3.7 OBTAINING DESCRIPTIVE SUMMARY MEASURES FROM A FREQUENCY DISTRIBUTION

It is sometimes the case that raw data are not available and the only source of data is a frequency distribution. In such a situation, approximations to descriptive summary measures such as the arithmetic mean and the standard deviation can be obtained.

Approximating the Arithmetic Mean and Standard Deviation

When data from a sample have been summarized into a frequency distribution, an approximation of the arithmetic mean can be computed by assuming that all values within each class interval are located at the midpoint of the class.

APPROXIMATING THE ARITHMETIC MEAN FROM A FREQUENCY DISTRIBUTION

$$\bar{X} = \frac{\sum_{j=1}^{c} m_j f_j}{n} \qquad \textbf{(3.17)}$$

where

$\bar{X}$ = sample arithmetic mean
n = number of observations or sample size
c = number of classes in the frequency distribution
m_j = midpoint of the jth class
f_j = frequencies tallied into the jth class

To calculate the standard deviation from a frequency distribution, all values within each class interval are also assumed to be located at the midpoint of the class.

APPROXIMATING THE STANDARD DEVIATION OF A FREQUENCY DISTRIBUTION

$$S = \sqrt{\frac{\sum_{j=1}^{c} (m_j - \bar{X})^2 f_j}{n - 1}} \qquad \textbf{(3.18)}$$

The computation of the arithmetic mean and the standard deviation from a frequency distribution is illustrated in Example 3.9.

EXAMPLE 3.9

APPROXIMATING THE ARITHMETIC MEAN AND STANDARD DEVIATION FROM A FREQUENCY DISTRIBUTION

Consider the frequency distribution of the 5-year annualized return of 158 growth mutual funds in Table 3.4. Approximate the arithmetic mean and standard deviation of the 5-year annualized return of 158 growth mutual funds.

TABLE 3.4
Frequency distribution of the 5-year annualized returns achieved by 158 growth funds

5-Year Annualized Return	Number of Funds
−10.0 but less than −5.0	1
−5.0 but less than 0.0	3
0.0 but less than 5.0	14
5.0 but less than 10.0	58
10.0 but less than 15.0	61
15.0 but less than 20.0	17
20.0 but less than 25.0	3
25.0 but less than 30.0	1
Total	158

SOLUTION

For these data,

Percentage Return	f_j Number of Funds	m_j Midpoint	$m_j f_j$	$(m_j - \bar{X})$	$(m_j - \bar{X})^2$	$(m_j - \bar{X})^2 f_j$
−10.0 but less than −5.0	1	−7.5	−7.5	−17.69	312.9361	312.9361
−5.0 but less than 0.0	3	−2.5	−7.5	−12.69	161.0361	483.1083
0.0 but less than 5.0	14	2.5	35.0	−7.69	59.1361	827.9054
5.0 but less than 10.0	58	7.5	435.0	−2.69	7.2361	419.6938
10.0 but less than 15.0	61	12.5	762.5	2.31	5.3361	325.5021
15.0 but less than 20.0	17	17.5	297.5	7.31	53.4361	908.4137
20.0 but less than 25.0	3	22.5	67.5	12.31	151.5361	454.6083
25.0 but less than 30.0	1	27.5	27.5	17.31	299.6361	299.6361
Total	158		1,610.0			4,031.8038

Using Equations (3.17) and (3.18) on page CD3-1,

$$\bar{X} = \frac{\sum_{j=1}^{c} m_j f_j}{n}$$

$$\bar{X} = \frac{1{,}610.0}{158} = 10.19$$

and

$$S = \sqrt{\frac{\sum_{j=1}^{c} (m_j - \bar{X})^2 f_j}{n - 1}}$$

$$S = \sqrt{\frac{4{,}031.8038}{158 - 1}} = 5.07$$

PROBLEMS FOR SECTION 3.7

Learning the Basics

3.76 Given the following frequency distribution for $n = 100$:

Class Intervals	Frequency
0—Under 10	10
10—Under 20	20
20—Under 30	40
30—Under 40	20
40—Under 50	10
	100

Approximate
a. the arithmetic mean.
b. the standard deviation.

3.77 Given the following frequency distribution for $n = 100$:

Class Intervals	Frequency
0—Under 10	40
10—Under 20	25
20—Under 30	15
30—Under 40	15
40—Under 50	5
	100

Approximate
a. the arithmetic mean.
b. the standard deviation.

Applying the Concepts

3.78 A wholesale appliance distributing firm wished to study its accounts receivable for two successive months. Two independent samples of 50 accounts were selected for each of the two months. The results are summarized in the following table:

Frequency Distributions for Accounts Receivable

Amount	March Frequency	April Frequency
\$0 to under \$2,000	6	10
\$2,000 to under \$4,000	13	14
\$4,000 to under \$6,000	17	13
\$6,000 to under \$8,000	10	10
\$8,000 to under \$10,000	4	0
\$10,000 to under \$12,000	0	3
Total	50	50

For each month, approximate the
a. arithmetic mean.
b. standard deviation.
c. On the basis of parts (a)–(b), do you think the arithmetic mean and standard deviation of the accounts receivable have changed substantially from March to April? Explain.

3.79 The following table contains the cumulative distributions and cumulative percentage distributions of braking distance (in feet) at 80 miles per hour for a sample of 25 U.S.-manufactured automobile models and for a sample of 72 foreign-made automobile models obtained in a recent year:

Cumulative Frequency and Percentage Distribution for Braking Distance (in Feet) at 80 mph for U.S.-Manufactured and Foreign-Made Automobile Models

	U.S.-Made Automobile Models "Less Than" Indicated Values		Foreign-Made Automobile Models "Less Than" Indicated Values	
Braking Distance (in ft)	Number	Percentage	Number	Percentage
210	0	0.0	0	0.0
220	1	4.0	1	1.4
230	2	8.0	4	5.6
240	3	12.0	19	26.4
250	4	16.0	32	44.4
260	8	32.0	54	75.0
270	11	44.0	61	84.7
280	17	68.0	68	94.4
290	21	84.0	68	94.4
300	23	92.0	70	97.2
310	25	100.0	71	98.6
320	25	100.0	72	100.0

For U.S.- and foreign-made automobiles
a. Construct a frequency distribution for each group.
b. On the basis of the results of (a), approximate the arithmetic mean of the braking distance.
c. On the basis of the results of (a), approximate the standard deviation of the braking distance.
d. On the basis of the results of (b) and (c), do U.S.- and foreign-made automobiles seem to differ in their braking distance? Explain.

3.80 The following data represent the distribution of the ages of employees within two different divisions of a publishing company.

Age of Employees (Years)	A Frequency	B Frequency
20—Under 30	8	15
30—Under 40	17	32
40—Under 50	11	20
50—Under 60	8	4
60—Under 70	2	0

For each of the two divisions (*A* and *B*), approximate the

a. arithmetic mean.

b. standard deviation.

c. On the basis of the results of (a) and (b), do you think there are differences in the age distribution between the two divisions? Explain.