experiment, with success on a trial corresponding to buying an extended warranty.]

b. Calculate $P(X = Y)$.

c. Determine the joint pmf of X and Y and then the marginal pmf of Y.

7. The joint probability distribution of the number X of cars and the number Y of buses per signal cycle at a proposed left-turn lane is displayed in the accompanying joint probability table.

$p(x, y)$		y = 0	1	2
	0	.025	.015	.010
	1	.050	.030	.020
	2	.125	.075	.050
x	3	.150	.090	.060
	4	.100	.060	.040
	5	.050	.030	.020

a. What is the probability that there is exactly one car and exactly one bus during a cycle?

b. What is the probability that there is at most one car and at most one bus during a cycle?

c. What is the probability that there is exactly one car during a cycle? Exactly one bus?

d. Suppose the left-turn lane is to have a capacity of five cars, and that one bus is equivalent to three cars. What is the probability of an overflow during a cycle?

e. Are X and Y independent rv's? Explain.

8. A stockroom currently has 30 components of a certain type, of which 8 were provided by supplier 1, 10 by supplier 2, and 12 by supplier 3. Six of these are to be randomly selected for a particular assembly. Let X = the number of supplier 1's components selected, Y = the number of supplier 2's components selected, and $p(x, y)$ denote the joint pmf of X and Y.

a. What is $p(3, 2)$? [*Hint*: Each sample of size 6 is equally likely to be selected. Therefore, $p(3, 2)$ = (number of outcomes with $X = 3$ and $Y = 2$)/(total number of outcomes). Now use the product rule for counting to obtain the numerator and denominator.]

b. Using the logic of part (a), obtain $p(x, y)$. (This can be thought of as a multivariate hypergeometric distribution—sampling without replacement from a finite population consisting of more than two categories.)

9. Each front tire on a particular type of vehicle is supposed to be filled to a pressure of 26 psi. Suppose the actual air pressure in each tire is a random variable—X for the right tire and Y for the left tire, with joint pdf

$$f(x, y) = \begin{cases} K(x^2 + y^2) & 20 \le x \le 30,\ 20 \le y \le 30 \\ 0 & \text{otherwise} \end{cases}$$

a. What is the value of K?

b. What is the probability that both tires are underfilled?

c. What is the probability that the difference in air pressure between the two tires is at most 2 psi?

d. Determine the (marginal) distribution of air pressure in the right tire alone.

e. Are X and Y independent rv's?

10. Annie and Alvie have agreed to meet between 5:00 P.M. and 6:00 P.M. for dinner at a local health-food restaurant. Let X = Annie's arrival time and Y = Alvie's arrival time. Suppose X and Y are independent with each uniformly distributed on the interval [5, 6].

a. What is the joint pdf of X and Y?

b. What is the probability that they both arrive between 5:15 and 5:45?

c. If the first one to arrive will wait only 10 min before leaving to eat elsewhere, what is the probability that they have dinner at the health-food restaurant? [*Hint*: The event of interest is $A = \{(x, y): |x - y| \le 1/6\}$.]

11. Two different professors have just submitted final exams for duplication. Let X denote the number of typographical errors on the first professor's exam and Y denote the number of such errors on the second exam. Suppose X has a Poisson distribution with parameter μ_1, Y has a Poisson distribution with parameter μ_2, and X and Y are independent.

a. What is the joint pmf of X and Y?

b. What is the probability that at most one error is made on both exams combined?

c. Obtain a general expression for the probability that the total number of errors in the two exams is m (where m is a nonnegative integer). [*Hint*: $A = \{(x, y): x + y = m\} = \{(m, 0), (m - 1, 1), \ldots, (1, m - 1), (0, m)\}$. Now sum the joint pmf over $(x, y) \in A$ and use the binomial theorem, which says that

$$\sum_{k=0}^{m} \binom{m}{k} a^k b^{m-k} = (a + b)^m$$

for any a, b.]

12. Two components of a minicomputer have the following joint pdf for their useful lifetimes X and Y:

$$f(x, y) = \begin{cases} xe^{-x(1+y)} & x \ge 0 \text{ and } y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

a. What is the probability that the lifetime X of the first component exceeds 3?

b. What are the marginal pdf's of X and Y? Are the two lifetimes independent? Explain.

c. What is the probability that the lifetime of at least one component exceeds 3?

13. You have two lightbulbs for a particular lamp. Let X = the lifetime of the first bulb and Y = the lifetime of the second bulb (both in 1000s of hours). Suppose that X and Y are independent and that each has an exponential distribution with parameter $\lambda = 1$.

a. What is the joint pdf of X and Y?

b. What is the probability that each bulb lasts at most 1000 hours (i.e., $X \le 1$ and $Y \le 1$)?

It can also be shown that the conditional distribution of Y given that $X = x$ is normal. This can be seen geometrically by slicing the density surface with a plane perpendicular to the (x, y) passing through the value x on that axis; the result is a normal curve sketched out on the slicing plane. The conditional mean value is $\mu_{Y \cdot x} = (\mu_2 - \rho\mu_1\sigma_2/\sigma_1) + \rho\sigma_2 x/\sigma_1$, a linear function of x, and the conditional variance is $\sigma^2_{Y \cdot x} = (1 - \rho^2)\sigma^2_2$. The closer the correlation coefficient is to 1 or -1, the less variability there is in the conditional distribution. Analogous results hold for the conditional distribution of X given that $Y = y$.

The bivariate normal distribution can be generalized to the *multivariate normal distribution*. Its density function is quite complicated, and the only way to write it compactly is to employ matrix notation. If a collection of variables has this distribution, then the marginal distribution of any single variable is normal, the conditional distribution of any single variable given values of the other variables is normal, the joint marginal distribution of any pair of variables is bivariate normal, and the joint marginal distribution of any subset of three or more of the variables is again multivariate normal.

EXERCISES **Section 5.2 (22–36)**

22. An instructor has given a short quiz consisting of two parts. For a randomly selected student, let $X =$ the number of points earned on the first part and $Y =$ the number of points earned on the second part. Suppose that the joint pmf of X and Y is given in the accompanying table.

$p(x, y)$		y: 0	5	10	15
x	0	.02	.06	.02	.10
	5	.04	.15	.20	.10
	10	.01	.15	.14	.01

a. If the score recorded in the grade book is the total number of points earned on the two parts, what is the expected recorded score $E(X + Y)$?

b. If the maximum of the two scores is recorded, what is the expected recorded score?

23. The difference between the number of customers in line at the express checkout and the number in line at the super-express checkout in Exercise 3 is $X_1 - X_2$. Calculate the expected difference.

24. Six individuals, including A and B, take seats around a circular table in a completely random fashion. Suppose the seats are numbered 1, . . . , 6. Let $X =$ A's seat number and $Y =$ B's seat number. If A sends a written message around the table to B in the direction in which they are closest, how many individuals (including A and B) would you expect to handle the message?

25. A surveyor wishes to lay out a square region with each side having length L. However, because of a measurement error, he instead lays out a rectangle in which the north–south sides both have length X and the east–west sides both have length Y. Suppose that X and Y are independent and that each is uniformly distributed on the interval $[L - A, L + A]$ (where $0 < A < L$). What is the expected area of the resulting rectangle?

26. Consider a small ferry that can accommodate cars and buses. The toll for cars is \$3, and the toll for buses is \$10. Let X and Y denote the number of cars and buses, respectively, carried on a single trip. Suppose the joint distribution of X and Y is as given in the table of Exercise 7. Compute the expected revenue from a single trip.

27. Annie and Alvie have agreed to meet for lunch between noon (0:00 P.M.) and 1:00 P.M. Denote Annie's arrival time by X, Alvie's by Y, and suppose X and Y are independent with pdf's

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

$$f_Y(y) = \begin{cases} 2y & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the expected amount of time that the one who arrives first must wait for the other person? [*Hint*: $h(X, Y) = |X - Y|$.]

28. Show that if X and Y are independent rv's, then $E(XY) = E(X) \cdot E(Y)$. Then apply this in Exercise 25. [*Hint*: Consider the continuous case with $f(x, y) = f_X(x) \cdot f_Y(y)$.]

29. Compute the correlation coefficient ρ for X and Y of Example 5.16 (the covariance has already been computed).

30. **a.** Compute the covariance for X and Y in Exercise 22.
b. Compute ρ for X and Y in the same exercise.

31. **a.** Compute the covariance between X and Y in Exercise 9.
b. Compute the correlation coefficient ρ for this X and Y.

32. Reconsider the minicomputer component lifetimes X and Y as described in Exercise 12. Determine $E(XY)$. What can be said about $\text{Cov}(X, Y)$ and ρ?

33. Use the result of Exercise 28 to show that when X and Y are independent, $\text{Cov}(X, Y) = \text{Corr}(X, Y) = 0$.

34. **a.** Recalling the definition of σ^2 for a single rv X, write a formula that would be appropriate for computing the variance of a function $h(X, Y)$ of two random variables. [*Hint*: Remember that variance is just a special expected value.]
b. Use this formula to compute the variance of the recorded score $h(X, Y)$ [$= \max(X, Y)$] in part (b) of Exercise 22.

35. **a.** Use the rules of expected value to show that $\text{Cov}(aX + b, cY + d) = ac\,\text{Cov}(X, Y)$.
b. Use part (a) along with the rules of variance and standard deviation to show that $\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y)$ when a and c have the same sign.
c. What happens if a and c have opposite signs?

36. Show that if $Y = aX + b$ $(a \neq 0)$, then $\text{Corr}(X, Y) = +1$ or -1. Under what conditions will $\rho = +1$?

5.3 Statistics and Their Distributions

The observations in a single sample were denoted in Chapter 1 by $x_1, x_2, \ldots, x_n$. Consider selecting two different samples of size n from the same population distribution. The x_i's in the second sample will virtually always differ at least a bit from those in the first sample. For example, a first sample of $n = 3$ cars of a particular type might result in fuel efficiencies $x_1 = 30.7$, $x_2 = 29.4$, $x_3 = 31.1$, whereas a second sample may give $x_1 = 28.8$, $x_2 = 30.0$, and $x_3 = 32.5$. Before we obtain data, there is uncertainty about the value of each x_i. Because of this uncertainty, *before* the data becomes available we now regard each observation as a random variable and denote the sample by $X_1, X_2, \ldots, X_n$ (uppercase letters for random variables).

This variation in observed values in turn implies that the value of any function of the sample observations—such as the sample mean, sample standard deviation, or sample fourth spread—also varies from sample to sample. That is, prior to obtaining $x_1, \ldots, x_n$, there is uncertainty as to the value of $\bar{x}$, the value of s, and so on.

EXAMPLE 5.20 Suppose that material strength for a randomly selected specimen of a particular type has a Weibull distribution with parameter values $\alpha = 2$ (shape) and $\beta = 5$ (scale). The corresponding density curve is shown in Figure 5.7. Formulas from Section 4.5 give

$$\mu = E(X) = 4.4311 \quad \tilde{\mu} = 4.1628 \quad \sigma^2 = V(X) = 5.365 \quad \sigma = 2.316$$

The mean exceeds the median because of the distribution's positive skew.

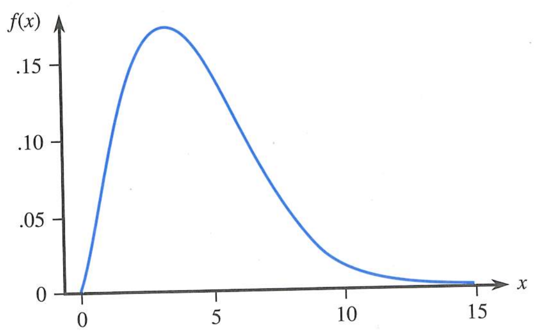


Figure 5.7 The Weibull density curve for Example 5.20

Unlike the normal case, these histograms all differ in shape. In particular, they become progressively less skewed as the sample size n increases. The average of the 500 $\bar{x}$ values for the four different sample sizes are all quite close to the mean value of the population distribution. If each histogram had been based on an unending sequence of $\bar{x}$ values rather than just 500, all four would have been centered at exactly 21.7584. Thus different values of n change the shape but not the center of the sampling distribution of $\bar{X}$. Comparison of the four histograms in Figure 5.14 also shows that as n increases, the spread of the histograms decreases. Increasing n results in a greater degree of concentration about the population mean value and makes the histogram look more like a normal curve. The histogram of Figure 5.14(d) and the normal probability plot in Figure 5.14(e) provide convincing evidence that a sample size of $n = 30$ is sufficient to overcome the skewness of the population distribution and give an approximately normal $\bar{X}$ sampling distribution. ■

EXERCISES Section 5.3 (37–45)

37. A particular brand of dishwasher soap is sold in three sizes: 25 oz, 40 oz, and 65 oz. Twenty percent of all purchasers select a 25-oz box, 50% select a 40-oz box, and the remaining 30% choose a 65-oz box. Let X_1 and X_2 denote the package sizes selected by two independently selected purchasers.

a. Determine the sampling distribution of $\bar{X}$, calculate $E(\bar{X})$, and compare to μ.

b. Determine the sampling distribution of the sample variance S^2, calculate $E(S^2)$, and compare to σ^2.

38. There are two traffic lights on a commuter's route to and from work. Let X_1 be the number of lights at which the commuter must stop on his way to work, and X_2 be the number of lights at which he must stop when returning from work. Suppose these two variables are independent, each with pmf given in the accompanying table (so X_1, X_2 is a random sample of size $n = 2$).

x_1	0	1	2
$p(x_1)$	.2	.5	.3

$\mu = 1.1, \sigma^2 = .49$

a. Determine the pmf of $T_o = X_1 + X_2$.

b. Calculate μ_{T_o}. How does it relate to μ, the population mean?

c. Calculate $\sigma^2_{T_o}$. How does it relate to σ^2, the population variance?

d. Let X_3 and X_4 be the number of lights at which a stop is required when driving to and from work on a second day assumed independent of the first day. With T_o = the sum of all four X_i's, what now are the values of $E(T_o)$ and $V(T_o)$?

e. Referring back to (d), what are the values of $P(T_o = 8)$ and $P(T_o \geq 7)$ [*Hint*: Don't even think of listing all possible outcomes!]

39. It is known that 80% of all brand A external hard drives work in a satisfactory manner throughout the warranty period (are "successes"). Suppose that $n = 15$ drives are randomly selected. Let X = the number of successes in the sample. The statistic X/n is the sample proportion (fraction) of successes. Obtain the sampling distribution of this statistic. [*Hint*: One possible value of X/n is .2, corresponding to $X = 3$. What is the probability of this value (what kind of rv is X)?]

40. A box contains ten sealed envelopes numbered 1, . . . , 10. The first five contain no money, the next three each contains $5, and there is a $10 bill in each of the last two. A sample of size 3 is selected *with* replacement (so we have a random sample), and you get the largest amount in any of the envelopes selected. If X_1, X_2, and X_3 denote the amounts in the selected envelopes, the statistic of interest is M = the maximum of X_1, X_2, and X_3.

a. Obtain the probability distribution of this statistic.

b. Describe how you would carry out a simulation experiment to compare the distributions of M for various sample sizes. How would you guess the distribution would change as n increases?

41. Let X be the number of packages being mailed by a randomly selected customer at a certain shipping facility. Suppose the distribution of X is as follows:

x	1	2	3	4
$p(x)$	.4	.3	.2	.1

a. Consider a random sample of size $n = 2$ (two customers), and let $\bar{X}$ be the sample mean number of packages shipped. Obtain the probability distribution of $\bar{X}$.

b. Refer to part (a) and calculate $P(\bar{X} \leq 2.5)$.

c. Again consider a random sample of size $n = 2$, but now focus on the statistic R = the sample range (difference between the largest and smallest values in the sample). Obtain the distribution of R. [*Hint*: Calculate

PROPOSITION

> Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution for which only positive values are possible $[P(X_i > 0) = 1]$. Then if n is sufficiently large, the product $Y = X_1X_2 \cdot \cdots \cdot X_n$ has approximately a lognormal distribution.

To verify this, note that

$$\ln(Y) = \ln(X_1) + \ln(X_2) + \cdots + \ln(X_n)$$

Since $\ln(Y)$ is a sum of independent and identically distributed rv's [the $\ln(X_i)$s], it is approximately normal when n is large, so Y itself has approximately a lognormal distribution. As an example of the applicability of this result, Bury (*Statistical Models in Applied Science,* Wiley, p. 590) argues that the damage process in plastic flow and crack propagation is a multiplicative process, so that variables such as percentage elongation and rupture strength have approximately lognormal distributions.

EXERCISES Section 5.4 (46–57)

46. The inside diameter of a randomly selected piston ring is a random variable with mean value 12 cm and standard deviation .04 cm.

a. If $\overline{X}$ is the sample mean diameter for a random sample of $n = 16$ rings, where is the sampling distribution of $\overline{X}$ centered, and what is the standard deviation of the $\overline{X}$ distribution?

b. Answer the questions posed in part (a) for a sample size of $n = 64$ rings.

c. For which of the two random samples, the one of part (a) or the one of part (b), is $\overline{X}$ more likely to be within .01 cm of 12 cm? Explain your reasoning.

47. Refer to Exercise 46. Suppose the distribution of diameter is normal.

a. Calculate $P(11.99 \le \overline{X} \le 12.01)$ when $n = 16$.

b. How likely is it that the sample mean diameter exceeds 12.01 when $n = 25$?

48. The National Health Statistics Reports dated Oct. 22, 2008, stated that for a sample size of 277 18-year-old American males, the sample mean waist circumference was 86.3 cm. A somewhat complicated method was used to *estimate* various population percentiles, resulting in the following values:

5th	10th	25th	50th	75th	90th	95th
69.6	70.9	75.2	81.3	95.4	107.1	116.4

a. Is it plausible that the waist size distribution is at least approximately normal? Explain your reasoning. If your answer is no, conjecture the shape of the population distribution.

b. Suppose that the population mean waist size is 85 cm and that the population standard deviation is 15 cm. How likely is it that a random sample of 277 individuals will result in a sample mean waist size of at least 86.3 cm?

c. Referring back to (b), suppose now that the population mean waist size in 82 cm. Now what is the (approximate) probability that the sample mean will be at least 86.3 cm? In light of this calculation, do you think that 82 cm is a reasonable value for μ?

49. There are 40 students in an elementary statistics class. On the basis of years of experience, the instructor knows that the time needed to grade a randomly chosen first examination paper is a random variable with an expected value of 6 min and a standard deviation of 6 min.

a. If grading times are independent and the instructor begins grading at 6:50 P.M. and grades continuously, what is the (approximate) probability that he is through grading before the 11:00 P.M. TV news begins?

b. If the sports report begins at 11:10, what is the probability that he misses part of the report if he waits until grading is done before turning on the TV?

50. The breaking strength of a rivet has a mean value of 10,000 psi and a standard deviation of 500 psi.

a. What is the probability that the sample mean breaking strength for a random sample of 40 rivets is between 9900 and 10,200?

b. If the sample size had been 15 rather than 40, could the probability requested in part (a) be calculated from the given information?

51. The time taken by a randomly selected applicant for a mortgage to fill out a certain form has a normal distribution with mean value 10 min and standard deviation 2 min. If five individuals fill out a form on one day and six on another, what is the probability that the sample average amount of time taken on each day is at most 11 min?

52. The lifetime of a certain type of battery is normally distributed with mean value 10 hours and standard deviation

1 hour. There are four batteries in a package. What lifetime value is such that the total lifetime of all batteries in a package exceeds that value for only 5% of all packages?

53. Rockwell hardness of pins of a certain type is known to have a mean value of 50 and a standard deviation of 1.2.

a. If the distribution is normal, what is the probability that the sample mean hardness for a random sample of 9 pins is at least 51?

b. Without assuming population normality, what is the (approximate) probability that the sample mean hardness for a random sample of 40 pins is at least 51?

54. Suppose the sediment density (g/cm) of a randomly selected specimen from a certain region is normally distributed with mean 2.65 and standard deviation .85 (suggested in "Modeling Sediment and Water Column Interactions for Hydrophobic Pollutants," *Water Research,* 1984: 1169–1174).

a. If a random sample of 25 specimens is selected, what is the probability that the sample average sediment density is at most 3.00? Between 2.65 and 3.00?

b. How large a sample size would be required to ensure that the first probability in part (a) is at least .99?

55. The number of parking tickets issued in a certain city on any given weekday has a Poisson distribution with parameter $\mu = 50$. What is the approximate probability that

a. Between 35 and 70 tickets are given out on a particular day? [*Hint*: When μ is large, a Poisson rv has approximately a normal distribution.]

b. The total number of tickets given out during a 5-day week is between 225 and 275?

56. A binary communication channel transmits a sequence of "bits" (0s and 1s). Suppose that for any particular bit transmitted, there is a 10% chance of a transmission error (a 0 becoming a 1 or a 1 becoming a 0). Assume that bit errors occur independently of one another.

a. Consider transmitting 1000 bits. What is the approximate probability that at most 125 transmission errors occur?

b. Suppose the same 1000-bit message is sent two different times independently of one another. What is the approximate probability that the number of errors in the first transmission is within 50 of the number of errors in the second?

57. Suppose the distribution of the time X (in hours) spent by students at a certain university on a particular project is gamma with parameters $\alpha = 50$ and $\beta = 2$. Because α is large, it can be shown that X has approximately a normal distribution. Use this fact to compute the approximate probability that a randomly selected student spends at most 125 hours on the project.

5.5 The Distribution of a Linear Combination

The sample mean $\overline{X}$ and sample total T_o are special cases of a type of random variable that arises very frequently in statistical applications.

DEFINITION

Given a collection of n random variables $X_1, \ldots, X_n$ and n numerical constants $a_1, \ldots, a_n$, the rv

$$Y = a_1X_1 + \cdots + a_nX_n = \sum_{i=1}^{n} a_iX_i \tag{5.7}$$

is called a **linear combination** of the X_i's.

For example, $4X_1 - 5X_2 + 8X_3$ is a linear combination of X_1, X_2, and X_3 with $a_1 = 4$, $a_2 = -5$, and $a_3 = 8$.

Taking $a_1 = a_2 = \cdots = a_n = 1$ gives $Y = X_1 + \cdots + X_n = T_o$, and $a_1 = a_2 = \cdots = a_n = \frac{1}{n}$ yields

$$Y = \frac{1}{n}X_1 + \cdots + \frac{1}{n}X_n = \frac{1}{n}(X_1 + \cdots + X_n) = \frac{1}{n}T_o = \overline{X}$$

Notice that we are not requiring the X_i's to be independent or identically distributed. All the X_i's could have different distributions and therefore different mean values and variances. We first consider the expected value and variance of a linear combination.

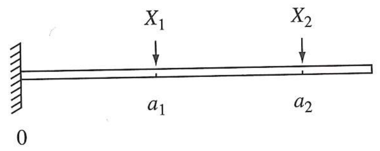
d. What are the expected value and variance of the difference between total morning waiting time and total evening waiting time for a particular week?

65. Suppose that when the pH of a certain chemical compound is 5.00, the pH measured by a randomly selected beginning chemistry student is a random variable with mean 5.00 and standard deviation .2. A large batch of the compound is subdivided and a sample given to each student in a morning lab and each student in an afternoon lab. Let $\overline{X}$ = the average pH as determined by the morning students and $\overline{Y}$ = the average pH as determined by the afternoon students.

a. If pH is a normal variable and there are 25 students in each lab, compute $P(-.1 \le \overline{X} - \overline{Y} \le .1)$. [*Hint*: $\overline{X} - \overline{Y}$ is a linear combination of normal variables, so is normally distributed. Compute $\mu_{\overline{X}-\overline{Y}}$ and $\sigma_{\overline{X}-\overline{Y}}$.]

b. If there are 36 students in each lab, but pH determinations are not assumed normal, calculate (approximately) $P(-.1 \le \overline{X} - \overline{Y} \le .1)$.

66. If two loads are applied to a cantilever beam as shown in the accompanying drawing, the bending moment at 0 due to the loads is $a_1X_1 + a_2X_2$.



a. Suppose that X_1 and X_2 are independent rv's with means 2 and 4 kips, respectively, and standard deviations .5 and 1.0 kip, respectively. If $a_1 = 5$ ft and $a_2 = 10$ ft, what is the expected bending moment and what is the standard deviation of the bending moment?

b. If X_1 and X_2 are normally distributed, what is the probability that the bending moment will exceed 75 kip-ft?

c. Suppose the positions of the two loads are random variables. Denoting them by A_1 and A_2, assume that these variables have means of 5 and 10 ft, respectively, that each has a standard deviation of .5, and that all A_i's and X_i's are independent of one another. What is the expected moment now?

d. For the situation of part (c), what is the variance of the bending moment?

e. If the situation is as described in part (a) except that $\text{Corr}(X_1, X_2) = .5$ (so that the two loads are not independent), what is the variance of the bending moment?

67. One piece of PVC pipe is to be inserted inside another piece. The length of the first piece is normally distributed with mean value 20 in. and standard deviation .5 in. The length of the second piece is a normal rv with mean and standard deviation 15 in. and .4 in., respectively. The amount of overlap is normally distributed with mean value 1 in. and standard deviation .1 in. Assuming that the lengths and amount of overlap are independent of one another, what is the probability that the total length after insertion is between 34.5 in. and 35 in.?

68. Two airplanes are flying in the same direction in adjacent parallel corridors. At time $t = 0$, the first airplane is 10 km ahead of the second one. Suppose the speed of the first plane (km/hr) is normally distributed with mean 520 and standard deviation 10 and the second plane's speed is also normally distributed with mean and standard deviation 500 and 10, respectively.

a. What is the probability that after 2 hr of flying, the second plane has not caught up to the first plane?

b. Determine the probability that the planes are separated by at most 10 km after 2 hr.

69. Three different roads feed into a particular freeway entrance. Suppose that during a fixed time period, the number of cars coming from each road onto the freeway is a random variable, with expected value and standard deviation as given in the table.

	Road 1	Road 2	Road 3
Expected value	800	1000	600
Standard deviation	16	25	18

a. What is the expected total number of cars entering the freeway at this point during the period? [*Hint*: Let X_i = the number from road *i*.]

b. What is the variance of the total number of entering cars? Have you made any assumptions about the relationship between the numbers of cars on the different roads?

c. With X_i denoting the number of cars entering from road *i* during the period, suppose that $\text{Cov}(X_1, X_2) = 80$, $\text{Cov}(X_1, X_3) = 90$, and $\text{Cov}(X_2, X_3) = 100$ (so that the three streams of traffic are not independent). Compute the expected total number of entering cars and the standard deviation of the total.

70. Consider a random sample of size *n* from a continuous distribution having median 0 so that the probability of any one observation being positive is .5. Disregarding the signs of the observations, rank them from smallest to largest in absolute value, and let W = the sum of the ranks of the observations having positive signs. For example, if the observations are $-.3$, $+.7$, $+2.1$, and -2.5, then the ranks of positive observations are 2 and 3, so $W = 5$. In Chapter 15, W will be called *Wilcoxon's signed-rank statistic*. W can be represented as follows:

$$W = 1 \cdot Y_1 + 2 \cdot Y_2 + 3 \cdot Y_3 + \cdots + n \cdot Y_n = \sum_{i=1}^{n} i \cdot Y_i$$

where the Y_i's are independent Bernoulli rv's, each with $p = .5$ ($Y_i = 1$ corresponds to the observation with rank *i* being positive).

a. Determine $E(Y_i)$ and then $E(W)$ using the equation for W. [*Hint*: The first n positive integers sum to $n(n+1)/2$.]

b. Determine $V(Y_i)$ and then $V(W)$. [*Hint*: The sum of the squares of the first n positive integers can be expressed as $n(n+1)(2n+1)/6$.]

71. In Exercise 66, the weight of the beam itself contributes to the bending moment. Assume that the beam is of uniform thickness and density so that the resulting load is uniformly distributed on the beam. If the weight of the beam is random, the resulting load from the weight is also random; denote this load by W (kip-ft).

a. If the beam is 12 ft long, W has mean 1.5 and standard deviation .25, and the fixed loads are as described in part (a) of Exercise 66, what are the expected value and variance of the bending moment? [*Hint*: If the load due to the beam were w kip-ft, the contribution to the bending moment would be $w\int_0^{12} x\,dx$.]

b. If all three variables (X_1, X_2, and W) are normally distributed, what is the probability that the bending moment will be at most 200 kip-ft?

72. I have three errands to take care of in the Administration Building. Let X_i = the time that it takes for the ith errand ($i = 1, 2, 3$), and let X_4 = the total time in minutes that I spend walking to and from the building and between each errand. Suppose the X_i's are independent, and normally distributed, with the following means and standard deviations: $\mu_1 = 15$, $\sigma_1 = 4$, $\mu_2 = 5$, $\sigma_2 = 1$, $\mu_3 = 8$, $\sigma_3 = 2$, $\mu_4 = 12$, $\sigma_4 = 3$. I plan to leave my office at precisely 10:00 A.M. and wish to post a note on my door that reads, "I will return by t A.M." What time t should I write down if I want the probability of my arriving after t to be .01?

73. Suppose the expected tensile strength of type-A steel is 105 ksi and the standard deviation of tensile strength is 8 ksi. For type-B steel, suppose the expected tensile strength and standard deviation of tensile strength are 100 ksi and 6 ksi, respectively. Let $\overline{X}$ = the sample average tensile strength of a random sample of 40 type-A specimens, and let $\overline{Y}$ = the sample average tensile strength of a random sample of 35 type-B specimens.

a. What is the approximate distribution of $\overline{X}$? Of $\overline{Y}$?

b. What is the approximate distribution of $\overline{X} - \overline{Y}$? Justify your answer.

c. Calculate (approximately) $P(-1 \le \overline{X} - \overline{Y} \le 1)$.

d. Calculate $P(\overline{X} - \overline{Y} \ge 10)$. If you actually observed $\overline{X} - \overline{Y} \ge 10$, would you doubt that $\mu_1 - \mu_2 = 5$?

74. In an area having sandy soil, 50 small trees of a certain type were planted, and another 50 trees were planted in an area having clay soil. Let X = the number of trees planted in sandy soil that survive 1 year and Y = the number of trees planted in clay soil that survive 1 year. If the probability that a tree planted in sandy soil will survive 1 year is .7 and the probability of 1-year survival in clay soil is .6, compute an approximation to $P(-5 \le X - Y \le 5)$ (do not bother with the continuity correction).

SUPPLEMENTARY EXERCISES (75–96)

75. A restaurant serves three fixed-price dinners costing \$12, \$15, and \$20. For a randomly selected couple dining at this restaurant, let X = the cost of the man's dinner and Y = the cost of the woman's dinner. The joint pmf of X and Y is given in the following table:

$p(x, y)$		y: 12	15	20
x	12	.05	.05	.10
	15	.05	.10	.35
	20	0	.20	.10

a. Compute the marginal pmf's of X and Y.

b. What is the probability that the man's and the woman's dinner cost at most \$15 each?

c. Are X and Y independent? Justify your answer.

d. What is the expected total cost of the dinner for the two people?

e. Suppose that when a couple opens fortune cookies at the conclusion of the meal, they find the message "You will receive as a refund the difference between the cost of the more expensive and the less expensive meal that you have chosen." How much would the restaurant expect to refund?

76. In cost estimation, the total cost of a project is the sum of component task costs. Each of these costs is a random variable with a probability distribution. It is customary to obtain information about the total cost distribution by adding together characteristics of the individual component cost distributions—this is called the "roll-up" procedure. For example, $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$, so the roll-up procedure is valid for mean cost. Suppose that there are two component tasks and that X_1 and X_2 are independent, normally distributed random variables. Is the roll-up procedure valid for the 75th percentile? That is, is the 75th percentile of the distribution of $X_1 + X_2$ the same as the sum of the 75th percentiles of the two individual distributions? If not, what is the relationship between the percentile of the sum and the sum of percentiles? For what percentiles is the roll-up procedure valid in this case?

Bonus problem:

Suppose that a point (X,Y) is chosen uniformly at random from the set $D = \{ (x,y) \mid x^2 + y^2 \leq 4 \}$. That is, the joint PDF of X and Y is given by

$$f(x,y) = \begin{cases} c & if\ (x,y) \in D \\ 0 & otherwise \end{cases}$$

(a) Find the constant c.

(b) Find $P(XY > 0)$, $P(X + Y > 0)$, $P(-1 \leq X \leq 1, -1 \leq Y \leq 1)$ and $P(XY > 0, X + Y < 2)$.

(c) Find the marginal supports and marginal PDFs of X and Y.

(d) Are X and Y independent?

(e) Let x be a fixed value from the support of X. What is the conditional support and conditional PDF of Y given X = x? Is this a member of a known distribution family?

(f) Find the conditional mean, standard deviation and third quartile of Y given X = x.

(g) Calculate Cov(X,Y) and $\rho_{X,Y}$. (Hint: Consider the value of $\int_{-c}^{c} f(x)dx$ for any odd function f and constant c > 0.)