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

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

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

11. A technician services mailing machines at companies in the Phoenix area. Depending on the type of malfunction, the service call can take 1, 2, 3, or 4 hours. The different types of malfunctions occur at about the same frequency.
a. Develop a probability distribution for the duration of a service call.
b. Draw a graph of the probability distribution.
c. Show that your probability distribution satisfies the conditions required for a discrete probability function.
d. What is the probability a service call will take 3 hours?
e. A service call has just come in, but the type of malfunction is unknown. It is 3:00 P.M. and service technicians usually get off at 5:00 P.M. What is the probability the service technician will have to work overtime to fix the machine today?

12. Time Warner Cable provides television and Internet service to millions of customers. Suppose that the management of Time Warner Cable subjectively assesses a probability distribution for the number of new subscribers next year in the state of New York as follows.

x	$f(x)$
100,000	.10
200,000	.20
300,000	.25
400,000	.30
500,000	.10
600,000	.05

a. Is this probability distribution valid? Explain.
b. What is the probability Time Warner will obtain more than 400,000 new subscribers?
c. What is the probability Time Warner will obtain fewer than 200,000 new subscribers?

13. A psychologist determined that the number of sessions required to obtain the trust of a new patient is either 1, 2, or 3. Let x be a random variable indicating the number of sessions required to gain the patient's trust. The following probability function has been proposed.

$$f(x) = \frac{x}{6} \qquad \text{for } x = 1, 2, \text{ or } 3$$

a. Is this probability function valid? Explain.
b. What is the probability that it takes exactly 2 sessions to gain the patient's trust?
c. What is the probability that it takes at least 2 sessions to gain the patient's trust?

14. The following table is a partial probability distribution for the MRA Company's projected profits (x = profit in \$1000s) for the first year of operation (the negative value denotes a loss).

x	$f(x)$
−100	.10
0	.20
50	.30
100	.25
150	.10
200	

a. What is the proper value for $f(200)$? What is your interpretation of this value?
b. What is the probability that MRA will be profitable?
c. What is the probability that MRA will make at least \$100,000?

5.3 Expected Value and Variance

Expected Value

The **expected value**, or mean, of a random variable is a measure of the central location for the random variable. The formula for the expected value of a discrete random variable x follows.

The expected value is a weighted average of the values of the random variable where the weights are the probabilities.

EXPECTED VALUE OF A DISCRETE RANDOM VARIABLE

$$E(x) = \mu = \sum xf(x) \tag{5.4}$$

Both the notations $E(x)$ and μ are used to denote the expected value of a random variable.

Equation (5.4) shows that to compute the expected value of a discrete random variable, we must multiply each value of the random variable by the corresponding probability $f(x)$ and then add the resulting products. Using the DiCarlo Motors automobile sales example from Section 5.2, we show the calculation of the expected value for the number of automobiles sold during a day in Table 5.5. The sum of the entries in the $xf(x)$ column shows that the expected value is 1.50 automobiles per day. We therefore know that although sales of 0, 1, 2, 3, 4, or 5 automobiles are possible on any one day, over time DiCarlo can anticipate selling an average of 1.50 automobiles per day. Assuming 30 days of operation during a month, we can use the expected value of 1.50 to forecast average monthly sales of 30(1.50) = 45 automobiles.

The expected value does not have to be a value the random variable can assume.

Variance

The expected value provides a measure of central tendency for a random variable, but we often also want a measure of variability, or dispersion. Just as we used the variance in Chapter 3 to summarize the variability in data, we now use **variance** to summarize the variability in the values of a random variable. The formula for the variance of a discrete random variable follows.

TABLE 5.5 CALCULATION OF THE EXPECTED VALUE FOR THE NUMBER OF AUTOMOBILES SOLD DURING A DAY AT DICARLO MOTORS

x	$f(x)$	$xf(x)$
0	.18	0(.18) = .00
1	.39	1(.39) = .39
2	.24	2(.24) = .48
3	.14	3(.14) = .42
4	.04	4(.04) = .16
5	.01	5(.01) = .05
		1.50

$E(x) = \mu = \sum xf(x)$

The variance is a weighted average of the squared deviations of a random variable from its mean. The weights are the probabilities.

VARIANCE OF A DISCRETE RANDOM VARIABLE

$$Var(x) = \sigma^2 = \sum(x - \mu)^2 f(x) \tag{5.5}$$

As equation (5.5) shows, an essential part of the variance formula is the deviation, $x - \mu$, which measures how far a particular value of the random variable is from the expected value, or mean, μ. In computing the variance of a random variable, the deviations are squared and then weighted by the corresponding value of the probability function. The sum of these weighted squared deviations for all values of the random variable is referred to as the *variance.* The notations $Var(x)$ and σ^2 are both used to denote the variance of a random variable.

The calculation of the variance for the probability distribution of the number of automobiles sold during a day at DiCarlo Motors is summarized in Table 5.6. We see that the variance is 1.25. The **standard deviation**, σ, is defined as the positive square root of the variance. Thus, the standard deviation for the number of automobiles sold during a day is

$$\sigma = \sqrt{1.25} = 1.118$$

The standard deviation is measured in the same units as the random variable ($\sigma = 1.118$ automobiles) and therefore is often preferred in describing the variability of a random variable. The variance σ^2 is measured in squared units and is thus more difficult to interpret.

TABLE 5.6 CALCULATION OF THE VARIANCE FOR THE NUMBER OF AUTOMOBILES SOLD DURING A DAY AT DICARLO MOTORS

x	$x - \mu$	$(x - \mu)^2$	$f(x)$	$(x - \mu)^2 f(x)$
0	0 − 1.50 = −1.50	2.25	.18	2.25(.18) = .4050
1	1 − 1.50 = −.50	.25	.39	.25(.39) = .0975
2	2 − 1.50 = .50	.25	.24	.25(.24) = .0600
3	3 − 1.50 = 1.50	2.25	.14	2.25(.14) = .3150
4	4 − 1.50 = 2.50	6.25	.04	6.25(.04) = .2500
5	5 − 1.50 = 3.50	12.25	.01	12.25(.01) = .1225
				1.2500

$\sigma^2 = \sum(x - \mu)^2 f(x)$

Using Excel to Compute the Expected Value, Variance, and Standard Deviation

The calculations involved in computing the expected value and variance for a discrete random variable can easily be made in an Excel worksheet. One approach is to enter the formulas necessary to make the calculations in Tables 5.4 and 5.5. An easier way, however, is to make use of Excel's SUMPRODUCT function. In this subsection we show how to use the SUMPRODUCT function to compute the expected value and variance for daily automobile sales at DiCarlo Motors. Refer to Figure 5.2 as we describe the tasks involved. The formula worksheet is in the background; the value worksheet is in the foreground.

Enter/Access Data: The data needed are the values for the random variable and the corresponding probabilities. Labels, values for the random variable, and the corresponding probabilities are entered in cells A1:B7.

Enter Functions and Formulas: The SUMPRODUCT function multiplies each value in one range by the corresponding value in another range and sums the products. To use the SUMPRODUCT function to compute the expected value of daily automobile sales at DiCarlo Motors, we entered the following formula into cell B9:

=SUMPRODUCT(A2:A7,B2:B7)

Note that the first range, A2:A7, contains the values for the random variable, daily automobile sales. The second range, B2:B7, contains the corresponding probabilities. Thus, the SUMPRODUCT function in cell B9 is computing A2*B2 + A3*B3 + A4*B4 + A5*B5 + A6*B6 + A7*B7; hence, it is applying the formula in equation (5.4) to compute the expected value. The result, shown in cell B9 of the value worksheet, is 1.5.

FIGURE 5.2 EXCEL WORKSHEET FOR EXPECTED VALUE, VARIANCE, AND STANDARD DEVIATION

	A	B	C	D
1	Sales	Probability	Sq Dev from Mean	
2	0	0.18	=(A2-B9)^2	
3	1	0.39	=(A3-B9)^2	
4	2	0.24	=(A4-B9)^2	
5	3	0.14	=(A5-B9)^2	
6	4	0.04	=(A6-B9)^2	
7	5	0.01	=(A7-B9)^2	
8				
9	Mean	=SUMPRODUCT(A2:A7,B2:B7)		
10				
11	Variance	=SUMPRODUCT(C2:C7,B2:B7)		
12				
13	Std Deviation	=SQRT(B11)		
14				

	A	B	C	D
1	Sales	Probability	Sq Dev from Mean	
2	0	0.18	2.25	
3	1	0.39	0.25	
4	2	0.24	0.25	
5	3	0.14	2.25	
6	4	0.04	6.25	
7	5	0.01	12.25	
8				
9	Mean	1.5		
10				
11	Variance	1.25		
12				
13	Std Deviation	1.118		
14				

The formulas in cells C2:C7 are used to compute the squared deviations from the expected value or mean of 1.5 (the mean is in cell B9). The results, shown in the value worksheet, are the same as the results shown in Table 5.5. The formula necessary to compute the variance for daily automobile sales was entered into cell B11. It uses the SUMPRODUCT function to multiply each value in the range C2:C7 by each corresponding value in the range B2:B7 and sums the products. The result, shown in the value worksheet, is 1.25. Because the standard deviation is the square root of the variance, we entered the formula =SQRT(B11) into cell B13 to compute the standard deviation for daily automobile sales. The result, shown in the value worksheet, is 1.118.

Exercises

Methods

15. The following table provides a probability distribution for the random variable x.

x	$f(x)$
3	.25
6	.50
9	.25

a. Compute $E(x)$, the expected value of x.
b. Compute σ^2, the variance of x.
c. Compute σ, the standard deviation of x.

SELF*test*

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

y	$f(y)$
2	.20
4	.30
7	.40
8	.10

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

Applications

17. During the summer of 2014, Coldstream Country Club in Cincinnati, Ohio collected data on 443 rounds of golf played from its white tees. The data for each golfer's score on the twelfth hole are contained in the DATAfile *Coldstream12*.

DATA *file*
Coldstream12

a. Construct an empirical discrete probability distribution for the player scores on the twelfth hole.
b. A *par* is the score that a good golfer is expected to get for the hole. For hole number 12, par is four. What is the probability of a player scoring less than or equal to par on hole number 12?
c. What is the expected score for hole number 12?
d. What is the variance for hole number 12?
e. What is the standard deviation for hole number 12?

18. The American Housing Survey reported the following data on the number of times that owner-occupied and renter-occupied units had a water supply stoppage lasting 6 or more hours over a 3-month period.

TABLE 5.4 PROBABILITY DISTRIBUTION FOR THE NUMBER OF AUTOMOBILES SOLD DURING A DAY AT DICARLO MOTORS

x	$f(x)$
0	.18
1	.39
2	.24
3	.14
4	.04
5	.01
Total	1.00

In probability function notation, $f(0)$ provides the probability of 0 automobiles sold, $f(1)$ provides the probability of 1 automobile sold, and so on. Because historical data show 54 of 300 days with 0 automobiles sold, we assign the relative frequency $54/300 = .18$ to $f(0)$, indicating that the probability of 0 automobiles being sold during a day is .18. Similarly, because 117 of 300 days had 1 automobile sold, we assign the relative frequency $117/300 = .39$ to $f(1)$, indicating that the probability of exactly 1 automobile being sold during a day is .39. Continuing in this way for the other values of the random variable, we compute the values for $f(2)$, $f(3)$, $f(4)$, and $f(5)$ as shown in Table 5.4.

A primary advantage of defining a random variable and its probability distribution is that once the probability distribution is known, it is relatively easy to determine the probability of a variety of events that may be of interest to a decision maker. For example, using the probability distribution for DiCarlo Motors as shown in Table 5.4, we see that the most probable number of automobiles sold during a day is 1 with a probability of $f(1) = .39$. In addition, the probability of selling 3 or more automobiles during a day is $f(3) + f(4) + f(5) = .14 + .04 + .01 = .19$. These probabilities, plus others the decision maker may ask about, provide information that can help the decision maker understand the process of selling automobiles at DiCarlo Motors.

In the development of a probability function for any discrete random variable, the following two conditions must be satisfied.

These conditions are the analogs to the two basic requirements for assigning probabilities to experimental outcomes presented in Chapter 4.

REQUIRED CONDITIONS FOR A DISCRETE PROBABILITY FUNCTION

$$f(x \geq 0 \qquad (5.1)$$

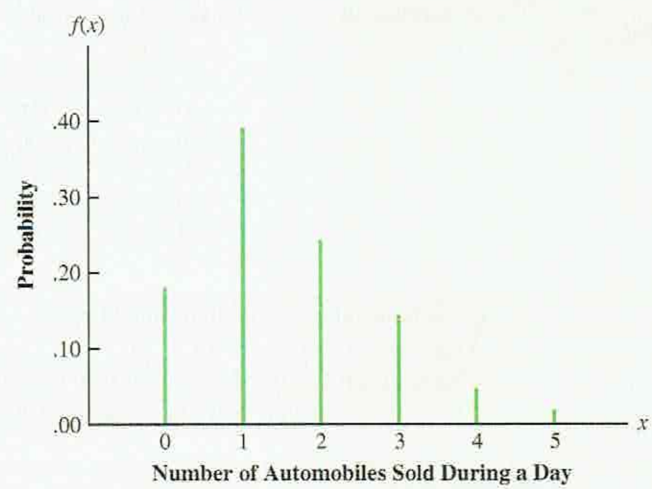
$$\Sigma f(x) = 1 \qquad (5.2)$$

Table 5.4 shows that the probabilities for the random variable x satisfy equation (5.1); $f(x)$ is greater than or equal to 0 for all values of x. In addition, because the probabilities sum to 1, equation (5.2) is satisfied. Thus, the DiCarlo Motors probability function is a valid discrete probability function.

We can also show the DiCarlo Motors probability distribution graphically. In Figure 5.1 the values of the random variable x for DiCarlo Motors are shown on the horizontal axis and the probability associated with these values is shown on the vertical axis.

In addition to the probability distributions shown in tables, a formula that gives the probability function, $f(x)$, for every value of x is often used to describe probability

FIGURE 5.1 GRAPHICAL REPRESENTATION OF THE PROBABILITY DISTRIBUTION FOR THE NUMBER OF AUTOMOBILES SOLD DURING A DAY AT DICARLO MOTORS



distributions. The simplest example of a discrete probability distribution given by a formula is the **discrete uniform probability distribution**. Its probability function is defined by equation (5.3).

DISCRETE UNIFORM PROBABILITY FUNCTION

$$f(x) = 1/n \tag{5.3}$$

where

$n =$ the number of values the random variable may assume

For example, consider again the experiment of rolling a die. We define the random variable x to be the number of dots on the upward face. For this experiment, $n = 6$ values are possible for the random variable; $x = 1, 2, 3, 4, 5, 6$. We showed earlier how the probability distribution for this experiment can be expressed as a table. Since the probabilities are equally likely, the discrete uniform probability function can also be used. The probability function for this discrete uniform random variable is

$$f(x) = 1/6 \quad x = 1, 2, 3, 4, 5, 6$$

Several widely-used discrete probability distributions are specified by formulas. Three important cases are the binomial, Poisson, and hypergeometric distributions; these distributions are discussed later in the chapter.