EXERCISES FOR SECTION 4-7

4-77. Suppose that X is a binomial random variable with $n = 200$ and $p = 0.4$.
(a) Approximate the probability that X is less than or equal to 70.
(b) Approximate the probability that X is greater than 70 and less than 90.
(c) Approximate the probability that $X = 80$.

4-78. Suppose that X is a Poisson random variable with $\lambda = 6$.
(a) Compute the exact probability that X is less than 4.
(b) Approximate the probability that X is less than 4 and compare to the result in part (a).
(c) Approximate the probability that $8 < X < 12$.

4-79. Suppose that X has a Poisson distribution with a mean of 64. Approximate the following probabilities:
(a) $P(X > 72)$
(b) $P(X < 64)$
(c) $P(60 < X \leq 68)$

4-80. The manufacturing of semiconductor chips produces 2% defective chips. Assume the chips are independent and that a lot contains 1000 chips.
(a) Approximate the probability that more than 25 chips are defective.
(b) Approximate the probability that between 20 and 30 chips are defective.

4-81. There were 49.7 million people with some type of long-lasting condition or disability living in the United States in 2000. This represented 19.3 percent of the majority of civilians aged five and over (http://factfinder. census.gov). A sample of 1000 persons is selected at random.
(a) Approximate the probability that more than 200 persons in the sample have a disability.
(b) Approximate the probability that between 180 and 300 people in the sample have a disability.

4-82. Phoenix water is provided to approximately 1.4 million people, who are served through more than 362,000 accounts (http://phoenix.gov/WATER/wtrfacts.html). All accounts are metered and billed monthly. The probability that an account has an error in a month is 0.001, and accounts can be assumed to be independent.
(a) What is the mean and standard deviation of the number of account errors each month?
(b) Approximate the probability of fewer than 350 errors in a month.
(c) Approximate a value so that the probability that the number of errors exceeds this value is 0.05.
(d) Approximate the probability of more than 400 errors per month in the next two months. Assume that results between months are independent.

4-83. An electronic office product contains 5000 electronic components. Assume that the probability that each component operates without failure during the useful life of the product is 0.999, and assume that the components fail independently. Approximate the probability that 10 or more of the original 5000 components fail during the useful life of the product.

4-84. A corporate Web site contains errors on 50 of 1000 pages. If 100 pages are sampled randomly, without replacement, approximate the probability that at least 1 of the pages in error is in the sample.

4-85. Suppose that the number of asbestos particles in a sample of 1 squared centimeter of dust is a Poisson random variable with a mean of 1000. What is the probability that 10 squared centimeters of dust contains more than 10,000 particles?

4-86. A high-volume printer produces minor print-quality errors on a test pattern of 1000 pages of text according to a Poisson distribution with a mean of 0.4 per page.
(a) Why are the numbers of errors on each page independent random variables?
(b) What is the mean number of pages with errors (one or more)?
(c) Approximate the probability that more than 350 pages contain errors (one or more).

4-87. Hits to a high-volume Web site are assumed to follow a Poisson distribution with a mean of 10,000 per day. Approximate each of the following:
(a) The probability of more than 20,000 hits in a day
(b) The probability of less than 9900 hits in a day
(c) The value such that the probability that the number of hits in a day exceeds the value is 0.01
(d) Approximate the expected number of days in a year (365 days) that exceed 10,200 hits.
(e) Approximate the probability that over a year (365 days) more than 15 days each have more than 10,200 hits.

4-88. An acticle in *Biometrics* ["Integrative Analysis of Transcriptomic and Proteomic Data of Desulfovibrio vulgaris: A Nonlinear Model to Predict Abundance of Undetected Proteins" (2009)] found that protein abundance from an operon (a set of biologically related genes) was less dispersed than from randomly selected genes. In the research, 1000 sets of genes were randomly constructed and 75% of these sets were more disperse than a specific opteron. If the probability that a random set is more disperse than this opteron is truly 0.5, approximate the probability that 750 or more random sets exceed the opteron. From this result, what do you conclude about the dispersion in the opteron versus random genes?

4-89. An article under review for *Air Quality, Atmosphere & Health* titled "Linking Particulate Matter (PM10) and Childhood Asthma in Central Phoenix" linked air quality to childhood asthma incidents. The study region in central Phoenix, Arizona, recorded 10,500 asthma incidents in children in a 21-month period. Assume that the number of asthma incidents follows a Poisson distribution.

(a) Approximate the probability of more than 550 asthma incidents in a month.
(b) Approximate the probability of 450 to 550 asthma incidents in a month.
(c) Approximate the number of asthma incidents exceeded with probability 5%.
(d) If the number of asthma incidents were greater during the winter than the summer, what would this imply about the Poisson distribution assumption?.

4-8 EXPONENTIAL DISTRIBUTION

The discussion of the Poisson distribution defined a random variable to be the number of flaws along a length of copper wire. The distance between flaws is another random variable that is often of interest. Let the random variable X denote the length from any starting point on the wire until a flaw is detected. As you might expect, the distribution of X can be obtained from knowledge of the distribution of the number of flaws. The key to the relationship is the following concept. The distance to the first flaw exceeds 3 millimeters if and only if there are no flaws within a length of 3 millimeters—simple, but sufficient for an analysis of the distribution of X.

In general, let the random variable N denote the number of flaws in x millimeters of wire. If the mean number of flaws is λ per millimeter, N has a Poisson distribution with mean λx. We assume that the wire is longer than the value of x. Now,

$$P(X > x) = P(N = 0) = \frac{e^{-\lambda x}(\lambda x)^0}{0!} = e^{-\lambda x}$$

Therefore,

$$F(x) = P(X \le x) = 1 - e^{-\lambda x}, \quad x \ge 0$$

is the cumulative distribution function of X. By differentiating $F(x)$, the probability density function of X is calculated to be

$$f(x) = \lambda e^{-\lambda x}, \quad x \ge 0$$

The derivation of the distribution of X depends only on the assumption that the flaws in the wire follow a **Poisson process.** Also, the starting point for measuring X doesn't matter because the probability of the number of flaws in an interval of a Poisson process depends only on the length of the interval, not on the location. For any Poisson process, the following general result applies.

Exponential Distribution

The random variable X that equals the distance between successive events of a Poisson process with mean number of events $\lambda > 0$ per unit interval is an **exponential random variable** with parameter λ. The probability density function of X is

$$f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad 0 \le x < \infty \tag{4-14}$$