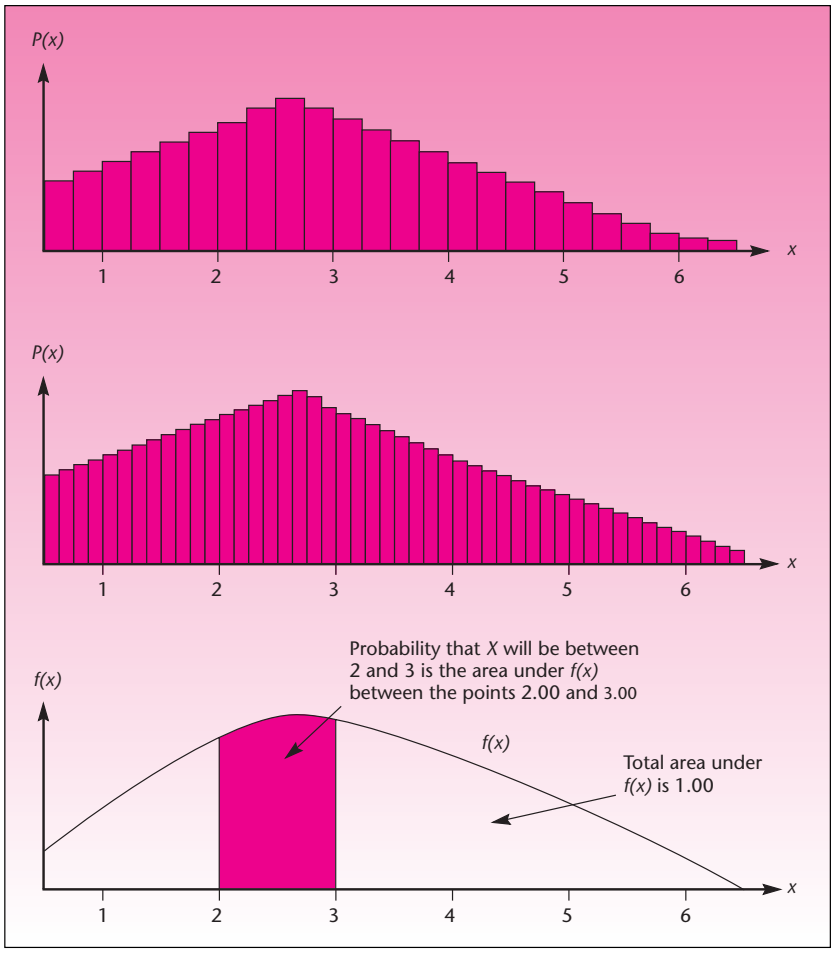


FIGURE 3–22 Histograms of the Distribution of Time to Complete a Task as Measurement Is Refined to Smaller and Smaller Intervals of Time, and the Limiting Density Function $f(x)$



X tends to a continuous probability distribution. The steplike surface formed by the tops of the rectangles in the histogram tends to a smooth function. This function is denoted by $f(x)$ and is called the **probability density function** of the continuous random variable X. Probabilities are still measured as areas under the curve. The probability that the task will be completed in 2 to 3 minutes is the area under $f(x)$ between the points $x = 2$ and $x = 3$. Histograms of the probability distribution of X with our measurement scale refined further and further are shown in Figure 3–22. Also shown is the density function $f(x)$ of the limiting continuous random variable X. The density function is the limit of the histograms as the number of rectangles approaches infinity and the width of each rectangle approaches zero.

Now that we have developed an intuitive feel for continuous random variables, and for probabilities of intervals of values as areas under a density function, we make some formal definitions.

A **continuous random variable** is a random variable that can take on any value in an interval of numbers.

The probabilities associated with a continuous random variable X are determined by the **probability density function** of the random variable. The function, denoted $f(x)$, has the following properties.

1. $f(x) \geq 0$ for all x.
2. The probability that X will be between two numbers a and b is equal to the area under $f(x)$ between a and b.
3. The total area under the entire curve of $f(x)$ is equal to 1.00.

When the sample space is continuous, the probability of any single given value is zero. For a continuous random variable, therefore, the probability of occurrence of any given value is zero. We see this from property 2, noting that the area under a curve between a point and itself is the area of a line, which is zero. *For a continuous random variable, nonzero probabilities are associated only with intervals of numbers.*

We define the cumulative distribution function $F(x)$ for a continuous random variable similarly to the way we defined it for a discrete random variable: $F(x)$ is the probability that X is less than (or equal to) x.

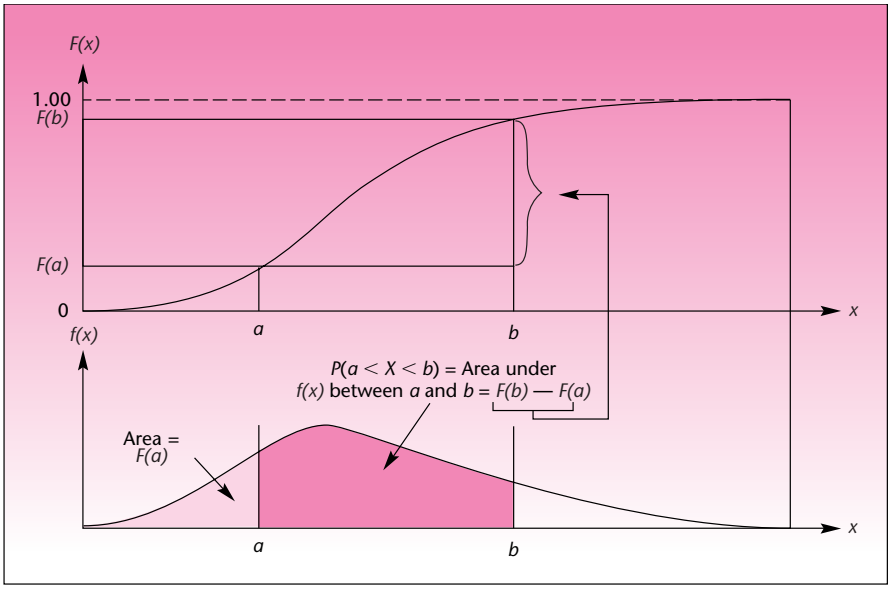
The **cumulative distribution function** of a continuous random variable:[4]

$F(x) = P(X \leq x) =$ area under $f(x)$ between the *smallest* possible value of X (often $-\infty$) and point x

The cumulative distribution function $F(x)$ is a smooth, nondecreasing function that increases from 0 to 1.00. The connection between $f(x)$ and $F(x)$ is demonstrated in Figure 3–23.

The expected value of a continuous random variable X, denoted by $E(X)$, and its variance, denoted by $V(X)$, require the use of calculus for their computation.[5]

FIGURE 3–23 Probability Density Function and Cumulative Distribution Function of a Continuous Random Variable



[4] If you are familiar with calculus, you know that the area under a curve of a function is given by the integral of the function. The probability that X will be between a and b is the definite integral of $f(x)$ between these two points: $P(a < X < b) = \int_a^b f(x)\,dx$. In calculus notation, we define the cumulative distribution function as $F(x) = \int_{-\infty}^{x} f(y)\,dy$.

[5] $E(X) = \int_{-\infty}^{\infty} x f(x)dx$; $V(X) = \int_{-\infty}^{\infty} [x - E(X)]^2 f(x)dx$.

3–11 The Uniform Distribution

The uniform distribution is the simplest of continuous distributions. The probability density function is

$$\begin{aligned} f(x) &= 1/(b-a) && a \le x \le b \\ &= 0 && \text{all other } x \end{aligned}$$

where a is the minimum possible value and b is the maximum possible value of X. The graph of $f(x)$ is shown in Figure 3–24. Because the curve of $f(x)$ is a flat line, the area under it between any two points x_1 and x_2, where $a \le x_1 < x_2 \le b$, will be a rectangle with height $1/(b-a)$ and width $(x_2 - x_1)$. Thus $P(x_1 \le X \le x_2) = (x_2 - x_1)/(b-a)$. If X is uniformly distributed between a and b, we shall write $X \sim \text{U}(a, b)$.

The mean of the distribution is the midpoint between a and b, which is $(a+b)/2$. By using integration, it can be shown that the variance is $(b-a)^2/12$. Because the shape of a uniform distribution is always a rectangle, the skewness and kurtosis are the same for all uniform distributions. The skewness is zero. (Why?) Because the shape is flat, the (relative) kurtosis is negative, always equal to -1.2.

The formulas for uniform distribution are summarized in the following box. Because the probability calculation is simple, there is no special spreadsheet function for uniform distribution. The box contains some sample calculations.

Uniform Distribution Formulas

If $X \sim \text{U}(a, b)$, then

$$\begin{aligned} f(x) &= 1/(b-a) && a \le x \le b \\ &= 0 && \text{all other } x \\ P(x_1 \le X \le x_2) &= (x_2 - x_1)/(b-a) && a \le x_1 < x_2 \le b \\ E(X) &= (a+b)/2 \\ V(X) &= (b-a)^2/12 \end{aligned}$$

For example, if $a = 10$ and $b = 20$, then

$$\begin{aligned} P(12 \le X \le 18) &= (18-12)/(20-10) = 0.6 \\ E(X) &= (10+20)/2 = 15 \\ V(X) &= (20-10)^2/12 = 8.3333 \end{aligned}$$

A common instance of uniform distribution is waiting time for a facility that goes in cycles. Two good examples are a shuttle bus and an elevator, which move, roughly, in cycles with some cycle time. If a user comes to a stop at a random time and waits till the facility arrives, the waiting time will be uniformly distributed between a minimum of zero and a maximum equal to the cycle time. In other words, if a shuttle bus has a cycle time of 20 minutes, the waiting time would be uniformly distributed between 0 and 20 minutes.

FIGURE 3–24 The Uniform Distribution

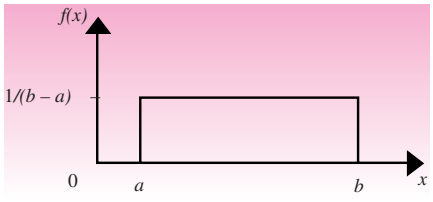
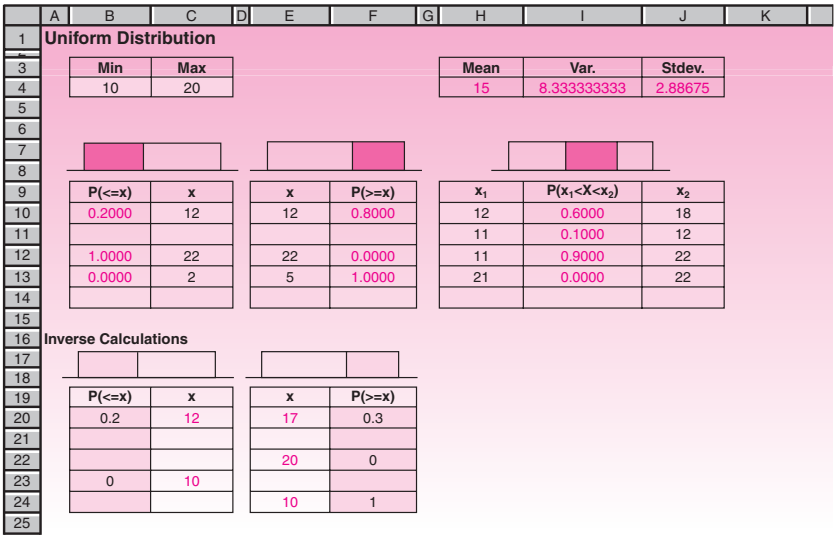


FIGURE 3–25 Template for the Uniform Distribution [Uniform.xls]



Problem Solving with the Template

Figure 3–25 shows the template for the uniform distributions. If $X \sim$ U(10, 20), what is $P(12 \leq X \leq 18)$? In the template, make sure the Min and Max are set to 10 and 20 in cells B4 and C4. Enter 12 and 18 in cells H10 and J10. The answer of 0.6 appears in cell I10.

What is the probability $P(X < 12)$? To answer this, enter 12 in cell C10. The answer 0.2 appears in cell B10. What is $P(X > 12)$? To answer this, enter 12 in cell E10. The answer 0.8 appears in F10.

Inverse calculations are possible in the bottom area of the template. Suppose you want to find x such that $P(X < x) = 0.2$. Enter 0.2 in cell B20. The answer, 12, appears in cell C20. To find x such that $P(X > x) = 0.3$, enter 0.3 in cell F20. The answer, 17, appears in cell E20.

As usual, you may also use facilities such as the Goal Seek command or the Solver tool in conjunction with this template.

3–12 The Exponential Distribution

Suppose an event occurs with an average frequency of λ occurrences per hour and this average frequency is constant in that the probability that the event will occur during any tiny duration t is λt. Suppose further we arrive at the scene at any given time and wait till the event occurs. The waiting time will then follow an **exponential distribution,** which is the continuous limit of the geometric distribution. Suppose our waiting time was x. For the event (or success) to occur at time x, every tiny duration t from time 0 to time x should be a failure and the interval x to $x + t$ must be a success. This is nothing but a geometric distribution. To get the continuous version, we take the limit of this process as t approaches zero.

The exponential distribution is fairly common in practice. Here are some examples.

1. The time between two successive breakdowns of a machine will be exponentially distributed. This information is relevant to maintenance engineers. The mean μ in this case is known as the **mean time between failures,** or **MTBF.**

2. The life of a product that fails by accident rather than by wear-and-tear follows an exponential distribution. Electronic components are good examples. This information is relevant to warranty policies.
3. The time gap between two successive arrivals to a waiting line, known as the **interarrival time,** will be exponentially distributed. This information is relevant to waiting line management.

When X is exponentially distributed with frequency λ, we shall write $X \sim E(\lambda)$. The probability density function $f(x)$ of the exponential distribution has the form

$$f(x) = \lambda e^{-\lambda x}$$

where λ is the frequency with which the event occurs. The frequency λ is expressed as so many times per unit time, such as 1.2 times per month. The mean of the distribution is $1/\lambda$ and the variance is $(1/\lambda)^2$. Just like the geometric distribution, the exponential distribution is positively skewed.

A Remarkable Property

The exponential distribution has a remarkable property. Suppose the time between two successive breakdowns of a machine is exponentially distributed with an MTBF of 100 hours, and we have just witnessed one breakdown. If we start a stopwatch as soon as it is repaired and put back into service so as to measure the time until the next failure, then that time will, of course, be exponentially distributed with a μ of 100 hours. What is remarkable is the following. Suppose we arrive at the scene at some random time and start the stopwatch (instead of starting it immediately after a breakdown); the time until next breakdown will still be exponentially distributed with the same μ of 100 hours. In other words, it is immaterial when the event occurred last and how much later we start the stopwatch. For this reason, an exponential process is known as a *memoryless process*. It does not depend on the past at all.

The Template

The template for this distribution is seen in Figure 3–26. The following box summarizes the formulas and provides example calculations.

Exponential Distribution Formulas

If $X \sim E(\lambda)$, then

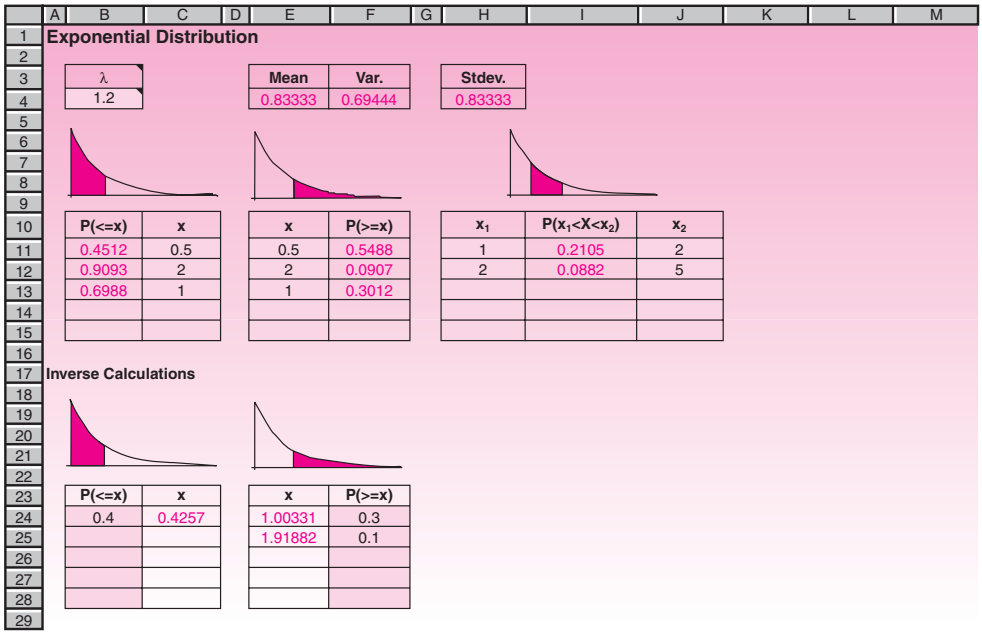
$$\begin{aligned} f(x) &= \lambda e^{-\lambda x} && x \geq 0 \\ P(X \leq x) &= 1 - e^{-\lambda x} && \text{for } x \geq 0 \\ P(X \geq x) &= e^{-\lambda x} && \text{for } x \geq 0 \\ P(x_1 \leq X \leq x_2) &= e^{-\lambda x_1} - e^{-\lambda x_2} && 0 \leq x_1 < x_2 \\ E(X) &= 1/\lambda \\ V(X) &= 1/\lambda^2 \end{aligned}$$

For example, if $\lambda = 1.2$, then

$$\begin{aligned} P(X \geq 0.5) &= e^{-1.2*0.5} = 0.5488 \\ P(1 \leq X \leq 2) &= e^{-1.2*1} - e^{-1.2*2} = 0.2105 \\ E(X) &= 1/1.2 = 0.8333 \\ V(X) &= 1/1.2^2 = 0.6944 \end{aligned}$$

To use the exponential distribution template seen in Figure 3–26, the value of λ must be entered in cell B4. At times, the mean μ rather than λ may be known, in which case its reciprocal $1/\mu$ is what should be entered as λ in cell B4. Note that λ is

FIGURE 3–26 Exponential Distribution Template [Exponential.xls]



Exponential Distribution

λ	Mean	Var.	Stdev.
1.2	0.83333	0.69444	0.83333

P(<=x)	x
0.4512	0.5
0.9093	2
0.6988	1

x	P(>=x)
0.5	0.5488
2	0.0907
1	0.3012

x_1	$P(x_1<X<x_2)$	x_2
1	0.2105	2
2	0.0882	5

Inverse Calculations

P(<=x)	x
0.4	0.4257

x	P(>=x)
1.00331	0.3
1.91882	0.1

the average number of occurrences of a rare event in unit time and μ is the average time gap between two successive occurrences. The shaded cells are the input cells and the rest are protected. As usual, the Goal Seek command and the Solver tool can be used in conjunction with this template to solve problems.

EXAMPLE 3–5

A particular brand of handheld computers fails following an exponential distribution with a μ of 54.82 months. The company gives a warranty for 6 months.

a. What percentage of the computers will fail within the warranty period?

b. If the manufacturer wants only 8% of the computers to fail during the warranty period, what should be the average life?

Solution

a. Enter the reciprocal of 54.82 = 0.0182 as λ in the template. (You may enter the formula "=1/54.82" in the cell. But then you will not be able to use the Goal Seek command to change this entry. The Goal Seek command requires that the changing cell contain a number rather than a formula.) The answer we are looking for is the area to the left of 6. Therefore, enter 6 in cell C11. The area to the left, 0.1037, appears in cell B11. Thus 10.37% of the computers will fail within the warranty period.

b. Enter 0.08 in cell B25. Invoke the Goal Seek command to set cell C25 to the value of 6 by changing cell B4. The λ value in cell B4 reaches 0.0139, which corresponds to a μ value of 71.96 months, as seen in cell E4. Therefore, the average life of the computers must be 71.96 months.

Value at Risk

When a business venture involves chances of large losses, a measure of risk that many companies use is the **value at risk.** Suppose the profit from a venture has a negatively

FIGURE 3–27 Distribution of Profit Showing Value at Risk

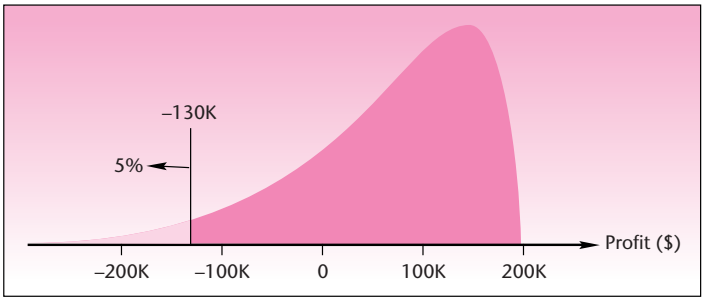
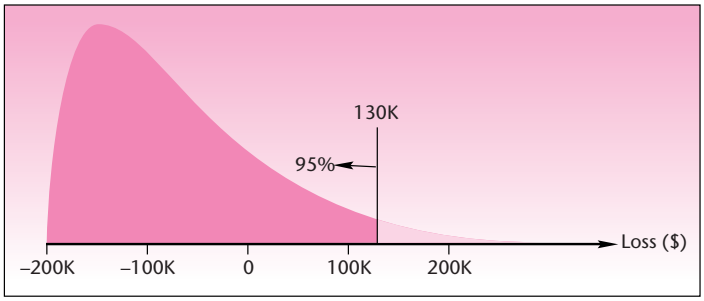


FIGURE 3–28 Distribution of Loss Showing Value at Risk



skewed distribution, shown in Figure 3–27. A negative profit signifies loss. The distribution shows that large losses are possible. A common definition of value at risk is the amount of loss at the 5th percentile of the distribution. In Figure 3–27, the 5th percentile is $–130,000, meaning a loss of $130,000. Thus the value at risk is $130,000.

If the profit is a discrete random variable, then the percentile used may be a convenient one closest to 5%.

If the distribution of loss rather than profit is plotted, then we will have the mirror image of Figure 3–27, which is shown in Figure 3–28. In this case, the value at risk is the 95th percentile.

Keep in mind that value at risk applies only to distributions of profit/loss where there exist small chances of large losses.

3–13 Using the Computer

Using Excel Formulas for Some Standard Distributions

Excel has built-in functions that you may use to calculate certain probabilities without using templates. These formulas are described in this section.

You can use **BINOMDIST** to obtain the individual term binomial distribution probability. In the formula `BINOMDIST(x, n, p, cumulative)`, x is the number of successes in trials, n is the number of independent trials, p is the probability of success on each trial, and cumulative is a logical value that determines the form of the function. If cumulative is `TRUE`, then BINOMDIST returns the cumulative distribution function, which is the probability that there are at most x successes; if `FALSE`, it returns the probability mass function, which is the probability that there are x successes.

NEGBINOMDIST returns the negative binomial distribution. By using the formula `NEGBINOMDIST(f, s, p)` you can obtain the probability that there will be f failures before the sth success, when the constant probability of a success is p. As we can see, the conventions for negative binomial distributions are slightly different in Excel. We have used the symbol x in this chapter to denote the total number of trials until the sth success is achieved, but in the Excel formula we count the total number of failures before the sth success. For example, `NEGBINOMDIST (3,2,0.5)` will return the probability of three failures before the 2nd success, which is the same as probability of 5 trials before the 2nd success. It returns the value 0.0922.

No function is available for geometric distribution per se, but the negative binomial formula can be used with `s=1`. For example, the geometric probability of 5 trials when `p=0.6` can be computed using the formula `NEGBINOMDIST(4,1,0.6)`. It returns the value of 0.2381.

HYPGEOMDIST returns the hypergeometric distribution. Using the formula `HYPGEOMDIST(x, n, s, N)` you can obtain the probability of x success in a random sample of size n, when the population has s success and size N. For example, the formula `HYPGEOMDIST(2,5,6,10)` will return a value of 0.2381.

POISSON returns the Poisson distribution. In the formula `POISSON(x, mean, cumulative)`, x is the number of events, *mean* is the expected numeric value and *cumulative* is a logical value that determines the form of the probability distribution returned. If cumulative is `TRUE`, POISSON returns the cumulative Poisson probability that the number of random events occurring will be between zero and x inclusive; if `FALSE`, it returns the Poisson probability mass function that the number of events occurring will be exactly x.

EXPONDIST returns the exponential distribution. In the formula `EXPONDIST (x, lambda, cumulative)`, x is the value of the function, *lambda* is the parameter value, and *cumulative* is a logical value that indicates which form of the exponential function to provide. If cumulative is `TRUE`, EXPONDIST returns the cumulative distribution function; if `FALSE`, it returns the probability density function. For example, `EXPONDIST (0.5, 1.2, TRUE)` will return the cumulative exponential probability $P(X < x)$, which is 0.4512, while `EXPONDIST(0.5, 1.2, FALSE)` will return the exponential probability density function $f(x)$, which we do not need for any practical purpose.

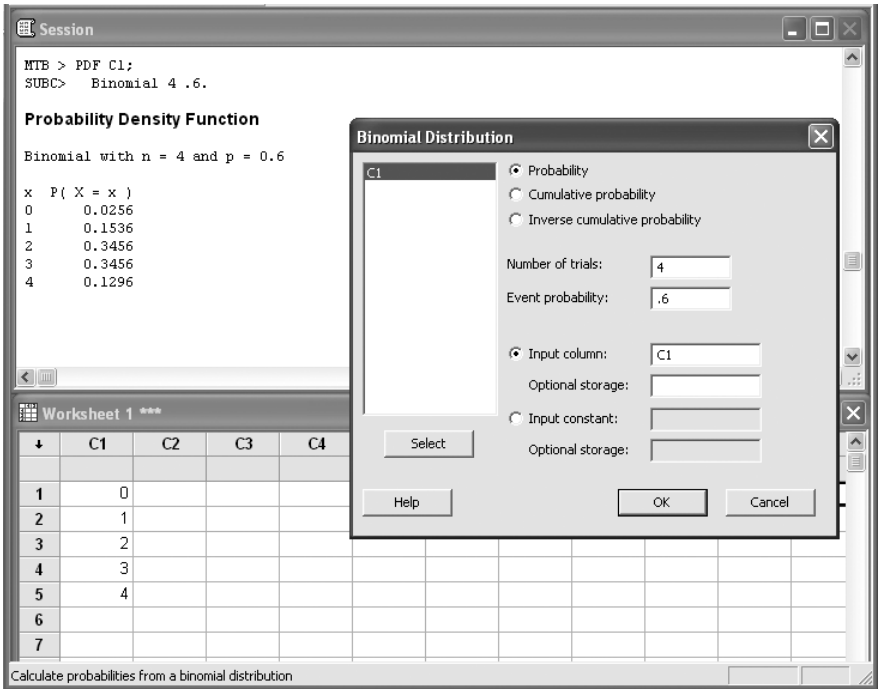
No probability function is available for the uniform distribution but the probability formulas are simple enough to compute manually.

Using MINITAB for Some Standard Distributions

In this section we will demonstrate how we can use MINITAB to obtain the probability density function or cumulative distribution function of various random variables.

Start by choosing Calc ► Probability Distributions from the menu. This option will display commands that allow you to compute probability densities and cumulative probabilities for continuous and discrete distributions. For example when you select Calc ► Probability Distributions ► Binomial, the Binomial Distribution dialog box will appear. From the items available in the dialog box, you can choose to calculate probabilities or cumulative probabilities. You also need to specify the parameters of the binomial distribution, which are number of trials and event probability. In the input section the values for which you aim to obtain probability densities or cumulative probabilities are specified. These values can be a constant or a set of values that have been defined in a column. Then press OK to observe the obtained result in the Session window. Figure 3–29 shows how MINITAB has been used for obtaining probability distributions for a binomial distribution with parameters 4 and 0.6. The final result and corresponding session commands are presented in the session window.

FIGURE 3–29 Using MINITAB for Generating a Binomial Distribution



3–14 Summary and Review of Terms

In this chapter we described several important standard random variables, the associated formulas, and problem solving with spreadsheets. In order to use a spreadsheet template, you need to know *which* template to use, but first you need to know the kind of random variable at hand. This summary concentrates on this question.

A **discrete random variable** X will follow a **binomial distribution** if it is the number of successes in n independent **Bernoulli trials.** Make sure that the probability of success, p, remains constant in all trials. X will follow a **negative binomial distribution** if it is the number of Bernoulli trials made to achieve a desired number of successes. It will follow a **geometric distribution** when the desired number of successes is one. X will follow a **hypergeometric distribution** if it is the number of successes in a random sample drawn from a finite pool of successes and failures. X will follow a **Poisson distribution** if it is the number of occurrences of a rare event during a finite period.

Waiting time for an event that occurs periodically is **uniformly distributed.** Waiting time for a rare event is **exponentially distributed.**

ADDITIONAL PROBLEMS

3–42. An investment portfolio has equal proportions invested in five stocks. The expected returns and standard deviations (both in percent per year) are (8, 3), (5, 2), (12, 8), (7, 9), (14, 15). What are average return and standard deviation for this portfolio?

3–43. A graduating student keeps applying for jobs until she has three offers. The probability of getting an offer at any trial is 0.48.

a. What is the expected number of applications? What is the variance?

b. If she has enough time to complete only six applications, how confident can she be of getting three offers within the available time?

c. If she wants to be at least 95% confident of getting three offers, how many applications should she prepare?

d. Suppose she has time for at most six applications. For what minimum value of p can she still have 95% confidence of getting three offers within the available time?

3–44. A real estate agent has four houses to sell before the end of the month by contacting prospective customers one by one. Each customer has an independent 0.24 probability of buying a house on being contacted by the agent.

a. If the agent has enough time to contact only 15 customers, how confident can she be of selling all four houses within the available time?

b. If the agent wants to be at least 70% confident of selling all the houses within the available time, at least how many customers should she contact? (If necessary, extend the template downward to more rows.)

c. What minimum value of p will yield 70% confidence of selling all four houses by contacting at most 15 customers?

d. To answer (*c*) above more thoroughly, tabulate the confidence for p values ranging from 0.2 to 0.6 in steps of 0.05.

3–45. A graduating student keeps applying for jobs until she gets an offer. The probability of getting an offer at any trial is 0.35.

a. What is the expected number of applications? What is the variance?

b. If she has enough time to complete at most four applications, how confident can she be of getting an offer within the available time?

c. If she wants to be at least 95% confident of getting an offer, how many applications should she prepare?

d. Suppose she has time for at most four applications. For what minimum value of p can she have 95% confidence of getting an offer within the available time?

3–46. A shipment of pins contains 25 good ones and 2 defective ones. At the receiving department, an inspector picks three pins at random and tests them. If any defective pin is found among the three that are tested, the shipment would be rejected.

a. What is the probability that the shipment would be accepted?

b. To increase the probability of acceptance to at least 90%, it is decided to do one of the following:

i. Add some good pins to the shipment.

ii. Remove some defective pins in the shipment.

For each of the two options, find out exactly how many pins should be added or removed.

3–47. A committee of 7 members is to be formed by selecting members at random from a pool of 14 candidates consisting of 5 women and 9 men.

a. What is the probability that there will be at least three women in the committee?

b. It is desired to increase the chance that there are at least three women in the committee to 80% by doing one of the following:

i. Adding more women to the pool.

ii. Removing some men from the pool.

For each of the two options, find out how many should be added or removed.

3–48. A mainframe computer in a university crashes on the average 0.71 time in a semester.

a. What is the probability that it will crash at least two times in a given semester?

b. What is the probability that it will not crash at all in a given semester?

c. The MIS administrator wants to increase the probability of no crash at all in a semester to at least 90%. What is the largest μ that will achieve this goal?

3–49. The number of rescue calls received by a rescue squad in a city follows a Poisson distribution with $\mu = 2.83$ per day. The squad can handle at most four calls a day.

a. What is the probability that the squad will be able to handle all the calls on a particular day?

b. The squad wants to have at least 95% confidence of being able to handle all the calls received in a day. At least how many calls a day should the squad be prepared for?

c. Assuming that the squad can handle at most four calls a day, what is the largest value of μ that would yield 95% confidence that the squad can handle all calls?

3–50. A student takes the campus shuttle bus to reach the classroom building. The shuttle bus arrives at his stop every 15 minutes but the actual arrival time at the stop is random. The student allows 10 minutes waiting time for the shuttle in his plan to make it in time to the class.

a. What is the expected waiting time? What is the variance?

b. What is the probability that the wait will be between four and six minutes?

c. What is the probability that the student will be in time for the class?

d. If he wants to be 95% confident of being on time for the class, how much time should he allow for waiting for the shuttle?

3–51. A hydraulic press breaks down at the rate of 0.1742 time per day.

a. What is the MTBF?

b. On a given day, what is the probability that it will break down?

c. If four days have passed without a breakdown, what is the probability that it will break down on the fifth day?

d. What is the probability that five consecutive days will pass without any breakdown?

3–52. Laptop computers produced by a company have an average life of 38.36 months. Assume that the life of a computer is exponentially distributed (which is a good assumption).

a. What is the probability that a computer will fail within 12 months?

b. If the company gives a warranty period of 12 months, what proportion of computers will fail during the warranty period?

c. Based on the answer to (*b*), would you say the company can afford to give a warranty period of 12 months?

d. If the company wants not more than 5% of the computers to fail during the warranty period, what should be the warranty period?

e. If the company wants to give a warranty period of three months and still wants not more than 5% of the computers to fail during the warranty period, what should be the minimum average life of the computers?

3–53. In most statistics textbooks, you will find cumulative binomial probability tables in the format shown below. These can be created using spreadsheets using the Binomial template and Data | Table commands.

$n = 5$		p								
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
x	0	0.5905	0.3277	0.1681	0.0778	0.0313	0.0102	0.0024	0.0003	0.0000
	1	0.9185	0.7373	0.5282	0.3370	0.1875	0.0870	0.0308	0.0067	0.0005
	2	0.9914	0.9421	0.8369	0.6826	0.5000	0.3174	0.1631	0.0579	0.0086
	3	0.9995	0.9933	0.9692	0.9130	0.8125	0.6630	0.4718	0.2627	0.0815
	4	1.0000	0.9997	0.9976	0.9898	0.9688	0.9222	0.8319	0.6723	0.4095

a. Create the above table.

b. Create a similar table for $n = 7$.

3–54. Look at the shape of the binomial distribution for various combinations of n and p. Specifically, let $n = 5$ and try $p = 0.2$, 0.5, and 0.8. Repeat the same for other values of n. Can you say something about how the skewness of the distribution is affected by p and n?

3–55. Try various values of s and p on the negative binomial distribution template and answer this question: How is the skewness of the negative binomial distribution affected by s and p values?

3–56. An MBA graduate keeps interviewing for jobs, one by one, and will stop interviewing on receiving an offer. In each interview he has an independent probability 0.2166 of getting the job.

a. What is the expected number of interviews? What is the variance?

b. If there is enough time for only six interviews, how confident can he be of getting a job within the available time?

c. If he wants to be at least 95% confident of getting a job, how many interviews should he be prepared for?

d. Suppose there is enough time for at most six interviews. For what minimum value of p can he have 95% confidence of getting a job within the available time?

e. In order to answer (*d*) more thoroughly, tabulate the confidence level for p values ranging from 0.1 to 0.5 in steps of 0.05.

3–57. A shipment of thousands of pins contains some percentage of defectives. To decide whether to accept the shipment, the consumer follows a sampling plan where 80 items are chosen at random from the sample and tested. If the number of defectives in the sample is at most three, the shipment is accepted. (The number 3 is known as the *acceptance number* of the sampling plan.)

a. Assuming that the shipment includes 3% defectives, what is the probability that the shipment will be accepted? (*Hint:* Use the binomial distribution.)

b. Assuming that the shipment includes 6% defectives, what is the probability that the shipment will be accepted?

c. Using the DataITable command, tabulate the probability of acceptance for defective percentage ranging from 0% to 15% in steps of 1%.

d. Plot a line graph of the table created in (*c*). (This graph is known as the *operating characteristic curve* of the sampling plan.)

3–58. A shipment of 100 pins contains some defectives. To decide whether to accept the shipment, the consumer follows a sampling plan where 15 items are chosen at random from the sample and tested. If the number of defectives in the sample is at most one, the shipment is accepted. (The number 1 is known as the *acceptance number* of the sampling plan.)

a. Assuming that the shipment includes 5% defectives, what is the probability that the shipment will be accepted? (*Hint:* Use the hypergeometric distribution.)

b. Assuming that the shipment includes 8% defectives, what is the probability that the shipment will be accepted?

c. Using the DataITable command, tabulate the probability of acceptance for defective percentage ranging from 0% to 15% in steps of 1%.

d. Plot a line graph of the table created in part (*c*) above. (This graph is known as the *operating characteristic curve* of the sampling plan.)

3–59. A recent study published in the *Toronto Globe and Mail* reveals that 25% of mathematics degrees from Canadian universities and colleges are awarded to women. If five recent graduates from Canadian universities and colleges are selected at random, what is the probability that

a. At least one would be a woman.

b. None of them would be a woman.

3–60. An article published in *Access* magazine states that according to a survey conducted by the American Management Association, 78% of major U.S. companies electronically monitor their employees. If five such companies are selected at random, find the probability that

a. At most one company monitors its employees electronically.

b. All of them monitor their employees electronically.

3–61. An article published in *BusinessWeek* says that according to a survey by a leading organization 45% of managers change jobs for intellectual challenge, 35% for pay, and 20% for long-term impact on career. If nine managers who recently changed jobs are randomly chosen, what is the probability that

a. Three changed for intellectual challenges.

b. Three changed for pay reasons.

c. Three changed for long-term impact.

3–62. Estimates published by the World Health Organization state that one out of every three workers may be toiling away in workplaces that make them sick. If seven workers are selected at random, what is the probability that a majority of them are made sick by their workplace?

3–63. Based on the survey conducted by a municipal administration in the Netherlands, Monday appeared to be managements' preferred day for laying off workers. Of the total number of workers laid off in a given period, 30% were on Monday, 25% on Tuesday, 20% on Wednesday, 13% on Thursday, and 12% on Friday. If a random sample of 15 layoffs is taken, what is the probability that

a. Five were laid off on Monday.

b. Four were laid off on Tuesday.

c. Three were laid off on Wednesday.

d. Two were laid off on Thursday.

e. One was laid off on Friday.

3–64. A recent survey published in *BusinessWeek* concludes that Gatorade commands an 83% share of the sports drink market versus 11% for Coca-Cola's PowerAde and 3% for Pepsi's All Sport. A market research firm wants to conduct a new taste test for which it needs Gatorade drinkers. Potential participants for the test are selected by random screening of drink users to find Gatorade drinkers. What is the probability that

a. The first randomly selected drinker qualifies.

b. Three soft drink users will have to be interviewed to find the first Gatorade drinker.

3–65. The time between customer arrivals at a bank has an exponential distribution with a mean time between arrivals of three minutes. If a customer just arrived, what is the probability that another customer will not arrive for at least two minutes?

3–66. Lightbulbs manufactured by a particular company have an exponentially distributed life with mean 100 hours.

a. What is the probability that the lightbulb I am now putting in will last at least 65 hours?

b. What is the standard deviation of the lifetime of a lightbulb?

3–67. The Bombay Company offers reproductions of classic 18th- and 19th-century English furniture pieces, which have become popular in recent years. The following table gives the probability distribution of the number of Raffles tables sold per day at a particular Bombay store.

Number of Tables	Probability
0	0.05
1	0.05
2	0.10
3	0.15
4	0.20
5	0.15
6	0.15
7	0.10
8	0.05

a. Show that the probabilities above form a proper probability distribution.

b. Find the cumulative distribution function of the number of Raffles tables sold daily.

c. Using the cumulative distribution function, find the probability that the number of tables sold in a given day will be at least three and less than seven.

d. Find the probability that at most five tables will be sold tomorrow.

e. What is the expected number of tables sold per day?

f. Find the variance and the standard deviation of the number of tables sold per day.

g. Use Chebyshev's theorem to determine bounds of at least 0.75 probability on the number of tables sold daily. Compare with the actual probability for these bounds using the distribution itself.

3–68. According to an article in *USA Today,* 90% of Americans will suffer from high blood pressure as they age. Out of 20 randomly chosen people what is the probability that at most 3 will suffer from high blood pressure?

3–69. The number of orders for installation of a computer information system arriving at an agency per week is a random variable X with the following probability distribution:

x	$P(x)$
0	0.10
1	0.20
2	0.30
3	0.15
4	0.15
5	0.05
6	0.05

a. Prove that $P(X)$ is a probability distribution.

b. Find the cumulative distribution function of X.

c. Use the cumulative distribution function to find probabilities $P(2 < X \leq 5)$, $P(3 \leq X \leq 6)$, and $P(X > 4)$.

d. What is the probability that either four or five orders will arrive in a given week?

e. Assuming independence of weekly orders, what is the probability that three orders will arrive next week and the same number of orders the following week?

f. Find the mean and the standard deviation of the number of weekly orders.

3–70. Consider the situation in the previous problem, and assume that the distribution holds for all weeks throughout the year and that weekly orders are independent from week to week. Let Y denote the number of weeks in the year in which no orders are received (assume a year of 52 weeks).

a. What kind of random variable is Y? Explain.

b. What is the expected number of weeks with no orders?

3–71. An analyst kept track of the daily price quotation for a given stock. The frequency data led to the following probability distribution of daily stock price:

Price x in Dollars	$P(x)$
17	0.05
17.125	0.05
17.25	0.10
17.375	0.15
17.5	0.20
17.625	0.15
17.75	0.10
17.875	0.05
18	0.05
18.125	0.05
18.25	0.05

Assume that the stock price is independent from day to day.

a. If 100 shares are bought today at 17 1/4 and must be sold tomorrow, by prearranged order, what is the expected profit, disregarding transaction costs?

b. What is the standard deviation of the stock price? How useful is this information?

c. What are the limitations of the analysis in part (*a*)? Explain.

3–72. In problem 3–69, suppose that the company makes \$1,200 on each order but has to pay a fixed weekly cost of \$1,750. Find the expected weekly profit and the standard deviation of weekly profits.

3–73. Out of 140 million cellular telephone subscribers in the United States, 36 million use Verizon.[6]

a. Ten wireless customers are chosen. Under what conditions is the number of Verizon customers a binomial random variable?

b. Making the required assumptions above, find the probability that at least two are Verizon customers.

3–74. An advertisement claims that two out of five doctors recommend a certain pharmaceutical product. A random sample of 20 doctors is selected, and it is found that only 2 of them recommend the product.

a. Assuming the advertising claim is true, what is the probability of the observed event?

b. Assuming the claim is true, what is the probability of observing two or fewer successes?

c. Given the sampling results, do you believe the advertisement? Explain.

d. What is the expected number of successes in a sample of 20?

3–75. Five percent of the many cars produced at a plant are defective. Ten cars made at the plant are sent to a dealership. Let X be the number of defective cars in the shipment.

a. Under what conditions can we assume that X is a binomial random variable?

b. Making the required assumptions, write the probability distribution of X.

c. What is the probability that two or more cars are defective?

d. What is the expected number of defective cars?

3–76. Refer to the situation in the previous problem. Suppose that the cars at the plant are checked one by one, and let X be the number of cars checked until the first defective car is found. What type of probability distribution does X have?

3–77. Suppose that 5 of a total of 20 company accounts are in error. An auditor selects a random sample of 5 out of the 20 accounts. Let X be the number of accounts in the sample that are in error. Is X binomial? If not, what distribution does it have? Explain.

3–78. The time, in minutes, necessary to perform a certain task has the uniform [5, 9] distribution.

a. Write the probability density function of this random variable.

b. What is the probability that the task will be performed in less than 8 minutes? Explain.

c. What is the expected time required to perform the task?

3–79. Suppose X has the following probability density function:

$$f(x) = \begin{cases} (1/8)(x - 3) & \text{for } 3 \le x \le 7 \\ 0 & \text{otherwise} \end{cases}$$

a. Graph the density function.

b. Show that $f(x)$ is a density function.

c. What is the probability that X is greater than 5.00?

3–80. Recently, the head of the Federal Deposit Insurance Corporation (FDIC) revealed that the agency maintains a secret list of banks suspected of being in financial trouble. The FDIC chief further stated that of the nation's 14,000 banks, 1,600 were on the list at the time. Suppose that, in an effort to diversify your savings, you randomly choose six banks and split your savings among them. What is the probability that no more than three of your banks are on the FDIC's suspect list?

---

[6]Matt Richtel and Andrew Ross Sorkin, "AT&T Wireless for Sale as a Shakeout Starts," *The New York Times,* January 21, 2004, p. C1.

3–81. Corporate raider Asher Adelman, teaching a course at Columbia University's School of Business, made the following proposal to his students. He would pay $100,000 to any student who would give him the name of an undervalued company, which Adelman would then buy.[7] Suppose that Adelman has 15 students in his class and that 5% of all companies in this country are undervalued. Suppose also that due to liquidity problems, Adelman can give the award to at most three students. Finally, suppose each student chooses a single company at random without consulting others. What is the probability that Adelman would be able to make good on his promise?

3–82. An applicant for a faculty position at a certain university is told by the department chair that she has a 0.95 probability of being invited for an interview. Once invited for an interview, the applicant must make a presentation and win the votes of a majority (at least 8) of the department's 14 current members. From previous meetings with four of these members, the candidate believes that three of them would certainly vote for her while one would not. She also feels that any member she has not yet met has a 0.50 probability of voting for her. Department members are expected to vote independently and with no prior consultation. What are the candidate's chances of getting the position?

3–83. The ratings of viewership for the three major networks during prime time recently were as follows. Also shown is the proportion of viewers watching each program.

Program	Network	Rating	Proportion
20/20	ABC	13.8	0.44
CSI	CBS	10.4	0.33
Law and Order	NBC	7.5	0.23

a. What is the mean rating given a program that evening?

b. How many standard deviations above or below the mean is the rating for each one of the programs?

3–84. A major ski resort in the eastern United States closes in late May. Closing day varies from year to year depending on when the weather becomes too warm for making and preserving snow. The day in May and the number of years in which closing occurred that day are reported in the table:

Day	Number of Years
21	2
22	5
23	1
24	3
25	3
26	1
27	2
28	1

a. Based only on this information, estimate the probability that you could ski at this resort after May 25 next year.

b. What is the average closing day based on history?

3–85. Ten percent of the items produced at a plant are defective. A random sample of 20 items is selected. What is the probability that more than three items in the sample are defective? If items are selected randomly until the first defective item is encountered, how many items, on average, will have to be sampled before the first defective item is found?

[7]Columbia has since questioned this offer on ethical grounds, and the offer has been retracted.

3–86. Lee Iacocca volunteered to drive one of his Chryslers into a brick wall to demonstrate the effectiveness of airbags used in these cars. Airbags are known to activate at random when the car decelerates anywhere from 9 to 14 miles per hour per second (mph/s). The probability distribution for the deceleration speed at which bags activate is given below.

mph/s	Probability
9	0.12
10	0.23
11	0.34
12	0.21
13	0.06
14	0.04

a. If the airbag activates at a deceleration of 12 mph/s or more, Iacocca would get hurt. What is the probability of his being hurt in this demonstration?

b. What is the mean deceleration at airbag activation moment?

c. What is the standard deviation of deceleration at airbag activation time?

3–87. In the previous problem, the time that it takes the airbag to completely fill up from the moment of activation has an exponential distribution with mean 1 second. What is the probability that the airbag will fill up in less than 1/2 second?

3–88. The time interval between two successive customers entering a store in a mall is exponentially distributed with a mean of 6.55 seconds.

a. What is the probability that the time interval is more than 10 seconds?

b. What is the probability that the time interval is between 10 and 20 seconds?

c. On a particular day a security camera is installed. Using an entry sensor, the camera takes pictures of every customer entering the shop. It needs 0.75 second after a picture is taken to get ready for the next picture. What is the probability that the camera will miss an entering customer?

d. How quick should the camera be if the store owner wants to photograph at least 95% of entering customers?

3–89. The Dutch consumer-electronics giant, Philips, is protected against takeovers by a unique corporate voting structure that gives power only to a few trusted shareholders. A decision of whether to sever Philips' links with the loss-producing German electronics firm Grundig had to be made. The decision required a simple majority of nine decision-making shareholders. If each is believed to have a 0.25 probability of voting yes on the issue, what is the probability that Grundig will be dumped?

3–90. According to a front-page article in *The Wall Street Journal,* 30% of all students in American universities miss classes due to drinking.[8] If 10 students are randomly chosen, what is the probability that at most 3 of them miss classes due to drinking?

3–91. According to an article in *USA Today,* 60% of 7- to 12-year-olds who use the Internet do their schoolwork on line.[9] If 8 kids within this age group who use the Internet are randomly chosen, what is the probability that 2 of them do their schoolwork on line? What is the probability that no more than 5 of them do their schoolwork on line?

[8]Bryan Gruley, "How One University Stumbled in Its Attack on Alcohol Abuse," *The Wall Street Journal,* October 14, 2003, p. 1A.

[9]Ruth Peters, "Internet: Boon or Bane for Kids?" *USA Today,* October 15, 2003, p. 19A.

3–92. The cafeteria in a building offers three different lunches. The demands for the three types of lunch on any given day are independent and Poisson distributed with means 4.85, 12.70, and 27.61. The cost of the three types are $12.00, $8.50, and $6.00, respectively. Find the expected value and variance of the total cost of lunches bought on a particular day.

3–93. The mean time between failures (MTBF) of a hydraulic press is to be estimated assuming that the time between failures (TBF) is exponentially distributed. A foreman observes that the chance that the TBF is more than 72 hours is 50%, and he quotes 72 hours as the MTBF.

a. Is the foreman right? If not, what is the MTBF?

b. If the MTBF is indeed 72 hours, 50% of the time the TBF will be more than how many hours?

c. Why is the mean of an exponential distribution larger than its median?

3–94. An operator needs to produce 4 pins and 6 shafts using a lathe which has 72% chance of producing a defect-free pin at each trial and 65% chance of producing a defect-free shaft at each trial. The operator will first produce pins one by one until he has 4 defect-free pins and then produce shafts one by one until he has 6 defect-free shafts.

a. What is the expected value and variance of the total number of trials that the operator will make?

b. Suppose each trial for pins takes 12 minutes and each trial for shafts takes 25 minutes. What is the expected value and variance of the total time required?

CASE 3 Concepts Testing

Hedge funds are institutions that invest in a wide variety of instruments, from stocks and bonds to commodities and real estate. One of the reasons for the success of this industry is that it manages expected return and risk better than other financial institutions. Using the concepts and ideas described in this chapter, discuss how a hedge fund might maximize expected return and minimize risk by investing in various financial instruments. Include in your discussion the concepts of means and variances of linear composites of random variables and the concept of independence.