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 $\overline{X}$, calculate $E(\overline{X})$, 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 zip drives work in a satisfactory manner throughout the warranty period (are "successes"). Suppose that $n = 10$ 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 .3, corresponding to $X = 3$. What is the probability of this value (what kind of random variable 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 $\overline{X}$ be the sample mean number of packages shipped. Obtain the probability distribution of $\overline{X}$.

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

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 the value of R for each outcome and use the probabilities from part (a).]

d. If a random sample of size $n = 4$ is selected, what is $P(\overline{X} \leq 1.5)$? [*Hint*: You should not have to list all possible outcomes, only those for which $\bar{x} \leq 1.5$.]

42. A company maintains three offices in a certain region, each staffed by two employees. Information concerning yearly salaries (1000s of dollars) is as follows:

Office	1	1	2	2	3	3
Employee	1	2	3	4	5	6
Salary	29.7	33.6	30.2	33.6	25.8	29.7

a. Suppose two of these employees are randomly selected from among the six (without replacement). Determine the sampling distribution of the sample mean salary $\overline{X}$.

b. Suppose one of the three offices is randomly selected. Let X_1 and X_2 denote the salaries of the two employees. Determine the sampling distribution of $\overline{X}$.

c. How does $E(\overline{X})$ from parts (a) and (b) compare to the population mean salary μ?

43. Suppose the amount of liquid dispensed by a certain machine is uniformly distributed with lower limit $A = 8$ oz and upper limit $B = 10$ oz. Describe how you would carry out simulation experiments to compare the sampling distribution of the (sample) fourth spread for sample sizes $n = 5$, 10, 20, and 30.

44. Carry out a simulation experiment using a statistical computer package or other software to study the sampling distribution of $\overline{X}$ when the population distribution is Weibull with $\alpha = 2$ and $\beta = 5$, as in Example 5.19. Consider the four sample sizes $n = 5$, 10, 20, and 30, and in each case use 1000 replications. For which of these sample sizes does the $\overline{X}$ sampling distribution appear to be approximately normal?

45. Carry out a simulation experiment using a statistical computer package or other software to study the sampling distribution of $\overline{X}$ when the population distribution is lognormal with $E(\ln(X)) = 3$ and $V(\ln(X)) = 1$. Consider the four sample sizes $n = 10$, 20, 30, and 50, and in each case use 1000 replications. For which of these sample sizes does the $\overline{X}$ sampling distribution appear to be approximately normal?

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 $Y = X_1 X_2 \cdot \cdots \cdot X_n$ 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

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
	12	.05	.05	.10
x	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?