By the alternative formula for variance, we have

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2 = \int_0^2 x^2 \cdot \frac{1}{2}xdx - \left(\frac{4}{3}\right)^2$$
$$= \frac{x^4}{8}\bigg|_0^2 - \frac{16}{9} = 2 - \frac{16}{9} = \frac{2}{9}$$

Thus, the standard deviation is

$$\sigma = \sqrt{\frac{2}{9}} = \frac{\sqrt{2}}{3}$$

Now Work Problem 5 ◁

We conclude this section by emphasizing that a density function for a continuous random variable must not be confused with a probability distribution function for a discrete random variable. Evaluating such a probability distribution function at a *point* gives a probability. But evaluating a density function at a point does not. Instead, the *area* under the density function curve over an *interval* is interpreted as a probability. That is, probabilities associated with a continuous random variable are given by integrals.

PROBLEMS 16.1

1. Suppose X is a continuous random variable with density function given by

$$f(x) = \begin{cases} \frac{1}{6}(x+1) & \text{if } 1 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $P(1 < X < 2)$. **(b)** Find $P(X < 2.5)$.
(c) Find $P(X \geq \frac{3}{2})$.
(d) Find c such that $P(X < c) = \frac{1}{2}$. Give your answer in radical form.

2. Suppose X is a continuous random variable with density function given by

$$f(x) = \begin{cases} \dfrac{1000}{x^2} & \text{if } x > 1000 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $P(1000 < X < 2000)$. **(b)** Find $P(X > 5000)$.

3. Suppose X is a continuous random variable that is uniformly distributed on $[1, 4]$.
(a) What is the formula of the density function for X? Sketch its graph.
(b) Find $P\left(\frac{3}{2} < X < \frac{7}{2}\right)$. **(c)** Find $P(0 < X < 1)$.
(d) Find $P(X \leq 3.5)$. **(e)** Find $P(X > 3)$.
(f) Find $P(X = 2)$. **(g)** Find $P(X < 5)$.
(h) Find μ. **(i)** Find σ.
(j) Find the cumulative distribution function F and sketch its graph. Use F to find $P(X < 2)$ and $P(1 < X < 3)$.

4. Suppose X is a continuous random variable that is uniformly distributed on $[0, 5]$.
(a) What is the formula of the density function for X? Sketch its graph.
(b) Find $P(1 < X < 3)$. **(c)** Find $P(4.5 \leq X < 5)$.
(d) Find $P(X = 4)$. **(e)** Find $P(X > 2)$.
(f) Find $P(X < 5)$. **(g)** Find $P(X > 5)$.
(h) Find μ. **(i)** Find σ.
(j) Find the cumulative distribution function F and sketch its graph. Use F to find $P(1 < X < 3.5)$.

5. If X is a random variable with density function f, then the expectation of X is given by $\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$. Now, we will also write $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$ and $E((X-\mu)^2) = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$. In the text it was claimed that the variance of X, $\text{Var}(X) = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$ is also given by $\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$, so that

$$E((X - E(X))^2) = E(X^2) - E(X)^2$$

Prove this.

6. Suppose X is a continuous random variable with density function given by

$$f(x) = \begin{cases} k & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $k = \dfrac{1}{b-a}$ and thus X is uniformly distributed.
(b) Find the cumulative distribution function F.

7. Suppose the random variable X is exponentially distributed with $k = 2$.
(a) Find $P(1 < X < 2)$. **(b)** Find $P(X < 3)$.
(c) Find $P(X > 5)$.
(d) Find $P(\mu - 2\sigma < X < \mu + 2\sigma)$.
(e) Find the cumulative distribution function F.

8. Suppose the random variable X is exponentially distributed with $k = 0.5$.
(a) Find $P(X > 4)$. **(b)** Find $P(0.5 < X < 2.6)$.
(c) Find $P(X < 5)$. **(d)** Find $P(X = 4)$.
(e) Find c such that $P(0 < X < c) = \frac{1}{2}$.

9. The density function for a random variable X is given by

$$f(x) = \begin{cases} kx & \text{if } 0 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find k.
(b) Find $P(2 < X < 3)$.
(c) Find $P(X > 2.5)$.
(d) Find $P(X > 0)$.
(e) Find μ.
(f) Find σ.
(g) Find c such that $P(X < c) = \frac{1}{2}$.
(h) Find $P(3 < X < 5)$.

10. The density function for a random variable X is given by

$$f(x) = \begin{cases} \dfrac{x}{16} + k & \text{if } 1 \le x \le 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find k.
(b) Find $P(X \ge 3)$.
(c) Find μ.
(d) Find $P(2 < X < \mu)$.

11. Waiting Time At a bus stop, the time X (in minutes) that a randomly arriving person must wait for a bus is uniformly distributed with density function $f(x) = \frac{1}{10}$, where $0 \le x \le 10$ and $f(x) = 0$ otherwise. What is the probability that a person must wait at most seven minutes? What is the average time that a person must wait?

12. Soft-Drink Dispensing An automatic soft-drink dispenser at a fast-food restaurant dispenses X ounces of cola in a 12-ounce drink. If X is uniformly distributed over $[11.92, 12.08]$, what is the probability that less than 12 ounces will be dispensed? What is the probability that exactly 12 ounces will be dispensed? What is the average amount dispensed?

13. Emergency Room Arrivals At a particular hospital, the length of time X (in hours) between successive arrivals at the emergency room is exponentially distributed with $k = 3$. What is the probability that more than one hour passes without an arrival?

14. Electronic Component Life The length of life, X (in years), of a computer component has an exponential distribution with $k = \frac{2}{5}$. What is the probability that such a component will fail within three years of use? What is the probability that it will last more than five years?

Objective

To discuss the normal distribution, standard units, and the table of areas under the standard normal curve (Appendix C).

16.2 The Normal Distribution

Quite often, measured data in nature—such as heights of individuals in a population—are represented by a random variable whose density function may be approximated by the bell-shaped curve in Figure 16.9. The curve extends indefinitely to the right and left and never touches the x-axis. This curve, called the **normal curve**, is the graph of the most important of all density functions, the *normal density function*.

Definition

A continuous random variable X is a *normal random variable*, and equivalently, has a **normal** (also called Gaussian[1]) **distribution**, if its density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)[(x-\mu)/\sigma]^2} \qquad -\infty < x < \infty$$

called the **normal density function**. The parameters μ and σ are the mean and standard deviation of X, respectively.

Observe in Figure 16.9 that $f(x) \to 0$ as $x \to \pm\infty$. That is, the normal curve has the x-axis as a horizontal asymptote. Also note that the normal curve is symmetric about the vertical line $x = \mu$. That is, the height of a point on the curve d units to the right of $x = \mu$ is the same as the height of the point on the curve that is d units to the left of $x = \mu$. Because of this symmetry and the fact that the area under the normal curve is 1, the area to the right (or left) of the mean must be $\frac{1}{2}$.

Each choice of values for μ and σ determines a different normal curve. The value of μ determines where the curve is "centered", and σ determines how "spread out" the curve is. The smaller the value of σ, the less spread out is the area near μ. For example, Figure 16.10 shows normal curves C_1, C_2, and C_3, where C_1 has mean μ_1 and standard deviation σ_1, C_2 has mean μ_2, and so on. Here C_1 and C_2 have the same mean but different standard deviations: $\sigma_1 > \sigma_2$. C_1 and C_3 have the same standard deviation but different means: $\mu_1 < \mu_3$. Curves C_2 and C_3 have different means and different standard deviations.

[1] After the German mathematician Carl Friedrich Gauss (1777–1855).

PROBLEMS 16.2

1. If Z is a standard normal random variable, find each of the following probabilities.

(a) $P(0 < Z < 2.3)$ **(b)** $P(0.35 < Z < 1.31)$
(c) $P(Z > -0.57)$ **(d)** $P(Z \leq 1.46)$
(e) $P(-2.38 < Z \leq 1.70)$ **(f)** $P(Z > 0.19)$

2. If Z is a standard normal random variable, find each of the following.

(a) $P(-1.96 < Z < 1.96)$ **(b)** $P(-2.11 < Z < -1.35)$
(c) $P(Z < -1.05)$ **(d)** $P(Z > 3\sigma)$
(e) $P(|Z| > 2)$ **(f)** $P(|Z| < \frac{1}{2})$

In Problems 3–8, find z_0 such that the given statement is true. Assume that Z is a standard normal random variable.

3. $P(Z < z_0) = 0.6368$ **4.** $P(Z < z_0) = 0.0668$
5. $P(Z > z_0) = 0.8599$ **6.** $P(Z > z_0) = 0.4286$
7. $P(-z_0 < Z < z_0) = 0.2662$ **8.** $P(|Z| > z_0) = 0.0456$

9. If X is normally distributed with $\mu = 16$ and $\sigma = 4$, find each of the following probabilities.

(a) $P(X < 27)$ **(b)** $P(X < 10)$
(c) $P(10.8 < X < 12.4)$

10. If X is normally distributed with $\mu = 200$ and $\sigma = 40$, find each of the following probabilities.

(a) $P(X > 150)$ **(b)** $P(210 < X < 250)$

11. If X is normally distributed with $\mu = 57$ and $\sigma = 10$, find $P(X > 80)$.

12. If X is normally distributed with $\mu = 0$ and $\sigma = 1.5$, find $P(X < 3)$.

13. If X is normally distributed with $\mu = 60$ and $\sigma^2 = 100$, find $P(50 < X \leq 75)$.

14. If X is normally distributed with $\mu = 8$ and $\sigma = 1$, find $P(X > \mu - \sigma)$.

15. If X is normally distributed such that $\mu = 40$ and $P(X > 54) = 0.0401$, find σ.

16. If X is normally distributed with $\mu = 62$ and $\sigma = 11$, find x_0 such that the probability that X is between x_0 and 62 is 0.4554.

17. Test Scores The scores on a national achievement test are normally distributed with mean 500 and standard deviation 100. What percentage of those who took the test had a score between 300 and 700?

18. Test Scores In a test given to a large group of people, the scores were normally distributed with mean 55 and standard deviation 10. What is the greatest whole-number score that a person could get and yet score in about the bottom 10%?

19. Adult Heights The heights (in inches) of adults in a large population are normally distributed with $\mu = 68$ and $\sigma = 3$. What percentage of the group is under 6 feet tall?

20. Income The yearly income for a group of 10,000 professional people is normally distributed with $\mu = \$60{,}000$ and $\sigma = \$5000$.

(a) What is the probability that a person from this group has a yearly income less than \$46,000?
(b) How many of these people have yearly incomes over \$75,000?

21. IQ The IQs of a large population of children are normally distributed with mean 100.4 and standard deviation 11.6.

(a) What percentage of the children have IQs greater than 120?
(b) About 95.05% of the children have IQs greater than what value?

22. Suppose X is a random variable with $\mu = 10$ and $\sigma = 2$. If $P(4 < X < 16) = 0.25$, can X be normally distributed?

Objective

To show the technique of estimating the binomial distribution by using the normal distribution.

16.3 The Normal Approximation to the Binomial Distribution

We conclude this chapter by bringing together the notions of a discrete random variable and a continuous random variable. Recall from Chapter 9 that if X is a binomial random variable (which is discrete), and if the probability of success on any trial is p, then for n independent trials, the probability of x successes is given by

$$P(X = x) = {}_nC_x p^x q^{n-x}$$

where $q = 1 - p$. Calculating probabilities for a binomial random variable can be time consuming when the number of trials is large. For example, ${}_{100}C_{40}(0.3)^{40}(0.7)^{60}$ is a lot of work to compute "by hand". Fortunately, we can approximate a binomial distribution like this by a normal distribution and then use a table of areas.

To show how this is done, let us take a simple example. Figure 16.22 gives a probability histogram for a binomial experiment with $n = 10$ and $p = 0.5$. The rectangles centered at $x = 0$ and $x = 10$ are not shown because their heights are very close to 0. Superimposed on the histogram is a normal curve, which approximates it. The approximation would be even better if n were larger. That is, as n gets larger, the width of each unit interval appears to get smaller, and the outline of the histogram tends to take on the appearance of a smooth curve. In fact, *it is not unusual to think of a density curve as*

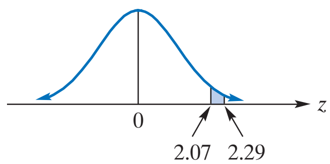

FIGURE 16.25 $P(2.07 \leq Z \leq 2.29)$.

This probability is the area under a standard normal curve between $z = 2.07$ and $z = 2.29$ (Figure 16.25). That area is the difference of the area between $z = 0$ and $z = 2.29$, which is $A(2.29)$, and the area between $z = 0$ and $z = 2.07$, which is $A(2.07)$. Thus,

$$\begin{aligned} P(X = 40) &\approx P(2.07 \leq Z \leq 2.29) \\ &= A(2.29) - A(2.07) \\ &= 0.4890 - 0.4808 = 0.0082 \qquad \text{from Appendix C} \end{aligned}$$

Now Work Problem 3 ◁

EXAMPLE 2 Quality Control

In a quality-control experiment, a sample of 500 items is taken from an assembly line. Customarily, 8% of the items produced are defective. What is the probability that more than 50 defective items appear in the sample?

Solution: If X is the number of defective items in the sample, then we will consider X to be binomial with $n = 500$ and $p = 0.08$. To find $P(X \geq 51)$, we use the normal approximation to the binomial distribution with

$$\mu = np = 500(0.08) = 40$$

and

$$\sigma = \sqrt{npq} = \sqrt{500(0.08)(0.92)} = \sqrt{36.8} \approx 6.07$$

Converting the corrected value 50.5 to a Z-value gives

$$z = \frac{50.5 - 40}{\sqrt{36.8}} \approx 1.73$$

Thus,

$$P(X \geq 51) \approx P(Z \geq 1.73)$$

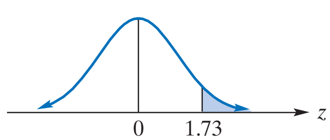

FIGURE 16.26 $P(Z \geq 1.73)$.

This probability is the area under a standard normal curve to the right of $z = 1.73$ (Figure 16.26). That area is the difference of the area to the right of $z = 0$, which is 0.5, and the area between $z = 0$ and $z = 1.73$, which is $A(1.73)$. Hence,

$$\begin{aligned} P(X \geq 51) &\approx P(Z \geq 1.73) \\ &= 0.5 - A(1.73) = 0.5 - 0.4582 = 0.0418 \end{aligned}$$

Now Work Problem 7 ◁

PROBLEMS 16.3

In Problems 1–4, X is a binomial random variable with the given values of n and p. Calculate the indicated probabilities by using the normal approximation.

1. $n = 150, p = 0.4$; $P(X \geq 52), P(X \geq 74)$

2. $n = 50, p = 0.3$; $P(X = 25), P(X \leq 20)$

3. $n = 200, p = 0.6$; $P(X = 125), P(110 \leq X \leq 135)$

4. $n = 50, p = 0.20$; $P(X \geq 10)$

5. Die Tossing Suppose a fair die is tossed 300 times. What is the probability that a 5 turns up between 45 and 60 times, inclusive?

6. Coin Tossing For a biased coin, $P(H) = 0.4$ and $P(T) = 0.6$. If the coin is tossed 200 times, what is the probability of getting between 90 and 100 heads, inclusive?

7. Taxis out of service A taxi company has a fleet of 100 cars. At any given time, the probability of a car being out of service due to factors such as breakdowns and maintenance is 0.1. What is the probability that 10 or more cars are out of service at any time?

8. Quality Control In a manufacturing plant, a sample of 200 items is taken from the assembly line. For each item in the sample, the probability of being defective is 0.05. What is the probability that there are 7 or more defective items in the sample?

9. True–False Exam In a true–false exam with 50 questions, what is the probability of getting at least 25 correct answers by just guessing on all the questions? If there are 100 questions instead of 50, what is the probability of getting at least 50 correct answers by just guessing?

10. Multiple-Choice Exam In a multiple-choice test with 50 questions, each question has four answers, only one of which is correct. If a student guesses on the last 20 questions, what is the probability of getting at least half of them correct?

11. Poker In a poker game, the probability of being dealt a hand consisting of three cards of one suit and two cards of another suit (in any order) is about 0.1. In 100 dealt hands, what is the probability that 16 or more of them will be as just described?

12. Taste Test An energy drink company sponsors a national taste test, in which subjects sample its drink as well as the best-selling brand. Neither drink is identified by brand. The subjects are then asked to choose the drink that tastes better. If each of the 49 subjects in a supermarket actually has no preference and arbitrarily chooses one of the drinks, what is the probability that 30 or more of the subjects choose the drink from the sponsoring company?

Chapter 16 Review

Important Terms and Symbols

		Examples
Section 16.1	**Continuous Random Variables**	
	continuous random variable uniform density function	Ex. 1, p. 716
	exponential density function exponential distribution	Ex. 3, p. 717
	cumulative distribution function (probability) density function	Ex. 4, p. 718
	mean, μ variance, σ^2 standard deviation, σ	Ex. 5, p. 719
Section 16.2	**The Normal Distribution**	
	normal distribution normal density function	Ex. 1, p. 722
	standard normal curve standard normal random variable	Ex. 2, p. 724
	standard normal distribution standard normal density function	Ex. 4, p. 725
Section 16.3	**The Normal Approximation to the Binomial Distribution**	
	continuity correction	Ex. 1, p. 728

Summary

A continuous random variable, X, can assume any value in an interval or intervals. A density function is a function that has the following properties:

1. $f(x) \geq 0$ **2.** $\int_{-\infty}^{\infty} f(x)dx = 1$

A density function is a density function *for the random variable X* if

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

which means that the probability that X assumes a value in the interval $[a, b]$ is to be given by the area under the graph of f and above the x-axis from $x = a$ to $x = b$. The probability that X assumes a particular value is 0.

The continuous random variable X has a uniform distribution over $[a, b]$ if its density function is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

X has an exponential density function, f, if

$$f(x) = \begin{cases} ke^{-kx} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

where k is a positive constant.

The cumulative distribution function, F, for the continuous random variable X with density function f is given by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

Geometrically, $F(x)$ represents the area under the density curve to the left of x. By using F, we are able to find $P(a \leq x \leq b)$:

$$P(a \leq x \leq b) = F(b) - F(a)$$

The mean μ of X (also called expectation of X) $E(X)$ is given by

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

provided that the integral is convergent. The variance is given by

$$\begin{aligned} \sigma^2 = \text{Var}(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx \\ &= \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2 \end{aligned}$$

provided that the integral is convergent. The standard deviation is given by

$$\sigma = \sqrt{\mathrm{Var}(X)}$$

The graph of the normal density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)((x-\mu)/\sigma)^2}$$

is called a normal curve and is bell shaped. If X has a normal distribution, then the probability that X lies within one standard deviation of the mean μ is (approximately) 0.68; within two standard deviations, the probability is 0.95; and within three standard deviations, it is 0.997. If Z is a normal random variable with $\mu = 0$ and $\sigma = 1$, then Z is called a standard normal random variable. The probability $P(0 < Z < z_0)$ is the area under the graph of the standard normal curve from $z = 0$ to $z = z_0$ and is denoted $A(z_0)$. Values of $A(z_0)$ appear in Appendix C.

If X is normally distributed with mean μ and standard deviation σ, then X can be transformed into a standard normal random variable by the change-of-variable formula

$$Z = \frac{X - \mu}{\sigma}$$

With this formula, probabilities for X can be found by using areas under the standard normal curve.

If X is a binomial random variable and the number, n, of independent trials is large, then the distribution of X can be approximated by using a normal random variable with mean np and standard deviation $\sqrt{npq}$, where p is the probability of success on any trial and $q = 1-p$. It is important that continuity corrections be considered when we estimate binomial probabilities by a normal random variable.

Review Problems

1. Suppose X is a continuous random variable with density function given by

$$f(x) = \begin{cases} \frac{1}{3} + kx^2 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find k. **(b)** Find $P(\frac{1}{2} < X < \frac{3}{4})$. **(c)** Find $P(X \ge \frac{1}{2})$.

(d) Find the cumulative distribution function.

2. Suppose X is exponentially distributed with $k = \frac{1}{3}$. Find $P(X > 2)$.

3. Suppose X is a random variable with density function given by

$$f(x) = \begin{cases} \frac{1}{8}x & \text{if } 0 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find μ. **(b)** Find σ.

4. Let X be uniformly distributed over the interval [2, 6]. Find $P(X < 5)$.

Let X be normally distributed with mean 20 and standard deviation 4. In Problems 5–10, determine the given probabilities.

5. $P(X > 25)$ **6.** $P(X < 21)$ **7.** $P(14 < X < 18)$

8. $P(X > 32)$ **9.** $P(X < 23)$ **10.** $P(21 < X < 31)$

In Problems 11 and 12, X is a binomial random variable with $n = 100$ and $p = 0.35$. Find the given probabilities by using the normal approximation.

11. $P(25 \le X \le 47)$ **12.** $P(X = 48)$

13. Heights of Individuals The heights in meters of individuals in a certain group are normally distributed with mean 1.73 and standard deviation 0.05. Find the probability that an individual from this group is taller than 1.83.

14. Coin Tossing If a fair coin is tossed 500 times, use the normal approximation to the binomial distribution to estimate the probability that a head comes up at least 215 times.