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.]

b. Construct a Weibull probability plot. Is the Weibull distribution family plausible?

93. Construct a probability plot that will allow you to assess the plausibility of the lognormal distribution as a model for the rainfall data of Exercise 83 in Chapter 1.

94. The accompanying observations are precipitation values during March over a 30-year period in Minneapolis-St. Paul.

.77	1.20	3.00	1.62	2.81	2.48
1.74	.47	3.09	1.31	1.87	.96
.81	1.43	1.51	.32	1.18	1.89
1.20	3.37	2.10	.59	1.35	.90
1.95	2.20	.52	.81	4.75	2.05

a. Construct and interpret a normal probability plot for this data set.

b. Calculate the square root of each value and then construct a normal probability plot based on this transformed data. Does it seem plausible that the square root of precipitation is normally distributed?

c. Repeat part (b) after transforming by cube roots.

95. Use a statistical software package to construct a normal probability plot of the tensile ultimate-strength data given in Exercise 13 of Chapter 1, and comment.

96. Let the *ordered* sample observations be denoted by $y_1, y_2, \ldots, y_n$ (y_1 being the smallest and y_n the largest). Our suggested check for normality is to plot the $(\Phi^{-1}((i - .5)/n), y_i)$ pairs. Suppose we believe that the observations come from a distribution with mean 0, and let $w_1, \ldots, w_n$ be the *ordered absolute* values of the x_i's. A **half-normal** plot is a probability plot of the w_i's. More specifically, since $P(|Z| \leq w) = P(-w \leq Z \leq w) = 2\Phi(w) - 1$, a half-normal plot is a plot of the $(\Phi^{-1}\{[(i - .5)/n + 1]/2\}, w_i)$ pairs. The virtue of this plot is that small or large outliers in the original sample will now appear only at the upper end of the plot rather than at both ends. Construct a half-normal plot for the following sample of measurement errors, and comment: −3.78, −1.27, 1.44, −.39, 12.38, −43.40, 1.15, −3.96, −2.34, 30.84.

97. The following failure time observations (1000s of hours) resulted from accelerated life testing of 16 integrated circuit chips of a certain type:

82.8	11.6	359.5	502.5	307.8	179.7
242.0	26.5	244.8	304.3	379.1	212.6
229.9	558.9	366.7	204.6		

Use the corresponding percentiles of the exponential distribution with $\lambda = 1$ to construct a probability plot. Then explain why the plot assesses the plausibility of the sample having been generated from *any* exponential distribution.

SUPPLEMENTARY EXERCISES (98–128)

98. Let X = the time it takes a read/write head to locate a desired record on a computer disk memory device once the head has been positioned over the correct track. If the disks rotate once every 25 millisec, a reasonable assumption is that X is uniformly distributed on the interval [0, 25].

a. Compute $P(10 \leq X \leq 20)$.

b. Compute $P(X \geq 10)$.

c. Obtain the cdf $F(X)$.

d. Compute $E(X)$ and σ_X.

99. A 12-in. bar that is clamped at both ends is to be subjected to an increasing amount of stress until it snaps. Let Y = the distance from the left end at which the break occurs. Suppose Y has pdf

$$f(y) = \begin{cases} \left(\frac{1}{24}\right)y\left(1 - \frac{y}{12}\right) & 0 \leq y \leq 12 \\ 0 & \text{otherwise} \end{cases}$$

Compute the following:

a. The cdf of Y, and graph it.

b. $P(Y \leq 4)$, $P(Y > 6)$, and $P(4 \leq Y \leq 6)$

c. $E(Y)$, $E(Y^2)$, and $V(Y)$

d. The probability that the break point occurs more than 2 in. from the expected break point.

e. The expected length of the shorter segment when the break occurs.

100. Let X denote the time to failure (in years) of a certain hydraulic component. Suppose the pdf of X is $f(x) = 32/(x + 4)^3$ for $x > 0$.

a. Verify that $f(x)$ is a legitimate pdf.

b. Determine the cdf.

c. Use the result of part (b) to calculate the probability that time to failure is between 2 and 5 years.

d. What is the expected time to failure?

e. If the component has a salvage value equal to $100/(4 + x)$ when its time to failure is x, what is the expected salvage value?

101. The completion time X for a certain task has cdf $F(x)$ given by

$$\begin{cases} 0 & x < 0 \\ \frac{x^3}{3} & 0 \leq x < 1 \\ 1 - \frac{1}{2}\left(\frac{7}{3} - x\right)\left(\frac{7}{4} - \frac{3}{4}x\right) & 1 \leq x \leq \frac{7}{3} \\ 1 & x > \frac{7}{3} \end{cases}$$

a. Obtain the pdf $f(x)$ and sketch its graph.
b. Compute $P(.5 \le X \le 2)$.
c. Compute $E(X)$.

102. The breakdown voltage of a randomly chosen diode of a certain type is known to be normally distributed with mean value 40 V and standard deviation 1.5 V.
a. What is the probability that the voltage of a single diode is between 39 and 42?
b. What value is such that only 15% of all diodes have voltages exceeding that value?
c. If four diodes are independently selected, what is the probability that at least one has a voltage exceeding 42?

103. The article "Computer Assisted Net Weight Control" (*Quality Progress,* 1983: 22–25) suggests a normal distribution with mean 137.2 oz and standard deviation 1.6 oz for the actual contents of jars of a certain type. The stated contents was 135 oz.
a. What is the probability that a single jar contains more than the stated contents?
b. Among ten randomly selected jars, what is the probability that at least eight contain more than the stated contents?
c. Assuming that the mean remains at 137.2, to what value would the standard deviation have to be changed so that 95% of all jars contain more than the stated contents?

104. When circuit boards used in the manufacture of compact disc players are tested, the long-run percentage of defectives is 5%. Suppose that a batch of 250 boards has been received and that the condition of any particular board is independent of that of any other board.
a. What is the approximate probability that at least 10% of the boards in the batch are defective?
b. What is the approximate probability that there are exactly 10 defectives in the batch?

105. The article "Characterization of Room Temperature Damping in Aluminum-Indium Alloys" (*Metallurgical Trans.,* 1993: 1611–1619) suggests that Al matrix grain size (μm) for an alloy consisting of 2% indium could be modeled with a normal distribution with a mean value 96 and standard deviation 14.
a. What is the probability that grain size exceeds 100?
b. What is the probability that grain size is between 50 and 80?
c. What interval (a, b) includes the central 90% of all grain sizes (so that 5% are below a and 5% are above b)?

106. The reaction time (in seconds) to a certain stimulus is a continuous random variable with pdf

$$f(x) = \begin{cases} \dfrac{3}{2} \cdot \dfrac{1}{x^2} & 1 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

a. Obtain the cdf.
b. What is the probability that reaction time is at most 2.5 sec? Between 1.5 and 2.5 sec?
c. Compute the expected reaction time.
d. Compute the standard deviation of reaction time.
e. If an individual takes more than 1.5 sec to react, a light comes on and stays on either until one further second has elapsed or until the person reacts (whichever happens first). Determine the expected amount of time that the light remains lit. [*Hint*: Let $h(X)$ = the time that the light is on as a function of reaction time X.]

107. Let X denote the temperature at which a certain chemical reaction takes place. Suppose that X has pdf

$$f(x) = \begin{cases} \dfrac{1}{9}(4 - x^2) & -1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

a. Sketch the graph of $f(x)$.
b. Determine the cdf and sketch it.
c. Is 0 the median temperature at which the reaction takes place? If not, is the median temperature smaller or larger than 0?
d. Suppose this reaction is independently carried out once in each of ten different labs and that the pdf of reaction time in each lab is as given. Let Y = the number among the ten labs at which the temperature exceeds 1. What kind of distribution does Y have? (Give the names and values of any parameters.)

108. The article "Determination of the MTF of Positive Photoresists Using the Monte Carlo Method" (*Photographic Sci. and Engr.,* 1983: 254–260) proposes the exponential distribution with parameter $\lambda = .93$ as a model for the distribution of a photon's free path length (μm) under certain circumstances. Suppose this is the correct model.
a. What is the expected path length, and what is the standard deviation of path length?
b. What is the probability that path length exceeds 3.0? What is the probability that path length is between 1.0 and 3.0?
c. What value is exceeded by only 10% of all path lengths?

109. The article "The Prediction of Corrosion by Statistical Analysis of Corrosion Profiles" (*Corrosion Science,* 1985: 305–315) suggests the following cdf for the depth X of the deepest pit in an experiment involving the exposure of carbon manganese steel to acidified seawater.

$$F(x; \alpha, \beta) = e^{-e^{-(x-\alpha)/\beta}} \qquad -\infty < x < \infty$$

The authors propose the values $\alpha = 150$ and $\beta = 90$. Assume this to be the correct model.
a. What is the probability that the depth of the deepest pit is at most 150? At most 300? Between 150 and 300?
b. Below what value will the depth of the maximum pit be observed in 90% of all such experiments?
c. What is the density function of X?
d. The density function can be shown to be unimodal (a single peak). Above what value on the measurement axis does this peak occur? (This value is the mode.)