The probability that headway time is at most 5 sec is

$$P(X \le 5) = \int_{-\infty}^{5} f(x)\,dx = \int_{.5}^{5} .15e^{-.15(x-.5)}\,dx$$

$$= .15e^{.075}\int_{.5}^{5} e^{-.15x}\,dx = .15e^{.075}\cdot\left(-\frac{1}{.15}e^{-.15x}\Big|_{x=.5}^{x=5}\right)$$

$$= e^{.075}(-e^{-.75} + e^{-.075}) = 1.078(-.472 + .928) = .491$$

$$= P(\text{less than 5 sec}) = P(X < 5)$$ ■

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:

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

   a. Graph the pdf and verify that the total area under the density curve is indeed 1.
   b. Calculate $P(X \le 4)$. How does this probability compare to $P(X < 4)$?
   c. Calculate $P(3.5 \le X \le 4.5)$ 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 $B = 5$.
   a. Compute $P(X < 0)$.
   b. Compute $P(-2.5 < X < 2.5)$.
   c. Compute $P(-2 \le X \le 3)$.
   d. 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

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

   a. Sketch the graph of $f(x)$.
   b. Compute $P(X > 0)$.
   c. Compute $P(-1 < X < 1)$.
   d. 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

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

   as a model for the X distribution.
   a. Verify that $f(x;\theta)$ is a legitimate pdf.
   b. 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?
   c. What is the probability that X is between 100 and 200 (again assuming $\theta = 100$)?
   d. Give an expression for $P(X \le x)$.

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

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

   a. Find the value of k and draw the corresponding density curve. [*Hint*: Total area under the graph of $f(x)$ is 1.]
   b. 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

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

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 μ 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

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

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 θ, both > 0, and the pdf is

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

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 $P(X \le b)$.

d. For $\theta < a < b$, obtain an expression for the probability $P(a \le X \le b)$.

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 $P(X \le x)$. It is obtained by summing the pmf $p(y)$ over all possible values y satisfying $y \le x$. The cdf of a continuous rv gives the same probabilities $P(X \le x)$ and is obtained by integrating the pdf $f(y)$ between the limits $-\infty$ and x.

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

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^3 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 $P(X \le .5)$ [i.e., $F(.5)$]?
c. Using the cdf from (a), what is $P(.25 < X \le .5)$? What is $P(.25 \le X \le .5)$?
d. What is the 75th percentile of the distribution?
e. Compute $E(X)$ and σ_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 σ_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 $|X| \le .5$, $Y = .5$ 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

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

[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. $P(X \le 1)$?
b. $P(1 \le X \le 3)$?
c. The pdf of X?

20. Consider the pdf for total waiting time Y for two buses

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

introduced in Exercise 8.
a. Compute and sketch the cdf of Y. [*Hint*: Consider separately $0 \le y < 5$ and $5 \le y \le 10$ 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

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

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

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

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.]

Bonus problem:

Let the distribution of a random variable X be geometric with parameter p with $0 < p < 1$.

a. Show that $P(X > m + n \mid X > m) = P(X > n)$, for $m, n \in \{1,2,3,\dots\}$. This is called the *memoryless property* of the geometric distribution.

b. Find $E(2^{-X})$.

c. Let $p = 1/3$ and $Y = | X - 5 |$. Find the support and PMF of Y.