Schaum's Solved Problems Series

Each title in this series is a complete and expert source of solved problems with solutions worked out in step-by-step detail.

Titles on the Current List Include:

3000 Solved Problems in Calculus
2500 Solved Problems in College Algebra and Trigonometry
2500 Solved Problems in Differential Equations
2000 Solved Problems in Discrete Mathematics
3000 Solved Problems in Linear Algebra
2000 Solved Problems in Numerical Analysis
3000 Solved Problems in Precalculus

Available at most college bookstores, or for a complete list of titles and prices, write to: Schaum Division
McGraw-Hill, Inc.
1221 Avenue of the Americas
New York, NY 10020

INTRODUCTION TO PROBABILITY AND STATISTICS

Principles and Applications for Engineering and the Computing Sciences

Third Edition

J. S. Milton

Radford University

Jesse C. Arnold

Virginia Polytechnic Institute and State University

McGraw-Hill, Inc.

New York St. Louis San Francisco Auckland Bogotá Caracas Lisbon London Madrid Mexico City Milan Montreal New Delhi San Juan Singapore Sydney Tokyo Toronto

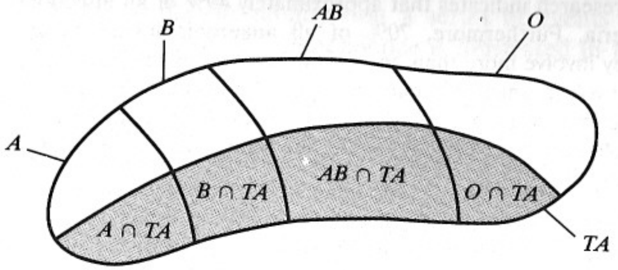


FIGURE 2.4
$TA = (A \cap TA) \cup (B \cap TA) \cup (AB \cap TA) \cup (O \cap TA)$.

Since the question asked is conditional, the first inclination is to try to apply the definition of conditional probability. Let us do so.

$$P[A|TA] = \frac{P[A \cap TA]}{P[TA]}$$

Unfortunately, neither $P[A \cap TA]$ nor $P[TA]$ is given. We must compute these quantities for ourselves. Each is easy to find. Note that by the multiplication rule,

$$\begin{aligned} P[A \cap TA] &= P[TA|A]P[A] \\ &= (.88)(.41) \\ &\doteq .36 \end{aligned}$$

Note also that the event TA can be partitioned into four mutually exclusive events, as shown in Fig. 2.4. That is,

$$TA = (A \cap TA) \cup (B \cap TA) \cup (AB \cap TA) \cup (O \cap TA)$$

By axiom 3,

$$P[TA] = P[A \cap TA] + P[B \cap TA] + P[AB \cap TA] + P[O \cap TA]$$

Applying the multiplication rule to each of the terms on the right-hand side of this equation, we obtain

$$\begin{aligned} P[TA] &= P[TA|A]P[A] + P[TA|B]P[B] \\ &\quad + P[TA|AB]P[AB] + P[TA|O]P[O] \\ &= (.88)(.41) + (.04)(.09) + (.10)(.04) + (.04)(.46) \\ &\doteq .39 \end{aligned}$$

Substituting .39 for $P[TA]$ yields

$$\begin{aligned} P[A|TA] &= \frac{P[A \cap TA]}{P[TA]} \\ &= \frac{.36}{.39} \\ &\doteq .92 \end{aligned}$$

The previous problem was solved using Bayes' rule. We now state the rule. By a partition of S we mean a collection of mutually exclusive events whose union is S.

Theorem 2.4.1 (Bayes' theorem). Let $A_1, A_2, A_3, \ldots, A_n$ be a collection of events which partition S. Let B be an event such that $P[B] \neq 0$. Then for any of the events A_j, $j = 1, 2, 3, \ldots, n$,

$$P[A_j|B] = \frac{P[B|A_j]P[A_j]}{\sum_{i=1}^{n} P[B|A_i]P[A_i]}$$

To see that Bayes' theorem was used in Example 2.4.1, make these notational changes:

$$A_1 = A \qquad A_3 = AB \qquad B = TA$$
$$A_2 = B \qquad A_4 = O$$

and use the theorem to find $P[A_1|B]$. Your answer will, of course, agree with that obtained earlier.

CHAPTER SUMMARY

In this chapter we presented some of the laws that govern the behavior of probabilities. We began with the axioms and from those we were able to derive the remaining laws. In particular, we derived the addition rule, which deals with the probability of the union of two events; the multiplication rule, which deals with the probability of the intersection of two events; and Bayes' theorem, which deals with conditional probability. We introduced and defined important terms that you should know. These are:

Conditional probability   Independent events

Care must be taken when using the concept of independence. In an applied problem, be sure that it is reasonable to assume that events A and B are independent before finding the probability of their joint occurrence via the definition $P[A \cap B] = P[A]P[B]$.

EXERCISES

Section 2.1

1. The probability that a wildcat well will produce oil is $1/13$. What is the probability that it will not be productive?
2. The theft of precious metals from companies in the United States is becoming a serious problem. The estimated probability that such a theft will involve a particular

metal is given below: (Based on data reported in "Materials Theft," *Materials Engineering*, February 1982, pp. 27–31.)

tin: 1/35	platinum: 1/35	nickel: 1/35
steel: 11/35	gold: 5/35	zinc: 1/35
copper: 8/35	aluminum: 2/35	silver: 4/35
titanium: 1/35		

(Note that these events are assumed to be mutually exclusive.)

(*a*) What is the probability that a theft of precious metal will involve gold, silver, or platinum?

(*b*) What is the probability that a theft will not involve steel?

3. Assuming the blood type distribution to be A: 41%, B: 9%, AB: 4%, O: 46%, what is the probability that the blood of a randomly selected individual will contain the A antigen? That it will contain the B antigen? That it will contain neither the A nor the B antigen?

4. Assume that the engine component of a spacecraft consists of two engines in parallel. If the main engine is 95% reliable, the backup is 80% reliable, and the engine component as a whole is 99% reliable, what is the probability that both engines will be operable? Use a Venn diagram to find the probability that the main engine will fail but the backup will be operable. Find the probability that the backup engine will fail but the main engine will be operable. What is the probability that the engine component will fail?

5. When an individual is exposed to radiation, death may ensue. Factors affecting the outcome are the size of the dose, the length and intensity of the exposure, and the biological makeup of the individual. The term LD_{50} is used to denote the dose that is usually lethal for 50% of the individuals exposed to it. Assume that in a nuclear accident 30% of the workers are exposed to the LD_{50} and die; 40% of the workers die; and 68% are exposed to the LD_{50} or die. What is the probability that a randomly selected worker is exposed to the LD_{50}? Use a Venn diagram to find the probability that a randomly selected worker is exposed to the LD_{50} but does not die. Find the probability that a randomly selected worker is not exposed to the LD_{50} but dies.

6. When a computer goes down, there is a 75% chance that it is due to an overload and a 15% chance that it is due to a software problem. There is an 85% chance that it is due to an overload or a software problem. What is the probability that both of these problems are at fault? What is the probability that there is a software problem but no overload?

***7.** Derive Theorem 2.1.1.

Hint: Note that $S = S \cup \varnothing$ and that S and $\varnothing$ are mutually exclusive. Apply axioms 3 and 1.

***8.** Derive Theorem 2.1.2.

Hint: Note that $S = A \cup A'$ and that A and A' are mutually exclusive. Apply axioms 3 and 1.

***9.** Let $A \subseteq B$. Show that $P[A] \leq P[B]$.

Hint: $B = A \cup (A' \cap B)$. Apply axioms 3 and 2.

***10.** Show that the probability of any event A is at most 1.

Hint: $A \subseteq S$. Apply Exercise 9 and axiom 1.

***11.** Derive the addition rule.

Hint: Note that

$$A_1 = (A_1 \cap A_2) \cup (A_1 \cap A_2')$$
$$A_2 = (A_1 \cap A_2) \cup (A_1' \cap A_2)$$
$$A_1 \cup A_2 = (A_1 \cap A_2) \cup (A_1 \cap A_2') \cup (A_1' \cap A_2)$$

Apply axiom 3 to each of these expressions.

12. Let A_1 and A_2 be mutually exclusive. By axiom 3 $P[A_1 \cup A_2] = P[A_1] + P[A_2]$. Show that the general addition rule yields the same result.

Section 2.2

13. Use the data of Exercise 5 to answer these questions.

(*a*) What is the probability that a randomly selected worker will die given that he is exposed to the lethal dose of radiation?

(*b*) What is the probability that a randomly selected worker will not die given that he is exposed to the lethal dose of radiation?

(*c*) What theorem allows you to determine the answer to (*b*) from knowledge of the answer to (*a*)?

(*d*) What is the probability that a randomly selected worker will die given that he is not exposed to the lethal dose?

(*e*) Is $P[\text{die}] = P[\text{die} \mid \text{exposed to lethal dose}]$? Did you expect these to be the same? Explain.

14. Use the data of Exercise 4 to answer these questions.

(*a*) What is the probability that in an engine system such as that described the backup engine will function given that the main engine fails?

(*b*) Is $P[\text{backup functions}] = P[\text{backup functions} \mid \text{main fails}]$? Did you expect these to be the same? Explain.

15. In a study of waters near power plants and other industrial plants that release wastewater into the water system it was found that 5% showed signs of chemical and thermal pollution, 40% showed signs of chemical pollution, and 35% showed evidence of thermal pollution. Assume that the results of the study accurately reflect the general situation. What is the probability that a stream that shows some thermal pollution will also show signs of chemical pollution? What is the probability that a stream showing chemical pollution will not show signs of thermal pollution?

16. A random digit generator on an electronic calculator is activated twice to simulate a random two-digit number. Theoretically, each digit from 0 to 9 is just as likely to appear on a given trial as any other digit.

(*a*) How many random two-digit numbers are possible?

(*b*) How many of these numbers begin with the digit 2?

(*c*) How many of these numbers end with the digit 9?

(*d*) How many of these numbers begin with the digit 2 and end with the digit 9?

(*e*) What is the probability that a randomly formed number ends with 9 given that it begins with a 2 Did you anticipate this result?

17. In studying the causes of power failures, these data have been gathered.

5% are due to transformer damage
80% are due to line damage
1% involve both problems

Based on these percentages, approximate the probability that a given power failure involves

(*a*) line damage given that there is transformer damage
(*b*) transformer damage given that there is line damage
(*c*) transformer damage but not line damage
(*d*) transformer damage given that there is no line damage
(*e*) transformer damage or line damage

Section 2.3

18. Let A_1 and A_2 be events such that $P[A_1] = .5$, $P[A_2] = .7$. What must $P[A_1 \cap A_2]$ equal for A_1 and A_2 to be independent?

19. Let A_1 and A_2 be events such that $P[A_1] = .6$, $P[A_2] = .4$ and $P[A_1 \cup A_2] = .8$. Are A_1 and A_2 independent?

20. Consider your answer to Exercise 14(*b*). Are the events A_1: the backup engine functions, and A_2: the main engine fails independent?

21. Studies in population genetics indicate that 39% of the available genes for determining the Rh blood factor are negative. Rh negative blood occurs if and only if the individual has two negative genes. One gene is inherited independently from each parent. What is the probability that a randomly selected individual will have Rh negative blood?

22. An individual's blood group (A, B, AB, O) is independent of the Rh classification. Find the probability that a randomly selected individual will have AB negative blood. *Hint:* See Example 2.1.1 and Exercise 21.

23. The use of plant appearance in prospecting for ore deposits is called geobotanical prospecting. One indicator of copper is a small mint with a mauve-colored flower. Suppose that, for a given region, there is a 30% chance that the soil has a high copper content and a 23% chance that the mint will be present there. If the copper content is high, there is a 70% chance that the mint will be present.

(*a*) Find the probability that the copper content will be high and the mint will be present.
(*b*) Find the probability that the copper content will be high given that the mint is present.

24. The most common water pollutants are organic. Since most organic materials are broken down by bacteria that require oxygen, an excess of organic matter may result in a depletion of available oxygen. In turn this can be harmful to other organisms living in the water. The demand for oxygen by the bacteria is called the biological oxygen demand (BOD). A study of streams located near an industrial complex revealed that 35% have a high BOD, 10% show high acidity, and 40% of streams with high acidity have a high BOD. Find the probability that a randomly selected stream will exhibit both characteristics.

25. A study of major flash floods that occurred over the last 15 years indicates that the probability that a flash flood warning will be issued is .5 and that the probability of dam failure during the flood is .33. The probability of dam failure given that a warning is issued is .17. Find the probability that a flash flood warning will be issued and a dam failure will occur. (Based on data reported in *McGraw-Hill Yearbook of Science and Technology*, 1980, pp. 185–186.)

***26.** The ability to observe and recall details is important in science. Unfortunately, the power of suggestion can distort memory. A study of recall is conducted as follows:

Subjects are shown a film in which a car is moving along a country road. There is no barn in the film. The subjects are then asked a series of questions concerning the film. Half the subjects are asked, "How fast was the car moving when it passed the barn?" The other half is not asked the question. Later each subject is asked; "Is there a barn in the film?" Of those asked the first question concerning the barn, 17% answer "yes"; only 3% of the others answer "yes." What is the probability that a randomly selected participant in this study claims to have seen the nonexistent barn? Is claiming to see the barn independent of being asked the first question about the barn? *Hint:*

$$P[\text{yes}] = P[\text{yes and asked about barn}] + P[\text{yes and not asked about barn}]$$

(Based on a study reported in *McGraw-Hill Yearbook of Science and Technology*, 1981, pp. 249–251.)

***27.** The probability that a unit of blood was donated by a paid donor is .67. If the donor was paid, the probability of contracting serum hepatitis from the unit is .0144. If the donor was not paid, this probability is .0012. A patient receives a unit of blood. What is the probability of the patient's contracting serum hepatitis from this source?

***28.** Show that the impossible event is independent of every other event.

***29.** Show that if A_1 and A_2 are independent, then A_1 and A'_2 are also independent. *Hint:* $A_1 = (A_1 \cap A_2) \cup (A_1 \cap A'_2)$.

***30.** Use Exercise 29 to show that if A_1 and A_2 are independent, then A'_1 and A'_2 are also independent.

***31.** It can be shown that the result of Exercise 30 holds for any collection of n independent events. That is, if $A_1, A_2, \ldots, A_n$ are independent, then $A'_1, A'_2, \ldots, A'_n$ are also independent. Use this result and the data of Example 2.3.4 to find the probability that at least one of the three computer systems will be operable at the time of the launch.

***32.** Let A_1 and A_2 be mutually exclusive events such that $P[A_1]P[A_2] > 0$. Show that these events are not independent.

***33.** Let A_1 and A_2 be independent events such that $P[A_1]P[A_2] > 0$. Show that these events are not mutually exclusive.

Section 2.4

34. Use the data of Example 2.4.1 to find the probability that an inductee who was typed as having type A blood actually had type B blood.

35. A test has been developed to detect a particular type of arthritis in individuals over 50 years old. From a national survey it is known that approximately 10% of the individuals in this age group suffer from this form of arthritis. The proposed test was given to individuals with confirmed arthritic disease, and a correct test result was obtained in 85% of the cases. When the test was administered to individuals of the same age group who were known to be free of the disease, 4% were reported to have the disease. What is the probability that an individual has this disease given that the test indicates its presence?

36. It is reported that 50% of all computer chips produced are defective. Inspection ensures that only 5% of the chips legally marketed are defective. Unfortunately, some chips are stolen before inspection. If 1% of all chips on the market are stolen, find the probability that a given chip is stolen given that it is defective.

37. As society becomes dependent on computers, data must be communicated via public communication networks such as satellites, microwave systems, and telephones. When a message is received, it must be authenticated. This is done by using a secret enciphering key. Even though the key is secret, there is always the probability that it will fall into the wrong hands, thus allowing an unauthentic message to appear to be authentic. Assume that 95% of all messages received are authentic. Furthermore, assume that only .1% of all unauthentic messages are sent using the correct key and that all authentic messages are sent using the correct key. Find the probability that a message is authentic given that the correct key is used.

REVIEW EXERCISES

38. A survey of engineering firms reveals that 80% have their own mainframe computer (M); 10% anticipate purchasing a mainframe computer in the near future (B); and 5% have a mainframe computer and anticipate buying another in the near future. Find the probability that a randomly selected firm:
    (*a*) has a mainframe computer or anticipates purchasing one in the near future
    (*b*) does not have a mainframe computer and does not anticipate purchasing one in the near future
    (*c*) anticipates purchasing a mainframe computer given that it does not currently have one
    (*d*) has a mainframe computer given that it anticipates purchasing one in the near future
39. In a simulation program, three random two-digit numbers will be generated independently of one another. These numbers assume the values $00, 01, 02, \ldots, 99$ with equal probability.
    (*a*) What is the probability that a given number will be less than 50?
    (*b*) What is the probability that each of the three numbers generated will be less than 50?
40. A power network involves three substations A, B, and C. Overloads at any of these substations might result in a blackout of the entire network. Past history has shown that if substation A alone experiences an overload, then there is a 1% chance of a network blackout. For stations B and C alone these percentages are 2% and 3%, respectively. Overloads at two or more substations simultaneously result in a blackout 5% of the time. During a heat wave there is a 60% chance that substation A alone will experience an overload. For stations B and C these percentages are 20 and 15%, respectively. There is a 5% chance of an overload at two or more substations simultaneously. During a particular heat wave a blackout due to an overload occurred. Find the probability that the overload occurred at substation A alone; substation B alone; substation C alone; two or more substations simultaneously.
41. A computer center has three printers A, B, and C, which print at different speeds. Programs are routed to the first available printer. The probability that a program is routed to printers A, B, and C are .6, .3, and .1, respectively. Occasionally a printer will jam and destroy a printout. The probability that printers A, B, and C will jam are .01, .05, and .04 respectively. Your program is destroyed when a printer jams. What is the probability that printer A is involved? Printer B is involved? Printer C is involved?

42. A chemical engineer is in charge of a particular process at an oil refinery. Past experience indicates that 10% of all shutdowns are due to equipment failure *alone*, 5% are due to a combination of equipment failure and operator error, and 40% involve operator error. A shutdown occurs. Find the probability that
    (*a*) equipment failure or operator error is involved
    (*b*) operator error alone is involved
    (*c*) neither operator error nor equipment failure is involved
    (*d*) operator error is involved given that equipment failure occurs
    (*e*) operator error is involved given that equipment failure does not occur
43. Assume that the probability that the air brakes on large trucks will fail on a particularly long downgrade is .001. Assume also that the emergency brakes on such trucks can stop a truck on this downgrade with probability .8. These braking systems operate independently of one another. Find the probability that
    (*a*) the air brakes fail but the emergency brakes can stop the truck
    (*b*) the air brakes fail and the emergency brakes cannot stop the truck
    (*c*) the emergency brakes cannot stop the truck given that the air brakes fail
44. Consider the problem of Example 1.2.3. Assume that sampling is independent and that at each stage the probability of obtaining a defective part when the process is working correctly is .01. If the process is working correctly, what is the probability that the first defective part will be obtained on the fourth sample? On or before the fourth sample?

TABLE 3.11

Time span, min	Random 3-digit number	Number of arrivals (x)	Number of departures (y)	Number on ground at end of time period (z)
1	015	0		100
	255		0	100
2	225	0		
	062		0	100
3	818	2		
	110		0	102
4	564	1		
	054		0	103
5	636	1		
	433		1	103

time, we could begin to answer such questions as: "On the average, how many planes are on the ground at a given time?" and "How much variability is there in the number of planes on the ground?"

CHAPTER SUMMARY

In this chapter we introduced the concept of a random variable and showed you how to distinguish a discrete random variable from one that is not discrete. We studied two functions, the density function and the cumulative distribution function, that are used to compute probabilities. The density gives the probability that X assumes a specific value x; the cumulative distribution gives the probability that X assumes a value less than or equal to x. The concept of expected value was introduced and used to define three important parameters, the mean (μ), the variance (σ^2), and the standard deviation (σ). The mean is a measure of the center of location of the distribution; the variance and standard deviation measure the variability of the random variable about its mean. The moment generating function was introduced as a means of finding the mean and variance of X. Special discrete distributions that find extensive use in all areas of application were presented. These are the geometric, hypergeometric, negative binomial, binomial, Bernoulli, uniform, and Poisson distributions. We also discussed briefly how to simulate a discrete distribution. We introduced and defined terms that you should know. These are:

Random variable	Variance
Discrete random variable	Standard deviation
Discrete density	Bernoulli trial
Cumulative distribution	Moment generating function
Expected value	Sampling with replacement
Mean	Sampling without replacement

EXERCISES

Section 3.1

In each of the following, identify the variable as discrete or not discrete.

1. T: the turnaround time for a computer job (the time it takes to run the program and receive the results).
2. M: the number of meteorites hitting a satellite per day.
3. N: the number of neutrons expelled per thermal neutron absorbed in fission of uranium-235.
4. Neutrons emitted as a result of fission are either prompt neutrons or delayed neutrons. Prompt neutrons account for about 99% of all neutrons emitted and are released within 10^{-14} s of the instant of fission. Delayed neutrons are emitted over a period of several hours. Let D denote the time at which a delayed neutron is emitted in a fission reaction.
5. Electrical resistance is the opposition which is offered by electrical conductors to the flow of current. The unit of resistance is the ohm. For example, a $2\frac{1}{2}$-inch electric bell will usually have a resistance somewhere between 1.5 and 3 ohms. Let O denote the actual resistance of a randomly selected bell of this type.
6. The number of power failures per month in the Tennessee Valley power network.

Section 3.2

7. Grafting, the uniting of the stem of one plant with the stem or root of another, is widely used commercially to grow the stem of one variety that produces fine fruit on the root system of another variety with a hardy root system. Most Florida sweet oranges grow on trees grafted to the root of a sour orange variety. The density for X, the number of grafts that fail in a series of five trials, is given by Table 3.12.
   (*a*) Find $f(5)$.
   (*b*) Find the table for F.
   (*c*) Use F to find the probability that at most three grafts fail; that at least two grafts fail.
   (*d*) Use F to verify that the probability of exactly three failures is .03.

TABLE 3.12

x	0	1	2	3	4	5
$f(x)$	.7	.2	.05	.03	.01	?

8. In blasting soft rock such as limestone, the holes bored to hold the explosives are drilled with a Kelly bar. This drill is designed so that the explosives can be packed into the hole before the drill is removed. This is necessary since in soft rock the hole often collapses as the drill is removed. The bits for these drills must be changed fairly often. Let X denote the number of holes that can be drilled per bit. The

density for X is given in Table 3.13. (Based on data reported in *The Explosives Engineer*, vol. 1, 1976, p. 12.)

(*a*) Find $f(8)$.
(*b*) Find the table for F.
(*c*) Use F to find the probability that a randomly selected bit can be used to drill between three and five holes inclusive.
(*d*) Find $P[X \le 4]$ and $P[X < 4]$. Are these probabilities the same?
(*e*) Find $F(-3)$ and $F(10)$. *Hint:* Express these in terms of the probabilities that they represent and their values will become obvious.

TABLE 3.13

x	1	2	3	4	5	6	7	8
$f(x)$	.02	.03	.05	.2	.4	.2	.07	?

9. Consider Example 1.2.1. Let X denote the number of computer systems operable at the time of the launch. Assume that the probability that each system is operable is .9.
(*a*) Use the tree of Fig. 1.2 to find the density table.
(*b*) There is a pattern to the probabilities in the density table. In particular,

$$f(x) = k(x)(.9)^x(.1)^{3-x}$$

where $k(x)$ gives the number of paths through the tree yielding a particular value for X. Use Exercise 23 of Chap. 1 to express $k(x)$ in terms of the number of computers available and the number operable.
(*c*) Find the table for F.
(*d*) Use F to find the probability that at least one system is operable at launch time.
(*e*) Use F to find the probability that at most one system is operable at the time of the launch.

10. It is known that the probability of being able to log on to a computer from a remote terminal at any given time is .7. Let X denote the number of attempts that must be made to gain access to the computer.
(*a*) Find the first four terms of the density table.
(*b*) Find a closed-form expression for $f(x)$.
(*c*) Find $P[X = 6]$.
(*d*) Find a closed-form expression for $F(x)$.
(*e*) Use F to find the probability that at most four attempts must be made to gain access to the computer.
(*f*) Use F to find the probability that at least five attempts must be made to gain access to the computer.

In parts (*c*), (*d*), and (*e*) of each of the next two exercises we point out the necessary and sufficient conditions for a function F to be a cumulative distribution function for a discrete random variable.

***11.** Even though there is no closed-form expression for the cumulative distribution function of Exercise 7, we can rewrite it as follows:

$$F(x) = \begin{cases} 0 & x < 0 \\ .70 & 0 \le x < 1 \\ .90 & 1 \le x < 2 \\ .95 & 2 \le x < 3 \\ .98 & 3 \le x < 4 \\ .99 & 4 \le x < 5 \\ 1.00 & x \ge 5 \end{cases}$$

(*a*) Draw the graph of this function. Recall from elementary calculus that graphs of this form are called step functions.
(*b*) Is F a continuous function?
(*c*) Is F a right continuous function?
(*d*) What is $\lim_{x \to \infty} F(x)$? What is $\lim_{x \to -\infty} F(x)$?
(*e*) Is F nondecreasing?

***12.** (*a*) Express the cumulative distribution function F of Exercise 8 in the manner shown in Exercise 11 and draw the graph of F.
(*b*) Is F a continuous function?
(*c*) Is F right continuous?
(*d*) What is $\lim_{x \to \infty} F(x)$? What is $\lim_{x \to -\infty} F(x)$?
(*e*) Is F nondecreasing?

***13.** State the conditions that are necessary and sufficient for a function F to be a cumulative distribution function for a discrete random variable.

Section 3.3

14. In an experiment to graft Florida sweet orange trees to the root of a sour orange variety, a series of five trials is conducted. Let X denote the number of grafts that fail. The density for X is given in Table 3.12.
(*a*) Find $E[X]$.
(*b*) Find μ_X.
(*c*) Find $E[X^2]$.
(*d*) Find Var X.
(*e*) Find σ_X^2.
(*f*) Find the standard deviation for X.
(*g*) What physical unit is associated with σ_X?

15. The density for X, the number of holes that can be drilled per bit while drilling into limestone is given in Table 3.13.
(*a*) Find $E[X]$ and $E[X^2]$.
(*b*) Find Var X and σ_X.
(*c*) What physical unit is associated with σ_X?

16. Use the density derived in Exercise 9 to find the expected value and variance for X, the number of computer systems operable at the time of the launch. Can you express $E[X]$ and Var X in terms of n, the number of systems available, and p, the probability that a given system will be operable?

***17.** The probability p of being able to log on to a computer from a remote terminal at any given time is .7. Let X denote the number of attempts that must be made to gain access to the computer. Find $E[X]$. Can you express $E[X]$ in terms of p?

Hint: The series $\sum_{x=1}^{\infty} x(.7)(.3)^{x-1} = E[X]$ is not geometric. To find $E[X]$, expand this series and the series $.3E[X]$. Subtract the two to form the series $.7E[X]$. Evaluate this *geometric* series and solve for $E[X]$.

*18. The probability that a cell will fuse in the presence of polyethylene glycol is 1/2. Let Y denote the number of cells exposed to antigen-carrying lymphocytes to obtain the first fusion. Use the method of Exercise 17 to find $E[Y]$.

*19. Let X be a discrete random variable with density f. Let c be any real number. Show that
(*a*) $E[c] = c$. *Hint:* Remember that constants can be factored from summations and that $\sum_{\text{all } x} f(x) = 1$.
(*b*) $E[cX] = cE[X]$.

*20. Use the rules for expectation to verify that $\text{Var}\, c = 0$ and $\text{Var}\, cX = c^2 \text{Var}\, X$ for any real number c. *Hint:* $\text{Var}\, c = E[c^2] - (E[c])^2$.

21. Let X and Y be independent random variables with $E[X] = 3$, $E[X^2] = 25$, $E[Y] = 10$ and $E[Y^2] = 164$.
(*a*) Find $E[3X + Y - 8]$.
(*b*) Find $E[2X - 3Y + 7]$.
(*c*) Find Var X.
(*d*) Find σ_X.
(*e*) Find Var Y.
(*f*) Find σ_Y.
(*g*) Find $\text{Var}[3X + Y - 8]$.
(*h*) Find $\text{Var}[2X - 3Y + 7]$.
(*i*) Find $E[(X - 3)/4]$ and $\text{Var}[(X - 3)/4]$.
(*j*) Find $E[(Y - 10)/8]$ and $\text{Var}[(Y - 10)/8]$.
(*k*) The results of parts (*i*) and (*j*) are not coincidental. Can you generalize and verify the conjecture suggested by these two exercises?

*22. Consider the function f defined by
$$f(x) = (1/2)2^{-|x|} \qquad x = \pm 1, \pm 2, \pm 3, \pm 4, \ldots$$
(*a*) Verify that this is the density for a discrete random variable X. *Hint:* Expand the series $\sum_{\text{all } x} f(x)$ for a few terms. A recognizable series will develop!
(*b*) Let $g(X) = (-1)^{|X|-1}[2^{|X|}/(2|X| - 1)]$. Show that $\sum_{\text{all } x} g(x)f(x) < \infty$. *Hint:* Expand the series for a few terms. You will obtain an alternating series that can be shown to converge.
(*c*) Show that $\sum_{\text{all } x} |g(x)| f(x)$ does not converge. This will show that $E[g(X)]$ does not exist. *Hint:* Expand the series for a few terms. You will obtain a series that is term by term larger than the diverging harmonic type series $(1/3)\sum_{x=1}^{\infty} 1/x$.

*23. (*An application to sort algorithms.*) In studying various sort algorithms in computer science, it is of interest to compare their efficiency by estimating the average number of interchanges needed to sort random arrays of various sizes. It is also of interest to compare these estimated averages to the "ideal" average, where by "ideal" we mean the expected minimum number of interchanges needed to sort the array. In this exercise you will derive this ideal average. (American Mathematical Association of Two-Year Colleges, "A Note on the Minimum Number of Interchanges Needed to Sort a Random Array," with T. McMillan, I. Liss, and J. Milton, Fall 1990.)
(*a*) Consider a random array of length n. When the positions of exactly two elements of the array are exchanged, we say that an "interchange" has taken place. Let X_n denote the minimum number of interchanges necessary to sort an array of size n. Note that
$$X_n = X_{n-1} + I$$
where $I = 0$ if the last element of the array is in the correct position and $I = 1$ otherwise. Argue that $P[I = 0] = 1/n$ and $P[I = 1] = 1 - (1/n)$.
(*b*) Show that
$$E[I] = 1 - \frac{1}{n}$$
(*c*) Argue that
$$E[X_n] = E[X_{n-1}] + 1 - \frac{1}{n}$$
$$E[X_{n-1}] = E[X_{n-2}] + 1 - \frac{1}{n-1}$$
$$E[X_{n-2}] = E[X_{n-3}] + 1 - \frac{1}{n-2}$$
$$\vdots$$
$$E[X_3] = E[X_2] + 1 - \frac{1}{3}$$
$$E[X_2] = E[X_1] + 1 - \frac{1}{2}$$
$$E[X_1] = 0$$
(*d*) Use a recursive argument to show that
$$E[X_n] = (n - 1) - \sum_{i=2}^{n} \frac{1}{i}$$
(*e*) Illustrate the expression given in part (*d*) by finding $E[X_5]$.
(*f*) Elementary calculus can be used to approximate $E[X_n]$ by noting that
$$\sum_{i=2}^{n} \frac{1}{i} \doteq \int_{1.5}^{n+.5} \frac{1}{t}\, dt$$
Use this idea to approximate $E[X_5]$ and to compare the result to the exact solution found in part (*e*).
(*g*) A random digit generator is used to generate sets of 100 different three-digit numbers lying between 0 and 1. What is the ideal average number of interchanges needed to sort such an array?

Section 3.4

24. The probability that a wildcat well will be productive is 1/13. Assume that a group is drilling wells in various parts of the country so that the status of one well has no bearing on that of any other. Let X denote the number of wells drilled to obtain the first strike.
(*a*) Verify that X is geometric and identify the value of the parameter p.
(*b*) What is the exact expression for the density for X?

(c) What is the exact expression for the moment generating function for X?
(d) What are the numerical values of $E[X]$, $E[X^2]$, σ^2, and σ?
(e) Find $P[X \geq 2]$.
(f) Suppose that 10 wells have been drilled and no strikes have occurred. What is the probability that at least 2 more wells must be drilled in order to obtain the first strike? What property can be used to answer this question?

(g) see 84.

25. The zinc-phosphate coating on the threads of steel tubes used in oil and gas wells is critical to their performance. To monitor the coating process, an uncoated metal sample with known outside area is weighed and treated along with the lot of tubing. This sample is then stripped and reweighed. From this it is possible to determine whether or not the proper amount of coating was applied to the tubing. Assume that the probability that a given lot is unacceptable is .05. Let X denote the number of runs conducted to produce an unacceptable lot. Assume that the runs are independent in the sense that the outcome of one run has no effect on that of any other. (Based on a report in *American Machinist*, November 1982, p. 81.)
(a) Verify that X is geometric. What is "success" in this experiment? What is the numerical value of p?
(b) What is the exact expression for the density for X?
(c) What is the exact expression for the moment generating function for X?
(d) What are the numerical values of $E[X]$, $E[X^2]$, σ^2, and σ?
(e) Find the probability that the number of runs required to produce an unacceptable lot is at least 3.
(f) Suppose that 19 lots have been inspected and that all 19 are deemed to be acceptable. What is the probability that at least 21 acceptable lots are produced prior to obtaining the first defective lot?

26. Let X be geometric with probability of success p. Prove that when x is a positive integer, $F(x) = 1 - q^x$. Verify that this result holds true for the density given in Example 3.2.4. Argue that, in general, $F(x) = 1 - q^{[x]}$.

27. Find the expression for the cumulative distribution function for the random variable of Exercise 25. Use this function to find the probability that at most three runs are required to produce an unacceptable lot.

28. A system used to read electric meters automatically requires the use of a 128-bit computer message. Occasionally random interference causes a digit reversal resulting in a transmission error. Assume that the probability of a digit reversal for each bit is $1/1000$. Let X denote the number of transmission errors per 128-bit message sent. Is X geometric? If not, what geometric property fails?

29. Verify that the random variable X of Exercise 17 is geometric. Use Theorem 3.4.3 to find $E[X]$ and compare your answer to that obtained in Exercise 17.

30. Verify that the random variable Y of Exercise 18 is geometric. Use Theorem 3.4.3 to find $E[Y]$ and compare your answer to that obtained in Exercise 18.

31. Consider the random variable X whose density is given by

$$f(x) = \frac{(x-3)^2}{5} \qquad x = 3, 4, 5$$

(a) Verify that this function is a density for a discrete random variable.
(b) Find $E[X]$ directly. That is, evaluate $\sum_{\text{all } x} xf(x)$.
(c) Find the moment generating function for X.

(d) Use the moment generating function to find $E[X]$, thus verifying your answer to part (b) of this exercise.
(e) Find $E[X^2]$ directly. That is, evaluate $\sum_{\text{all } x} x^2 f(x)$.
(f) Use the moment generating function to find $E[X^2]$, thus verifying your answer to part (e) of this exercise.
(g) Find σ^2 and σ.

32. A discrete random variable has moment generating function

$$m_X(t) = e^{2(e^t - 1)}$$

(a) Find $E[X]$.
(b) Find $E[X^2]$.
(c) Find σ^2 and σ.

*33. Let X have a geometric distribution with parameter p.
(a) Show that the probability that X is odd is $p/(1 - q^2)$, where $q = 1 - p$. *Hint:* If x is odd, then x can be expressed in the form $x = 2m - 1$ for $m = 1, 2, 3, \ldots$.
(b) Show that the probability that X is odd is never $1/2$ regardless of the value chosen for p.

34. (*Discrete uniform distribution.*) A discrete random variable is said to be *uniformly distributed* if it assumes a finite number of values with each value occurring with the same probability. If we consider the generation of a single random digit, then Y, the number generated, is uniformly distributed with each possible digit occurring with probability $1/10$. In general, the density for a uniformly distributed random variable is given by

$$f(x) = 1/n \qquad \begin{array}{l} n \text{ a positive integer} \\ x = x_1, x_2, x_3, \ldots, x_n \end{array}$$

(a) Find the moment generating function for a discrete uniform random variable.
(b) Use the moment generating function to find $E[X]$, $E[X^2]$, and σ^2.
(c) Find the mean and variance for the random variable Y, the number obtained when a random digit generator is activated once. *Hint:* The sum of the first n positive integers is $n(n + 1)/2$; the sum of the squares of the first n positive integers is $n(n + 1)(2n + 1)/6$.

35. Let the density for X be given by

$$f(x) = ce^{-x} \qquad x = 1, 2, 3, \ldots$$

(a) Find the value of c that makes this a density.
(b) Find the moment generating function for X.
(c) Use $m_X(t)$ to find $E[X]$.

Section 3.5

36. Let X be binomial with parameters $n = 15$ and $p = .2$.
(a) Find the expression for the density for X.
(b) Find the expression for the moment generating function for X.
(c) Find $E[X]$ and Var X.
(d) Find $E[X]$, $E[X^2]$, and Var X using the moment generating function, thus verifying your answer to part (c) of this exercise.

(*e*) Find $P[X \le 1]$ by evaluating the density directly. Compare your answer to that given in Table I of App. A.

(*f*) Draw dot diagrams similar to that of Fig. 3.2 to illustrate each of these probabilities, and find the probabilities using Table I of App. A.

$$P[X \le 5] \qquad P[X \ge 3]$$
$$P[X < 5] \qquad F(9)$$
$$P[2 \le X \le 7] \qquad F(20)$$
$$P[2 \le X < 7] \qquad P[X = 10]$$

37. Albino rats used to study the hormonal regulation of a metabolic pathway are injected with a drug that inhibits body synthesis of protein. The probability that a rat will die from the drug before the experiment is over is .2. If 10 animals are treated with the drug, how many are expected to die before the experiment ends? What is the probability that at least eight will survive? Would you be surprised if at least five died during the course of the experiment? Explain, based on the probability of this occurring.

38. Consider Example 1.2.1. The random variable X is the number of computer systems operable at the time of a space launch. The systems are assumed to operate independently. Each is operable with probability .9.

(*a*) Argue that X is binomial and find its density. Compare your answer to that obtained in Exercise 9(*b*).

(*b*) Find $E[X]$ and Var X.

39. In humans, geneticists have identified two sex chromosomes, R and Y. Every individual has an R chromosome, and the presence of a Y chromosome distinguishes the individual as male. Thus the two sexes are characterized as *RR* (female) and *RY* (male). Color blindness is caused by a recessive allele on the R chromosome, which we denote by r. The Y chromosome has no bearing on color blindness. Thus relative to color blindness, there are three genotypes for females and two for males:

Female	Male
RR (normal)	*RY* (normal)
Rr (carrier)	*rY* (color-blind)
rr (color-blind)	

A child inherits one sex chromosome randomly from each parent.

(*a*) A carrier of color blindness parents a child with a normal male. Construct a tree to represent the possible genotypes for the child. Use the tree to find the probability that a given child will be a color-blind male.

(*b*) If the couple has five children, what is the expected number of color-blind males? What is the probability that three or more will be color-blind males?

40. In scanning electron microscopy photography, a specimen is placed in a vacuum chamber and scanned by an electron beam. Secondary electrons emitted from the specimen are collected by a detector and an image is displayed on a cathode-ray tube. This image is photographed. In the past a 4-× 5-inch camera has been used. It is thought that a 35-millimeter (mm) camera can obtain the same clarity. This type of camera is faster and more economical than the 4-× 5-inch variety. (Based on a report entitled "Adaptation of a Thirty-Five Millimeter Photographic System for a Scanning Electron Microscope," E. A. Lawton, *Biological Photography*, vol. 50, no. 3, July 1982, p. 65.)

(*a*) Photographs of 15 specimens are made using each camera system. These unmarked photographs are judged for clarity by an impartial judge. The judge is asked to select the better of the two photographs from each pair. Let X denote the number selected taken by a 35-mm camera. If there is really no difference in clarity and the judge is randomly selecting photographs, what is the expected value of X?

(*b*) Would you be surprised if the judge selected 12 or more photographs taken by the 35-mm camera? Explain, based on the probability involved.

(*c*) If $X \ge 12$, do you think that there is reason to suspect that the judge is not selecting the photographs at random?

41. It has been found that 80% of all printers used on home computers operate correctly at the time of installation. The rest require some adjustment. A particular dealer sells 10 units during a given month.

(*a*) Find the probability that at least nine of the printers operate correctly upon installation.

(*b*) Consider 5 months in which 10 units are sold per month. What is the probability that at least 9 units operate correctly in each of the 5 months?

42. It is possible for a computer to pick up an erroneous signal that does not show up as an error on the screen. The error is called a silent error. A particular terminal is defective, and when using the system word processor, it introduces a silent paging error with probability .1. The word processor is used 20 times during a given week.

(*a*) Find the probability that no silent paging errors occur.

(*b*) Find the probability that at least one such error occurs.

(*c*) Would it be unusual for more than four such errors to occur? Explain, based on the probability involved.

43. (*a*) Find the moment generating function for a binomial random variable with parameters n and p. *Hint:* Let

$$\binom{n}{x} e^{tx} p^x (1-p)^{n-x} = \binom{n}{x} (pe^t)^x (1-p)^{n-x}$$

and apply the binomial theorem.

(*b*) Use $m_X(t)$ to show that $E[X] = np$.

(*c*) Use $m_X(t)$ to show that $E[X^2] = n^2p^2 - np^2 + np$.

(*d*) Show that Var $X = npq$, where $q = 1 - p$.

***44.** Find the mean value for a binomial random variable with parameters n and p from the definition. That is, evaluate

$$\sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

Hint:

$$\sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

Now let $z = x - 1$ and evaluate

$$\sum_{z=0}^{n-1} (z+1)\binom{n}{z+1} p^{z+1}(1-p)^{n-(z+1)}$$

45. (*Point binomial or Bernoulli distribution.*) Assume that an experiment is conducted and that the outcome is considered to be either a success or a failure. Let p denote the probability of success. Define X to be 1 if the experiment is a success and 0 if it is a failure. X is said to have a *point binomial* or a *Bernoulli* distribution with parameter p.
(*a*) Argue that X is a binomial random variable with $n = 1$.
(*b*) Find the density for X.
(*c*) Find the moment generating function for X.
(*d*) Find the mean and variance for X.
(*e*) In DNA replication errors can occur that are chemically induced. Some of these errors are "silent" in that they do not lead to an observable mutation. Growing bacteria are exposed to a chemical that has probability .14 of inducing an observable error. Let X be 1 if an observable mutation results and let X be 0 otherwise. Find $E[X]$.

46. A binomial random variable has mean 5 and variance 4. Find the values of n and p that characterize the distribution of this random variable.

Section 3.6

***47.** In this exercise you will show that the density for the negative binomial distribution sums to 1.
(*a*) Show that

$$\sum_{x=r}^{\infty} \binom{x-1}{r-1} q^{x-r} p^r = \sum_{z=0}^{\infty} \binom{z+r-1}{r-1} q^z p^r$$

where $q = 1 - p$. *Hint:* Let $z = x - r$.
(*b*) Show that the Taylor series expansion for $h(q) = 1/(1-q)^r$ about 0 is given by

$$1 + rq + \frac{r(r+1)q^2}{2} + \frac{r(r+1)(r+2)q^3}{3!} + \cdots = 1/(1-q)^r$$

(*c*) Show that

$$\sum_{z=0}^{\infty} \binom{z+r-1}{r-1} q^z = 1/(1-q)^r$$

(*d*) Show that

$$\sum_{z=0}^{\infty} \binom{z+r-1}{r-1} q^z p^r = 1$$

Hint: Factor the term p^r from the expression and use part (*c*) with $q = 1 - p$.

***48.** In this exercise you will derive the moment generating function for the negative binomial distribution with parameters r and p.
(*a*) Show that

$$m_X(t) = \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} q^{x-r} p^r = (pe^t)^r \sum_{z=0}^{\infty} \binom{z+r-1}{r-1} (qe^t)^z$$

Hint: Let $z = x - r$.

(*b*) Use the idea given in Exercise 47(*c*) to show that

$$\sum_{z=0}^{\infty} \binom{z+r+1}{r-1} (qe^t)^z = 1/(1-qe^t)^r$$

(*c*) Show that

$$m_X(t) = \frac{(pe^t)^r}{(1-qe^t)^r}$$

***49.** Use the moment generating function to show that the mean of a negative binomial distribution with parameters r and p is r/p.

***50.** Use the moment generating function to show that $E[X^2] = (r^2 + rq)/p^2$ and that $\text{Var } X = rq/p^2$ for the negative binomial distribution with parameters r and p.

***51.** Show that the geometric distribution is a special case of the negative binomial distribution with $r = 1$. Find the mean and variance of a geometric random variable with parameter p using Exercises 49 and 50. Compare your answer with the results of Theorem 3.4.3.

***52.** A vaccine for desensitizing patients to bee stings is to be packed with three vials in each box. Each vial is checked for strength before packing. The probability that a vial meets specifications is .9. Let X denote the number of vials that must be checked to fill a box. Find the density for X and its mean and variance. Would you be surprised if seven or more vials have to be tested to find three that meet specifications? Explain, based on the probability of this occurrence.

***53.** Some characteristics in animals are said to be sex-influenced. For example, the production of horns in sheep is governed by a pair of alleles, H and h. The allele H for the production of horns is dominant in males but recessive in females. The allele h for hornlessness is dominant in females and recessive in males. Thus, given a heterozygous male (Hh) and a heterozygous female (Hh), the male will have horns but the female will be hornless. Assume that two such animals mate and the offspring is just as likely to be male as female. The lamb inherits one gene for horns randomly from each parent. Use a tree diagram to show that the probability that a lamb will be a hornless female is 3/8. Find the average number of lambs born to obtain the second hornless female. Would you be surprised if at most five lambs were born to obtain the second hornless female? Explain.

Section 3.7

54. Suppose that X is hypergeometric with $N = 20$, $r = 17$, and $n = 5$. What are the possible values for X? What is $E[X]$ and $\text{Var } X$?

55. Suppose that X is hypergeometric with $N = 20$, $r = 3$, and $n = 5$. What are the possible values for X? What is $E[X]$ and $\text{Var } X$?

56. Suppose that X is hypergeometric with $N = 20$, $r = 10$, and $n = 5$. What are the possible values for X? What is $E[X]$ and $\text{Var } X$?

57. Twenty microprocessor chips are in stock. Three have etching errors that cannot be detected by the naked eye. Five chips are selected and installed in field equipment.
(*a*) Find the density for X, the number of chips selected that have etching errors.
(*b*) Find $E[X]$ and $\text{Var } X$.
(*c*) Find the probability that no chips with etching errors will be selected.
(*d*) Find the probability that at least one chip with an etching error will be chosen.

58. Production line workers assemble 15 automobiles per hour. During a given hour, four are produced with improperly fitted doors. Three automobiles are selected at random and inspected. Let X denote the number inspected that have improperly fitted doors.
(*a*) Find the density for X.
(*b*) Find $E[X]$ and Var X.
(*c*) Find the probability that at most one will be found with improperly fitted doors.

59. A distributor of computer software wants to obtain some customer feedback concerning its newest package. Three thousand customers have purchased the package. Assume that 600 of these customers are dissatisfied with the product. Twenty customers are randomly sampled and questioned about the package. Let X denote the number of dissatisfied customers sampled.
(*a*) Find the density for X.
(*b*) Find $E[X]$ and Var X.
(*c*) Set up the calculations needed to find $P[X \leq 3]$.
(*d*) Use the binomial tables to approximate $P[X \leq 3]$.

60. A random telephone poll is conducted to ascertain public opinion concerning the construction of a nuclear power plant in a particular community. Assume that there are 150,000 numbers listed for private individuals and that 90,000 of these would elicit a negative response if contacted. Let X denote the number of negative responses obtained in 15 calls.
(*a*) Find the density for X.
(*b*) Find $E[X]$ and Var X.
(*c*) Set up the calculations needed to find $P[X \geq 6]$.
(*d*) Use the binomial tables to approximate $P[X \geq 6]$.

Section 3.8

61. Let X be a Poisson random variable with parameter $k = 10$.
(*a*) Find $E[X]$.
(*b*) Find Var X.
(*c*) Find σ_X.
(*d*) Find the expression for the density for X.
(*e*) Find $P[X \leq 4]$.
(*f*) Find $P[X < 4]$.
(*g*) Find $P[X = 4]$.
(*h*) Find $P[X \geq 4]$.
(*i*) Find $P[4 \leq X \leq 9]$.

62. A particular nuclear plant releases a detectable amount of radioactive gases twice a month on the average. Find the probability that there will be at most four such emissions during a month. What is the expected number of emissions during a 3-month period? If, in fact, 12 or more emissions are detected during a 3-month period, do you think that there is a reason to suspect the reported average figure of twice a month? Explain, on the basis of the probability involved.

63. Geophysicists determine the age of a zircon by counting the number of uranium fission tracks on a polished surface. A particular zircon is of such an age that the average number of tracks per square centimeter is five. What is the probability that a 2-centimeter-square sample of this zircon will reveal at most three tracks, thus leading to an underestimation of the age of the material?

64. California is hit by approximately 500 earthquakes that are large enough to be felt every year. However, those of destructive magnitude occur on the average once every year. Find the probability that California will experience at least one earthquake of this magnitude during a 6-month period. Would it be unusual to have 3 or more earthquakes of destructive magnitude in a 6-month period? Explain, based on the probability of this occurring. (Based on data presented in Robert Iacopi, *Earthquake Country*, Lana Books, Menlo Park, Calif., 1971.)

65. Load-bearing structures in underground mines are often required to carry additional loads while mining operations are in progress. As the structures adjust to this new weight small-scale displacements take place that result in the release of seismic and acoustic energy, called *rock noise*. This energy can be detected using special geophysical equipment. Assume that in a particular mine the average number of rock noises recorded during normal activity is 3 per hour. Would you consider it unusual if more than 10 were detected in a 2-hour period? Explain, based on the probability involved. (Based on "A multichannel rock noise monitoring system," T. Gowd and M. S. Rao, *Journal of Mines, Metals and Fuels*, September 1981, pp. 288–290.)

66. A burr is a thin ridge or rough area that occurs when shaping a metal part. These must be removed by hand or by means of some newer method such as water jets, thermal energy, or electrochemical processing before the part can be used. Assume that a part used in automatic transmissions typically averages two burrs each. What is the probability that the total number of burrs found on seven randomly selected parts will be at most four? (Based on "Advances in Deburring," B. Hignett, *Production Engineering*, December 1982, pp. 44–47.)

67. Cast iron is an alloy composed primarily of iron together with smaller amounts of other elements, including carbon, silicon, sulfur, and phosphorus. The carbon occurs as graphite, which is soft, or iron carbide, which is very hard and brittle. The type of cast iron produced is determined by the amount and distribution of carbon in the iron. Five types of cast iron are identifiable. These are gray, compacted graphite, ductile, malleable, and white. In malleable cast iron the carbon is present as discrete graphite particles. Assume that in a particular casting these particles average 20 per square inch. Would it be unusual to see a 1/4-inch-square area of this casting with fewer than two graphite particles? Explain, based on the probability involved. (Based on "Space Age Metal: Cast Iron," J. Lalich, *Mines Magazine*, February 1982, pp. 2–6.)

68. A Poisson random variable is such that it assumes the values 0 and 1 with equal probability. Find the value of the Poisson parameter k for this variable.

69. Prove Theorem 3.8.1. *Hint:* Note that

$$m_X(t) = E[e^{tX}] = \sum_{x=0}^{\infty} \frac{e^{tx}e^{-k}k^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-k}(ke^t)^x}{x!}$$

and use the Maclaurin series.

70. Let X be binomial with $n = 20$ and $p = .05$. Find $P[X = 0]$ using the binomial density and compare your answer to that obtained by using the Poisson approximation to this probability. Do you think that the error in the approximation is large?

71. In *Escherichia coli*, a bacterium often found in the human digestive tract, 1 cell in every 10^9 will mutate from streptomycin sensitivity to streptomycin resistance. This

mutation can cause the individual involved to become resistant to the antibiotic streptomycin. In observing 2 billion (2×10^9) such cells, what is the probability that none will mutate? What is the probability that at least one will mutate?

72. The spontaneous flipping of a bit stored in a computer memory is called a "soft fail." Soft fails are rare, averaging only one per million hours per chip. However, the probability of a soft fail is increased when the chip is exposed to α particles (helium nuclei), which occur naturally in the environment. Assume that the probability of a soft fail under these conditions is $1/1000$. If a chip containing 6000 bits is exposed to α particles, what is the probability that there will be at least one soft fail? Would you be surprised if there were more than five soft fails? Explain, based on the probability of this occurring. (*McGraw-Hill Yearbook of Science and Technology*, 1981, p. 142.)

Section 3.9

73. Use Table II of App. A to simulate the arrival and departure of planes to the airport described in Example 3.9.2 for 10 more 1-minute periods. Based on these data, approximate the average number of planes on the ground at a given time by finding the arithmetic average of the values of Z simulated in the experiment.

74. Consider the random variable X, the number of runs conducted to produce an unacceptable lot when coating steel tubes (see Exercise 25.) X is geometric with $p = .05$. Divide the 100 possible two-digit numbers into two categories, with numbers 00–04 denoting the production of an unacceptable lot and the remaining numbers denoting the production of an acceptable lot. Simulate the experiment of producing lots until an unacceptable one is obtained 10 times. Record the value obtained for X in each simulation. Based on these data, approximate the average value of X. Does your approximate value lie close to the theoretical mean value of 20? If not, run the simulation 10 more times. Is the arithmetic average of your observed values for X closer to 20 this time?

REVIEW EXERCISES

75. A large microprocessor chip contains multiple copies of circuits. If a circuit fails, the chip knows it and knows how to select the proper logic to repair itself. The average number of defects per chip is 300. What is the probability that 10 or fewer defects will be found in a randomly selected region that comprises 5% of the total surface area? What is the probability that more than 10 defects are found? ("Self-Repairing Chips," *Datamation*, May 1983, p. 68.)

76. When a program is submitted to the computer in a time-sharing system, it is processed on a space-available basis. Past experience shows that a program submitted to one such system is accepted for processing within 1 minute with probability .25. Assume that during the course of a day five programs are submitted with enough time between submissions to ensure independence. Let X denote the number of programs accepted for processing within 1 minute.
    (*a*) Find $E[X]$ and Var X.
    (*b*) Find the probability that none of these programs will be accepted for processing within 1 minute.
    (*c*) Five programs are submitted on each of two consecutive days. What is the probability that no programs will be accepted for processing within 1 minute during this two-day period?

77. A new type of brake lining is being studied. It is thought that the lining will last for at least 70,000 miles on 90% of the cars in which it is used. Laboratory trials are conducted to simulate the driving experience of 100 cars in which this lining is used. Let X denote the number of cars whose brakes must be relined before the 70,000-mile mark.
    (*a*) What is the distribution of X? What is $E[X]$?
    (*b*) What distribution can be used to approximate probabilities for X?
    (*c*) Suppose that we agree that the 90% figure is too high if 17 or more of the 100 cars require a relinement prior to the 70,000-mile mark. What is the probability that we will come to this conclusion by chance even though the 90% figure is correct?

78. A bank of guns fires on a target one after the other. Each has probability $1/4$ of hitting the target on a given shot. Find the probability that the second hit comes before the seventh gun fires.

79. In a video game the player attempts to capture a treasure lying behind one of five doors. The location of the treasure varies randomly in such a way that at any given time it is just as likely to be behind one door as any other. When the player knocks on a given door, the treasure is his if it lies behind that door. Otherwise he must return to his original starting point and approach the doors through a dangerous maze again. Once the treasure is captured, the game ends. Let X denote the number of trials needed to capture the treasure. Find the average number of trials needed to capture the treasure. Find $P[X \le 3]$. Find $P[X > 3]$.

80. An automobile repair shop has 10 rebuilt transmissions in stock. Three are not in correct working order and have an internal defect that will cause trouble within the first 1000 miles of operation. Four of these transmissions are randomly selected and installed in customers' cars. Find the probability that no defective transmissions are installed. Find the probability that exactly one defective transmission is installed.

81. A computer terminal can pick up an erroneous signal from the keyboard that does not show up on the screen. This creates a silent error that is difficult to detect. Assume that for a particular keyboard the probability that this will occur per entry is $1/1000$. In 12,000 entries find the probability that no silent errors occur. Find the probability of at least one silent error.

82. It is thought that 1 of every 10 cars on the road has a speedometer that is miscalibrated to the extent that it reads at least 5 miles per hour low. During the course of a day 15 drivers are stopped and charged with exceeding the speed limit by at least 5 miles per hour. Would you be surprised to find that at least 5 of the cars involved have miscalibrated speedometers? Explain, based on the probability of observing a result this unusual by chance.

83. Let

$$f(x) = \frac{x^2}{14} \qquad x = 1, 2, 3$$

    (*a*) Show that f is the density for a discrete random variable.
    (*b*) Find $E[X]$ and $E[X^2]$ from the definition of these terms.
    (*c*) Find $m_X(t)$.
    (*d*) Use $m_X(t)$ to verify your answers to part (*b*).
    (*e*) Find Var X and σ.

84. Find the expression for the cumulative distribution function for the random variable of Exercise 24. Use this function to find the probability that at least three wells must be drilled to obtain the first strike.

85. Consider the moment generating function given below. In each case, state the name of the distribution involved and the numerical value of the parameters that identify the distribution. For example, if the distribution is binomial, state the value of n and p; if geometric, give the value of p.

(*a*) $(.2 + .8e^t)^{10}$

(*b*) $e^{5(e^t 1)}$

(*c*) $(.7 + .3e^t)$

(*d*) $\dfrac{.6e^t}{1 - .4e^t}$

(*e*) $\dfrac{(.3e^t)^5}{(1 - .7e^t)^5}$

(*f*) $e^{e^t 1}$

86. For each of the distribution in Exercise 85, give the numerical values of the mean, variance, and standard deviation.

87. Consider the problem of Example 1.2.3. Assume that sampling is independent and that at each stage the probability of obtaining a defective part when the process is working correctly is .01. Let X denote the number of samples taken to obtain the first defective part.

(*a*) Find the density for X.

(*b* What is the average value of X?

(*c*) What is the equation for the cumulative distribution function for X? Use F to find the probability that the first defective part will be found on or before the 90th sample.

CHAPTER 4

CONTINUOUS DISTRIBUTIONS

In Chap. 3 we learned to distinguish a discrete random variable from one that is not discrete. In this chapter we consider a large class of nondiscrete random variables. In particular, we consider random variables that are called *continuous*. We first study the general properties of variables of the continuous type and then present some important families of continuous random variables.

4.1 CONTINUOUS DENSITIES

In Chap. 3 we considered the random variable T, the time of the peak demand for electricity at a particular power plant. We agreed that this random variable is not discrete since, "a priori"—before the fact—we cannot limit the set of possible values for T to some finite or countably infinite collection of times. Time is measured continuously and T can conceivably assume any value in the time interval $[0, 24)$, where 0 denotes 12 midnight one day and 24 denotes 12 midnight the next day. Furthermore, if we ask *before* the day begins, What is the probability that the peak demand will occur exactly 12.013 278 650 931 271? The answer is 0. It is virtually impossible for the peak load to occur at this split second in time, not the slightest bit earlier or later. These two properties, possible values occurring as intervals and the a priori probability of assuming any specific value being 0, are the characteristics that identify a random variable as being continuous. This leads us to our next definition.

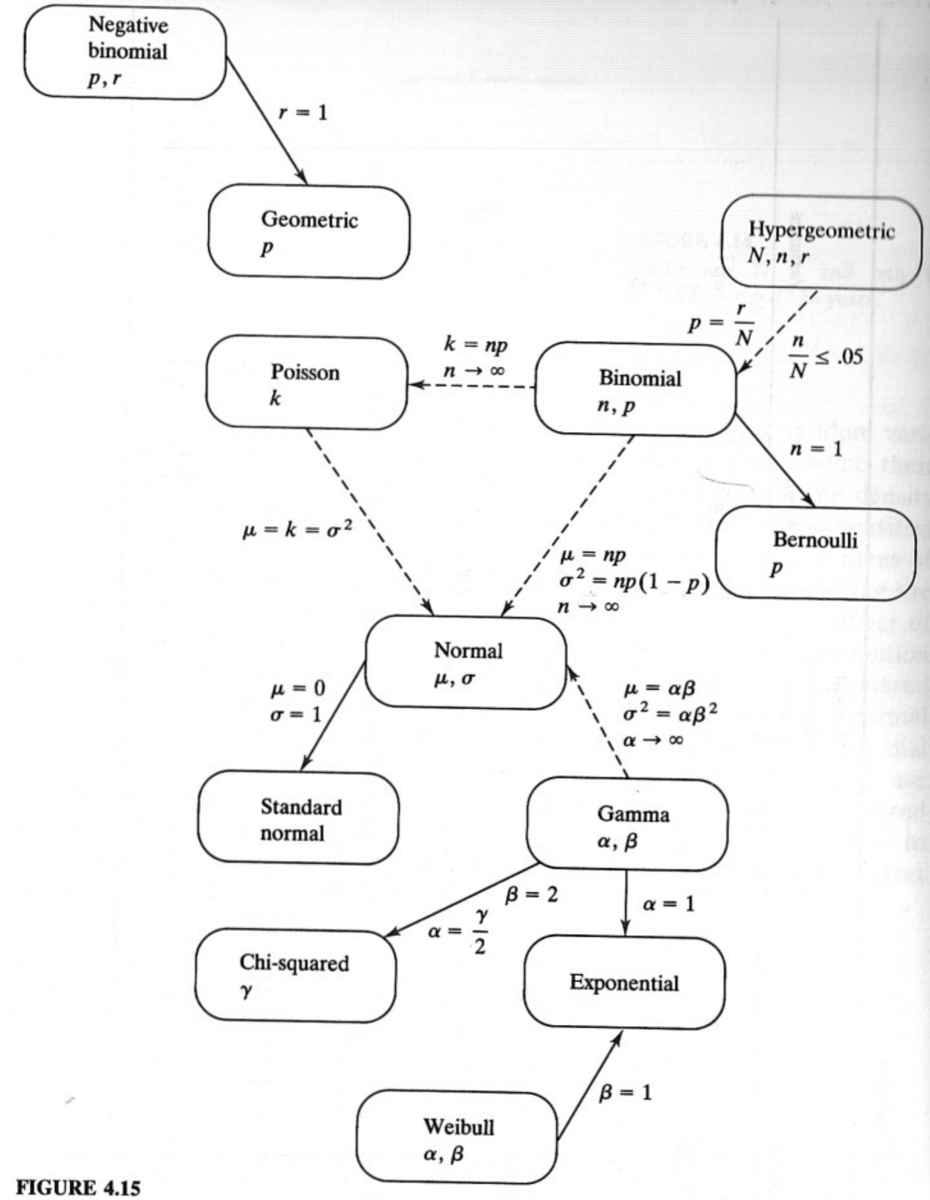


FIGURE 4.15
Some interrelationships among common distributions.

EXERCISES

Section 4.1

1. Consider the function

$$f(x) = kx \qquad 2 \leq x \leq 4$$

(*a*) Find the value of k that makes this a density for a continuous random variable.
(*b*) Find $P[2.5 \leq X \leq 3]$.
(*c*) Find $P[X = 2.5]$.
(*d*) Find $P[2.5 < X \leq 3]$.

2. Consider the areas shown in Fig. 4.16. In each case, state what probability is being depicted. What is the relationship between the areas depicted in Figs. 4.16(*a*) and (*b*)? Between those in Figs. 4.16(*d*) and (*e*)

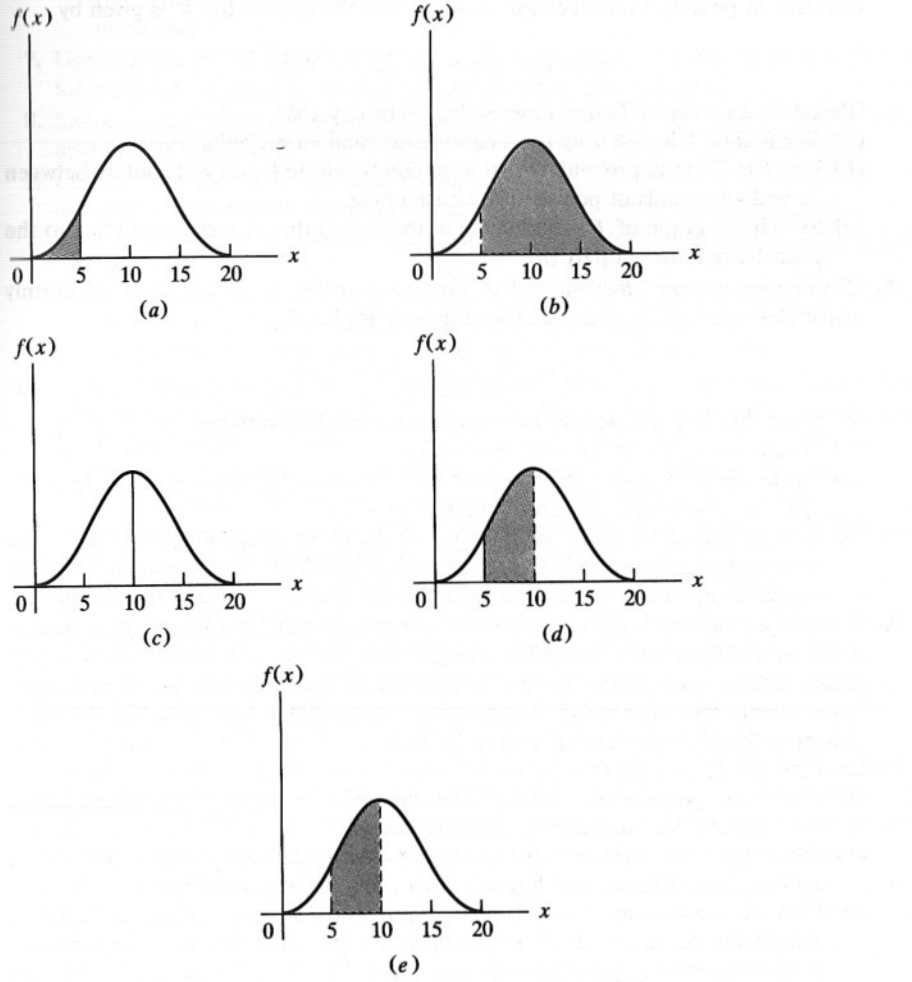


FIGURE 4.16

3. Let X denote the length in minutes of a long-distance telephone conversation. Assume that the density for X is given by

$$f(x) = (1/10)e^{-x/10} \qquad x > 0$$

(a) Verify that f is a density for a continuous random variable.
(b) Assuming that f adequately describes the behavior of the random variable X, find the probability that a randomly selected call will last at most 7 minutes; at least 7 minutes; exactly 7 minutes.
(c) Would it be unusual for a call to last between 1 and 2 minutes? Explain, based on the probability of this occurring.
(d) Sketch the graph of f and indicate in the sketch the area corresponding to each of the probabilities found in part (b).

4. Some plastics in scrapped cars can be stripped out and broken down to recover the chemical components. The greatest success has been in processing the flexible polyurethane cushioning found in these cars. Let X denote the amount of this material, in pounds, found per car. Assume that the density for X is given by

$$f(x) = \frac{1}{\ln 2}\frac{1}{x} \qquad 25 \le x \le 50$$

(Based on a report in *Design Engineering*, February 1982, p. 7.)
(a) Verify that f is a density for a continuous random variable.
(b) Use f to find the probability that a randomly selected auto will contain between 30 and 40 pounds of polyurethane cushioning.
(c) Sketch the graph of f and indicate in the sketch the area corresponding to the probability found in part (b).

5. (*Continuous uniform distribution.*) A random variable X is said to be uniformly distributed over an interval (a, b) if its density is given by

$$f(x) = \frac{1}{b-a} \qquad a < x < b$$

(a) Show that this is a density for a continuous random variable.
(b) Sketch the graph of the uniform density.
(c) Shade the area in the graph of part (b) that represents $P[X \le (a+b)/2]$.
(d) Find the probability pictured in part (c).
(e) Let (c, d) and (e, f) be subintervals of (a, b) of equal length. What is the relationship between $P[c \le X \le d]$ and $P[e \le X \le f]$? Generalize the idea suggested by this example, thus justifying the name "uniform" distribution.

6. When a pair of coils is placed around a homing pigeon and a magnetic field applied that reverses the earth's field, it is thought that the bird will become disoriented. Under these circumstances it is just as likely to fly in one direction as in any other. Let θ denote the direction in radians of the bird's initial flight. See Fig. 4.17. θ is uniformly distributed over the interval $[0, 2\pi]$.
(a) Find the density for θ.
(b) Sketch the graph of the density. The uniform distribution is sometimes called the "rectangular" distribution. Do you see why?
(c) Shade the area corresponding to the probability that a bird will orient within $\pi/4$ radians of home, and find this area using plane geometry.
(d) Find the probability that a bird will orient within $\pi/4$ radians of home by integrating the density over the appropriate region(s), and compare your answer to that obtained in part (c).

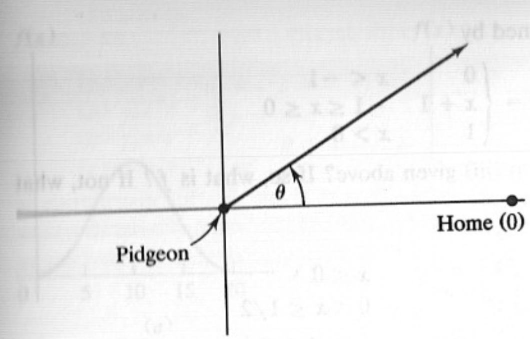


FIGURE 4.17
θ = direction of the initial flight of a homing pigeon measured in radians.

(e) If 10 birds are released independently and at least seven orient within $\pi/4$ radians of home, would you suspect that perhaps the coils are not disorienting the birds to the extent expected? Explain, based on the probability of this occurring.

7. Use Definition 4.1.2 to show that for a continuous random variable X, $P[X = a] = 0$ for every real number a. *Hint:* Write $P[X = a]$ as $P[a \le X \le a]$.

8. Express each of the probabilities depicted in Fig. 4.16 in terms of the cumulative distribution function F.

9. Consider the random variable of Exercise 1.
(a) Find the cumulative distribution function F.
(b) Use F to find $P[2.5 \le X \le 3]$, and compare your answer to that obtained previously.
*(c) Sketch the graph of F. Is F right continuous? Is F continuous? What is $\lim_{x \to -\infty} F(x)$? What is $\lim_{x \to \infty} F(x)$? Is F nondecreasing?
*(d) Find $dF(x)/dx$ for $x \in (2, 4)$. Does your answer look familiar?

10. (*Uniform distribution.*) Find the general expression for the cumulative distribution function for a random variable X that is uniformly distributed over the interval (a, b). See Exercise 5.

11. (*Uniform distribution.*) Consider the random variable of Exercise 6.
(a) Use Exercise 10 to find the cumulative distribution function F.
*(b) Sketch the graph of F. Is F right continuous? Is F continuous? What is $\lim_{x \to -\infty} F(x)$? What is $\lim_{x \to \infty} F(x)$? Is F nondecreasing?
*(c) Find $dF(x)/dx$ for $x \in (0, 2\pi)$. Does your answer look familiar?

12. Find the cumulative distribution function for the random variable of Exercise 3. Use F to find $P[1 \le X \le 2]$, and compare your answer to that obtained previously.

13. Find the cumulative distribution function for the random variable of Exercise 4. Use F to find $P[30 \le X \le 40]$, and compare your answer to that obtained previously.

*14. In Exercise 13 of Chap. 3 the mathematical properties of the cumulative distribution function for discrete random variables were pointed out. In the continuous case similar properties hold. The results of Exercises 9 and 11 are not coincidental! It can be shown that the cumulative distribution function F for any continuous random variable has these characteristics:
(i) F is continuous.
(ii) $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.
(iii) F is nondecreasing.
(iv) $dF(x)/dx = f(x)$ for all values of x for which this derivative exists.

(*a*) Consider the function F defined by

$$F(x) = \begin{cases} 0 & x < -1 \\ x+1 & -1 \le x \le 0 \\ 1 & x > 0 \end{cases}$$

Does F satisfy properties (i) to (iii) given above? If so, what is f? If not, what property fails?

(*b*) Consider the function defined by

$$F(x) = \begin{cases} 0 & x \le 0 \\ x^2 & 0 < x \le 1/2 \\ (1/2)x & 1/2 < x \le 1 \\ 1 & x > 1 \end{cases}$$

Does F satisfy properties (i) to (iii) given above? If so, what is f? If not, what property fails?

Section 4.2

15. Consider the random variable X with density

$$f(x) = (1/6)x \qquad 2 \le x \le 4$$

(*a*) Find $E[X]$.
(*b*) Find $E[X^2]$.
(*c*) Find σ^2 and σ.

16. Let X denote the amount in pounds of polyurethane cushioning found in a car. (See Exercise 4.) The density for X is given by

$$f(x) = \frac{1}{\ln 2}\frac{1}{x} \qquad 25 \le x \le 50$$

Find the mean, variance, and standard deviation for X.

17. Let X denote the length in minutes of a long-distance telephone conversation. The density for X is given by

$$f(x) = (1/10)e^{-x/10} \qquad x > 0$$

(*a*) Find the moment generating function, $m_X(t)$.
(*b*) Use $m_X(t)$ to find the average length of such a call.
(*c*) Find the variance and standard deviation for X.

18. (*Uniform distribution.*) The density for a random variable X distributed uniformly over (a, b) is

$$f(x) = \frac{1}{b-a} \qquad a < x < b$$

Use Definition 4.2.1 to show that

$$E[X] = \frac{a+b}{2} \qquad \text{and} \qquad \text{Var } X = \frac{(b-a)^2}{12}$$

19. (*Uniform distribution.*) Let θ denote the direction in radians of the flight of a bird whose sense of direction has been disoriented as described in Exercise 6. Assume that θ is uniformly distributed over the interval $[0, 2\pi]$. Use the results of Exercise 18 to find the mean, variance, and standard deviation of θ.

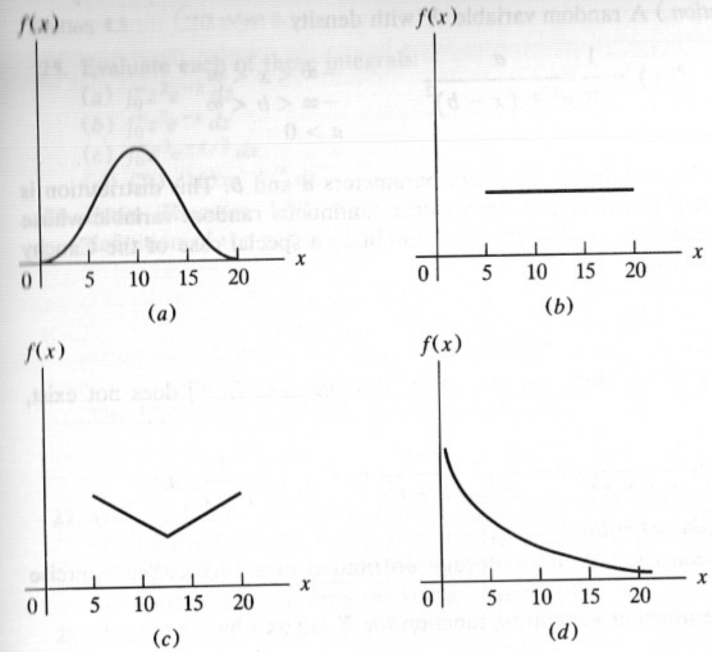


FIGURE 4.18

20. Let X be continuous with density f. Imagine cutting out of a piece of thin rigid metal the region bounded by the graph of f and the x axis, and attempting to balance this region on a knife-edge held parallel to the vertical axis. The point at which the region would balance, if such a point exists, is the mean of X. Thus μ_X is a "location" parameter in that it indicates the position of the center of the density along the x axis. Figure 4.18 gives the graphs of the densities of four continuous random variables whose means do exist. In each case, approximate the value of μ_X from the graph.

21. In the continuous case variance is a "shape" parameter in the sense that a random variable with small variance will have a compact density; one with a large variance will have a density that is rather spread out or flat. Consider the two densities given in Fig. 4.19. What is μ_X? What is μ_Y? Which random variable has the larger variance?

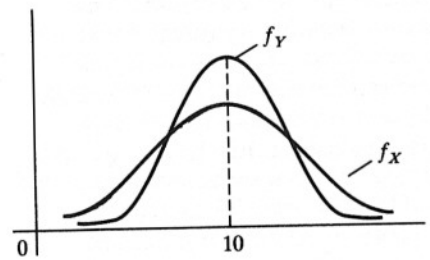


FIGURE 4.19

★22. (*Cauchy distribution.*) A random variable X with density

$$f(x) = \frac{1}{\pi}\frac{a}{a^2 + (x-b)^2} \qquad \begin{array}{l} -\infty < x < \infty \\ -\infty < b < \infty \\ a > 0 \end{array}$$

is said to have a Cauchy distribution with parameters a and b. This distribution is interesting in that it provides an example of a continuous random variable whose mean does not exist. Let $a = 1$ and $b = 0$ to obtain a special case of the Cauchy distribution with density

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2} \qquad -\infty < x < \infty$$

Show that $\int_{-\infty}^{\infty}|x| f(x)\,dx$ does not exist, thus showing that $E[X]$ does not exist. *Hint:* Write

$$\int_{-\infty}^{\infty} |x| \frac{1}{\pi}\frac{1}{1+x^2}\,dx = \frac{1}{\pi}\int_{-\infty}^{0}\frac{-x}{1+x^2}\,dx + \frac{1}{\pi}\int_{0}^{\infty}\frac{x}{1+x^2}\,dx$$

and recall that $\int (du/u) = \ln|u|$.

★23. (*Uniform distribution.*) Let X be uniformly distributed over (a, b). (See Exercise 18.)

(*a*) Show that the moment generating function for X is given by

$$m_X(t) = \begin{cases} \dfrac{e^{tb} - e^{ta}}{t(b-a)} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

Hint: When $t = 0$, $m_X(t) = E[e^{0 \cdot X}]$.

(*b*) Use $m_X(t)$ to find $E[X]$. *Hint:* Find

$$\frac{d}{dt}\left[\frac{e^{tb} - e^{ta}}{t(b-a)}\right]$$

and take the limit of this derivative as $t \to 0$ using L'Hospital's rule.

★24. Let the density for X be given by

$$f(x) = ce^{-|x|} \qquad -\infty < x < \infty$$

(*a*) Find the value of c that makes this a density.

(*b*) Show that

$$\frac{1}{2}\int_{-\infty}^{\infty} |x| e^{-|x|}\,dx$$

exists.

(*c*) Find $E[X]$.

(*d*) Show that

$$m_X(t) = \frac{-1}{t^2 - 1} \qquad -1 < t < 1$$

(*e*) Use $m_X(t)$ to find $E[X]$ and $E[X^2]$.

(*f*) Find Var X.

Section 4.3

25. Evaluate each of these integrals:

(*a*) $\int_0^\infty z^2 e^{-z}\,dz$

(*b*) $\int_0^\infty z^7 e^{-z}\,dz$

(*c*) $\int_0^\infty x^3 e^{-x/2}\,dx$

(*d*) $\int_0^\infty (1/16) x e^{-x/4}\,dx$

26. Prove Theorem 4.3.1. *Hint:* To prove part 1, evaluate $\Gamma(1)$ directly from the definition of the gamma function. To prove part 2, use integration by parts with

$$u = z^{\alpha-1} \qquad dv = \int e^{-z}\,dz$$

$$du = (\alpha - 1)z^{\alpha-2}\,dz \qquad v = -e^{-z}$$

Use L'Hospital's rule repeatedly to show that

$$-z^{\alpha-1}e^{-z}\Big|_0^\infty = 0$$

27. (*a*) Use Theorem 4.3.1 to evaluate $\Gamma(2)$, $\Gamma(3)$, $\Gamma(4)$, $\Gamma(5)$, and $\Gamma(6)$.

(*b*) Can you generalize the pattern suggested in part (*a*)?

(*c*) Does the result of part (*b*) hold even if $n = 1$?

(*d*) Evaluate $\Gamma(15)$ using the result of part (*b*).

28. Show that for $\alpha > 0$ and $\beta > 0$,

$$\int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}\,dx = 1$$

thereby showing that the function given in Definition 4.3.2 is a density for a continuous random variable. *Hint:* Change the variable by letting $z = x/\beta$.

29. Let X be a gamma random variable with $\alpha = 3$ and $\beta = 4$.

(*a*) What is the expression for the density for X?

(*b*) What is the moment generating function for X?

(*c*) Find μ, σ^2, and σ.

30. Let X be a gamma random variable with parameters α and β. Use the moment generating function to find $E[X]$ and $E[X^2]$. Use these expectations to show that Var $X = \alpha\beta^2$.

31. Let X be a gamma random variable with parameters α and β.

(*a*) Use Definition 4.2.1, the definition of expected value, to find $E[X]$ and $E[X^2]$ directly. *Hint:* $z^\alpha = z^{(\alpha+1)-1}$ and $z^{\alpha+1} = z^{(\alpha+2)-1}$.

(*b*) Use the results of (*a*) to verify that Var $X = \alpha\beta^2$.

32. Show that the graph of the density for a gamma random variable with parameters α and β assumes its maximum value at $x = \beta(\alpha - 1)$ for $\alpha > 1$. Sketch a rough graph of the density for a gamma random variable with $\alpha = 3$ and $\beta = 4$. *Hint:* Find the first derivative of the density, set this derivative equal to 0, and solve for x.

33. Let X be an exponential random variable with parameter β. Find general expressions for the moment generating function, mean, and variance for X.

34. A particular nuclear plant releases a detectable amount of radioactive gases twice a month on the average. Find the probability that at least 3 months will elapse before the release of the first detectable emission. What is the average time that one must wait to observe the first emission?

35. The "forgetfulness" property of the exponential distribution says that the probability that we must wait a total of $w_1 + w_2$ units before the occurrence of an event given that we have already waited w_1 units is the same as the probability that we must wait w_2 units at the outset. That is,

$$P[W > w_1 + w_2 | W > w_1] = P[W > w_2]$$

Verify this statement for any exponential random variable W with parameter β.

36. Rock noise in an underground mine occurs at an average rate of three per hour. (See Exercise 65, Chap. 3.)
(*a*) Find the probability that no rock noise will be recorded for at least 30 minutes.
(*b*) Suppose that no rock noise has been heard for 15 minutes. Find the probability that another 30 minutes will elapse before the first rock noise is detected.

37. California is hit every year by approximately 500 earthquakes that are large enough to be felt. However, those of destructive magnitude occur, on the average, once a year.
(*a*) Find the probability that at least 3 months elapse before the first earthquake of destructive magnitude occurs. (See Exercise 64, Chap. 3.)
(*b*) Suppose that no destructive quake has occurred for 4 months. Find the probability that an additional 3 months will elapse before a destructive quake occurs.

38. Consider a chi-squared random variable with 15 degrees of freedom.
(*a*) What is the mean of X^2_{15}? What is its variance?
(*b*) What is the expression for the density for X^2_{15}?
(*c*) What is the expression for the moment generating function for X^2_{15}?
(*d*) Use Table IV of App. A to find each of the following:

$$P[X^2_{15} \le 5.23] \qquad P[6.26 \le X^2_{15} \le 27.5] \qquad \chi^2_{.05}$$

$$P[X^2_{15} \ge 22.3] \qquad \chi^2_{.01} \qquad \chi^2_{.95}$$

Section 4.4

39. Use Table V of App. A to find each of the following:
(*a*) $P[Z \le 1.57]$. (*b*) $P[Z < 1.57]$.
(*c*) $P[Z = 1.57]$. (*d*) $P[Z > 1.57]$.
(*e*) $P[-1.25 \le Z \le 1.75]$. (*f*) $z_{.10}$.
(*g*) $z_{.90}$.
(*h*) The point z such that $P[-z \le Z \le z] = .95$.
(*i*) The point z such that $P[-z \le Z \le z] = .90$.

40. The bulk density of soil is defined as the mass of dry solids per unit bulk volume. A high bulk density implies a compact soil with few pores. Bulk density is an important factor in influencing root development, seedling emergence, and aeration. Let X denote the bulk density of Pima clay loam. Studies show that X is normally distributed with $\mu = 1.5$ and $\sigma = .2$ g/cm^3. (*McGraw-Hill Yearbook of Science and Technology*, 1981, p. 361.)
(*a*) What is the density for X? Sketch a graph of the density function. Indicate on this graph the probability that X lies between 1.1 and 1.9. Find this probability.
(*b*) Find the probability that a randomly selected sample of Pima clay loam will have bulk density less than .9 g/cm^3.

(*c*) Would you be surprised if a randomly selected sample of this type of soil has a bulk density in excess of 2.0 g/cm^3? Explain, based on the probability of this occurring.
(*d*) What point has the property that only 10% of the soil samples have bulk density this high or higher?
(*e*) What is the moment generating function for X?

41. Most galaxies take the form of a flattened disc with the major part of the light coming from this very thin fundamental plane. The degree of flattening differs from galaxy to galaxy. In the Milky Way Galaxy most gases are concentrated near the center of the fundamental plane. Let X denote the perpendicular distance from this center to a gaseous mass. X is normally distributed with mean 0 and standard deviation 100 parsecs. (A parsec is equal to approximately 19.2 trillion miles.) (*McGraw-Hill Encyclopedia of Science and Technology*, vol. 6, 1971, p. 10.)
(*a*) Sketch a graph of the density for X. Indicate on this graph the probability that a gaseous mass is located within 200 parsecs of the center of the fundamental plane. Find this probability.
(*b*) Approximately what percentage of the gaseous masses are located more than 250 parsecs from the center of the plane?
(*c*) What distance has the property that 20% of the gaseous masses are at least this far from the fundamental plane?
(*d*) What is the moment generating function for X?

42. Among diabetics, the fasting blood glucose level X may be assumed to be approximately normally distributed with mean 106 milligrams per 100 milliliters and standard deviation 8 milligrams per 100 milliliters.
(*a*) Sketch a graph of the density for X. Indicate on this graph the probability that a randomly selected diabetic will have a blood glucose level between 90 and 122 mg/100 ml. Find this probability.
(*b*) Find $P[X \le 120 \text{ mg}/100 \text{ ml}]$.
(*c*) Find the point that has the property that 25% of all diabetics have a fasting glucose level of this value or lower.
(*d*) If a randomly selected diabetic is found to have fasting blood glucose level in excess of 130, do you think there is cause for concern? Explain, based on the probability of this occurring naturally.

43. (*a*) Find the density for the standard normal random variable Z.
(*b*) Find $f'(z)$. Show that the only critical point for f occurs at $z = 0$. Use the first derivative test to show that f assumes its maximum value at $z = 0$.
(*c*) Find $f''(z)$. Show that the possible inflection points occur at $z = \pm 1$. Use the second derivative to show that f changes concavity at $z = \pm 1$, implying that the inflection points do occur when $z = \pm 1$.
(*d*) Let X be normal with parameters μ and σ. Let $(X - \mu)/\sigma = Z$. Use the results of parts (*b*) and (*c*) to verify that, in general, a normal curve assumes its maximum value at $x = \mu$ and has points of inflection at $x = \mu \pm \sigma$.

44. Let X be normal with parameters μ and σ. Use the moment generating function to find $E[X^2]$. Find Var X, thus completing the proof of Theorem 4.4.2.

*45. (*Log-normal distribution.*) The log-normal distribution is the distribution of a random variable whose natural logarithm follows a normal distribution. Thus if X is a normal random variable then $Y = e^X$ follows a log-normal distribution. Complete the argument below, thus deriving the density for a log-normal random variable.

Let X be normal with mean μ and variance σ^2. Let G denote the cumulative distribution function for $Y = e^X$, and let F denote the cumulative distribution function for X.

(*a*) Show that $G(y) = F(\ln y)$.

(*b*) Show that $G'(y) = F'(\ln y)/y$.

(*c*) Use Exercise 14 part (iv) to show that the density for Y is given by

$$g(y) = \frac{1}{\sqrt{2\pi}\sigma y} \exp\left[-\frac{1}{2}\frac{(\ln y - \mu)^2}{\sigma^2}\right] \qquad \begin{array}{l} -\infty < \mu < \infty \\ \sigma > 0 \\ y > 0 \end{array}$$

Note that μ and σ are the mean and standard deviation of the underlying normal distribution; they are not the mean and standard deviation of Y itself.

***46.** Let Y denote the diameter in millimeters of Styrofoam pellets used in packing. Assume that Y has a log-normal distribution with parameters $\mu = .8$ and $\sigma = .1$.

(*a*) Find the probability that a randomly selected pellet has a diameter that exceeds 2.7 millimeters.

(*b*) Between what two values will Y fall with probability approximately .95?

Section 4.5

47. Verify the normal probability rule.

48. The number of Btu's of petroleum and petroleum products used per person in the United States in 1975 was normally distributed with mean 153 million Btu and standard deviation 25 million Btu. Approximately what percentage of the population used between 128 and 178 million Btu during that year? Approximately what percentage of the population used in excess of 228 million Btu?

49. Reconsider Exercises 40(*a*), 41(*a*), and 42(*a*) in light of the normal probability rule.

50. For a normal random variable, $P[|X - \mu| < 3\sigma] \doteq .99$. What value is assigned to this probability via Chebyshev's inequality? Are the results consistent? Which rule gives a stronger statement in the case of a normal variable?

51. Animals have an excellent spatial memory. In an experiment to confirm this statement an eight-armed maze such as that shown in Fig. 4.20 is used. At the beginning of a test, one pellet of food is placed at the end of each arm. A hungry animal is placed at the center of the maze and is allowed to choose freely from among the arms. The optimal strategy is to run to the end of each arm exactly once. This requires that the animal remember where it has been. Let X denote the number of correct arms (arms still containing food) selected among its first eight choices. Studies indicate that $\mu = 7.9$. (*McGraw-Hill Encyclopedia of Science and Technology*, 1980.)

(*a*) Is X normally distributed?

(*b*) State and interpret Chebyshev's inequality in the context of this problem for $k = .5, 1, 2$, and 3. At what point does the inequality begin to give us some practical information?

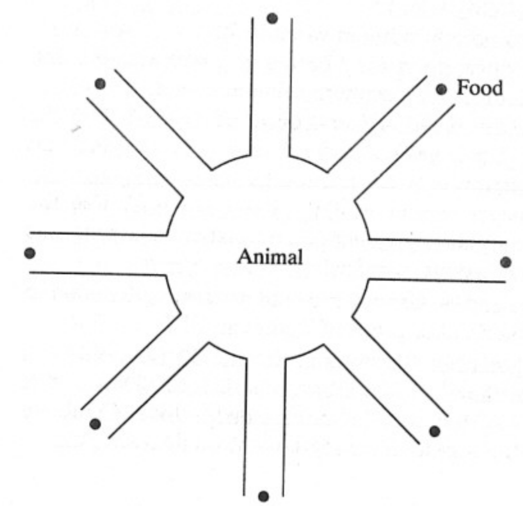


FIGURE 4.20
An eight-armed maze.

Section 4.6

52. Let X be binomial with $n = 20$ and $p = .3$. Use the normal approximation to approximate each of the following. Compare your results with the values obtained from Table I of App. A.

(*a*) $P[X \leq 3]$.

(*b*) $P[3 \leq X \leq 6]$.

(*c*) $P[X \geq 4]$.

(*d*) $P[X = 4]$.

53. Although errors are likely when taking measurements from photographic images, these errors are often very small. For sharp images with negligible distortion, errors in measuring distances are often no larger than .0004 inch. Assume that the probability of a serious measurement error is .05. A series of 150 independent measurements are made. Let X denote the number of serious errors made.

(*a*) In finding the probability of making at least one serious error, is the normal approximation appropriate? If so, approximate the probability using this method.

(*b*) Approximate the probability that at most three serious errors will be made.

54. A chemical reaction is run in which the usual yield is 70%. A new process has been devised that should improve the yield. Proponents of the new process claim that it produces better yields than the old process more than 90% of the time. The new process is tested 60 times. Let X denote the number of trials in which the yield exceeds 70%.

(*a*) If the probability of an increased yield is .9, is the normal approximation appropriate?

(*b*) If $p = .9$, what is $E[X]$?

(*c*) If $p > .9$ as claimed, then, on the average, more than 54 of every 60 trials will result in an increased yield. Let us agree to accept the claim if X is at least 59. What is the probability that we will accept the claim if p is really only .9?

(*d*) What is the probability that we shall not accept the claim ($X \leq 58$) if it is true, and p is really .95?

55. Opponents of a nuclear power project claim that the majority of those living near a proposed site are opposed to the project. To justify this statement, a random sample of 75 residents is selected and their opinions are sought. Let X denote the number opposed to the project.

(*a*) If the probability that an individual is opposed to the project is .5, is the normal approximation appropriate?

(*b*) If $p = .5$, what is $E[X]$?

(*c*) If $p > .5$ as claimed, then, on the average, more than 37.5 of every 75 individuals are opposed to the project. Let us agree to accept the claim if X is

at least 46. What is the probability that we shall accept the claim if p is really only .5?
(*d*) What is the probability that we shall not accept the claim ($X \le 45$) even though it is true and p is really .7?

56. (*Normal approximation to the Poisson distribution.*) Let X be Poisson with parameter λs. Then for large values of λs, X is approximately normal with mean λs and variance λs. (The proof of this theorem is also based on the Central Limit Theorem and will be considered in Chap. 7.) Let X be a Poisson random variable with parameter $\lambda s = 15$. Find $P[X \le 12]$ from Table II of App. A. Approximate this probability using a normal curve. Be sure to employ the half-unit correction factor.

57. The average number of jets either arriving at or departing from O'Hare Airport is one every 40 seconds. What is the approximate probability that at least 75 such flights will occur during a randomly selected hour? What is the probability that fewer than 100 such flights will take place in an hour?

Section 4.7 Weibull

58. The length of time in hours that a rechargeable calculator battery will hold its charge is a random variable. Assume that this variable has a Weibull distribution with $\alpha = .01$ and $\beta = 2$.
(*a*) What is the density for X?
(*b*) What are the mean and variance for X? *Hint:* It can be shown that $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for any $\alpha > 1$. Furthermore, $\Gamma(1/2) = \sqrt{\pi}$.
(*c*) What is the reliability function for this random variable?
(*d*) What is the reliability of such a battery at $t = 3$ hours? At $t = 12$ hours? At $t = 20$ hours?
(*e*) What is the hazard rate function for these batteries?
(*f*) What is the failure rate at $t = 3$ hours? At $t = 12$ hours? At $t = 20$ hours?
(*g*) Is the hazard rate function an increasing or a decreasing function? Does this seem to be reasonable from a practical point of view? Explain.

59. Computer chips do not "wear out" in the ordinary sense. Assuming that defective chips have been removed from the market by factory inspection, it is reasonable to assume that these chips exhibit a constant hazard rate. Let the hazard rate be given by $\rho(t) = .02$. (Time is in years.)
(*a*) In a practical sense, what are the main causes of failure of these chips?
(*b*) What is the reliability function for chips of this type?
(*c*) What is the reliability of a chip 20 years after it has been put into use?
(*d*) What is the failure density for these chips?
(*e*) What type of random variable is X, the time to failure of a chip?
(*f*) What is the mean and variance for X?
(*g*) What is the probability that a chip will be operable for at least 30 years?

60. The random variable X, the time to failure (in thousands of miles driven) of the signal lights on an automobile has a Weibull distribution with $\alpha = .04$ and $\beta = 2$.
(*a*) Find the density, mean, and variance for X.
(*b*) Find the reliability function for X.
(*c*) What is the reliability of these lights at 5000 miles? At 10,000 miles?
(*d*) What is the hazard rate function?
(*e*) What is the hazard rate at 5000 miles? At 10,000 miles?
(*f*) What is the probability that the lights will fail during the first 3000 miles driven?

61. Show that for $\alpha > 0$ and $\beta > 0$,

$$\int_0^\infty \alpha\beta x^{\beta-1}e^{-\alpha x^\beta}\,dx = 1$$

thereby showing that the nonnegative function given in Definition 4.7.1 is a density for a continuous random variable. *Hint:* Let $z = \alpha x^\beta$.

62. Let X be a Weibull random variable with parameters α and β. Show that $E[X^2] = \alpha^{-2/\beta}\Gamma(1 + 2/\beta)$. *Hint:* In evaluating

$$\int_0^\infty x^2\alpha\beta x^{\beta-1}e^{-\alpha x^\beta}\,dx$$

let $z = \alpha x^\beta$. Evaluate the integral in a manner similar to that used in the proof of Theorem 4.7.1.

63. Use the result of Exercise 62 to find $\operatorname{Var} X$ for a Weibull random variable with parameters α and β, thus completing the proof of Theorem 4.7.1.

64. Consider the hazard rate function

$$\rho(t) = \alpha\beta t^{\beta-1} \qquad t > 0,\ \alpha > 0,\ \beta > 0$$

(*a*) Show that $\rho(t)$ is constant if $\beta = 1$.
(*b*) Find $\rho'(t)$. Argue that $\rho'(t) > 0$ if $\beta > 1$, thus producing an increasing hazard rate. Argue that $\rho'(t) < 0$ if $\beta < 1$, thus producing a decreasing hazard rate.

65. A system has eight components connected as shown in Fig. 4.21.
(*a*) Find the reliability of each of the parallel assemblies.
(*b*) Find the system reliability.
(*c*) Suppose that assembly II is replaced by two identical components in parallel, each with reliability .98. What is the reliability of the new assembly?
(*d*) What is the new system reliability after making the change suggested in part (*c*)?
(*e*) Make changes analogous to that of part (*c*) in each of the remaining single component assemblies. Compute the new system reliability.

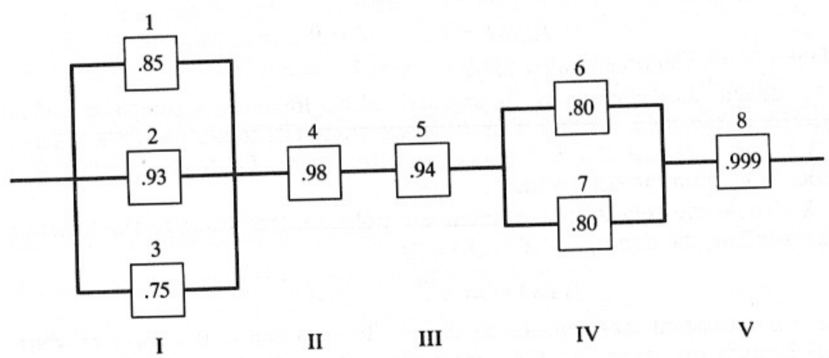


FIGURE 4.21

66. A system consists of two independent components connected in series. The life span of the first component follows a Weibull distribution with $\alpha = .006$ and $\beta = .5$; the second has a life span that follows the exponential distribution with $\beta = .00004$.
(*a*) Find the reliability of the system at 2500 hours.
(*b*) Find the probability that the system will fail before 2000 hours.
(*c*) If the two components are connected in parallel, what is the system reliability at 2500 hours?

67. Suppose that a missile can have several independent and identical computers, each with reliability .9 connected in parallel so that the system will continue to function as long as at least one computer is operating. If it is desired to have a system reliability of at least .999, how many computers should be connected in parallel?

68. Three independent and identical components, each with a reliability of .9, are to be used in an assembly.
(*a*) The assembly will function if at least one of the components is operable. Find the system reliability.
(*b*) The assembly will function if at least two of the components are operable. Find the reliability of the system.
(*c*) The assembly will function only if all three of the components are operable. Find the reliability of the system.

Section 4.8

69. Prove Theorem 4.8.1 in the case in which g is strictly increasing.

70. Let X be a random variable with density

$$f_X(x) = (1/4)x \qquad 0 \le x \le \sqrt{8}$$

and let $Y = X + 3$.
(*a*) Find $E[X]$, and then use the rules for expectation to find $E[Y]$.
(*b*) Find the density for Y.
(*c*) Use the density for Y to find $E[Y]$, and compare your answer to that found in part (*a*).

71. Let X be a random variable with density

$$f_X(x) = (1/4)xe^{-x/2} \qquad x \ge 0$$

and let $Y = (-1/2)X + 2$. Find the density for Y.

72. Let X be a random variable with density

$$f_X(x) = e^{-x} \qquad x > 0$$

and let $Y = e^X$. Find the density for Y.

73. Let C denote the temperature in degrees Celsius to which a computer will be subjected in the field. Assume that C is uniformly distributed over the interval $(15, 21)$. Let F denote the field temperature in degrees Fahrenheit so that $F = (9/5)C + 32$. Find the density for F.

74. Let X denote the velocity of a random gas molecule. According to the Maxwell-Boltzmann law, the density for X is given by

$$f_X(x) = cx^2e^{-\beta x^2} \qquad x > 0$$

Here c is a constant that depends on the gas involved and β is a constant whose value depends on the mass of the molecule and its absolute temperature. The kinetic energy of the molecule, Y, is given by $Y = (1/2)mX^2$ where $m > 0$. Find the density for Y.

75. Let X be a continuous random variable with density f_X, and let $Y = X^2$.
(*a*) Show that for $y \ge 0$,

$$F_Y(y) = P\left[-\sqrt{y} \le X \le \sqrt{y}\right]$$

(*b*) Show that for $y \ge 0$,

$$F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

(*c*) Use the technique given in the proof of Theorem 4.8.1 to show that

$$f_Y(y) = 1/(2\sqrt{y})\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right]$$

(*d*) Use the technique illustrated in Example 4.8.2 to show that

$$f_Y(y) = 1/(2\sqrt{y})\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right]$$

76. Let Z be a standard normal random variable and let $Y = Z^2$.
(*a*) Show that $\Gamma(1/2) = \int_0^\infty x^{-1/2}e^{-x}\,dx$.
(*b*) Show that $\Gamma(1/2) = \sqrt{\pi}$. *Hint:* Use the results of part (*a*) with $x = t^2/2$ and make use of the fact that the standard normal density integrates to 1 when integrated over the set of real numbers.
(*c*) Use the results of Exercise 75 to find f_Y.
(*d*) Argue that Y follows a chi-squared distribution with 1 degree of freedom.

77. Let X be normally distributed with mean μ and variance σ^2. Let $Y = e^X$. Show that Y follows the log-normal distribution. (See Exercise 45.)

78. Let Z be a standard normal random variable and let $Y = 2Z^2 - 1$. Find the density for Y.

Section 4.9

79. Use Table III of App. A to generate nine more observations on the random variable X, the time to failure of a computer chip. (See Example 4.9.1.) Based on these data, approximate the average time to failure by finding the arithmetic average of the values of X simulated in the experiment. Does this value agree well with the theoretical mean value of 50 years?

80. Simulate 20 observations on the random variable X, the time to failure of the signal lights on an automobile. (See Exercise 60.) Approximate the average time to failure for these lights based on the simulated data. Does this value agree well with the theoretical mean value for X?

81. A satellite has malfunctioned and is expected to reenter the earth's atmosphere sometime during a 4-hour period. Let X denote the time of reentry. Assume that X is uniformly distributed over the interval $[0, 4]$. Simulate 20 observations on X. (See Exercise 18.)

REVIEW EXERCISES

82. Let X be a continuous random variable with density

$$f(x) = cx^2 \qquad -3 \le x \le 3$$

(*a*) Assuming that $f(x) = 0$ elsewhere, find the value of c that makes this a density.
(*b*) Find $E[X]$ and $E[X^2]$ from the definitions of these terms.

(c) Find Var X and σ.
(d) Find $P[X \le 2]$; $P[-1 \le X \le 2]$; $P[X > 1]$ by direct integration.
(e) Find the closed-form expression for the cumulative distribution function F.
(f) Use F to find each of the probabilities of part (d), and compare your answers to those obtained earlier.

83. Find $\int_0^\infty z^{10}e^{-z}\,dz$.

84. A computer firm introduces a new home computer. Past experience shows that the random variable X, the time of peak demand measured in months after its introduction, follows a gamma distribution with variance 36.
(a) If the expected value of X is 18 months, find α and β.
(b) Find $P[X \le 7.01]$; $P[X \ge 26]$; $P[13.7 \le X \le 31.5]$.

85. Let X denote the lag time in a printing queue at a particular computer center. That is, X denotes the difference between the time that a program is placed in the queue and the time at which printing begins. Assume that X is normally distributed with mean 15 minutes and variance 25.
(a) Find the expression for the density for X.
(b) Find the probability that a program will reach the printer within 3 minutes of arriving in the queue.
(c) Would it be unusual for a program to stay in the queue between 10 and 20 minutes? Explain, based on the approximate probability of this occurring. You do not have to use the Z table to answer this question!
(d) Would you be surprised if it took longer than 30 minutes for the program to reach the printer? Explain, based on the probability of this occurring.

86. A computer center maintains a telephone consulting service to troubleshoot for its users. The service is available from 9 a.m. to 5 p.m. each working day. Past experience shows that the random variable X, the number of calls received per day, follows a Poisson distribution with $\lambda = 50$. For a given day, find the probability that the first call of the day will be received by 9:15 a.m.; after 3 p.m.; between 9:30 a.m. and 10 a.m.

87. Let $H(X) = X^2 + 3X + 2$. Find $E[H(X)]$ if
(a) X is normally distributed with mean 3 and variance 4.
(b) X has a gamma distribution with $\alpha = 2$ and $\beta = 4$.
(c) X has a chi-squared distribution with 10 degrees of freedom.
(d) X has an exponential distribution with $\beta = 5$.
(e) X has a Weibull distribution with $\alpha = 2$ and $\beta = 1$.

88. Let the density for the continuous random variable X be given by

$$f(x) = 1/2e^{-|x|} \qquad -\infty < x < \infty$$

(a) Show that $\int_{-\infty}^{\infty} f(x)\,dx = 1$.
(b) Show that

$$m_X(t) = (1/2)[1/(t+1) - 1/(t-1)] \qquad -1 < t < 1$$

(c) Use $m_X(t)$ to show that $E[X] = 0$.

89. Let X denote the time to failure in years of a telephone modem used to access a mainframe computer from a remote terminal. Assume that the hazard rate function for X is given by

$$\rho(t) = \alpha\beta t^{\beta-1}$$

where $\alpha = 2$ and $\beta = 1/5$.

(a) Find the failure density for X.
(b) Find the expected value of X.
(c) Find the reliability function for X.
(d) Find the probability that the modem will last for at least 2 years.
(e) What is the hazard rate at $t = 1$ year?
(f) Describe roughly the theoretical pattern in the causes of failure in these modems.

90. Past evidence shows that when a customer complains of an out-of-order phone there is an 8% chance that the problem is with the inside wiring. During a 1-month period, 100 complaints are lodged. Assume that there have been no wide-scale problems that could be expected to affect many phones at once, and that, for this reason, these failures are considered to be independent. Find the expected number of failures due to a problem with the inside wiring. Find the probability that at least 10 failures are due to a problem with the inside wiring. Would it be unusual if at most 5 were due to problems with the inside wiring? Explain, based on the probability of this occurring.

91. The cumulative distribution function for a continuous random variable X is defined by

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^3 + x^2}{2} & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

Find the density for X.

92. The density for a continuous random variable is given by

$$f(x) = xe^{-x} \qquad 0 < x < \infty$$

(a) Show that $\int_0^\infty xe^{-x}\,dx = 1$. *Hint:* Use the gamma function.
(b) Find $E[X]$, $E[X^2]$, and Var X.
(c) Show that $m_X(t) = 1/(1-t)^2$, where $t < 1$.
(d) Use $m_X(t)$ to find $E[X]$.

93. An electronic counter records the number of vehicles exiting the interstate at a particular point. Assume that the average number of vehicles leaving in a 5-minute period is 10. Approximate the probability that between 100 and 120 vehicles inclusive will exit at this point in a 1-hour period.

94. Consider the following moment generating functions. In each case, identify the distribution involved completely. Be sure to specify the numerical value of all parameters that identify the distribution. For example, if X is normal, give the numerical value of μ and σ^2; if gamma, state α and β.
(a) $e^{3t+16t^2/2}$
(b) $(1-3t)^{-7}$
(c) $(1-2t)^{-12}$
(d) $\dfrac{e^{3t} - e^t}{2t}$
(e) $e^{t^2/2}$
(f) $(1-7t)^{-1}$
(g) $e^{3t+t^2/2}$

95. For each random variable in Exercise 94, state the numerical value of the average for X and its variance.

$U = X - Y$ must lie between -3 and 2 and that $V = X + Y$ must lie between 0 and 5. Furthermore, U and V must satisfy the inequalities

$$0 < (v + u)/2 < 2$$
$$0 < (v - u)/2 < 3$$

or

$$0 < v + u < 4$$
$$0 < v - u < 6$$

Solving these inequalities simultaneously yields the region R shown in Fig. 5.6(b). Thus the density for (U, V) is given by

$$f_{UV}(u, v) = 1/12 \qquad (u, v) \in R$$

We leave it to you to verify that f_{UV} is, in fact, a valid density.

Other transformation theorems can be derived from Theorem 5.5.1. Some of these are given in Exercises 48, 50, and 51. For a more detailed discussion of this topic, please see [49].

CHAPTER SUMMARY

In this chapter we considered random variables of more than one dimension. Emphasis was on random variables of two dimensions. The joint density was defined by extending the notion of a density for a single variable in a logical way. This function was used to calculate probabilities associated with two-dimensional random variables (X, Y). We saw how to obtain the marginal densities for both X and Y from the joint density. These marginal densities are the usual densities for X or Y when considered alone. The correlation coefficient ρ was introduced as a measure of linearity between X and Y. The notion of independence between X and Y was defined formally and its relationship to ρ was investigated. We saw how to define the conditional densities for X given Y and Y given X from knowledge of the joint density for (X, Y) and the marginal densities for X and Y. The conditional densities were used to find the equations for the curves of regression of Y on X and X on Y. These regression curves are the graphs of the mean value of Y as a function of X or vice versa. We saw that these curves may be linear or nonlinear.

We introduced and defined important terms that you should know. These are:

Two-dimensional discrete random variable
Two-dimensional continuous random variable
Discrete joint density
Discrete marginal density
Independent random variables
Covariance
Perfect positive correlation
Uncorrelated
Curve of regression
n-dimensional discrete random variable
n-dimensional continuous random variable
Bivariate normal distribution
Continuous joint density
Continuous marginal density
Expected value of $H(X, Y)$
Correlation coefficient
Perfect negative correlation
Conditional density

EXERCISES

Section 5.1

1. Use Table 5.2 to find each of these probabilities:
   (a) The probability that exactly two defective welds and one improperly tightened bolt will be produced by the robots.
   (b) The probability that at least one defective weld and at least one improperly tightened bolt will be produced.
   (c) The probability that at most one defective weld will be produced.
   (d) The probability that at least two improperly tightened bolts will be produced.
2. In conducting an experiment in the laboratory, temperature gauges are to be used at four junction points in the equipment setup. These four gauges are randomly selected from a bin containing seven such gauges. Unknown to the scientist, three of the seven gauges give improper temperature readings. Let X denote the number of defective gauges selected and Y the number of nondefective gauges selected. The joint density for (X, Y) is given in Table 5.5.
   (a) The values given in Table 5.5 can be derived by realizing that the random variable X is hypergeometric. Use the results of Sec. 3.7 to verify the values given in Table 5.5.
   (b) Find the marginal densities for both X and Y. What type of random variable is Y?
   (c) Intuitively speaking, are X and Y independent? Justify your answer mathematically.
3. The joint density for (X, Y) is given by
   $$f_{XY}(x, y) = 1/n^2 \qquad x = 1, 2, 3, \ldots, n \quad y = 1, 2, 3, \ldots, n$$
   (a) Verify that $f_{XY}(x, y)$ satisfies the conditions necessary to be a density.
   (b) Find the marginal densities for X and Y.
   (c) Are X and Y independent?

*4. The joint density for (X, Y) is given by
   $$f_{XY}(x, y) = 2/n(n + 1) \qquad 1 \le y \le x \le n \qquad n \text{ a positive integer}$$
   (a) Verify that $f_{XY}(x, y)$ satisfies the conditions necessary to be a density. *Hint:* The sum of the first n integers is given by $n(n + 1)/2$.
   (b) Find the marginal densities for X and Y. *Hint:* Draw a picture of the region over which (X, Y) is defined.
   (c) Are X and Y independent?
   (d) Assume that $n = 5$. Use the joint density to find $P[X \le 3 \text{ and } Y \le 2]$. Find $P[X \le 3]$ and $P[Y \le 2]$. *Hint:* Draw a picture of the region over which (X, Y) is defined.

TABLE 5.5

x/y	0	1	2	3	4
0	0	0	0	0	1/35
1	0	0	0	12/35	0
2	0	0	18/35	0	0
3	0	4/35	0	0	0

TABLE 5.6

x/y	0	1	2	3
0	.400	.100	.020	.005
1	.300	.040	.010	.004
2	.040	.010	.009	.003
3	.009	.008	.007	.003
4	.008	.007	.005	.002
5	.005	.002	.002	.001

5. The two most common types of errors made by programmers are syntax errors and errors in logic. For a simple language such as BASIC the number of such errors is usually small. Let X denote the number of syntax errors and Y the number of errors in logic made on the first run of a BASIC program. Assume that the joint density for (X, Y) is as shown in Table 5.6.
   (*a*) Find the probability that a randomly selected program will have neither of these types of errors.
   (*b*) Find the probability that a randomly selected program will contain at least one syntax error and at most one error in logic.
   (*c*) Find the marginal densities for X and Y.
   (*d*) Find the probability that a randomly selected program contains at least two syntax errors.
   (*e*) Find the probability that a randomly selected program contains one or two errors in logic.
   (*f*) Are X and Y independent?
6. Consider Example 5.1.5. Verify that $P[X \leq 30 \text{ and } Y \leq 28] \doteq .15$ by integrating the joint density first with respect to y, then with respect to x.
7. (*a*) Use the joint density of Example 5.1.5 to find the probability that the inside pressure on the roof will be greater than 30, whereas the outside pressure is less than 32.
   (*b*) Use the marginal density for X to find $P[X \leq 28]$.
   (*c*) Use the marginal density for Y to find $P[Y > 30]$.
8. Let X denote the temperature (°C) and let Y denote the time in minutes that it takes for the diesel engine on an automobile to get ready to start. Assume that the joint density for (X, Y) is given by

$$f_{XY}(x, y) = c(4x + 2y + 1) \qquad \begin{array}{l} 0 \leq x \leq 40 \\ 0 \leq y \leq 2 \end{array}$$

   (*a*) Find the value of c that makes this a density.
   (*b*) Find the probability that on a randomly selected day the air temperature will exceed 20°C and it will take at least 1 minute for the car to be ready to start.
   (*c*) Find the marginal densities for X and Y.
   (*d*) Find the probability that on a randomly selected day it will take at least one minute for the car to be ready to start.
   (*e*) Find the probability that on a randomly selected day the air temperature will exceed 20°C.
   (*f*) Are X and Y independent? Explain on a mathematical basis.

9. An engineer is studying early morning traffic patterns at a particular intersection. The observation period begins at 5:30 a.m. Let X denote the time of arrival of the first vehicle from the north-south direction; let Y denote the first arrival time from the east-west direction. Time is measured in fractions of an hour after 5:30 a.m. Assume that the density for (X, Y) is given by

$$f_{XY}(x, y) = 1/x \qquad 0 < y < x < 1$$

   (*a*) Verify that this is a joint density for a two-dimensional random variable.
   (*b*) Find $P[X \leq .5 \text{ and } Y \leq .25]$.
   (*c*) Find $P[X > .5 \text{ or } Y > .25]$.
   (*d*) Find $P[X \geq .5 \text{ and } Y \geq .5]$.
   (*e*) Find the marginal densities for X and Y.
   (*f*) Find $P[X \leq .5]$.
   (*g*) Find $P[Y \leq .25]$.
   (*h*) Are X and Y independent? Explain.
10. The joint density for (X, Y) is given by

$$f_{XY}(x, y) = x^3 y^3/16 \qquad 0 \leq x \leq 2, 0 \leq y \leq 2$$

   (*a*) Find the marginal densities for X and Y.
   (*b*) Are X and Y independent?
   (*c*) Find $P[X \leq 1]$.
   (*d*) If it is known that $Y = 1$, what is $P[X \leq 1]$? (Do not use any computation to answer this question!)
11. Economic conditions cause fluctuations in the prices of raw commodities as well as in finished products. Let X denote the price paid for a barrel of crude oil by the initial carrier and let Y denote the price paid by the refinery purchasing the product from the carrier. Assume that the joint density for (X, Y) is given by

$$f_{XY}(x, y) = c \qquad 20 < x < y < 40$$

   (*a*) Find the value of c that makes this a joint density for a two-dimensional random variable.
   (*b*) Find the probability that the carrier will pay at least \$25 per barrel and the refinery will pay at most \$30 per barrel for the oil.
   (*c*) Find the probability that the price paid by the refinery exceeds that of the carrier by at least \$10 per barrel.
   (*d*) Find the marginal densities for X and Y.
   (*e*) Find the probability that the price paid by the carrier is at least \$25.
   (*f*) Find the probability that the price paid by the refinery is at most \$30.
   (*g*) Are X and Y independent? Explain.

*12. (*n-dimensional discrete random variables.*) Random variables of dimension $n > 2$ can be defined and studied by extending the definitions presented in the two-dimensional case in a logical way. For example, an n-tuple $(X_1, X_2, X_3, \ldots, X_n)$ in which each of the random variables $X_1, X_2, X_3, \ldots, X_n$ is a discrete random variable is called an n-dimensional discrete random variable. The density for such a random variable is given by

$$f(x_1, x_2, x_3, \ldots, x_n) = P[X_1 = x_1, X_2 = x_2, X_3 = x_3, \ldots, X_n = x_n]$$

This problem entails the use of a three-dimensional random variable.

Items coming off an assembly line are classed as being either nondefective, defective but salvageable, or defective and nonsalvageable. The probabilities of observing items in each of these categories are .9, .08, and .02, respectively. The probabilities do not change from trial to trial. Twenty items are randomly selected and classified. Let X_1 denote the number of nondefective items obtained, X_2 the number of defective but salvageable items obtained, and X_3 the number of defective and nonsalvageable items obtained.

(*a*) Find $P[X_1 = 15, X_2 = 3, X_3 = 2]$. *Hint:* Use the formula for the number of permutations of indistinguishable objects, Exercise 23, Chap. 1, to count the number of ways to get this sort of split in a sequence of 20 trials.

(*b*) Find the general formula for the density for (X_1, X_2, X_3).

***13.** (*n-dimensional continuous random variables.*) An n-tuple $(X_1, X_2, X_3, \ldots, X_n)$, where each of the random variables $X_1, X_2, \ldots, X_n$ is continuous, is called an n-dimensional continuous random variable. The density for an n-dimensional continuous random variable is defined by extending Definition 5.1.3 in a natural way. State the three properties that identify a function as a density for $(X_1, X_2, X_3, \ldots, X_n)$.

***14.** Let $f(x_1, x_2, x_3) = c(x_1 \cdot x_2 \cdot x_3)$ for $0 \le x_1 \le 1$, $0 \le x_2 \le 1$, $0 \le x_3 \le 1$. Find the value of c that makes this a density for the three-dimensional random variable (X_1, X_2, X_3).

Section 5.2

15. Four temperature gauges are randomly selected from a bin containing three defective and four nondefective gauges. Let X denote the number of defective gauges selected and Y the number of nondefective gauges selected. (See Exercise 2.) The joint density for (X, Y) is given in Table 5.5.

(*a*) From the physical description of the problem, should $\mathrm{Cov}(X, Y)$ be positive or negative?

(*b*) Find $E[X]$, $E[Y]$, $E[XY]$, and $\mathrm{Cov}(X, Y)$.

16. Let X denote the number of syntax errors and Y the number of errors in logic made on the first run of a BASIC program. (See Exercise 5.) The joint density for (X, Y) is given in Table 5.6.

(*a*) X and Y are not independent. Does this give any indication of the value of the covariance?

(*b*) Find $E[X]$, $E[Y]$, $E[XY]$, and $\mathrm{Cov}(X, Y)$. Give a rough physical interpretation of the covariance.

(*c*) Find $E[X + Y]$. What is the practical interpretation of this expectation?

17. Consider the random variable (X, Y) of Exercise 3. Without doing any additional computation, find $\mathrm{Cov}(X, Y)$.

18. Use the marginal densities given in Table 5.3 to compute $E[X]$ and $E[Y]$. Compare your results to those obtained in Example 5.2.1.

19. The joint density for (X, Y), where X is the inside and Y is the outside barometric pressure on an air support roof (see Example 5.1.5), is given by

$$f_{XY}(x, y) = c/x \qquad 27 \le y \le x \le 33 \qquad c = 1/(6 - 27 \ln 33/27) = 1.72$$

(*a*) Find $E[X]$, $E[Y]$, $E[XY]$, and $\mathrm{Cov}(X, Y)$.

(*b*) Find $E[X - Y]$. What is the practical physical interpretation of this expectation?

20. The joint density for (X, Y), where X is the temperature and Y is the time that it takes for a diesel engine on an automobile to get ready to start (see Exercise 8), is given by

$$f_{XY}(x, y) = (1/6640)(4x + 2y + 1) \qquad 0 \le x \le 40 \quad 0 \le y \le 2$$

(*a*) From a physical standpoint, do you think $\mathrm{Cov}(X, Y)$ should be positive or negative?

(*b*) Find $E[X]$, $E[Y]$, $E[XY]$, and $\mathrm{Cov}(X, Y)$.

21. The joint density for (X, Y), where X is the arrival time of the first vehicle from the north-south direction and Y is the arrival time of the first vehicle from the east-west direction at an intersection (see Exercise 9), is given by

$$f_{XY}(x, y) = 1/x \qquad 0 < y < x < 1$$

Find $E[X]$, $E[Y]$, $E[XY]$, and $\mathrm{Cov}(X, Y)$.

22. Find the covariance between the random variables X and Y of Exercise 10.

23. Let X denote the price paid for a barrel of crude oil by the initial carrier and let Y denote the price paid by the refinery purchasing the oil. (See Exercise 11.) The joint density for (X, Y) is given by

$$f_{XY}(x, y) = 1/200 \qquad 20 < x < y < 40$$

(*a*) From a physical standpoint, should $\mathrm{Cov}(X, Y)$ be positive or negative?

(*b*) Find $E[X]$, $E[Y]$, $E[XY]$, and $\mathrm{Cov}(X, Y)$.

(*c*) Find $E[Y - X]$. Interpret this expectation in a practical sense.

24. Show that $\mathrm{Cov}(XY) = E[XY] - E[X]E[Y]$. *Hint:* By definition, $\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$. Expand this product and apply the rules for expectation (Theorem 3.3.1). Remember that $\mu_X = E[X]$ and $\mu_Y = E[Y]$.

25. Prove that $\mathrm{Var}(X + Y) = \mathrm{Var}\, X + \mathrm{Var}\, Y + 2\,\mathrm{Cov}(X, Y)$. *Hint:* $\mathrm{Var}(X + Y) = E[(X + Y)^2] - (E[X + Y])^2$. Square these terms and apply the rules for expectation. (Theorem 3.3.1.)

26. Use the result of Exercise 25 to show that if X and Y are independent, then $\mathrm{Var}(X + Y) = \mathrm{Var}\, X + \mathrm{Var}\, Y$. This proves the third rule for variance. (Theorem 3.3.3.)

27. Show that if $X = Y$, then $\mathrm{Cov}(X, Y) = \mathrm{Var}\, X = \mathrm{Var}\, Y$.

***28.** Let the joint density for (X, Y) be given by

$$f(x, y) = \frac{1}{2(e - 1)}\left[\frac{1}{x} + \frac{1}{y}\right] \qquad 1 \le x \le e \qquad 1 \le y \le e$$

(*a*) Show that $\int_1^e \int_1^e f(x, y)\, dy\, dx = 1$.

(*b*) Find $E[X]$ and $E[Y]$.

(*c*) Find $E[XY]$.

(*d*) Are X and Y independent? Explain, based on your answers to parts (*b*) and (*c*) and Theorem 5.2.2.

Section 5.3

29. The joint density for (X, Y), where X denotes the number of defective and Y the number of nondefective temperature gauges selected from a bin containing three defective and four nondefective gauges, is given in Table 5.5. (See Exercise 2.)

(*a*) From the physical interpretation of the problem, should ρ_{XY} be positive or negative? Should ρ_{XY} be $+1$ or -1? Explain.

(*b*) Find $E[X^2]$ and $E[Y^2]$. Use the information from Exercise 15 to find ρ_{XY}.

In Exercises 30 to 34, find $E[X^2]$, $E[Y^2]$, Var X, Var Y, and ρ_{XY} for the random variables in the exercises referenced. In each case decide whether or not you would expect the graph of Y versus X to exhibit a strong linear trend.

30. Exercise 16.
31. Exercise 19.
32. Exercise 20.
33. Exercise 21.
34. Exercise 23.
35. Assume that $Y = \beta_0 + \beta_1 X$, $\beta_1 \neq 0$.
(*a*) Show that $\text{Cov}(X, Y) = \beta_1 \text{ Var } X$. *Hint:* $\text{Cov}(X, Y) = E[X(\beta_0 + \beta_1 X)] - E[X] \times E[\beta_0 + \beta_1 X]$. Use the rules for expectation.
(*b*) Show that Var $Y = \beta_1^2$ Var X. *Hint:* Use the rules for variance. (Theorem 3.3.3.)
(*c*) Find ρ_{XY}.
(*d*) Argue that $\rho_{XY} = 1$ if β_1, the slope of the line $Y = \beta_0 + \beta_1 X$, is positive and that $\rho_{XY} = -1$ if the slope of this line is negative.
36. Prove that if X and Y are independent, then $\rho_{XY} = 0$. Can we conclude that if X and Y are uncorrelated, then they are independent? Explain.
37. Without doing any additional computation, find ρ_{XY} for the random variables of Exercise 3.
38. What is the correlation between the random variables X and Y of Exercise 10?

Section 5.4

39. Consider Example 5.4.3.
(*a*) What is the expected value of X when $y = 31$?
(*b*) What is the expected value of Y when $x = 30$?
40. Consider Example 5.1.4.
(*a*) Find $f_{X|y}$. Note that $f_{X|y} = f_X$. From a physical standpoint, can you explain why these densities are the same?
(*b*) Find $f_{Y|x}$. Is $f_{Y|x} = f_Y$?
(*c*) Find the curve of regression of X on Y and the curve of regression of Y on X. Are these curves linear?
***41.** Consider the random variable (X, Y) of Exercise 4.
(*a*) Find the curve of regression of X on Y. Is the regression linear?
(*b*) Assume that $n = 10$ and find the mean value of X when $y = 4$.
(*c*) Find the curve of regression of Y on X. Is the regression linear?
(*d*) Assume that $n = 10$ and find the mean value of Y when $x = 4$.
42. Consider the random variable (X, Y) of Exercise 9.
(*a*) Find the curve of regression of X on Y. Is the regression linear?
(*b*) Find the mean value of X when $y = .5$.
(*c*) Find the curve of regression of Y on X. Is the regression linear?
(*d*) Find the mean value of Y when $x = .75$.
43. Consider Exercise 11.
(*a*) Find the curve of regression of X on Y. Is the regression linear?
(*b*) Find the mean price paid by the carrier for a barrel of crude oil given that the refinery price is \$30 per barrel.

(*c*) Find the curve of regression of Y on X. Is the regression linear?
(*d*) Find the mean price paid by the refinery for a barrel of crude oil given that the carrier paid \$35 per barrel.
44. Note that if $|\rho| = 1$, then $Y = \beta_0 + \beta_1 X$. For fixed values of X, $Y|x = \beta_0 + \beta_1 x$. Argue that $\mu_{Y|x}$ is a linear function of x. That is, argue that if X and Y are perfectly correlated, then the curve of regression of Y on X is linear. Is the converse true? Explain.

Section 5.5

45. Consider the linear transformation T defined by

$$T: u = 2x + y$$
$$v = x + 3y$$

(*a*) Is this transformation invertible? If so, find the defining equations for T^{-1}.
(*b*) Find the Jacobian for T^{-1}.
46. Consider the linear transformation T defined by

$$T: u = 3x + 2y$$
$$v = x - y$$

(*a*) Is this transformation invertible? If so, find the defining equations for T^{-1}.
(*b*) Find the Jacobian for T^{-1}.
47. Assume that X and Y are independent and uniformly distributed over $(0, 1)$ and $(0, 2)$, respectively. Find the joint density for (U, V), where U and V are as defined in Exercise 45.
48. (*Distribution of one function of two continuous random variables.*) Let X and Y be continuous random variables with joint density f_{XY}. Let $U = X + Y$. Prove that f_U, the density for $X + Y$, is given by

$$f_U(u) = \int_{-\infty}^{\infty} f_{XY}(u - v, v)\, dv$$

Hint: Define a transformation T by

$$u = g_1(x, y) = x + y$$
$$v = g_2(x, y) = y$$

Follow the procedure given in Theorem 5.5.1 to obtain the joint density for (U, V). Integrate the joint density to obtain the marginal density for U.
49. Let X and Y be independent standard normal random variables. Let $U = X + Y$. Use Exercise 48 to prove that U follows a normal distribution with mean 0 and variance 2. *Hint:* In integrating over v, complete the square in the exponent and remember that a normal density integrated over the real line is equal to 1.
50. Let X and Y be continuous random variables with joint density f_{XY}. Let $U = XY$. Prove that f_U, the density for XY, is given by

$$f_U(u) = \int_{-\infty}^{\infty} f_{XY}(u/v, v)|1/v|\, dv$$

Hint: Let $u = g_1(x, y) = xy$ and $v = y$, and apply Theorem 5.5.1.

51. Let X and Y be continuous random variables with joint density f_{XY}. Let $U = X/Y$. Prove that f_U, the density for X/Y, is given by

$$f_U(u) = \int_{-\infty}^{\infty} f_{XY}(uv, v)|v|\,dv$$

Hint: Let $u = g_1(x, y) = x/y$ and $v = y$, and apply Theorem 5.5.1.

52. Let X and Y be independent exponentially distributed random variables with parameters β_1 and β_2, respectively.
(*a*) Find the joint density for (X, Y).
(*b*) Let $U = X + Y$, and verify that

$$f_U(u) = \int_0^u f_{XY}(u - v, v)\,dv$$

Hint: Remember that $0 < x < \infty$ and that $x = u - v$.
(*c*) Assume that $\beta_1 = 3$ and $\beta_2 = 1$. Show that

$$f_U(u) = e^{-u/3} - e^{-u/2} \qquad 0 < u < \infty$$

53. Let X and Y be independent uniformly distributed random variables over the intervals $(0, 2)$ and $(0, 3)$, respectively.
(*a*) Let $U = XY$ and find f_U.
(*b*) Let $U = X/Y$ and find f_U.

REVIEW EXERCISES

54. An electronic device is designed to switch house lights on and off at random times after it has been activated. Assume that the device is designed in such a way that it will be switched on and off exactly once in a 1-hour period. Let Y denote the time at which the lights are turned on and X the time at which they are turned off. Assume that the joint density for (X, Y) is given by

$$f_{XY}(x, y) = 8xy \qquad 0 < y < x < 1$$

(*a*) Verify that f_{XY} satisfies the conditions necessary to be a density.
(*b*) Find $E[XY]$.
(*c*) Find the probability that the lights will be switched on within 1/2 hour after being activated and then switched off again within 15 minutes.
(*d*) Find the marginal density for X. Find $E[X]$ and $E[X^2]$.
(*e*) Find the marginal density for Y. Find $E[Y]$ and $E[Y^2]$.
(*f*) Are X and Y independent?
(*g*) Find the conditional distribution of X given Y.
(*h*) Find the probability that the lights will be switched off within 45 minutes of the system being activated given that they were switched on 10 minutes after the system was activated.
(*i*) Find the curve of regression of X on Y. Is the regression linear?
(*j*) Find the expected time that the lights will be turned off given that they were turned on 10 minutes after the system was activated.
(*k*) Based on the physical description of the problem, would you expect ρ to be positive, negative, or 0? Explain. Verify by computing ρ.

TABLE 5.7

x/y	1	2	3	4
0	.059	.100	.050	.001
1	.093	.120	.082	.003
2	.065	.102	.100	.010
3	.050	.075	.070	.020

55. Verify that

$$f_{XY}(x, y) = xye^{-x}e^{-y} \qquad x > 0 \qquad y > 0$$

satisfies the conditions necessary to be a density for a continuous random variable (X, Y). Find the marginal densities for X and Y. Are X and Y independent? Find ρ_{XY}.

56. Let X denote the number of "do loops" in a Fortran program and Y the number of runs needed for a novice to debug the program. Assume that the joint density for (X, Y) is given in Table 5.7.
(*a*) Find the probability that a randomly selected program contains at most one "do loop" and requires at least two runs to debug the program.
(*b*) Find $E[XY]$.
(*c*) Find the marginal densities for X and Y. Use these to find the mean and variance for both X and Y.
(*d*) Find the probability that a randomly selected program requires at least two runs to debug given that it contains exactly one "do loop."
(*e*) Find Cov(X, Y). Find the correlation between X and Y. Based on the observed value of ρ, can you claim that X and Y are not independent? Explain.

57. Vehicles arrive at a highway toll booth at random instances from both the south and north. Assume that they arrive at average rates of five and three per 5-minute period, respectively. Let X denote the number arriving from the south during a 5-minute period and let Y denote the number arriving from the north during this same time. Assume that X and Y are independent.
(*a*) Find the joint density for (X, Y).
(*b*) Find the probability that a total of four vehicles arrives during a five-minute time period.
(*c*) Find the correlation between X and Y.
(*d*) Find the conditional density for X given $Y = y$.

*58. (*Bivariate normal distribution.*) A random variable (X, Y) is said to have a bivariate normal distribution if its joint density is given by

$$f_{XY}(x, y) = \frac{\exp\left\{\dfrac{1}{-2(1-\rho^2)}\left[\left(\dfrac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\dfrac{x-\mu_X}{\sigma_X}\right)\left(\dfrac{y-\mu_Y}{\sigma_Y}\right) + \left(\dfrac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

where x and y can assume any real value. The parameters $\mu_X, \mu_Y, \sigma_X, \sigma_Y$ denote the respective means and standard deviations for X and Y. The parameter ρ is the correlation coefficient. The name of this distribution comes from the fact that the marginal densities for X and Y are both normal. Show that in the case of a bivariate normal distribution, if $\rho = 0$, then X and Y are independent.

(*f*) The results of part (*e*) indicate that the test as designed cannot distinguish well between $p = .1$ and $p = .2$. Keeping the sample size fixed at $n = 20$, can you suggest a way to modify the test that will lower β and to increase the power for detecting a failure rate of .1? Will α still be small enough to be acceptable? If not, can you suggest a way to redesign the experiment that will make both α and β low enough to be acceptable?

29. In Example 8.3.3 we test

$$H_0: p \leq .5$$

$$H_1: p > .5 \qquad \text{(majority of automobiles in operation have misaimed headlights)}$$

at the $\alpha = .0577$ level by agreeing to reject H_0 if at least 14 of the 20 cars sampled have misaimed headlights. We claim that values of X that are too large to occur by chance when $p = .5$ are also too large to occur by chance when $p < .5$. That is, if these values are rare when $p = .5$, they are even more rare when $p < .5$. To help see that this is true, find $P[X \geq 14]$ when $p = .4; .3; .2; .1$. Are each of these probabilities less than .0577 as expected?

30. A sample of size 9 from a normal distribution with $\sigma^2 = 25$ is used to test

$$H_0: \mu = 20$$

$$H_1: \mu = 28$$

The test statistic used is the sample mean, $\bar{X}$. Let us agree to reject H_0 in favor of H_1 if the observed value of $\bar{X}$ is greater than 25.

(*a*) If H_0 is true, what is the distribution of $\bar{X}$?
(*b*) In the diagram of Fig. 8.15, shade the region whose area is α.
(*c*) Find α. Remember that α is computed under the assumption that H_0 is true.
(*d*) If H_1 is true, what is the distribution of $\bar{X}$?
(*e*) In the diagram of Fig. 8.15, shade the region whose area is β. Remember that β is computed under the assumption that H_1 is true.
(*f*) Find β.
(*g*) Find the power of the test.
(*h*) If the sample size is increased, the standard deviation of $\bar{X}$ will decrease. What is the geometric effect of this on the two curves of Fig. 8.15?
(*i*) If the sample size is increased but the critical point is not changed, what will be the effect on α and β?

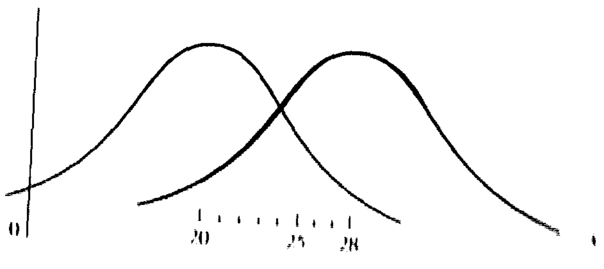


FIGURE 8.15

Section 8.4

31. Whenever a motorist encounters braking problems, especially an unpredictable pulling to one side, the villain is always held to be the brake pad. Trace elements, especially titanium, can combine with other elements to form minute particles of titanium carbonitride which alters the degree of friction between the pad and disc and leads to unequal wear. The percentage of titanium in a brake pad should not exceed 5%. A study is conducted to detect a situation in which the mean percentage of titanium in the brake pads being produced by a particular manufacturer exceeds 5%. (*Design Engineering*, February 1982, p. 24.)
(*a*) Set up the appropriate null and alternative hypotheses.
(*b*) Discuss the practical consequences of making a Type I and a Type II error.
(*c*) A sample of 100 brake pads yields a mean percentage of $\bar{x} = .051$. Assume that $\sigma = .008$. Find the P value for the test. Do you think that H_0 should be rejected? Explain. To what type error are you now subject?

32. The current particulate standard for diesel car emission is .6 g/mi. It is hoped that a new engine design has reduced the emissions to a level below this standard. (*Design Engineering*, February 1982, p. 13.)
(*a*) Set up the appropriate null and alternative hypotheses for confirming that the new engine has a mean emission level below the current standard.
(*b*) Discuss the practical consequences of making a Type I and a Type II error.
(*c*) A sample of 64 engines tested yields a mean emission level of $\bar{x} = .5$ g/mi. Assume that $\sigma = .4$. Find the P value of the test. Do you think that H_0 should be rejected? Explain. To what type error are you now subject?

33. It is thought that more than 15% of the furnaces used to produce steel in the United States are still open-hearth furnaces. To verify this contention, a random sample of 40 furnaces is selected and examined.
(*a*) Set up the appropriate null and alternative hypotheses required to support the stated contention.
(*b*) When the data are gathered, it is found that 9 of the 40 furnaces inspected are open-hearth furnaces. Use the normal approximation to the binomial distribution (Sec. 4.6) to find the P value for the test. Do you think that H_0 should be rejected? Explain. To what type error are you now subject?

34. It is known that defective items will be produced even on automated assembly lines. A particular process typically produces 5% defectives. If the proportion of defectives exceeds 5%, then the line must be shut down and adjusted.
(*a*) Set up the null and alternative hypotheses needed to detect a situation in which the proportion of defectives produced exceeds .05.
(*b*) Discuss the practical consequences of committing a Type I and a Type II error.
(*c*) A random sample of 100 items is selected and tested. Of these, 7 are found to be defective. Use the normal approximation to the binomial distribution to find the P value of the test. Do you think that H_0 should be rejected?

Section 8.5

35. Find the critical point(s) for conducting a hypothesis test on the mean with σ^2 unknown for
(*a*) a left-tailed test with $n = 25$; $\alpha = .05$
(*b*) a left-tailed test with $n = 150$; $\alpha = .10$
(*c*) a right-tailed test with $n = 20$; $\alpha = .025$

(*d*) a right-tailed test with $n = 16$; $\alpha = .01$
(*e*) a two-tailed test with $n = 20$; $\alpha = .10$
(*f*) a two-tailed test with $n = 30$; $\alpha = .05$

36. A new 8-bit microcomputer chip has been developed that can be reprogrammed without removal from the microcomputer. It is claimed that a byte of memory can be programmed in less than 14 seconds. (*Design News*, April 1983, p. 26.)
(*a*) Set up the null and alternative hypotheses needed to verify this claim.
(*b*) What is the critical point for an $\alpha = .05$ level test based on a sample of size 15?
(*c*) These data are obtained on X, the time required to reprogram a byte of memory:

11.6	14.7	12.9	13.3	13.2
13.1	14.2	15.1	12.5	15.3
13.3	13.4	13.0	13.8	12.3

Construct a stem-and-leaf diagram for these data. Does the normality assumption look reasonable?
(*d*) Test the null hypothesis. Can H_0 be rejected at the $\alpha = .05$ level? Interpret your result in a practical sense. To what type error are you now subject?

37. Ozone is a component of smog that can injure sensitive plants even at low levels. In 1979 a federal ozone standard of .12 ppm was set. It is thought that the ozone level in air currents over New England exceeds this level. To verify this contention, air samples are obtained from 30 monitoring stations set up across the region. ("Air Pollution Stress and Energy Policy," F. Bormann, *Ambio*, vol. XI, 1982, pp. 188–194.)
(*a*) Set up the appropriate null and alternative hypotheses for verifying the contention.
(*b*) What is the critical point for an $\alpha = .01$ level test based on a sample of size 30?
(*c*) When the data are analyzed, a sample mean of .135 and a sample standard deviation of .03 are obtained. Use these data to test H_0. Can H_0 be rejected at the $\alpha = .01$ level? What does this mean in a practical sense?
(*d*) What assumption are you making concerning the distribution of the random variable X, the ozone level in the air?

38. A model of Saudi Arabia's oil export strategy has been devised based on interviews with informed economists. The model is to be used to estimate the mean number of barrels of oil produced per day by this country. The usefulness of the model is to be partially checked by comparing the predicted mean for the year 1980 to its known value for that year, namely, 9.5 million barrels per day. ("Simulating Saudi Arabia's Oil Export Strategy," A. Picardi and A. Shorb, *Simulation*, January 1983, pp. 20–27.)
(*a*) Find the critical points for testing

$$H_0: \mu = 9.5$$
$$H_1: \mu \neq 9.5$$

at the $\alpha = .05$ level based on a sample of 50 simulations.
(*b*) For the data collected $\bar{x} = 9.8$ and $s = 1.2$. Test H_0. Can H_0 be rejected at the $\alpha = .05$ level? Based on these data, is there evidence that the model is not adequate? To what type error are you now subject?

39. A low-noise transistor for use in computing products is being developed. It is claimed that the mean noise level will be below the 2.5-dB level of products currently in use. (*Journal of Electronic Engineering*, March 1983, p. 17.)
(*a*) Set up the appropriate null and alternative hypotheses for verifying the claim.
(*b*) A sample of 16 transistors yields $\bar{x} = 1.8$ with $s = .8$. Find the P value for the test. Do you think that H_0 should be rejected? What assumption are you making concerning the distribution of the random variable X, the noise level of a transistor?
(*c*) Explain, in the context of this problem, what conclusion can be drawn concerning the noise level of these transistors. If you make a Type I error, what will have occurred? What is the probability that you are making such an error?

40. The Elbe River is important in the ecology of central Europe as it drains much of this region. Due to increased industrialization, it is feared that the mineral content in the soil is being depleted. This will be reflected in an increase in the level of certain minerals in the water of the Elbe. A study of the river conducted in 1982 indicated that the mean silicon level was 4.6 mg/l. ("Natural and Anthropogenic Flux of Major Elements from Central Europe." T. Paces, *Ambio*, vol. XI, November 1982, pp. 206–208.)
(*a*) Set up the null and alternative hypotheses needed to gain evidence to support the contention that the mean silicon concentration in the river has increased.
(*b*) A sample of size 28 yields $\bar{x} = 5.2$ and $s = 1.6$. Find the P value for the test. Do you think that H_0 should be rejected?
(*c*) What practical conclusion can be drawn from these data?

41. Coal-handling maintenance is a very young technology. The emission standard for coal-burning plants is 4.8 pounds SO_2/per million Btu/per 24-hour average. In an attempt to get emissions below this level engineers are experimenting with burning a blend of high- and low-sulfur coal. ("Upgrading and Maintaining Coal Handling," R. Rittenhouse, *Power Engineering*, March 1983, pp. 42–50.)
(*a*) Set up the null and alternative hypotheses needed to support the contention that the new mixture falls below the emission standard set by the government.
(*b*) Find the P value for the test if a sample of 200 readings yields a sample mean of 4.7 with a sample standard deviation of .5. Do you think that H_0 should be rejected? What does this mean in a practical sense?

42. Lasers are now used to detect structural movement in bridges and large buildings. These lasers must be extremely accurate. In laboratory testing of one such laser measurements of the error made by the device are taken. The data obtained are used to test

$$H_0: \mu = 0$$
$$H_1: \mu \neq 0$$

A sample of 25 measurements yields $\bar{x} = .03$ millimeter over 100 meters and $s = .1$. Find the P value for this two-tailed test. Do you think H_0 should be rejected? Interpret your result in a practical sense.

43. Clams, mussels, and other organisms that adhere to the water intake tunnels of electrical power plants are called *macrofoulants*. These organisms can, if left unchecked, inhibit the flow of water through the tunnel. Various techniques have been tried to control this problem, among them increasing the flow rate and coating the tunnel with Teflon, wax, or grease. In a year's time at a particular plant an times, yielding total organic carbon amounts removed of 38.8, 53.6, 39.0, 51.6, 40.1, 46.9, 40.9, 44.9, 41.0, and 43.2.

(*a*) What is $E[W]$?

(*b*) Using the signed-rank test, is there evidence that the new process removes significantly more total organic carbon than the standard process at the .05 level?

56. In an attempt to determine how many consultants are needed to answer questions of users at a computer center, these data are collected on X, the time in minutes required to answer a telephone inquiry:

1.5	1.0	5.0	1.9	3.0
1.3	2.1	1.7	6.5	4.2
6.3	5.6	5.1	2.5	6.9

(*a*) What is $E[W]$?

(*b*) Based on the signed-rank test, can we conclude that the median time required is less than 5 minutes? Explain, based on the P value of your test. (A zero score should be given the lowest rank and should be assigned the algebraic sign least conducive to rejecting the null hypothesis.)

***57.** A study of the expansion joints used in bridge beds is conducted. It is thought that these joints are expanding more than they were designed to expand, thus creating cracks in the pavement near the joint. The median design expansion at 95°F is 2 inches. Laboratory tests of 100 such joints are conducted at this temperature.

(*a*) What is $E[W]$?

(*b*) What is $\mathrm{Var}[W]$?

(*c*) Set up the appropriate null and alternative hypotheses.

(*d*) If $|W_-| = 1600$, can H_0 be rejected? Explain, based on the P value of the test.

REVIEW EXERCISES

58. A consumer group wants to estimate the mean cost of the base system for a personal computer with certain specifications. It is thought that these computers range in price from \$2390 to \$4000.

(*a*) How large a sample should be taken to estimate μ to within \$100 with 90% confidence?

(*b*) A random sample of size 50 yields these data (data in thousands of dollars):

2.43	2.86	2.74	2.75	2.69	2.64	2.91
2.89	3.18	3.00	3.21	3.07	3.72	3.24
3.17	3.57	3.37	3.56	3.30	2.32	3.09
2.99	3.20	3.25	3.70	3.45	2.82	2.88
2.71	3.25	2.86	2.93	3.45	3.11	3.86
2.96	3.00	2.88	3.19	3.56	3.21	3.33
3.39	3.14	2.90	3.49	3.02	3.56	2.87
2.32						

Construct a stem-and-leaf chart for these data. Use the digits 2 and 3 as stems 5 times each. Graph numbers beginning 2.0 and 2.1 on the first stem, those beginning 2.2 and 2.3 on the second stem, and so forth. Does the stem-and-leaf chart lead you to suspect that these data are not drawn from a distribution that is at least approximately normal?

(*c*) Find unbiased estimates for μ and σ^2 based on these data. Estimate σ. Is the estimate for σ unbiased?

(*d*) Find 90% confidence intervals on σ^2 and σ.

(*e*) Find a 90% confidence interval on μ.

59. Researchers are experimenting with a new compound used to bond Teflon to steel. The compounds currently in use require an average drying time of 3 minutes. It is thought that the new compound dries in a shorter length of time.

(*a*) Set up the null and alternative hypotheses needed to support the claim that the new compound dries faster than those currently in use.

(*b*) Discuss the practical consequences of making a Type I error; a Type II error.

(*c*) A pilot study shows that $\hat{\sigma} = .5$. Suppose that the new product is worth marketing if the average drying time can be shown to be 2.5 minutes or less. How large a sample is required to detect this situation with probability .95 with α set at .05?

(*d*) When the experiment is conducted, these data are obtained:

1.4	2.1	2.8	.9
2.4	1.7	3.7	2.7
2.6	1.9	2.8	2.8
2.2	2.2	3.4	1.9

Test the null hypothesis of part (*a*) at the $\alpha = .05$ level. Would you suggest marketing this new product?

60. It is thought that a majority of the procedures used in a statistical computer package run in less than .1 second. To verify this contention, a random sample of 20 programs which entail exactly one procedure is to be examined.

(*a*) Set up the appropriate null and alternative hypotheses needed to verify the claim.

(*b*) Let X denote the number of programs in which the procedure used runs in less than .1 second. Find the critical region for an $\alpha \doteq .025$ level test.

(*c*) When the test is conducted, 14 programs are found in which the procedure used runs in less than .1 second. Will H_0 be rejected? To what type error are you now subject?

(*d*) Find β if $p = .6$; if $p = .7$; if $p = .8$; if $p = .9$.

(*e*) Find the power of the test if $p = .6$; if $p = .7$; if $p = .8$; if $p = .9$.

61. Nickel powders are used in coatings used to shield electronic equipment from electromagnetic interference. It is thought that the mean size of the individual nickel particles in one such coating is less than 3 micrometers. Do these data support this contention? Explain, based on the P value of the appropriate test.

3.26	3.07	2.46	1.76
1.89	2.95	3.35	3.82
2.42	1.39	1.56	2.42
2.03	3.06	1.79	2.96

62. We want to test

$$H_0: \mu = 5$$
$$H_1: \mu > 5$$

based on a random sample of size 25. The sample standard deviation is 2 and the observed value of the sample mean is 5.5. What is the P value for the test?

TABLE 15.15

Fire	Suspect make: Yes	No	
Yes	9	31	
No	16	144	
			200

survey of 500 randomly selected customers of a bank that has been offering computerized banking for over a year. Is there evidence of an association between these two variables? Explain, based on the P value of your test and inspection of the table.

9. It is suspected that the tendency of an automobile to catch fire in a rear end collision is not independent of the make of the car. To support this contention, a random sample of 200 cars involved in rear end collisions is selected from past records. Each car is classified as to make and whether or not it is one of the cars suspected of being especially susceptible to fire under these circumstances. The data gathered is shown in Table 15.15. Is there evidence of an association between this make of car and the presence of fire when involved in a rear end collision? Explain.

10. A study is conducted to test for independence between air quality and air temperature. These data are obtained from records on 200 randomly selected days over the last few years. (See Table 15.16.) Do these data indicate an association between these variables? Explain, based on the P value of the test.

11. In a study of the association between color and the effectiveness of a graphical display 100 graphs are randomly selected from among current scientific journals. Each is classified as to whether or not color is used. Each is also rated as to its effectiveness in making its point. Resulting data are given in Table 15.17. Is there evidence that the effectiveness of a graphical display is not independent of color? Explain, based on the P value of the test.

12. It is suspected that there is an association between the day of the week on which an item is produced and the quality of the item. To support this contention, a random sample of 500 items is selected from stock and each item is classified as to the day on which it was produced via its lot number. The item is also rated for quality. The data gathered are shown in Table 15.18.

(*a*) Our guideline on expected cell frequencies states that no more than 20% can be less than 5 and none can be less than 1. Is this criterion satisfied in this case?

(*b*) To satisfy the criterion, combine the quality of categories "Fair" and "Poor" to form a new table with three rows and five columns. Use this table to test for independence.

(*c*) Has an association between quality and day of production been established? Explain.

TABLE 15.16

Temperature	Air quality: Poor	Fair	Good	
Below average	1	3	24	
Average	12	28	76	
Above average	12	14	30	
				200

TABLE 15.17

Effective	Color present: Yes	No	
Excellent	7	4	
Good	10	19	
Fair	9	26	
Poor	4	21	
			100

TABLE 15.18

Quality	Day produced: M	T	W	Th	F	
Excellent	44	74	79	72	31	
Good	14	25	27	24	10	
Fair	15	20	20	23	9	
Poor	3	5	5	0	0	
						500

Section 15.4

13. A study is conducted to assess the effectiveness of a new computerized system of filling orders in a particular industry. Random samples of 100 customers served via the old system and 100 served via the new system are selected. Each customer is contacted to determine whether or not the order was filled satisfactorily within 2 weeks. Table 15.19 gives the results of the study. Test the null hypothesis that the proportion of satisfied customers among those served by the new system is the same as that among those served by the old system at the $\alpha = .05$ level.

14. Although many jobs in the airline industry entail stress, it is thought that air traffic controllers are particularly susceptible to stress-related disorders such as heart problems, high blood pressure, and ulcers. To support this contention, a random sample of 500 air traffic controllers is selected and surveyed. For comparative purposes a sample of 700 workers from other areas of the airline industry is also

(*b*) {(0, 12, −1), (1, 12, 0), (2, 17, 0), (3, 17, 0), (4, 17, 0), (5, 17, 0), (6, 17, 0), (7, 17, 0), (8, 17, 0), (9, 17, 0)}
(*c*) yes
(*d*) 5/10
(*e*) 1/10
(*f*) 9/10
(*g*) 1

38. $180!/\underbrace{5!5!5!\cdots 5!}_{36 \text{ terms}}$

Section 2.1

1. 12/13
3. .45; .13; .46
5. .58; .28; .1

Section 2.2

13. (*a*) 30/58
(*b*) 28/58
(*c*) Theorem 2.1.2
(*d*) 10/42
(*e*) no; exposure to the lethal dose should increase the probability of death
15. 5/35; 35/40
17. (*a*) 1/5
(*b*) 1/80
(*c*) .04
(*d*) 4/20
(*e*) .84

Section 2.3

19. no, $P[A_1 \cap A_2] = .2 \neq P[A_1]P[A_2]$
21. $(.39)^2 \doteq .15$
23. (*a*) .21
(*b*) 21/23
25. .085
27. .0144(.67) + .0012(.33) = .010044

Section 2.4

35. .85(.10)/[.85(.10) + .04(.90)] = .7025
37. .9999

Chapter 2 review exercises

38. (*a*) .85
(*b*) .15
(*c*) 5/20
(*d*) 5/10

39. (*a*) 1/2
(*b*) 1/8
40. .3529; .2353; .2647; .1471
41. .24; .6; .16
42. (*a*) .5
(*b*) .35
(*c*) .50
(*d*) 1/3
(*e*) 35/85
43. (*a*) .0008
(*b*) .0002
(*c*) .2
44. $(.99)^3(.01) = .00970299$; $.01 + .99(.01) + (.99)^2(.01) + (.99)^3(.01)$

Section 3.1

1. not discrete
3. discrete
5. not discrete

Section 3.2

7. (*a*) .01
(*b*)

x	0	1	2	3	4	5
$F(x)$	.7	.9	.95	.98	.99	1.00

(*c*) .98; .1
(*d*) .03

9. (*a*)

x	0	1	2	3
$f(x)$	$(.1)^3$	$3(.9)(.1)^2$	$3(.9)^2(.1)$	$(.9)^3$

(*b*) $k(x) = \dfrac{3!}{x!(3-x)!}$
(*c*)

x	0	1	2	3
$F(x)$	.001	.028	.271	1.00

(*d*) .999
(*e*) .028

11. (*b*) no
(*c*) yes
(*d*) 1; 0
(*e*) yes
13. right continuous, nondecreasing, $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$

Section 3.3

15. (*a*) 4.96; 26.34
(*b*) 1.7384; 1.3185
(*c*) holes per bit

17. $1/.7 = 10/7$; $1/p$

21. (*a*) 11
(*b*) −17
(*c*) 16
(*d*) 4
(*e*) 64
(*f*) 8
(*g*) 208
(*h*) 640
(*i*) 0; 1
(*j*) 0; 1
(*k*) $E\left[\dfrac{X-\mu}{\sigma}\right] = 0$ and $\text{Var}\left[\dfrac{X-\mu}{\sigma}\right] = 1$

***23.** (*e*) $163/60 = 2.7167$
(*f*) 2.7007
(*g*) $E[X_{100}] = 94.7953$

Section 3.4

25. (*a*) success is selecting an unacceptable lot; .05
(*b*) $f(x) = (.95)^{x-1}(.05)$
(*c*) $m_X(t) = \dfrac{.05e^t}{1 - .95e^t}$
(*d*) 20; 780; 380; 19.4936
(*e*) .9025
(*f*) .9025

27. $F(x) = 1 - (.95)^x$ for x, a positive integer; .142625

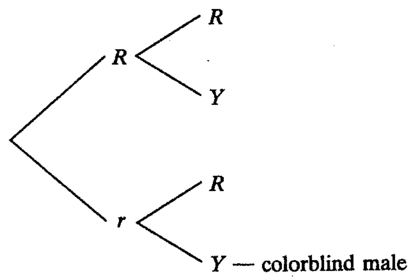
31. (*b*) $24/5$
(*c*) $(1/5)e^{4t} + (4/5)e^{5t}$
(*e*) $116/5$
(*g*) $4/25$; $2/5$

35. (*a*) $e - 1$
(*b*) $m_X(t) = \dfrac{(e-1)(e^{t-1})}{1 - e^{t-1}}$
(*c*) $E[X] = \dfrac{e^{-1}(e-1)}{(1-e^{-1})^2} = \dfrac{e}{e-1}$

Section 3.5

37. 2; .6778; yes, $P[X \geq 5] = .0328$

39. (*a*) $p = 1/4$



(*b*) 1.25; .1035

41. (*a*) .3758
(*b*) $(.3758)^5$

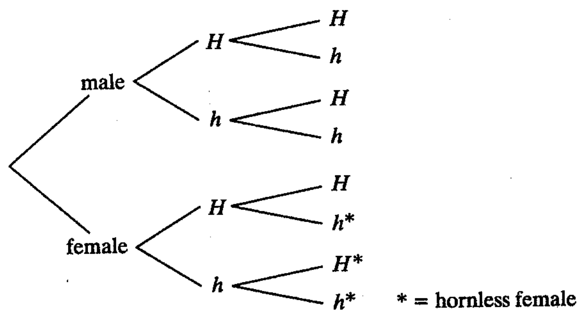
43. (*a*) $m_X(t) = (pe^t + q)^n$

45. (*a*) X = number of successes in one trial

$$f(x) = \binom{1}{x} p^x q^{1-x} \qquad x = 0, 1$$

(*b*) $f(x) = p^x q^{1-x} \qquad x = 0, 1$
(*c*) $m_X(t) = pe^t + q$
(*d*) p; pq
(*e*) .14

***53.**



when $r = 2$ and $p = 3/8$,

$E[X] = 16/3$; no; $P[X \leq 5] \doteq .6185$

Section 3.7

55. $x = 0, 1, 2, 3$; $15/20$; $5\left(\frac{3}{20}\right)\left(\frac{17}{20}\right)\left(\frac{15}{19}\right) = .5033$

57. (*a*) $f(x) = \dfrac{\binom{3}{x}\binom{17}{5-x}}{\binom{20}{5}} \qquad x = 0, 1, 2, 3$

(*b*) .75; .5033
(*c*) .3991
(*d*) .6009

59. (*a*) $f(x) = \dfrac{\binom{600}{x}\binom{2400}{20-x}}{\binom{3000}{20}} \qquad x = 0, 1, 2, \ldots, 20$

(*b*) 4; $20\left(\frac{600}{3000}\right)\left(\frac{2400}{3000}\right)\left(\frac{2980}{2999}\right)$

(*c*) $\displaystyle\sum_{x=0}^{3} \frac{\binom{600}{x}\binom{2400}{20-x}}{\binom{3000}{20}}$

(*d*) .4114

Section 3.8

59. (*a*) 10
(*b*) 10
(*c*) $\sqrt{10}$
(*d*) $f(x) = \dfrac{e^{-10}10^x}{x!} \qquad x = 0, 1, 2, \ldots$
(*e*) .029
(*f*) .010
(*g*) .019
(*h*) .99
(*i*) .448

63. .010

65. yes, $P[X > 10] = .043$

67. yes, $P[X < 2] = .04$

71. .135; .865

Chapter 3 review exercises

75. .118; .882

76. (*a*) 5/4; 15/16
(*b*) $(3/4)^5$
(*c*) $(3/4)^{10}$

77. (*a*) X is binomial with $n = 100$ and $p = .1$; $E[X] = 10$
(*b*) Poisson
(*c*) $P[X \geq 17] \doteq .027$

78. $P[X \leq 6] \doteq .4662$

79. 5; 61/125; 64/125

80. 35/210; 105/210

81. $(999/1000)^{12,000} \doteq e^{-12}12^0/0! \doteq .000006$; .999994

82. yes; $P[X \geq 5] = .0127$

83. (*b*) 36/14; 98/14
(*c*) $m_X(t) = (e^t + 4e^{2t} + 9e^{3t})/14$
(*e*) 76/196; $\sqrt{76/196}$

84. $F(x) = 1 - \left(\frac{12}{13}\right)^x \qquad x$ a positive integer;

$$P[X \geq 3] = 1 - P[X \leq 2] = 1 - F(2) = \left(\frac{12}{13}\right)^2$$

85. (*a*) binomial, $n = 10$, $p = .8$
(*b*) Poisson, $k = 5$
(*c*) point binomial, $p = .3$ or binomial, $n = 1$, $p = .3$
(*d*) geometric, $p = .6$
(*e*) negative binomial, $n = 5$, $p = .3$
(*f*) Poisson, $k = 1$

86. (*a*) 8; 1.6; $\sqrt{1.6}$
(*b*) 5; 5; $\sqrt{5}$
(*c*) .3; .21; $\sqrt{.21}$
(*d*) 10/6; 10/9; $\sqrt{10/9}$
(*e*) 50/3; $5(.7)/(.3)^2 = 38.89$; $\sqrt{38.89}$
(*f*) 1; 1; 1

87. (*a*) $f(x) = (.99)^{x-1}(.01) \qquad x = 1, 2, 3, \ldots$
(*b*) 100
(*c*) $F(x) = 1 - .99^x \qquad x$ a positive integer
(*d*) $P[X \leq 90] = 1 - (.99)^{90} \doteq .595$

Section 4.1

1. (*a*) 1/6
(*b*) 11/48
(*c*) 0
(*d*) 11/48

3. (*b*) $1 - e^{-.7} = .5034$; .4966; 0
(*c*) yes, $P[1 \leq X \leq 2] = e^{-.1} - e^{-.2} \doteq .086$

5. (*d*) .5
(*e*) equal; probabilities are constant or "uniform" over intervals of equal lengths

9. (*a*)

$$F(x) = \begin{cases} 0 & x < 2 \\ x^2/12 - 1/3 & 2 \leq x \leq 4 \\ 1 & x > 4 \end{cases}$$

(c) yes; yes; 0; 1; yes

(d) $\dfrac{dF(x)}{dx} = f(x)$

11. (a)

$$F(x) = \begin{cases} 0 & x < 0 \\ x/2\pi & 0 \le x < 2\pi \\ 1 & x \ge 2\pi \end{cases}$$

(b) yes; yes; 0; 1; yes

(c) $\dfrac{dF(x)}{dx} = f(x)$

13.

$$F(x) = \begin{cases} 0 & x < 25 \\ (\ln x - \ln 25)/\ln 2 & 25 \le x \le 50 \\ 1 & x > 50 \end{cases}$$

Section 4.2

15. (a) 56/18
(b) 10
(c) 104/324; $\sqrt{104/324}$

17. (a) $m_X(t) = (1 - 10t)^{-1} \qquad t < 1/10$
(b) 10 minutes
(c) 100; 10 minutes

19. π; $\pi^2/3$; $\pi/\sqrt{3}$

21. 10; 10; X

Section 4.3

25. (a) 2
(b) 5040
(c) 96
(d) 1

27. (a) 1; 2; 6; 24; 120
(b) $\Gamma(n) = (n-1)! \qquad n > 1$
(c) yes, $\Gamma(1) = 1 = 0!$
(d) 14!

29. (a) $f(x) = (1/128)x^2 e^{-x/4} \qquad x > 0$
(b) $m_X(t) = (1 - 4t)^{-3} \qquad t < 1/4$
(c) 12; 48; $\sqrt{48}$

33. $m_X(t) = (1 - \beta t)^{-1} \qquad t < 1/\beta$; β; β^2

37. (a) $e^{-.25} \doteq .779$
(b) $e^{-7/12}/e^{-4/12} = e^{-3/12}$; same as (a)

Section 4.4

39. (a) .9418
(b) .9418
(c) 0

(d) .0582
(e) .8543
(f) 1.28
(g) − 1.28
(h) 1.96
(i) 1.645

41. (a) .9544
(b) 1.24%
(c) 128 parsecs
(d) $m_X(t) = e^{5000t^2}$

43. (a) $f(z) = \dfrac{1}{\sqrt{2\pi}} e^{-(1/2)z^2} \qquad -\infty < z < \infty$

(b) $f'(z) = -\dfrac{1}{\sqrt{2\pi}} z e^{-(1/2)z^2}$

(c) $f''(z) = -\dfrac{1}{\sqrt{2\pi}} e^{-(1/2)z^2}[1 - z^2]$

Section 4.5

49. $P[-2\sigma < X - \mu < 2\sigma] = .95$

51. (a) no
(b) $P[|X - \mu| < .5\sigma] \ge -3$, $P[|X - \mu| < 1\sigma] \ge 0$,
$P[|X - \mu| < 2\sigma] \ge .75$, $P[|X - \mu| < 3\sigma] \ge .89$; $\qquad k = 2$

Section 4.6

53. (a) yes; .9956
(b) .0668

55. (a) yes
(b) 37.5
(c) .0322
(d) .0392

57. .9484; .8413

Section 4.7

59. (a) random factors
(b) $R(t) = e^{-.02t}$
(c) $R(20) = .6703$
(d) $f(x) = .02e^{-.02x} \qquad x > 0$
(e) exponential
(f) 50; 2500
(g) .5488

65. (a) .9974; .96
(b) .8812
(c) .9996
(d) .8988
(e) III: .9964, V: .999999; .9537

67. 3

Section 8.2

9. (*a*) 1.86
(*b*) −1.86
(*c*) −2.179
(*d*) 2.179
(*e*) 1.645
(*f*) 1.645
(*g*) 1.708
(*h*) 2.060
(*i*) 1.753
(*j*) 1.325
(*k*) 1.746
(*l*) 1.310

11. (*a*) 1.2896; .0000123; .0035
(*b*) 1.2896 ± .0016
(*c*) no, 1.29 is contained in the confidence interval

13. (*a*) 2.35; .89
(*b*) 2.35 ± .45
(*c*) yes, we are 99% confident that the new mean time is at most 2.80 seconds

15. (*a*)

2	9
3	8 5 9
4	2 1 7 8
5	1 3 5 6 7 5
6	1 8 1 0 2 5 7 3
7	9 7 3 2
8	1 0
9	2
10	0

(*b*) $\tilde{x} = 605$ $f_1 = 120$ no outliers
$q_1 = 480$ $f_3 = 1080$
$q_3 = 720$
$iqr = 240$
(*c*) $\bar{x} = 602.3$; $s = 169.1$; 602.3 ± 85.1
(*d*) lower the confidence

***19.** (*b*) 385
(*c*) 153

Section 8.3

21. (*a*) H_0: $\mu \geq .08$
H_1: $\mu < .08$
(*b*) we shall conclude that the average percentage of metal in household wastes has been reduced when, in fact, it has not been reduced
(*c*) we shall be unable to detect the fact that the mean percentage of metal in household waste has been reduced
(*d*) we have a 5% chance of having committed a Type I error

23. (*a*) we shall conclude that the model is not credible when, in fact, it is a valid model
(*b*) we shall be unable to detect the fact that the proposed model is not credible

25. (*a*) $C = \{10, 11, 12, 13, 14, 15\}$, $\alpha = .0338$
(*b*) yes

27. (*a*) H_0: $p \leq .5$
H_1: $p > .5$
(*b*) 7.5
(*c*) .0592
(*d*) .7827; .4845; .1642; .0127
(*e*) .2173; .5155; .8358; .9873
(*f*) yes; Type I
(*g*) no; Type II

29. .0065; .0003; 0; 0; yes

Section 8.4

31. (*a*) H_0: $\mu \leq .05$
H_1: $\mu > .05$
(*b*) we shall assume that the percentage titanium exceeds 5% when, in fact, it does not; we shall be unable to detect a situation in which the percentage titanium exceeds 5%
(*c*) .1056; debatable, a P value of .1056 might be considered small by some and large by others

33. (*a*) H_0: $p \leq .15$
H_1: $p > .15$
(*b*) .1335; no, this probability is not unusually small; Type II

Section 8.5

35. (*a*) −1.711
(*b*) −1.282
(*c*) 2.093
(*d*) 2.602
(*e*) ± 1.729
(*f*) ± 2.045

37. (*a*) H_0: $\mu = .12$
H_1: $\mu > .12$
(*b*) 2.462
(*c*) yes, $t = 2.738$
(*d*) that X is at least approximately normal

39. (*a*) H_0: $\mu = 2.5$
H_1: $\mu < 2.5$
(*b*) $t = -3.5$; $.001 < P < .005$; yes, P seems to be small; at least approximate normality
(*c*) conclude that the mean noise level is below 2.5 db; we shall assume that the new product reduces noise when, in fact, it does not

41. (*a*) H_0: $\mu = 4.8$
H_1: $\mu < 4.8$
(*b*) $t = -2.828$; $.001 < P < .005$; yes

43. (*a*) H_1: $\mu < 5$
(*b*) $t = -3.47$; $.001 < P < .005$; reject H_0
(*c*) yes, because $P < .05$
(*d*) probably not since $\bar{x} = 4.28$; there is still a large accumulation there

Section 8.6

47. (*a*) unable to reject H_0
(*b*) $t = 1.154$, critical point $= \pm 2.145$; unable to reject H_0
(*c*) $\chi^2 = 17.81$, critical point $= 23.7$, unable to reject H_0
49. (*a*) unable to reject H_0
(*b*) $t = 1.22$, critical point $= 1.729$, unable to reject H_0
(*c*) $\chi^2 = 9.297$, critical point $= 11.7$, reject H_0; yes

Section 8.7

51. (*a*) yes, .0037
(*b*) yes, .0207
(*c*) yes, .0207
(*d*) yes, .0107
(*e*) no, .0547
(*f*) yes, .0074
(*g*) yes, .0352
53. no, $P = .3770$
55. (*a*) 110/4
(*b*) yes, $|W_-| = 8.5$, critical point $= 11$
57. (*a*) 100(101)/4
(*b*) 100(101)(201)/24
(*c*) H_0: $M = 2$
H_1: $M > 2$
(*d*) $P \doteq .0007$, yes

Chapter 8 review exercises

58. (*a*) 44
(*b*) no
(*c*) 3.10; .1213; .348, no
(*d*) $\chi^2_{.95} \doteq 33.65$; [.08998, .1766]; [.298, .418]
$\chi^2_{.05} \doteq 66.05$
(*e*) $3.10 \pm .082$
59. (*a*) H_0: $\mu = 3$
H_1: $\mu < .3$
(*c*) 13
(*d*) $t_{.95} = -1.753$; $t \doteq -3.71$; reject H_0; yes, the product should be marketed
60. (*a*) H_0: $p \leq .5$
H_1: $p > .5$
(*b*) $C = \{15, 16, 17, 18, 19, 20\}$
(*c*) no; Type II
(*d*) .8744; .5836; .1958; .0113
(*e*) .1256; .4164; .8042; .9887
61. yes; $t = -2.69$; $.005 < P < .01$
62. $.10 < P < .25$
63. (*a*)

−.3	0
−.2	0 0 0 0
−.1	0 0 0 0 0 0 0 0
0	0 0 0 0 0 0 0 0 0 0 0
.1	0 0 0 0 0 0
.2	0 0 0 0
.3	0
.4	
.5	
.6	
.7	
.8	0

yes; .8 looks like an outlier
(*b*) $\tilde{x} = 0$ $\quad f_1 = -.4$
$q_1 = -.1$ $\quad f_3 = .4$
$q_3 = .1$ $\quad F_1 = -.7$
$iqr = .2$ $\quad F_3 = .7$
.8 is an extreme outlier
(*c*) $\bar{x} = -.008$; $s = .198$; $-.008 \pm .066$
64. (*a*) the number of trials is at most 3000; in a geometric setting there is no "a priori" number of trials
(*b*) 20
(*c*) exceed
(*d*) 59; $P[X \geq 59] = 0.51$; see Sec. 3.4
(*e*) no; Type II; we think that the system will not crash when, in fact, it will
(*f*) reject H_0 and conclude that the system will crash; Type I; we shall stop the system unnecessarily
65. (*b*) $8.1 \pm .8$
(*c*) [4.36, 10.56]
(*d*) 3.3 ± 10.7; [.80, 3.17]; [.90, 1.8]
(*e*) no
(*f*) no
66. (*a*) H_1: hollow arrows are faster than those made of solid aluminum
H_1: $\mu < 0$
(*b*) H_0: hollow arrows are no faster than those made of solid aluminum
H_0: $\mu = 0$
(*c*) $x = -20.41$; $s = .89$; negative sign means that the time with the hollow arrow was better than that with the solid arrow
(*d*) $t = -102.5$; $P < .0005$

Section 9.1

1. (*a*) .9
(*b*) $.9 \pm .069$
(*c*) 609
3. (*a*) $.6 \pm .025$
(*b*) $.9 \pm .019$
5. 1068
7. 6766

(*d*) yes

(*e*) 3

47. (*a*) $\mu_{Y|x_1,x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

(*b*) $$X = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & -1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & 1 & -1 \\ \cdot & 1 & -1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & -1 & 1 \\ \cdot & -1 & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & 1 & 1 \\ \cdot & 1 & 1 \\ \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot \\ 1 & \cdot & \cdot \end{bmatrix}$$

(*c*) $$X'X = \begin{bmatrix} 2^2n & 0 & 0 \\ 0 & 2^2n & 0 \\ 0 & 0 & 2^2n \end{bmatrix}$$

(*d*) $$X'Y = \begin{bmatrix} (1) + a + b + ab \\ -(1) + a - b + ab \\ -(1) - a + b + ab \end{bmatrix}$$

(*e*) $$\hat{\beta} = \begin{bmatrix} [(1) + a + b + ab]/2^2n \\ [-(1) + a - b + ab]/2^2n \\ [-(1) - a + b + ab]/2^2n \end{bmatrix}$$

(*f*) $\text{SSR} = [-(1) + a - b + ab]^2/2^2n + [-(1) - a + b + ab]^2/2^2n$

(*g*) $\text{SSE}_r = S_{yy} - [-(1) + a - b + ab]^2/2^2n - [-(1) - a + b + ab]^2/2^2n$

$\text{SSE}_f = S_{yy} - [-(1) + a - b + ab]^2/2^2n - [-(1) - a + b + ab]^2/2^2n - [(1) - a - b + ab]^2/2^2n$

(*h*) $\text{SSE}_r - SS_f = [(1) - a - b + ab]^2/2^2n$

48. (*a*) $Y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + E_{ijkl}$

(*b*)

Source of variation	Degrees of freedom
A	4
B	3
C	2
AB	12
AC	8
BC	6
Error	24
Total	59

49. (*a*) (1), *ab*, *ac*, *ad*, *ae*, *bc*, *bd*, *be*, *cd*, *ce*, *de*, *abcd*, *abce*, *abde*, *acde*, *bcde*

(*b*)

Source of variation	Degrees of freedom
A	1
B	1
C	1
D	1
E	1
Error	10
Total	15

50. (*a*) $B \equiv AC \equiv BCD \equiv AD$
$C \equiv AB \equiv D \equiv ABCD$
$D \equiv ABCD \equiv C \equiv AB$

(*b*) no, the main effect *C* is aliased with the main effect *D*

Section 15.1

1. 60, 45, 30, 15; yes, these seem to differ quite a bit from the expected numbers

3. 10; questionable, the observed values are not drastically different from those expected

Section 15.2

5. $\chi^2 = 22.66$; reject H_0, $P < .005$ based on the X_4^2 distribution

Section 15.3

9. $\chi^2 = 4.57$; reject H_0, $.025 < P < .05$ based on X_1^2 distribution

11. $\chi^2 = 8.84$; reject H_0, $.025 < P < .05$ based on X_3^2 distribution

Section 15.4

13. $\chi^2 = 3.95$; reject H_0, critical point = 3.84

15. H_0: $p_{11} = p_{21} = p_{31}$
$p_{12} = p_{22} = p_{32}$
$p_{13} = p_{23} = p_{33}$
$p_{14} = p_{24} = p_{34}$

$\chi^2 = 14.72$; reject H_0, $.01 < P < .025$ based on the X_6^2 distribution