Remark. When all the p_i are equal to p, X is a binomial random variable. Hence, the preceding inequality shows that, for any set of nonnegative integers A,

$$\left|\sum_{i\in A}\binom{n}{i}p^i(1-p)^{n-i} - \sum_{i\in A}\frac{e^{-np}(np)^i}{i!}\right| \le np^2 \qquad \blacksquare$$

SUMMARY

Two useful probability bounds are provided by the *Markov* and *Chebyshev* inequalities. The Markov inequality is concerned with nonnegative random variables and says that, for X of that type,

$$P\{X \ge a\} \le \frac{E[X]}{a}$$

for every positive value a. The Chebyshev inequality, which is a simple consequence of the Markov inequality, states that if X has mean μ and variance σ^2, then, for every positive k,

$$P\{|X-\mu| \ge k\sigma\} \le \frac{1}{k^2}$$

The two most important theoretical results in probability are the *central limit theorem* and the *strong law of large numbers*. Both are concerned with a sequence of independent and identically distributed random variables. The central limit theorem says that if the random variables have a finite mean μ and a finite variance σ^2, then the distribution of the sum of the first n of them is, for large n, approximately that of a normal random variable with mean $n\mu$ and variance $n\sigma^2$. That is, if $X_i, i \ge 1$, is the sequence, then the central limit theorem states that, for every real number a,

$$\lim_{n\to\infty} P\left\{\frac{X_1+\cdots+X_n-n\mu}{\sigma\sqrt{n}} \le a\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-x^2/2}\,dx$$

The *strong law of large numbers* requires only that the random variables in the sequence have a finite mean μ. It states that, with probability 1, the average of the first n of them will converge to μ as n goes to infinity. This implies that if A is any specified event of an experiment for which independent replications are performed, then the limiting proportion of experiments whose outcomes are in A will, with probability 1, equal $P(A)$. Therefore, if we accept the interpretation that "with probability 1" means "with certainty," we obtain the theoretical justification for the long-run relative frequency interpretation of probabilities.

PROBLEMS

8.1. Suppose that X is a random variable with mean and variance both equal to 20. What can be said about $P\{0 < X < 40\}$?

8.2. From past experience, a professor knows that the test score of a student taking her final examination is a random variable with mean 75.

(a) Give an upper bound for the probability that a student's test score will exceed 85. Suppose, in addition, that the professor knows that the variance of a student's test score is equal to 25.

(b) What can be said about the probability that a student will score between 65 and 85?

(c) How many students would have to take the examination to ensure, with probability at least .9, that the class average would be within 5 of 75? Do not use the central limit theorem.

8.3. Use the central limit theorem to solve part (c) of Problem 2.

8.4. Let $X_1, \ldots, X_{20}$ be independent Poisson random variables with mean 1.

(a) Use the Markov inequality to obtain a bound on

$$P\left\{\sum_{1}^{20} X_i > 15\right\}$$

(b) Use the central limit theorem to approximate

$$P\left\{\sum_{1}^{20} X_i > 15\right\}.$$

8.5. Fifty numbers are rounded off to the nearest integer and then summed. If the individual round-off errors are uniformly distributed over $(-.5, .5)$, approximate the probability that the resultant sum differs from the exact sum by more than 3.

8.6. A die is continually rolled until the total sum of all rolls exceeds 300. Approximate the probability that at least 80 rolls are necessary.

8.7. A person has 100 light bulbs whose lifetimes are independent exponentials with mean 5 hours. If the bulbs are used one at a time, with a failed bulb being replaced immediately by a new one, approximate the probability that there is still a working bulb after 525 hours.

8.8. In Problem 7, suppose that it takes a random time, uniformly distributed over (0, .5), to replace a failed bulb. Approximate the probability that all bulbs have failed by time 550.

8.9. If X is a gamma random variable with parameters $(n, 1)$, approximately how large need n be so that

$$P\left\{\left|\frac{X}{n} - 1\right| > .01\right\} < .01?$$

8.10. Civil engineers believe that W, the amount of weight (in units of 1000 pounds) that a certain span of a bridge can withstand without structural damage resulting, is normally distributed with mean 400 and standard deviation 40. Suppose that the weight (again, in units of 1000 pounds) of a car is a random variable with mean 3 and standard deviation .3. Approximately how many cars would have to be on the bridge span for the probability of structural damage to exceed .1?

8.11. Many people believe that the daily change of price of a company's stock on the stock market is a random variable with mean 0 and variance σ^2. That is, if Y_n represents the price of the stock on the nth day, then

$$Y_n = Y_{n-1} + X_n \quad n \ge 1$$

where $X_1, X_2, \ldots$ are independent and identically distributed random variables with mean 0 and variance σ^2. Suppose that the stock's price today is 100. If $\sigma^2 = 1$, what can you say about the probability that the stock's price will exceed 105 after 10 days?

8.12. We have 100 components that we will put in use in a sequential fashion. That is, component 1 is initially put in use, and upon failure, it is replaced by component 2, which is itself replaced upon failure by component 3, and so on. If the lifetime of component i is exponentially distributed with mean $10 + i/10, i = 1, \ldots, 100$, estimate the probability that the total life of all components will exceed 1200. Now repeat when the life distribution of component i is uniformly distributed over $(0, 20 + i/5), i = 1, \ldots, 100$.

8.13. Student scores on exams given by a certain instructor have mean 74 and standard deviation 14. This instructor is about to give two exams, one to a class of size 25 and the other to a class of size 64.

(a) Approximate the probability that the average test score in the class of size 25 exceeds 80.

(b) Repeat part (a) for the class of size 64.

(c) Approximate the probability that the average test score in the larger class exceeds that of the other class by over 2.2 points.

(d) Approximate the probability that the average test score in the smaller class exceeds that of the other class by over 2.2 points.

8.14. A certain component is critical to the operation of an electrical system and must be replaced immediately upon failure. If the mean lifetime of this type of component is 100 hours and its standard deviation is 30 hours, how many of these components must be in stock so that the probability that the system is in continual operation for the next 2000 hours is at least .95?

8.15. An insurance company has 10,000 automobile policyholders. The expected yearly claim per policyholder is \$240, with a standard deviation of \$800. Approximate the probability that the total yearly claim exceeds \$2.7 million.

8.16. A.J. has 20 jobs that she must do in sequence, with the times required to do each of these jobs being independent random variables with mean 50 minutes and standard deviation 10 minutes. M.J. has 20 jobs that he must do in sequence, with the times required to do each of these jobs being independent random variables with mean 52 minutes and standard deviation 15 minutes.

(a) Find the probability that A.J. finishes in less than 900 minutes.

(b) Find the probability that M.J. finishes in less than 900 minutes.

(c) Find the probability that A.J. finishes before M.J.

8.17. Redo Example 5b under the assumption that the number of man–woman pairs is (approximately) normally distributed. Does this seem like a reasonable supposition?

8.18. Repeat part (a) of Problem 2 when it is known that the variance of a student's test score is equal to 25.

8.19. A lake contains 4 distinct types of fish. Suppose that each fish caught is equally likely to be any one of these types. Let Y denote the number of fish that need be caught to obtain at least one of each type.

(a) Give an interval (a, b) such that $P\{a \le Y \le b\} \ge .90$.

(b) Using the one-sided Chebyshev inequality, how many fish need we plan on catching so as to be at least 90 percent certain of obtaining at least one of each type.

8.20. If X is a nonnegative random variable with mean 25, what can be said about

(a) $E[X^3]$?

(b) $E[\sqrt{X}]$?

(c) $E[\log X]$?

(d) $E[e^{-X}]$?

8.21. Let X be a nonnegative random variable. Prove that

$$E[X] \le (E[X^2])^{1/2} \le (E[X^3])^{1/3} \le \cdots$$

8.22. Would the results of Example 5f change if the investor were allowed to divide her money and invest the fraction $\alpha, 0 < \alpha < 1$, in the risky proposition and invest the remainder in the risk-free venture? Her return for such a split investment would be $R = \alpha X + (1 - \alpha)m$.

8.23. Let X be a Poisson random variable with mean 20.

(a) Use the Markov inequality to obtain an upper bound on

$$p = P\{X \ge 26\}$$

(b) Use the one-sided Chebyshev inequality to obtain an upper bound on p.

(c) Use the Chernoff bound to obtain an upper bound on p.

(d) Approximate p by making use of the central limit theorem.

(e) Determine p by running an appropriate program.

THEORETICAL EXERCISES

8.1. If X has variance σ^2, then σ, the positive square root of the variance, is called the *standard deviation*. If X has mean μ and standard deviation σ, show that

$$P\{|X - \mu| \ge k\sigma\} \le \frac{1}{k^2}$$

8.2. If X has mean μ and standard deviation σ, the ratio $r \equiv |\mu|/\sigma$ is called the *measurement signal-to-noise ratio* of X. The idea is that X can be expressed as $X = \mu + (X - \mu)$, with μ representing the signal and $X - \mu$ the noise. If we define $|(X - \mu)/\mu| \equiv D$ as the relative deviation of X from its signal (or mean) μ, show that, for $\alpha > 0$,

$$P\{D \le \alpha\} \ge 1 - \frac{1}{r^2\alpha^2}$$

8.3. Compute the measurement signal-to-noise ratio—that is, $|\mu|/\sigma$, where $\mu = E[X]$ and $\sigma^2 = \text{Var}(X)$—of the following random variables:

(a) Poisson with mean λ;

(b) binomial with parameters n and p;

(c) geometric with mean $1/p$;

(d) uniform over (a, b);

(e) exponential with mean $1/\lambda$;

(f) normal with parameters μ, σ^2.

8.4. Let $Z_n, n \ge 1$, be a sequence of random variables and c a constant such that, for each $\varepsilon > 0, P\{|Z_n - c| > \varepsilon\} \to 0$ as $n \to \infty$. Show that, for any bounded continuous function g,

$$E[g(Z_n)] \to g(c) \quad \text{as} \quad n \to \infty$$

8.5. Let $f(x)$ be a continuous function defined for $0 \le x \le 1$. Consider the functions

$$B_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k(1 - x)^{n-k}$$

(called *Bernstein polynomials*) and prove that

$$\lim_{n \to \infty} B_n(x) = f(x)$$

Hint: Let $X_1, X_2, \ldots$ be independent Bernoulli random variables with mean x. Show that

$$B_n(x) = E\left[f\left(\frac{X_1 + \cdots + X_n}{n}\right)\right]$$

and then use Theoretical Exercise 4.

Since it can be shown that the convergence of $B_n(x)$ to $f(x)$ is uniform in x, the preceding reasoning provides a probabilistic proof of the famous Weierstrass theorem of analysis, which states that any continuous function on a closed interval can be approximated arbitrarily closely by a polynomial.

8.6. **(a)** Let X be a discrete random variable whose possible values are $1, 2, \ldots$. If $P\{X = k\}$ is nonincreasing in $k = 1, 2, \ldots$, prove that

$$P\{X = k\} \le 2\frac{E[X]}{k^2}$$

(b) Let X be a nonnegative continuous random variable having a nonincreasing density function. Show that

$$f(x) \le \frac{2E[X]}{x^2} \quad \text{for all} \quad x > 0$$

8.7. Suppose that a fair die is rolled 100 times. Let X_i be the value obtained on the ith roll. Compute an approximation for

$$P\left\{\prod_{1}^{100} X_i \le a^{100}\right\} \quad 1 < a < 6$$

8.8. Explain why a gamma random variable with parameters (t, λ) has an approximately normal distribution when t is large.

8.9. Suppose a fair coin is tossed 1000 times. If the first 100 tosses all result in heads, what proportion of heads would you expect on the final 900 tosses? Comment on the statement "The strong law of large numbers swamps, but does not compensate."

8.10. If X is a Poisson random variable with mean λ, show that for $i < \lambda$,

$$P\{X \le i\} \le \frac{e^{-\lambda}(e\lambda)^i}{i^i}$$

8.11. Let X be a binomial random variable with parameters n and p. Show that, for $i > np$,

(a) $\underset{t>0}{\text{minimum}}\; e^{-ti}E[e^{tX}]$ occurs when t is such that $e^t = \frac{iq}{(n-i)p}$, where $q = 1 - p$.

(b) $P\{X \ge i\} \le \frac{n^n}{i^i(n-i)^{n-i}}p^i(1 - p)^{n-i}$.

8.12. The Chernoff bound on a standard normal random variable Z gives $P\{Z > a\} \le e^{-a^2/2}, a > 0$. Show, by considering the density of Z, that the right side of the inequality can be reduced by the factor 2. That is, show that

$$P\{Z > a\} \le \frac{1}{2}e^{-a^2/2} \quad a > 0$$

8.13. Show that if $E[X] < 0$ and $\theta \ne 0$ is such that $E[e^{\theta X}] = 1$, then $\theta > 0$.

SELF-TEST PROBLEMS AND EXERCISES

8.1. The number of automobiles sold weekly at a certain dealership is a random variable with expected value 16. Give an upper bound to the probability that

(a) next week's sales exceed 18;

(b) next week's sales exceed 25.

8.2. Suppose in Problem 1 that the variance of the number of automobiles sold weekly is 9.

(a) Give a lower bound to the probability that next week's sales are between 10 and 22, inclusively.

(b) Give an upper bound to the probability that next week's sales exceed 18.

8.3. If

$$E[X] = 75 \quad E[Y] = 75 \quad \text{Var}(X) = 10$$
$$\text{Var}(Y) = 12 \quad \text{Cov}(X, Y) = -3$$

give an upper bound to

(a) $P\{|X - Y| > 15\}$;

(b) $P\{X > Y + 15\}$;

(c) $P\{Y > X + 15\}$.

8.4. Suppose that the number of units produced daily at factory A is a random variable with mean 20 and standard deviation 3 and the number produced at factory B is a random variable with mean 18 and standard deviation 6. Assuming independence, derive an upper bound for the probability that more units are produced today at factory B than at factory A.

8.5. The amount of time that a certain type of component functions before failing is a random variable with probability density function

$$f(x) = 2x \quad 0 < x < 1$$

Once the component fails, it is immediately replaced by another one of the same type. If we let X_i denote the lifetime of the ith component to be put in use, then $S_n = \sum_{i=1}^{n} X_i$ represents the time of the nth failure. The long-term rate at which failures occur, call it r, is defined by

$$r = \lim_{n\to\infty} \frac{n}{S_n}$$

Assuming that the random variables $X_i, i \ge 1$, are independent, determine r.

8.6. In Self-Test Problem 5, how many components would one need to have on hand to be approximately 90 percent certain that the stock will last at least 35 days?

8.7. The servicing of a machine requires two separate steps, with the time needed for the first step being an exponential random variable with mean .2 hour and the time for the second step being an independent exponential random variable with mean .3 hour. If a repair person has 20 machines to service, approximate the probability that all the work can be completed in 8 hours.

8.8. On each bet, a gambler loses 1 with probability .7, loses 2 with probability .2, or wins 10 with probability .1. Approximate the probability that the gambler will be losing after his first 100 bets.

8.9. Determine t so that the probability that the repair person in Self-Test Problem 7 finishes the 20 jobs within time t is approximately equal to .95.

8.10. A tobacco company claims that the amount of nicotine in one of its cigarettes is a random variable with mean 2.2 mg and standard deviation .3 mg. However, the average nicotine content of 100 randomly chosen cigarettes was 3.1 mg. Approximate the probability that the average would have been as high as or higher than 3.1 if the company's claims were true.

8.11. Each of the batteries in a collection of 40 batteries is equally likely to be either a type A or a type B battery. Type A batteries last for an amount of time that has mean 50 and standard deviation 15; type B batteries last for an amount of time that has mean 30 and standard deviation 6.

(a) Approximate the probability that the total life of all 40 batteries exceeds 1700.

(b) Suppose it is known that 20 of the batteries are type A and 20 are type B. Now approximate the probability that the total life of all 40 batteries exceeds 1700.

8.12. A clinic is equally likely to have 2, 3, or 4 doctors volunteer for service on a given day. No matter how may volunteer doctors there are on a given day, the numbers of patients seen by these doctors are independent Poisson random variables with mean 30. Let X denote the number of patients seen in the clinic on a given day.

(a) Find $E[X]$.

(b) Find $\text{Var}(X)$.

(c) Use a table of the standard normal probability distribution to approximate $P\{X > 65\}$.

8.13. The strong law of large numbers states that, with probability 1, the successive arithmetic averages of a sequence of independent and identically distributed random variables converge to their common mean μ. What do the successive geometric averages converge to? That is, what is

$$\lim_{n\to\infty} \left(\prod_{i=1}^{n} X_i\right)^{1/n}$$