Applying the Concepts 6–1

Assessing Normality

Many times in statistics it is necessary to see if a set of data values is approximately normally distributed. There are special techniques that can be used. One technique is to draw a histogram for the data and see if it is approximately bell-shaped. (*Note:* It does not have to be exactly symmetric to be bell-shaped.)

The numbers of branches of the 50 top libraries are shown.

67	84	80	77	97	59	62	37	33	42
36	54	18	12	19	33	49	24	25	22
24	29	9	21	21	24	31	17	15	21
13	19	19	22	22	30	41	22	18	20
26	33	14	14	16	22	26	10	16	24

Source: The World Almanac and Book of Facts.

1. Construct a frequency distribution for the data.
2. Construct a histogram for the data.
3. Describe the shape of the histogram.
4. Based on your answer to question 3, do you feel that the distribution is approximately normal?

In addition to the histogram, distributions that are approximately normal have about 68% of the values fall within 1 standard deviation of the mean, about 95% of the data values fall within 2 standard deviations of the mean, and almost 100% of the data values fall within 3 standard deviations of the mean. (See Figure 6–5.)

5. Find the mean and standard deviation for the data.
6. What percent of the data values fall within 1 standard deviation of the mean?
7. What percent of the data values fall within 2 standard deviations of the mean?
8. What percent of the data values fall within 3 standard deviations of the mean?
9. How do your answers to questions 6, 7, and 8 compare to 68, 95, and 100%, respectively?
10. Does your answer help support the conclusion you reached in question 4? Explain.

(More techniques for assessing normality are explained in Section 6–2.)
See pages 367 and 368 for the answers.

Exercises 6–1

1. What are the characteristics of a normal distribution?

2. Why is the standard normal distribution important in statistical analysis?

3. What is the total area under the standard normal distribution curve?

4. What percentage of the area falls below the mean? Above the mean?

5. About what percentage of the area under the normal distribution curve falls within 1 standard deviation above and below the mean? 2 standard deviations? 3 standard deviations?

6. What are two other names for a normal distribution?

For Exercises 7 through 26, find the area under the standard normal distribution curve.

7. Between $z = 0$ and $z = 0.98$

8. Between $z = 0$ and $z = 1.77$

9. Between $z = 0$ and $z = -2.14$

10. Between $z = 0$ and $z = -0.32$

11. To the right of $z = 0.29$

12. To the right of $z = 2.01$

13. To the left of $z = -1.39$

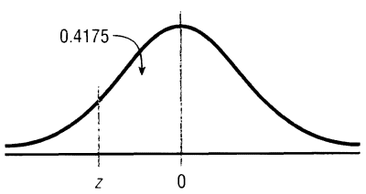
14. To the left of $z = -0.75$
15. Between $z = 1.09$ and $z = 1.83$
16. Between $z = 1.23$ and $z = 1.90$
17. Between $z = -1.56$ and $z = -1.83$
18. Between $z = -0.96$ and $z = -0.36$
19. Between $z = -1.46$ and $z = -1.98$
20. Between $z = 0.24$ and $z = -1.12$
21. To the left of $z = 2.22$
22. To the left of $z = 1.31$
23. To the right of $z = -0.12$
24. To the right of $z = -1.92$
25. To the right of $z = 1.92$ and to the left of $z = -0.44$
26. To the left of $z = -2.15$ and to the right of $z = 1.62$

In Exercises 27 through 40, find the probabilities for each, using the standard normal distribution.

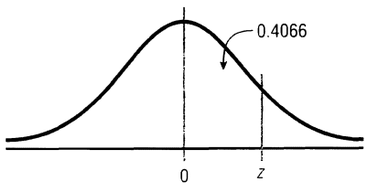
27. $P(0 < z < 0.92)$
28. $P(0 < z < 1.96)$
29. $P(-1.43 < z < 0)$
30. $P(-1.23 < z < 0)$
31. $P(z > 2.51)$
32. $P(z > 0.82)$
33. $P(z < -1.46)$
34. $P(z < -1.77)$
35. $P(-2.07 < z < 1.88)$
36. $P(-0.20 < z < 1.56)$
37. $P(1.51 < z < 2.17)$
38. $P(1.12 < z < 1.43)$
39. $P(z < 1.42)$
40. $P(z > -1.43)$

For Exercises 41 through 46, find the z value that corresponds to the given area.

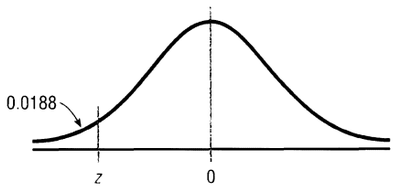
41.



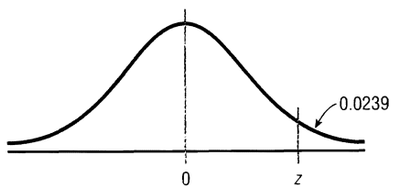
42.



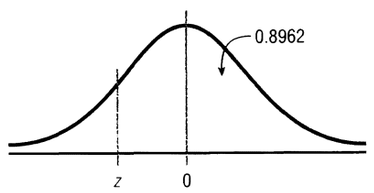
43.



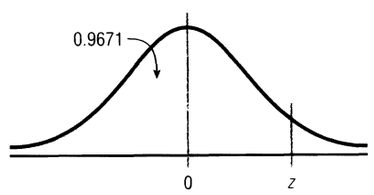
44.



45.



46.



47. Find the z value to the left of the mean so that
    a. 98.87% of the area under the distribution curve lies to the right of it.
    b. 82.12% of the area under the distribution curve lies to the right of it.
    c. 60.64% of the area under the distribution curve lies to the right of it.

48. Find the z value to the right of the mean so that
    a. 54.78% of the area under the distribution curve lies to the left of it.
    b. 69.85% of the area under the distribution curve lies to the left of it.
    c. 88.10% of the area under the distribution curve lies to the left of it.

female full professor at the same kind of institution, the salary is $90,330. If the standard deviation for the salaries of both genders is approximately $5200 and the salaries are normally distributed, find the 80th percentile salary for male professors and for female professors.

Source: *World Almanac.*

17. **Professors' Salaries** The average annual professor's salary at a doctoral level at a private, independent institution is $159,964 for men and $147,702 for women. Consider the women's salaries. Assume that they are normally distributed with a standard deviation of $8900. What is the probability that a woman professor makes more than the men's average salary?

Source: *World Almanac 2012.*

18. **Itemized Charitable Contributions** The average charitable contribution itemized per income tax return in Pennsylvania is $792. Suppose that the distribution of contributions is normal with a standard deviation of $103. Find the limits for the middle 50% of contributions.

Source: IRS. *Statistics of Income Bulletin.*

19. **New Home Sizes** A contractor decided to build homes that will include the middle 80% of the market. If the average size of homes built is 1810 square feet, find the maximum and minimum sizes of the homes the contractor should build. Assume that the standard deviation is 92 square feet and the variable is normally distributed.

Source: Michael D. Shook and Robert L. Shook, *The Book of Odds.*

20. **New-Home Prices** If the average price of a new one-family home is $246,300 with a standard deviation of $15,000, find the minimum and maximum prices of the houses that a contractor will build to satisfy the middle 80% of the market. Assume that the variable is normally distributed.

Source: *New York Times Almanac.*

21. **Cost of Personal Computers** The average price of a personal computer (PC) is $949. If the computer prices are approximately normally distributed and $\sigma = \$100$, what is the probability that a randomly selected PC costs more than $1200? The least expensive 10% of personal computers cost less than what amount?

Source: *New York Times Almanac.*

22. **Reading Improvement Program** To help students improve their reading, a school district decides to implement a reading program. It is to be administered to the bottom 5% of the students in the district, based on the scores on a reading achievement exam. If the average score for the students in the district is 122.6, find the cutoff score that will make a student eligible for the program. The standard deviation is 18. Assume the variable is normally distributed.

23. **Used Car Prices** An automobile dealer finds that the average price of a previously owned vehicle is $8256. He decides to sell cars that will appeal to the middle 60% of the market in terms of price. Find the maximum and minimum prices of the cars the dealer will sell. The standard deviation is $1150, and the variable is normally distributed.

24. **Ages of Amtrak Passenger Cars** The average age of Amtrak passenger train cars is 19.4 years. If the distribution of ages is normal and 20% of the cars are older than 22.8 years, find the standard deviation.

Source: *New York Times Almanac.*

25. **Lengths of Hospital Stays** The average length of a hospital stay for all diagnoses is 4.8 days. If we assume that the lengths of hospital stays are normally distributed with a variance of 2.1, then 10% of hospital stays are longer than how many days? Thirty percent of stays are less than how many days?

Source: www.cdc.gov

26. **High School Competency Test** A mandatory competency test for high school sophomores has a normal distribution with a mean of 400 and a standard deviation of 100.

   a. The top 3% of students receive $500. What is the minimum score you would need to receive this award?

   b. The bottom 1.5% of students must go to summer school. What is the minimum score you would need to stay out of this group?

27. **Product Marketing** An advertising company plans to market a product to low-income families. A study states that for a particular area, the average income per family is $24,596 and the standard deviation is $6256. If the company plans to target the bottom 18% of the families based on income, find the cutoff income. Assume the variable is normally distributed.

28. **Bottled Drinking Water** Americans drank an average of 23.2 gallons of bottled water per capita in 2008. If the standard deviation is 2.7 gallons and the variable is normally distributed, find the probability that a randomly selected American drank more than 25 gallons of bottled water. What is the probability that the selected person drank between 22 and 30 gallons?

Source: www.census.gov

29. **Wristwatch Lifetimes** The mean lifetime of a wristwatch is 25 months, with a standard deviation of 5 months. If the distribution is normal, for how many months should a guarantee be made if the manufacturer does not want to exchange more than 10% of the watches? Assume the variable is normally distributed.

30. **Police Academy Acceptance Exams** To qualify for a police academy, applicants are given a test of physical fitness. The scores are normally distributed with a mean of 64 and a standard deviation of 9. If only the top 20% of the applicants are selected, find the cutoff score.

Exercises 6–3

1. If samples of a specific size are selected from a population and the means are computed, what is this distribution of means called?

2. Why do most of the sample means differ somewhat from the population mean? What is this difference called?

3. What is the mean of the sample means?

4. What is the standard deviation of the sample means called? What is the formula for this standard deviation?

5. What does the central limit theorem say about the shape of the distribution of sample means?

6. What formula is used to gain information about an individual data value when the variable is normally distributed?

For Exercises 7 through 25, assume that the sample is taken from a large population and the correction factor can be ignored.

7. **Unemployment Benefits** The average weekly unemployment benefit in Montana is $272. Suppose that the benefits are normally distributed with a standard deviation of $43. A random sample of 15 benefits is chosen in Montana. What is the probability that the mean for this sample is greater than the U.S. average, which is $299? Is the normal distribution appropriate here since the sample size is only 15? Explain.

   Source: *World Almanac*.

8. **Glass Garbage Generation** A survey found that the American family generates an average of 17.2 pounds of glass garbage each year. Assume the standard deviation of the distribution is 2.5 pounds. Find the probability that the mean of a sample of 55 families will be between 17 and 18 pounds.

   Source: Michael D. Shook and Robert L. Shook, *The Book of Odds*.

9. **College Costs** The mean undergraduate cost for tuition, fees, room, and board for four-year institutions was $26,489 for a recent academic year. Suppose that σ = $3204 and that 36 four-year institutions are randomly selected. Find the probability that the sample mean cost for these 36 schools is

   a. Less than $25,000

   b. Greater than $26,000

   c. Between $24,000 and $26,000

   Source: www.nces.ed.gov

10. **Teachers' Salaries in Connecticut** The average teacher's salary in Connecticut (ranked first among states) is $57,337. Suppose that the distribution of salaries is normal with a standard deviation of $7500.

    a. What is the probability that a randomly selected teacher makes less than $52,000 per year?

    b. If we sample 100 teachers' salaries, what is the probability that the sample mean is less than $56,000?

    Source: *New York Times Almanac*.

11. **Serum Cholesterol Levels** The mean serum cholesterol level of a large population of overweight children is 220 milligrams per deciliter (mg/dl), and the standard deviation is 16.3 mg/dl. If a random sample of 35 overweight children is selected, find the probability that the mean will be between 220 and 222 mg/dl. Assume the serum cholesterol level variable is normally distributed.

12. **Teachers' Salaries in North Dakota** The average teacher's salary in North Dakota is $37,764. Assume a normal distribution with σ = $5100.

    a. What is the probability that a randomly selected teacher's salary is greater than $45,000?

    b. For a sample of 75 teachers, what is the probability that the sample mean is greater than $38,000?

    Source: *New York Times Almanac*.

13. **Movie Ticket Prices** In a recent year the average movie ticket cost $7.89. In a random sample of 50 movie tickets from various areas, what is the probability that the mean cost exceeds $8.00, given that the population standard deviation is $1.39?

    Source: *World Almanac*.

14. **SAT Scores** The national average SAT score (for Verbal and Math) is 1028. Suppose that nothing is known about the shape of the distribution and that the standard deviation is 100. If a random sample of 200 scores were selected and the sample mean were calculated to be 1050, would you be surprised? Explain.

    Source: *New York Times Almanac*.

15. **Cost of Overseas Trip** The average overseas trip cost is $2708 per visitor. If we assume a normal distribution with a standard deviation of $405, what is the probability that the cost for a randomly selected trip is more than $3000? If we select a random sample of 30 overseas trips and find the mean of the sample, what is the probability that the mean is greater than $3000?

    Source: *World Almanac*.

16. **Cell Phone Lifetimes** A recent study of the lifetimes of cell phones found the average is 24.3 months. The standard deviation is 2.6 months. If a company provides its 33 employees with a cell phone, find the probability that the mean lifetime of these phones will be less than 23.8 months. Assume cell phone life is a normally distributed variable.

17. **Water Use** The *Old Farmer's Almanac* reports that the average person uses 123 gallons of water daily. If the standard deviation is 21 gallons, find the probability that the mean of a randomly selected sample of 15 people will be between 120 and 126 gallons. Assume the variable is normally distributed.

Exercises 6–4

1. Explain why a normal distribution can be used as an approximation to a binomial distribution.

2. What conditions must be met to use the normal distribution to approximate the binomial distribution?

3. Why is a correction for continuity necessary?

4. When is the normal distribution not a good approximation for the binomial distribution?

5. Use the normal approximation to the binomial to find the probabilities for the specific value(s) of X.

   a. $n = 30, p = 0.5, X = 18$
   b. $n = 50, p = 0.8, X = 44$
   c. $n = 100, p = 0.1, X = 12$

6. Use the normal approximation to find the probabilities for the specific value(s) of X.

   a. $n = 10, p = 0.5, X \geq 7$
   b. $n = 20, p = 0.7, X \leq 12$
   c. $n = 50, p = 0.6, X \leq 40$

7. Check each binomial distribution to see whether it can be approximated by a normal distribution (i.e., are $np \geq 5$ and $nq \geq 5$?).

   a. $n = 20, p = 0.5$
   b. $n = 10, p = 0.6$
   c. $n = 40, p = 0.9$

8. Check each binomial distribution to see whether it can be approximated by a normal distribution (i.e., are $np \geq 5$ and $nq \geq 5$?).

   a. $n = 50, p = 0.2$
   b. $n = 30, p = 0.8$
   c. $n = 20, p = 0.85$

9. **People Who Smoke** In a recent year, 23.3% of Americans smoked cigarettes. What is the probability that in a random sample of 200 Americans, more than 50 smoke?

   Source: *World Almanac.*

10. **School Enrollment** Of all 3- to 5-year-old children, 56% are enrolled in school. If a sample of 500 such children is randomly selected, find the probability that at least 250 will be enrolled in school.

    Source: *Statistical Abstract of the United States.*

11. **Home Ownership** In a recent year, the rate of U.S. home ownership was 65.9%. Choose a random sample of 120 households across the United States. What is the probability that 65 to 85 (inclusive) of them live in homes that they own?

    Source: *World Almanac.*

12. **Mail Order** A mail order company has an 8% success rate. If it mails advertisements to 600 people, find the probability of getting fewer than 40 sales.

13. **Health Insurance** In a recent year, 56% of employers offered a consumer-directed health plan (CDHP). This type of plan typically combines a high deductible with a health savings plan. Choose 80 employers at random. What is the probability that more than 50 will offer a CDHP?

    Source: *USA TODAY.*

14. **Household Computers** According to recent surveys, 60% of households have personal computers. If a random sample of 180 households is selected, what is the probability that more than 60 but fewer than 100 have a personal computer?

    Source: *New York Times Almanac.*

15. **Youth Smoking** Two out of five adult smokers acquired the habit by age 14. If 400 smokers are randomly selected, find the probability that 170 or fewer acquired the habit by age 14.

    Source: *Harper's Index.*

16. **Population of College Cities** College students often make up a substantial portion of the population of college cities and towns. State College, Pennsylvania, ranks first with 71.1% of its population made up of college students. What is the probability that in a random sample of 150 people from State College, more than 50 are not college students?

    Source: www.infoplease.com

17. **Voter Preference** A political candidate estimates that 30% of the voters in her party favor her proposed tax reform bill. If there are 400 people at a rally, find the probability that at least 100 voters will favor her tax bill. Based on your answer, is it likely that 100 or more people will favor the bill?

18. **Telephone Answering Devices** Seventy-eight percent of U.S. homes have a telephone answering device. In a random sample of 250 homes, what is the probability that fewer than 50 do not have a telephone answering device?

    Source: *New York Times Almanac.*

19. **Female Americans Who Have Completed 4 Years of College** The percentage of female Americans 25 years old and older who have completed 4 years of college or more is 26.1. In a random sample of 200 American women who are at least 25, what is the probability that at most 50 have completed 4 years of college or more?

    Source: *New York Times Almanac.*

20. **Residences of U.S. Citizens** According to the U.S. Census, 67.5% of the U.S. population were born in their state of residence. In a random sample of