(a) Find the mgf of the sum $Y = X_1 + X_2 + X_3$.

(b) How is Y distributed?

(c) Compute $P(3 \le Y \le 9)$.

5.4-4. Generalize Exercise 5.4-3 by showing that the sum of n independent Poisson random variables with respective means $\mu_1, \mu_2, \ldots, \mu_n$ is Poisson with mean

$$\mu_1 + \mu_2 + \cdots + \mu_n.$$

5.4-5. Let $Z_1, Z_2, \ldots, Z_7$ be a random sample from the standard normal distribution $N(0,1)$. Let $W = Z_1^2 + Z_2^2 + \cdots + Z_7^2$. Find $P(1.69 < W < 14.07)$.

5.4-6. Let X_1, X_2, X_3, X_4, X_5 be a random sample of size 5 from a geometric distribution with $p = 1/3$.

(a) Find the mgf of $Y = X_1 + X_2 + X_3 + X_4 + X_5$.

(b) How is Y distributed?

5.4-7. Let X_1, X_2, X_3 denote a random sample of size 3 from a gamma distribution with $\alpha = 7$ and $\theta = 5$.

(a) Find the mgf of $Y = X_1 + X_2 + X_3$.

(b) How is Y distributed?

5.4-8. Let $W = X_1 + X_2 + \cdots + X_h$, a sum of h mutually independent and identically distributed exponential random variables with mean θ. Show that W has a gamma distribution with parameters $\alpha = h$ and θ, respectively.

5.4-9. Let X and Y, with respective pmfs $f(x)$ and $g(y)$, be independent discrete random variables, each of whose support is a subset of the nonnegative integers $0, 1, 2, \ldots$. Show that the pmf of $W = X + Y$ is given by the **convolution formula**

$$h(w) = \sum_{x=0}^{w} f(x)g(w-x), \qquad w = 0, 1, 2, \ldots.$$

HINT: Argue that $h(w) = P(W = w)$ is the probability of the $w + 1$ mutually exclusive events $(x, y = w - x)$, $x = 0, 1, \ldots, w$.

5.4-10. Let X equal the outcome when a fair four-sided die that has its faces numbered 0, 1, 2, and 3 is rolled. Let Y equal the outcome when a fair four-sided die that has its faces numbered 0, 4, 8, and 12 is rolled.

(a) Define the mgf of X.

(b) Define the mgf of Y.

(c) Let $W = X + Y$, the sum when the pair of dice is rolled. Find the mgf of W.

(d) Give the pmf of W; that is, determine $P(W = w)$, $w = 0, 1, \ldots, 15$, from the mgf of W.

5.4-11. Let X and Y equal the outcomes when two fair six-sided dice are rolled. Let $W = X + Y$. Assuming independence, find the pmf of W when

(a) The first die has three faces numbered 0 and three faces numbered 2, and the second die has its faces numbered 0, 1, 4, 5, 8, and 9.

(b) The faces on the first die are numbered 0, 1, 2, 3, 4, and 5, and the faces on the second die are numbered 0, 6, 12, 18, 24, and 30.

5.4-12. Let X and Y be the outcomes when a pair of fair eight-sided dice is rolled. Let $W = X + Y$. How should the faces of the dice be numbered so that W has a uniform distribution on $0, 1, \ldots, 15$?

5.4-13. Let $X_1, X_2, \ldots, X_8$ be a random sample from a distribution having pmf $f(x) = (x+1)/6$, $x = 0, 1, 2$.

(a) Use Exercise 5.4-9 to find the pmf of $W_1 = X_1 + X_2$.

(b) What is the pmf of $W_2 = X_3 + X_4$?

(c) Now find the pmf of $W = W_1 + W_2 = X_1 + X_2 + X_3 + X_4$.

(d) Find the pmf of $Y = X_1 + X_2 + \cdots + X_8$.

(e) Construct probability histograms for X_1, W_1, W, and Y. Are these histograms skewed or symmetric?

5.4-14. The number of accidents in a period of one week follows a Poisson distribution with mean 2. The numbers of accidents from week to week are independent. What is the probability of exactly seven accidents in a given three weeks? HINT: See Exercise 5.4-4.

5.4-15. Given a fair four-sided die, let Y equal the number of rolls needed to observe each face at least once.

(a) Argue that $Y = X_1 + X_2 + X_3 + X_4$, where X_i has a geometric distribution with $p_i = (5-i)/4$, $i = 1, 2, 3, 4$, and X_1, X_2, X_3, X_4 are independent.

(b) Find the mean and variance of Y.

(c) Find $P(Y = y)$, $y = 4, 5, 6, 7$.

5.4-16. The number X of sick days taken during a year by an employee follows a Poisson distribution with mean 2. Let us observe four such employees. Assuming independence, compute the probability that their total number of sick days exceeds 10.

5.4-17. In a study concerning a new treatment of a certain disease, two groups of 25 participants in each were followed for five years. Those in one group took the old treatment and those in the other took the new treatment. The theoretical dropout rate for an individual was 50% in both groups over that 5-year period. Let X be the number that dropped out in the first group and Y the number in the second group. Assuming independence where needed, give the sum that equals the probability that $Y \ge X + 2$. HINT: What is the distribution of $Y - X + 25$?

5.4-18. The number of cracks on a highway averages 0.5 per mile and follows a Poisson distribution.

(a) Sketch, on the same set of axes, the graphs of the pdfs of X and of $\overline{X}$.

(b) Let S^2 be the sample variance of the nine weights. Find constants a and b so that $P(a \le S^2 \le b) = 0.90$.

HINT: Because $8S^2/0.16$ is $\chi^2(8)$ and $P(a \le S^2 \le b)$ is equivalent to $P(8a/0.16 \le 8S^2/0.16 \le 8b/0.16)$, you can find $8a/0.16$ and $8b/0.16$ in Table IV in Appendix B.

5.5-6. At a heat-treating company, iron castings and steel forgings are heat-treated to achieve desired mechanical properties and machinability. One steel forging is annealed to soften the part for each machining. Two lots of this part, made of 1020 steel, are heat-treated in two different furnaces. The specification for this part is 36-66 on the Rockwell G scale. Let X_1 and X_2 equal the respective hardness measurements for parts selected randomly from furnaces 1 and 2. Assume that the distributions of X_1 and X_2 are $N(47.88, 2.19)$ and $N(43.04, 14.89)$, respectively.

(a) Sketch the pdfs of X_1 and X_2 on the same graph.

(b) Compute $P(X_1 > X_2)$, assuming independence of X_1 and X_2.

5.5-7. Suppose that the distribution of the weight of a prepackaged "1-pound bag" of carrots is $N(1.18, 0.07^2)$ and the distribution of the weight of a prepackaged "3-pound bag" of carrots is $N(3.22, 0.09^2)$. Selecting bags at random, find the probability that the sum of three 1-pound bags exceeds the weight of one 3-pound bag. HINT: First determine the distribution of Y, the sum of the three, and then compute $P(Y > W)$, where W is the weight of the 3-pound bag.

5.5-8. Let X denote the wing length in millimeters of a male gallinule and Y the wing length in millimeters of a female gallinule. Assume that X is $N(184.09, 39.37)$ and Y is $N(171.93, 50.88)$ and that X and Y are independent. If a male and a female gallinule are captured, what is the probability that X is greater than Y?

5.5-9. Suppose that the length of life in hours (say, X) of a light bulb manufactured by company A is $N(800, 14400)$ and the length of life in hours (say, Y) of a light bulb manufactured by company B is $N(850, 2500)$. One bulb is randomly selected from each company and is burned until "death."

(a) Find the probability that the length of life of the bulb from company A exceeds the length of life of the bulb from company B by at least 15 hours.

(b) Find the probability that at least one of the bulbs "lives" for at least 920 hours.

5.5-10. A consumer buys n light bulbs, each of which has a lifetime that has a mean of 800 hours, a standard deviation of 100 hours, and a normal distribution. A light bulb is replaced by another as soon as it burns out. Assuming independence of the lifetimes, find the smallest n so that the succession of light bulbs produces light for at least 10,000 hours with a probability of 0.90.

5.5-11. A marketing research firm suggests to a company that two possible competing products can generate incomes X and Y (in millions) that are $N(3, 1)$ and $N(3.5, 4)$, respectively. Clearly, $P(X < Y) > 1/2$. However, the company would prefer the one with the smaller variance if, in fact, $P(X > 2) > P(Y > 2)$. Which product does the company select?

5.5-12. Let the independent random variables X_1 and X_2 be $N(0, 1)$ and $\chi^2(r)$, respectively. Let $Y_1 = X_1/\sqrt{X_2/r}$ and $Y_2 = X_2$.

(a) Find the joint pdf of Y_1 and Y_2.

(b) Determine the marginal pdf of Y_1 and show that Y_1 has a t distribution. (This is another, equivalent, way of finding the pdf of T.)

5.5-13. Let Z_1, Z_2, and Z_3 have independent standard normal distributions, $N(0, 1)$.

(a) Find the distribution of

$$W = \frac{Z_1}{\sqrt{(Z_2^2 + Z_3^2)/2}}.$$

(b) Show that

$$V = \frac{Z_1}{\sqrt{(Z_1^2 + Z_2^2)/2}}$$

has pdf $f(v) = 1/\left(\pi\sqrt{2 - v^2}\right)$, $-\sqrt{2} < v < \sqrt{2}$.

(c) Find the mean of V.

(d) Find the standard deviation of V.

(e) Why are the distributions of W and V so different?

5.5-14. Let T have a t distribution with r degrees of freedom. Show that $E(T) = 0$ provided that $r \ge 2$, and $\mathrm{Var}(T) = r/(r-2)$ provided that $r \ge 3$, by first finding $E(Z)$, $E(1/\sqrt{U})$, $E(Z^2)$, and $E(1/U)$.

5.5-15. Let the distribution of T be $t(17)$. Find

(a) $t_{0.01}(17)$.

(b) $t_{0.95}(17)$.

(c) $P(-1.740 \le T \le 1.740)$.

5.5-16. Let $n = 9$ in the T statistic defined in Equation 5.5-2.

(a) Find $t_{0.025}$ so that $P(-t_{0.025} \le T \le t_{0.025}) = 0.95$.

(b) Solve the inequality $[-t_{0.025} \le T \le t_{0.025}]$ so that μ is in the middle.

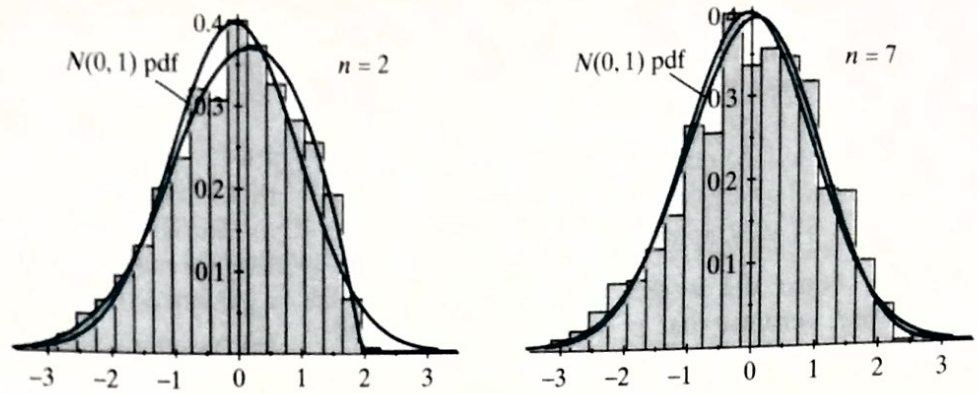

Figure 5.6-3 pdfs of $(\overline{X} - \mu)/(\sigma/\sqrt{n})$, underlying distribution triangular

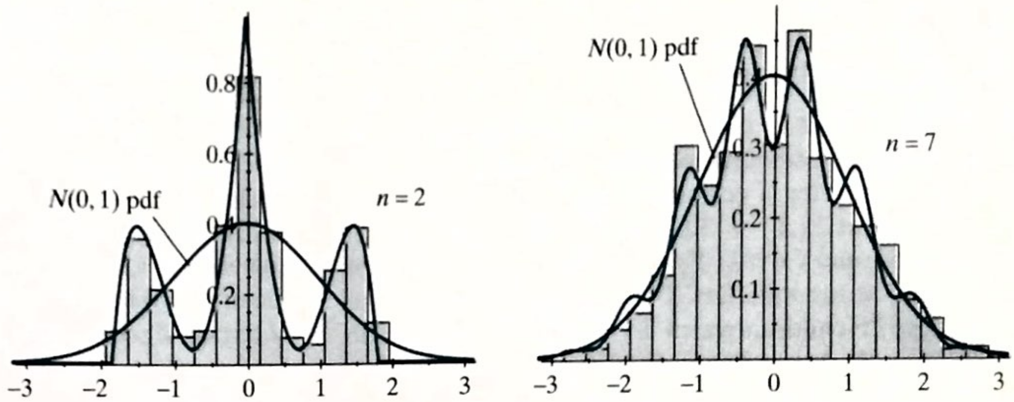

Figure 5.6-4 pdfs of $(\overline{X} - \mu)/(\sigma/\sqrt{n})$, underlying distribution U-shaped

So far, all the illustrations have concerned distributions of the continuous type. However, the hypotheses for the central limit theorem do not require the distribution to be continuous. We shall consider applications of the central limit theorem for discrete-type distributions in the next section.

Exercises

5.6-1. Let $\overline{X}$ be the mean of a random sample of size 12 from the uniform distribution on the interval $(0, 1)$. Approximate $P(1/2 \le \overline{X} \le 2/3)$.

5.6-2. Let $Y = X_1 + X_2 + \cdots + X_{15}$ be the sum of a random sample of size 15 from the distribution whose pdf is $f(x) = (3/2)x^2$, $-1 < x < 1$. Using the pdf of Y, we find that $P(-0.3 \le Y \le 1.5) = 0.22788$. Use the central limit theorem to approximate this probability.

5.6-3. Let $\overline{X}$ be the mean of a random sample of size 36 from an exponential distribution with mean 3. Approximate $P(2.5 \le \overline{X} \le 4)$.

5.6-4. Approximate $P(39.75 \le \overline{X} \le 41.25)$, where $\overline{X}$ is the mean of a random sample of size 32 from a distribution with mean $\mu = 40$ and variance $\sigma^2 = 8$.

5.6-5. Let $X_1, X_2, \ldots, X_{18}$ be a random sample of size 18 from a chi-square distribution with $r = 1$. Recall that $\mu = 1$ and $\sigma^2 = 2$.

(a) How is $Y = \sum_{i=1}^{18} X_i$ distributed?

(b) Using the result of part (a), we see from Table IV in Appendix B that

$$P(Y \le 9.390) = 0.05 \quad \text{and} \quad P(Y \le 34.80) = 0.99.$$

Compare these two probabilities with the approximations found with the use of the central limit theorem.

5.6-6. A random sample of size $n = 18$ is taken from the distribution with pdf $f(x) = 1 - x/2$, $0 \le x \le 2$.

(a) Find μ and σ^2. **(b)** Find $P(2/3 \le \overline{X} \le 5/6)$, approximately.

5.6-7. Let X equal the maximal oxygen intake of a human on a treadmill, where the measurements are in milliliters of oxygen per minute per kilogram of weight. Assume that, for a particular population, the mean of X is $\mu = 54.030$ and the standard deviation is $\sigma = 5.8$. Let $\overline{X}$ be the sample mean of a random sample of size $n = 47$. Find $P(52.761 \le \overline{X} \le 54.453)$, approximately.

5.6-8. Let X equal the weight in grams of a miniature candy bar. Assume that $\mu = E(X) = 24.43$ and $\sigma^2 = \text{Var}(X) = 2.20$. Let $\overline{X}$ be the sample mean of a random sample of $n = 30$ candy bars. Find

(a) $E(\overline{X})$. **(b)** $\text{Var}(\overline{X})$. **(c)** $P(24.17 \le \overline{X} \le 24.82)$, approximately.

5.6-9. In Example 5.6-4, compute $P(1.7 \le Y \le 3.2)$ with $n = 4$ and compare your answer with the normal approximation of this probability.

5.6-10. Let X and Y equal the respective numbers of hours a randomly selected child watches movies or cartoons on TV during a certain month. From experience, it is known that $E(X) = 30$, $E(Y) = 50$, $\text{Var}(X) = 52$, $\text{Var}(Y) = 64$, and $\text{Cov}(X, Y) = 14$. Twenty-five children are selected at random. Let Z equal the total number of hours these 25 children watch TV movies or cartoons in the next month. Approximate $P(1970 < Z < 2090)$. HINT: Use the remark after Theorem 5.3-2.

5.6-11. A company has a one-year group life policy that divides its employees into two classes as follows:

Class	Probability of Death	Benefit	Number in Class
A	0.01	\$20,000	1000
B	0.03	\$10,000	500

The insurance company wants to collect a premium that equals the 90th percentile of the distribution of the total claims. What should that premium be?

5.6-12. At certain times during the year, a bus company runs a special van holding 10 passengers from Iowa City to Chicago. After the opening of sales of the tickets, the time (in minutes) between sales of tickets for the trip has a gamma distribution with $\alpha = 3$ and $\theta = 2$.

(a) Assuming independence, record an integral that gives the probability of being sold out within one hour.

(b) Approximate the answer in part (a).

5.6-13. The tensile strength X of paper, in pounds per square inch, has $\mu = 30$ and $\sigma = 3$. A random sample of size $n = 100$ is taken from the distribution of tensile strengths. Compute the probability that the sample mean $\overline{X}$ is greater than 29.5 pounds per square inch.

5.6-14. Suppose that the sick leave taken by the typical worker per year has $\mu = 10$, $\sigma = 2$, measured in days. A firm has $n = 20$ employees. Assuming independence, how many sick days should the firm budget if the financial officer wants the probability of exceeding the number of days budgeted to be less than 20%?

5.6-15. Let X_1, X_2, X_3, X_4 represent the random times in days needed to complete four steps of a project. These times are independent and have gamma distributions with common $\theta = 2$ and $\alpha_1 = 3, \alpha_2 = 2, \alpha_3 = 5, \alpha_4 = 3$, respectively. One step must be completed before the next can be started. Let Y equal the total time needed to complete the project.

(a) Find an integral that represents $P(Y \le 25)$.

(b) Using a normal distribution, approximate the answer to part (a). Is this approach justified?

5.7 APPROXIMATIONS FOR DISCRETE DISTRIBUTIONS

In this section, we illustrate how the normal distribution can be used to approximate probabilities for certain discrete-type distributions. One of the more important discrete distributions is the binomial distribution. To see how the central limit theorem can be applied, recall that a binomial random variable can be described as the sum of Bernoulli random variables. That is, let $X_1, X_2, \ldots, X_n$ be a random sample from a Bernoulli distribution with mean $\mu = p$ and variance $\sigma^2 = p(1-p)$, where $0 < p < 1$. Then $Y = \sum_{i=1}^{n} X_i$ is $b(n, p)$. The central limit theorem states that the distribution of

$$W = \frac{Y - np}{\sqrt{np(1-p)}} = \frac{\overline{X} - p}{\sqrt{p(1-p)/n}}$$

is $N(0, 1)$ in the limit as $n \to \infty$. Thus, if n is "sufficiently large," the distribution of Y is approximately $N[np, np(1-p)]$, and probabilities for the binomial distribution