

The probability that headway time is at most 5 sec is

$$\begin{aligned} P(X \leq 5) &= \int_{-\infty}^5 f(x) dx = \int_{.5}^5 .15e^{-.15(x-.5)} dx \\ &= .15e^{.075} \int_{.5}^5 e^{-.15x} dx = .15e^{.075} \cdot \left( -\frac{1}{.15} e^{-.15x} \Big|_{x=.5}^{x=5} \right) \\ &= e^{.075} (-e^{-.75} + e^{-.075}) = 1.078(-.472 + .928) = .491 \\ &= P(\text{less than 5 sec}) = P(X < 5) \end{aligned}$$

Unlike discrete distributions such as the binomial, hypergeometric, and negative binomial, the distribution of any given continuous rv cannot usually be derived using simple probabilistic arguments. Instead, one must make a judicious choice of pdf based on prior knowledge and available data. Fortunately, there are some general families of pdf's that have been found to be sensible candidates in a wide variety of experimental situations; several of these are discussed later in the chapter.

Just as in the discrete case, it is often helpful to think of the population of interest as consisting of  $X$  values rather than individuals or objects. The pdf is then a model for the distribution of values in this numerical population, and from this model various population characteristics (such as the mean) can be calculated.

### EXERCISES Section 4.1 (1–10)

1. The current in a certain circuit as measured by an ammeter is a continuous random variable  $X$  with the following density function:

$$f(x) = \begin{cases} .075x + .2 & 3 \leq x \leq 5 \\ 0 & \text{otherwise} \end{cases}$$

- Graph the pdf and verify that the total area under the density curve is indeed 1.
  - Calculate  $P(X \leq 4)$ . How does this probability compare to  $P(X < 4)$ ?
  - Calculate  $P(3.5 \leq X \leq 4.5)$  and also  $P(4.5 < X)$ .
2. Suppose the reaction temperature  $X$  (in °C) in a certain chemical process has a uniform distribution with  $A = -5$  and  $B = 5$ .
- Compute  $P(X < 0)$ .
  - Compute  $P(-2.5 < X < 2.5)$ .
  - Compute  $P(-2 \leq X \leq 3)$ .
  - For  $k$  satisfying  $-5 < k < k + 4 < 5$ , compute  $P(k < X < k + 4)$ .
3. The error involved in making a certain measurement is a continuous rv  $X$  with pdf

$$f(x) = \begin{cases} .09375(4 - x^2) & -2 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

- Sketch the graph of  $f(x)$ .
- Compute  $P(X > 0)$ .
- Compute  $P(-1 < X < 1)$ .
- Compute  $P(X < -.5 \text{ or } X > .5)$ .

4. Let  $X$  denote the vibratory stress (psi) on a wind turbine blade at a particular wind speed in a wind tunnel. The article "Blade Fatigue Life Assessment with Application to VAWTS" (*J. of Solar Energy Engr.*, 1982: 107–111) proposes the Rayleigh distribution, with pdf

$$f(x; \theta) = \begin{cases} \frac{x}{\theta^2} \cdot e^{-x^2/(2\theta^2)} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

as a model for the  $X$  distribution.

- Verify that  $f(x; \theta)$  is a legitimate pdf.
  - Suppose  $\theta = 100$  (a value suggested by a graph in the article). What is the probability that  $X$  is at most 200? Less than 200? At least 200?
  - What is the probability that  $X$  is between 100 and 200 (again assuming  $\theta = 100$ )?
  - Give an expression for  $P(X \leq x)$ .
5. A college professor never finishes his lecture before the end of the hour and always finishes his lectures within 2 min after the hour. Let  $X =$  the time that elapses between the end of the hour and the end of the lecture and suppose the pdf of  $X$  is

$$f(x) = \begin{cases} kx^2 & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

- Find the value of  $k$  and draw the corresponding density curve. [*Hint:* Total area under the graph of  $f(x)$  is 1.]
- What is the probability that the lecture ends within 1 min of the end of the hour?

- c. What is the probability that the lecture continues beyond the hour for between 60 and 90 sec?
  - d. What is the probability that the lecture continues for at least 90 sec beyond the end of the hour?
6. The actual tracking weight of a stereo cartridge that is set to track at 3 g on a particular changer can be regarded as a continuous rv  $X$  with pdf

$$f(x) = \begin{cases} k[1 - (x - 3)^2] & 2 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

- a. Sketch the graph of  $f(x)$ .
  - b. Find the value of  $k$ .
  - c. What is the probability that the actual tracking weight is greater than the prescribed weight?
  - d. What is the probability that the actual weight is within .25 g of the prescribed weight?
  - e. What is the probability that the actual weight differs from the prescribed weight by more than .5 g?
7. The time  $X$  (min) for a lab assistant to prepare the equipment for a certain experiment is believed to have a uniform distribution with  $A = 25$  and  $B = 35$ .
- a. Determine the pdf of  $X$  and sketch the corresponding density curve.
  - b. What is the probability that preparation time exceeds 33 min?
  - c. What is the probability that preparation time is within 2 min of the mean time? [*Hint*: Identify  $\mu$  from the graph of  $f(x)$ .]
  - d. For any  $a$  such that  $25 < a < a + 2 < 35$ , what is the probability that preparation time is between  $a$  and  $a + 2$  min?
8. In commuting to work, a professor must first get on a bus near her house and then transfer to a second bus. If the waiting time (in minutes) at each stop has a uniform distribution with  $A = 0$  and  $B = 5$ , then it can be shown that the total waiting time  $Y$  has the pdf

$$f(y) = \begin{cases} \frac{1}{25}y & 0 \leq y < 5 \\ \frac{2}{5} - \frac{1}{25}y & 5 \leq y \leq 10 \\ 0 & y < 0 \text{ or } y > 10 \end{cases}$$

- a. Sketch a graph of the pdf of  $Y$ .
  - b. Verify that  $\int_{-\infty}^{\infty} f(y) dy = 1$ .
  - c. What is the probability that total waiting time is at most 3 min?
  - d. What is the probability that total waiting time is at most 8 min?
  - e. What is the probability that total waiting time is between 3 and 8 min?
  - f. What is the probability that total waiting time is either less than 2 min or more than 6 min?
9. Consider again the pdf of  $X =$  time headway given in Example 4.5. What is the probability that time headway is
- a. At most 6 sec?
  - b. More than 6 sec? At least 6 sec?
  - c. Between 5 and 6 sec?
10. A family of pdf's that has been used to approximate the distribution of income, city population size, and size of firms is the Pareto family. The family has two parameters,  $k$  and  $\theta$ , both  $> 0$ , and the pdf is

$$f(x; k, \theta) = \begin{cases} \frac{k \cdot \theta^k}{x^{k+1}} & x \geq \theta \\ 0 & x < \theta \end{cases}$$

- a. Sketch the graph of  $f(x; k, \theta)$ .
- b. Verify that the total area under the graph equals 1.
- c. If the rv  $X$  has pdf  $f(x; k, \theta)$ , for any fixed  $b > \theta$ , obtain an expression for  $P(X \leq b)$ .
- d. For  $\theta < a < b$ , obtain an expression for the probability  $P(a \leq X \leq b)$ .

## 4.2 Cumulative Distribution Functions and Expected Values

Several of the most important concepts introduced in the study of discrete distributions also play an important role for continuous distributions. Definitions analogous to those in Chapter 3 involve replacing summation by integration.

### The Cumulative Distribution Function

The cumulative distribution function (cdf)  $F(x)$  for a discrete rv  $X$  gives, for any specified number  $x$ , the probability  $P(X \leq x)$ . It is obtained by summing the pmf  $p(y)$  over all possible values  $y$  satisfying  $y \leq x$ . The cdf of a continuous rv gives the same probabilities  $P(X \leq x)$  and is obtained by integrating the pdf  $f(y)$  between the limits  $-\infty$  and  $x$ .

$$f(x) = \begin{cases} \frac{k}{x^4} & x > 1 \\ 0 & x \leq 1 \end{cases}$$

- a. Determine the value of  $k$  for which  $f(x)$  is a legitimate pdf.
  - b. Obtain the cumulative distribution function.
  - c. Use the cdf from (b) to determine the probability that headway exceeds 2 sec and also the probability that headway is between 2 and 3 sec.
  - d. Obtain the mean value of headway and the standard deviation of headway.
  - e. What is the probability that headway is within 1 standard deviation of the mean value?
14. The article "Modeling Sediment and Water Column Interactions for Hydrophobic Pollutants" (*Water Research*, 1984: 1169–1174) suggests the uniform distribution on the interval (7.5, 20) as a model for depth (cm) of the bioturbation layer in sediment in a certain region.
- a. What are the mean and variance of depth?
  - b. What is the cdf of depth?
  - c. What is the probability that observed depth is at most 10? Between 10 and 15?
  - d. What is the probability that the observed depth is within 1 standard deviation of the mean value? Within 2 standard deviations?
15. Let  $X$  denote the amount of space occupied by an article placed in a 1-ft<sup>3</sup> packing container. The pdf of  $X$  is

$$f(x) = \begin{cases} 90x^8(1-x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

- a. Graph the pdf. Then obtain the cdf of  $X$  and graph it.
  - b. What is  $P(X \leq .5)$  [i.e.,  $F(.5)$ ]?]
  - c. Using the cdf from (a), what is  $P(.25 < X \leq .5)$ ? What is  $P(.25 \leq X \leq .5)$ ?
  - d. What is the 75th percentile of the distribution?
  - e. Compute  $E(X)$  and  $\sigma_X$ .
  - f. What is the probability that  $X$  is more than 1 standard deviation from its mean value?
16. Answer parts (a)–(f) of Exercise 15 with  $X$  = lecture time past the hour given in Exercise 5.
17. Let  $X$  have a uniform distribution on the interval  $[A, B]$ .
- a. Obtain an expression for the  $(100p)$ th percentile.
  - b. Compute  $E(X)$ ,  $V(X)$ , and  $\sigma_X$ .
  - c. For  $n$ , a positive integer, compute  $E(X^n)$ .
18. Let  $X$  denote the voltage at the output of a microphone, and suppose that  $X$  has a uniform distribution on the interval from  $-1$  to  $1$ . The voltage is processed by a "hard limiter" with cutoff values  $-.5$  and  $.5$ , so the limiter output is a random variable  $Y$  related to  $X$  by  $Y = X$  if  $|X| \leq .5$ ,  $Y = .5$  if  $X > .5$ , and  $Y = -.5$  if  $X < -.5$ .
- a. What is  $P(Y = .5)$ ?
  - b. Obtain the cumulative distribution function of  $Y$  and graph it.

19. Let  $X$  be a continuous rv with cdf

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x}{4} \left[ 1 + \ln\left(\frac{4}{x}\right) \right] & 0 < x \leq 4 \\ 1 & x > 4 \end{cases}$$

[This type of cdf is suggested in the article "Variability in Measured Bedload-Transport Rates" (*Water Resources Bull.*, 1985: 39–48) as a model for a certain hydrologic variable.] What is

- a.  $P(X \leq 1)$ ?
  - b.  $P(1 \leq X \leq 3)$ ?
  - c. The pdf of  $X$ ?
20. Consider the pdf for total waiting time  $Y$  for two buses

$$f(y) = \begin{cases} \frac{1}{25}y & 0 \leq y < 5 \\ \frac{2}{5} - \frac{1}{25}y & 5 \leq y \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

introduced in Exercise 8.

- a. Compute and sketch the cdf of  $Y$ . [Hint: Consider separately  $0 \leq y < 5$  and  $5 \leq y \leq 10$  in computing  $F(y)$ . A graph of the pdf should be helpful.]
  - b. Obtain an expression for the  $(100p)$ th percentile. [Hint: Consider separately  $0 < p < .5$  and  $.5 < p < 1$ .]
  - c. Compute  $E(Y)$  and  $V(Y)$ . How do these compare with the expected waiting time and variance for a single bus when the time is uniformly distributed on  $[0, 5]$ ?
21. An ecologist wishes to mark off a circular sampling region having radius 10 m. However, the radius of the resulting region is actually a random variable  $R$  with pdf

$$f(r) = \begin{cases} \frac{3}{4} [1 - (10 - r)^2] & 9 \leq r \leq 11 \\ 0 & \text{otherwise} \end{cases}$$

What is the expected area of the resulting circular region?

22. The weekly demand for propane gas (in 1000s of gallons) from a particular facility is an rv  $X$  with pdf

$$f(x) = \begin{cases} 2 \left( 1 - \frac{1}{x^2} \right) & 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

- a. Compute the cdf of  $X$ .
- b. Obtain an expression for the  $(100p)$ th percentile. What is the value of  $\tilde{\mu}$ ?
- c. Compute  $E(X)$  and  $V(X)$ .
- d. If 1.5 thousand gallons are in stock at the beginning of the week and no new supply is due in during the week, how much of the 1.5 thousand gallons is expected to be left at the end of the week? [Hint: Let  $h(x)$  = amount left when demand =  $x$ .]

The exact probabilities are .2622 and .8348, respectively, so the approximations are quite good. In the last calculation, the probability  $P(5 \leq X \leq 15)$  is being approximated by the area under the normal curve between 4.5 and 15.5—the continuity correction is used for both the upper and lower limits. ■

When the objective of our investigation is to make an inference about a population proportion  $p$ , interest will focus on the sample proportion of successes  $X/n$  rather than on  $X$  itself. Because this proportion is just  $X$  multiplied by the constant  $1/n$ , it will also have approximately a normal distribution (with mean  $\mu = p$  and standard deviation  $\sigma = \sqrt{pq/n}$ ) provided that both  $np \geq 10$  and  $nq \geq 10$ . This normal approximation is the basis for several inferential procedures to be discussed in later chapters.

### EXERCISES Section 4.3 (28–58)

28. Let  $Z$  be a standard normal random variable and calculate the following probabilities, drawing pictures wherever appropriate.
- |                                |                                |
|--------------------------------|--------------------------------|
| a. $P(0 \leq Z \leq 2.17)$     | b. $P(0 \leq Z \leq 1)$        |
| c. $P(-2.50 \leq Z \leq 0)$    | d. $P(-2.50 \leq Z \leq 2.50)$ |
| e. $P(Z \leq 1.37)$            | f. $P(-1.75 \leq Z)$           |
| g. $P(-1.50 \leq Z \leq 2.00)$ | h. $P(1.37 \leq Z \leq 2.50)$  |
| i. $P(1.50 \leq Z)$            | j. $P( Z  \leq 2.50)$          |
29. In each case, determine the value of the constant  $c$  that makes the probability statement correct.
- |                           |                                 |
|---------------------------|---------------------------------|
| a. $\Phi(c) = .9838$      | b. $P(0 \leq Z \leq c) = .291$  |
| c. $P(c \leq Z) = .121$   | d. $P(-c \leq Z \leq c) = .668$ |
| e. $P(c \leq  Z ) = .016$ |                                 |
30. Find the following percentiles for the standard normal distribution. Interpolate where appropriate.
- |         |        |         |
|---------|--------|---------|
| a. 91st | b. 9th | c. 75th |
| d. 25th | e. 6th |         |
31. Determine  $z_\alpha$  for the following:
- |                     |                   |
|---------------------|-------------------|
| a. $\alpha = .0055$ | b. $\alpha = .09$ |
| c. $\alpha = .663$  |                   |
32. Suppose the force acting on a column that helps to support a building is a normally distributed random variable  $X$  with mean value 15.0 kips and standard deviation 1.25 kips. Compute the following probabilities by standardizing and then using Table A.3.
- |                         |                           |
|-------------------------|---------------------------|
| a. $P(X \leq 15)$       | b. $P(X \leq 17.5)$       |
| c. $P(X \geq 10)$       | d. $P(14 \leq X \leq 18)$ |
| e. $P( X - 15  \leq 3)$ |                           |
33. Mopeds (small motorcycles with an engine capacity below 50 cm<sup>3</sup>) are very popular in Europe because of their mobility, ease of operation, and low cost. The article “Procedure to Verify the Maximum Speed of Automatic Transmission Mopeds in Periodic Motor Vehicle Inspections” (*J. of Automobile Engr.*, 2008: 1615–1623) described a rolling bench test for determining maximum vehicle speed. A normal distribution with mean value 46.8 km/h and standard deviation 1.75 km/h is postulated. Consider randomly selecting a single such moped.
- What is the probability that maximum speed is at most 50 km/h?
  - What is the probability that maximum speed is at least 48 km/h?
  - What is the probability that maximum speed differs from the mean value by at most 1.5 standard deviations?
34. The article “Reliability of Domestic-Waste Biofilm Reactors” (*J. of Envir. Engr.*, 1995: 785–790) suggests that substrate concentration (mg/cm<sup>3</sup>) of influent to a reactor is normally distributed with  $\mu = .30$  and  $\sigma = .06$ .
- What is the probability that the concentration exceeds .25?
  - What is the probability that the concentration is at most .10?
  - How would you characterize the largest 5% of all concentration values?
35. Suppose the diameter at breast height (in.) of trees of a certain type is normally distributed with  $\mu = 8.8$  and  $\sigma = 2.8$ , as suggested in the article “Simulating a Harvester-Forwarder Softwood Thinning” (*Forest Products J.*, May 1997: 36–41).
- What is the probability that the diameter of a randomly selected tree will be at least 10 in.? Will exceed 10 in.?
  - What is the probability that the diameter of a randomly selected tree will exceed 20 in.?
  - What is the probability that the diameter of a randomly selected tree will be between 5 and 10 in.?
  - What value  $c$  is such that the interval  $(8.8 - c, 8.8 + c)$  includes 98% of all diameter values?
  - If four trees are independently selected, what is the probability that at least one has a diameter exceeding 10 in.?
36. Spray drift is a constant concern for pesticide applicators and agricultural producers. The inverse relationship between droplet size and drift potential is well known. The

- paper "Effects of 2,4-D Formulation and Quinclorac on Spray Droplet Size and Deposition" (*Weed Technology*, 2005: 1030–1036) investigated the effects of herbicide formulation on spray atomization. A figure in the paper suggested the normal distribution with mean  $1050 \mu\text{m}$  and standard deviation  $150 \mu\text{m}$  was a reasonable model for droplet size for water (the "control treatment") sprayed through a 760 ml/min nozzle.
- a. What is the probability that the size of a single droplet is less than  $1500 \mu\text{m}$ ? At least  $1000 \mu\text{m}$ ?
  - b. What is the probability that the size of a single droplet is between  $1000$  and  $1500 \mu\text{m}$ ?
  - c. How would you characterize the smallest 2% of all droplets?
  - d. If the sizes of five independently selected droplets are measured, what is the probability that at least one exceeds  $1500 \mu\text{m}$ ?
37. Suppose that blood chloride concentration (mmol/L) has a normal distribution with mean 104 and standard deviation 5 (information in the article "Mathematical Model of Chloride Concentration in Human Blood," *J. of Med. Engr. and Tech.*, 2006: 25–30, including a normal probability plot as described in Section 4.6, supports this assumption).
    - a. What is the probability that chloride concentration equals 105? Is less than 105? Is at most 105?
    - b. What is the probability that chloride concentration differs from the mean by more than 1 standard deviation? Does this probability depend on the values of  $\mu$  and  $\sigma$ ?
    - c. How would you characterize the most extreme .1% of chloride concentration values?
  38. There are two machines available for cutting corks intended for use in wine bottles. The first produces corks with diameters that are normally distributed with mean 3 cm and standard deviation .1 cm. The second machine produces corks with diameters that have a normal distribution with mean 3.04 cm and standard deviation .02 cm. Acceptable corks have diameters between 2.9 cm and 3.1 cm. Which machine is more likely to produce an acceptable cork?
  39.
    - a. If a normal distribution has  $\mu = 30$  and  $\sigma = 5$ , what is the 91st percentile of the distribution?
    - b. What is the 6th percentile of the distribution?
    - c. The width of a line etched on an integrated circuit chip is normally distributed with mean  $3.000 \mu\text{m}$  and standard deviation .140. What width value separates the widest 10% of all such lines from the other 90%?
  40. The article "Monte Carlo Simulation—Tool for Better Understanding of LRFD" (*J. of Structural Engr.*, 1993: 1586–1599) suggests that yield strength (ksi) for A36 grade steel is normally distributed with  $\mu = 43$  and  $\sigma = 4.5$ .
    - a. What is the probability that yield strength is at most 40? Greater than 60?
    - b. What yield strength value separates the strongest 75% from the others?
  41. The automatic opening device of a military cargo parachute has been designed to open when the parachute is 200 m above the ground. Suppose opening altitude actually has a normal distribution with mean value 200 m and standard deviation 30 m. Equipment damage will occur if the parachute opens at an altitude of less than 100 m. What is the probability that there is equipment damage to the payload of at least one of five independently dropped parachutes?
  42. The temperature reading from a thermocouple placed in a constant-temperature medium is normally distributed with mean  $\mu$ , the actual temperature of the medium, and standard deviation  $\sigma$ . What would the value of  $\sigma$  have to be to ensure that 95% of all readings are within .1° of  $\mu$ ?
  43. The distribution of resistance for resistors of a certain type is known to be normal, with 10% of all resistors having a resistance exceeding 10.256 ohms and 5% having a resistance smaller than 9.671 ohms. What are the mean value and standard deviation of the resistance distribution?
  44. If bolt thread length is normally distributed, what is the probability that the thread length of a randomly selected bolt is
    - a. Within 1.5 SDs of its mean value?
    - b. Farther than 2.5 SDs from its mean value?
    - c. Between 1 and 2 SDs from its mean value?
  45. A machine that produces ball bearings has initially been set so that the true average diameter of the bearings it produces is .500 in. A bearing is acceptable if its diameter is within .004 in. of this target value. Suppose, however, that the setting has changed during the course of production, so that the bearings have normally distributed diameters with mean value .499 in. and standard deviation .002 in. What percentage of the bearings produced will not be acceptable?
  46. The Rockwell hardness of a metal is determined by impressing a hardened point into the surface of the metal and then measuring the depth of penetration of the point. Suppose the Rockwell hardness of a particular alloy is normally distributed with mean 70 and standard deviation 3. (Rockwell hardness is measured on a continuous scale.)
    - a. If a specimen is acceptable only if its hardness is between 67 and 75, what is the probability that a randomly chosen specimen has an acceptable hardness?
    - b. If the acceptable range of hardness is  $(70 - c, 70 + c)$ , for what value of  $c$  would 95% of all specimens have acceptable hardness?
    - c. If the acceptable range is as in part (a) and the hardness of each of ten randomly selected specimens is independently determined, what is the expected number of acceptable specimens among the ten?
    - d. What is the probability that at most eight of ten independently selected specimens have a hardness of less than

73.84? [Hint:  $Y$  = the number among the ten specimens with hardness less than 73.84 is a binomial variable; what is  $p$ ?]

47. The weight distribution of parcels sent in a certain manner is normal with mean value 12 lb and standard deviation 3.5 lb. The parcel service wishes to establish a weight value  $c$  beyond which there will be a surcharge. What value of  $c$  is such that 99% of all parcels are at least 1 lb under the surcharge weight?
48. Suppose Appendix Table A.3 contained  $\Phi(z)$  only for  $z \geq 0$ . Explain how you could still compute
- $P(-1.72 \leq Z \leq -.55)$
  - $P(-1.72 \leq Z \leq .55)$
- Is it necessary to tabulate  $\Phi(z)$  for  $z$  negative? What property of the standard normal curve justifies your answer?
49. Consider babies born in the "normal" range of 37–43 weeks gestational age. Extensive data supports the assumption that for such babies born in the United States, birth weight is normally distributed with mean 3432 g and standard deviation 482 g. [The article "Are Babies Normal?" (*The American Statistician*, 1999: 298–302) analyzed data from a particular year; for a sensible choice of class intervals, a histogram did not look at all normal, but after further investigations it was determined that this was due to some hospitals measuring weight in grams and others measuring to the nearest ounce and then converting to grams. A modified choice of class intervals that allowed for this gave a histogram that was well described by a normal distribution.]
- What is the probability that the birth weight of a randomly selected baby of this type exceeds 4000 g? Is between 3000 and 4000 g?
  - What is the probability that the birth weight of a randomly selected baby of this type is either less than 2000 g or greater than 5000 g?
  - What is the probability that the birth weight of a randomly selected baby of this type exceeds 7 lb?
  - How would you characterize the most extreme .1% of all birth weights?
  - If  $X$  is a random variable with a normal distribution and  $a$  is a numerical constant ( $a \neq 0$ ), then  $Y = aX$  also has a normal distribution. Use this to determine the distribution of birth weight expressed in pounds (shape, mean, and standard deviation), and then recalculate the probability from part (c). How does this compare to your previous answer?
50. In response to concerns about nutritional contents of fast foods, McDonald's has announced that it will use a new cooking oil for its french fries that will decrease substantially trans fatty acid levels and increase the amount of more beneficial polyunsaturated fat. The company claims that 97 out of 100 people cannot detect a difference in taste between the new and old oils. Assuming that this figure is correct (as a long-run proportion), what is the approximate probability that in a random sample of 1000 individuals who have purchased fries at McDonald's,
- At least 40 can taste the difference between the two oils?
  - At most 5% can taste the difference between the two oils?
51. Chebyshev's inequality, (see Exercise 44, Chapter 3), is valid for continuous as well as discrete distributions. It states that for any number  $k$  satisfying  $k \geq 1$ ,  $P(|X - \mu| \geq k\sigma) \leq 1/k^2$  (see Exercise 44 in Chapter 3 for an interpretation). Obtain this probability in the case of a normal distribution for  $k = 1, 2$ , and 3, and compare to the upper bound.
52. Let  $X$  denote the number of flaws along a 100-m reel of magnetic tape (an integer-valued variable). Suppose  $X$  has approximately a normal distribution with  $\mu = 25$  and  $\sigma = 5$ . Use the continuity correction to calculate the probability that the number of flaws is
- Between 20 and 30, inclusive.
  - At most 30. Less than 30.
53. Let  $X$  have a binomial distribution with parameters  $n = 25$  and  $p$ . Calculate each of the following probabilities using the normal approximation (with the continuity correction) for the cases  $p = .5, .6$ , and  $.8$  and compare to the exact probabilities calculated from Appendix Table A.1.
- $P(15 \leq X \leq 20)$
  - $P(X \leq 15)$
  - $P(20 \leq X)$
54. Suppose that 10% of all steel shafts produced by a certain process are nonconforming but can be reworked (rather than having to be scrapped). Consider a random sample of 200 shafts, and let  $X$  denote the number among these that are nonconforming and can be reworked. What is the (approximate) probability that  $X$  is
- At most 30?
  - Less than 30?
  - Between 15 and 25 (inclusive)?
55. Suppose only 75% of all drivers in a certain state regularly wear a seat belt. A random sample of 500 drivers is selected. What is the probability that
- Between 360 and 400 (inclusive) of the drivers in the sample regularly wear a seat belt?
  - Fewer than 400 of those in the sample regularly wear a seat belt?
56. Show that the relationship between a general normal percentile and the corresponding  $z$  percentile is as stated in this section.
57. a. Show that if  $X$  has a normal distribution with parameters  $\mu$  and  $\sigma$ , then  $Y = aX + b$  (a linear function of  $X$ ) also has a normal distribution. What are the parameters of the distribution of  $Y$  [i.e.,  $E(Y)$  and  $V(Y)$ ]? [Hint: Write the cdf of  $Y$ ,  $P(Y \leq y)$ , as an integral involving the pdf of  $X$ , and then differentiate with respect to  $y$  to get the pdf of  $Y$ .]