統計學

Fall 2025

政治大學統計系余清祥

2025年10月7日

第六章：連續機率分配

http://csyue.nccu.edu.tw

Chapter Contents

Introduction

In this chapter, we study continuous random variables and discuss three continuous probability distributions: the uniform, the normal, and the exponential.

Because of its wide applicability and its extensive use in statistical inference, the normal distribution is the continuous probability distribution of major importance.

Probabilities are computed differently for discrete and continuous random variables.

- with discrete random variables, the probability function $f(x)$ provides the probability that the random variable assumes a particular value.
- with continuous random variables,
    - the probability that x assumes a particular value is zero;
    - the probability that x assumes a value in a given interval is the corresponding area under the graph of $f(x)$;
    - $f(x)$ is called the **probability density function**.

We begin our study of continuous probability distributions with the uniform distribution.

6.1 Uniform Probability Density Function

A random variable is described by a **uniform probability distribution** whenever the probability is proportional to the interval's length.

The *uniform probability density function* is:

$$\boldsymbol{f(x)} = \begin{cases} \dfrac{\mathbf{1}}{\boldsymbol{b-a}} & \textbf{for } \boldsymbol{a \le x \le b} \\ \mathbf{0} & \textbf{elsewhere} \end{cases}$$

Where

a = smallest value the variable x can assume

b = largest value the variable x can assume

A uniform continuous probability distribution has expected value and variance (*see notes)

$$\boldsymbol{E(x) = \frac{a+b}{2}} \qquad \boldsymbol{Var(x) = \frac{(b-a)^2}{12}}$$

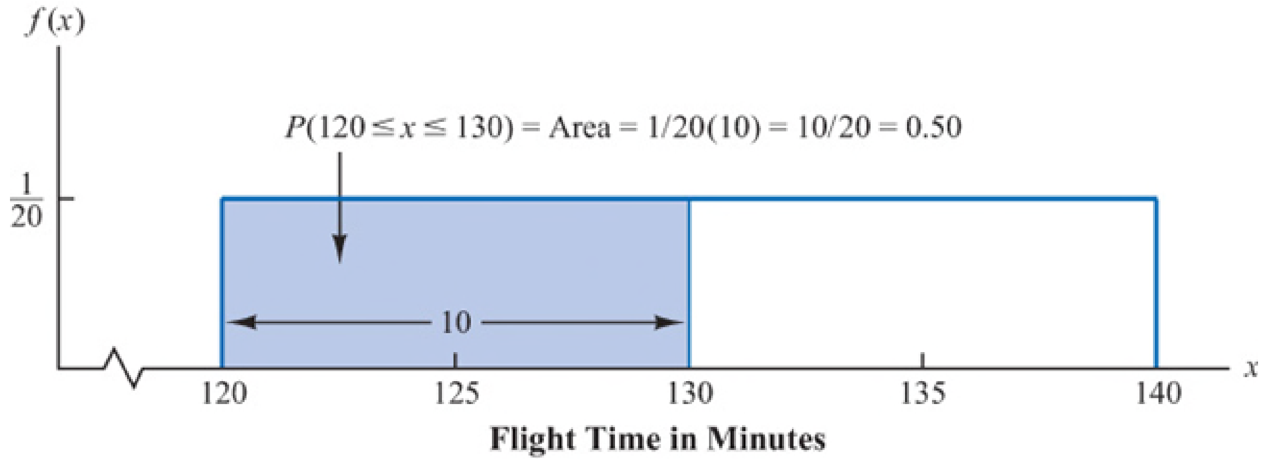
6.1 Area as a Measure of Probability

The flight time of an airplane traveling from Chicago to New York can be described by a continuous random variable that can assume any value in the interval from 120 to 140 minutes.

Let us assume that sufficient actual flight data are available to conclude that the probability of a flight time within any 1-minute interval is the same from 120 to 140 minutes.

Thus, the random variable x is said to have the following uniform probability distribution

$$f(x) = \begin{cases} 1/20 & \text{for } 120 \leq x \leq 140 \\ 0 & \text{elsewhere} \end{cases}$$



What is the probability that the flight time is between 120 and 130 minutes?

As shown in the graph, the area is:

$$P(120 \leq x \leq 130) = 1/20(10) = 0.50$$

6.1 Properties of a Uniform Probability Distribution

All continuous probability distributions have the following properties:

1. $f(x) \geq 0$ for any value of x
2. the area under the graph of $f(x)$ is equal to 1

For example, in the airplane traveling time problem seen in the previous slide

$$P(120 \leq x \leq 140) = 1/20(140 - 120) = 1/20(20) = 1$$

Applying the formulas for expected value and standard deviation (as the square root of the variance) to the uniform distribution for flight times from Chicago to New York, we obtain

$$E(x) = \frac{a+b}{2} = \frac{120+140}{2} = \frac{260}{2} = 130 \text{ minutes}$$

$$\sigma = \sqrt{Var(x)} = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(140-120)^2}{12}} = \sqrt{\frac{400}{12}} = \sqrt{33.33} = 5.77 \text{ minutes}$$

6.2 Normal Probability Density Function

The **normal probability distribution** is the most common probability continuous distribution.

The normal distribution, used in a wide variety of practical applications, is widely applied to statistical inference, which is the major topic of the remainder of this book.

The normal distribution assumes the characteristic bell-shaped curve and provides a description of the likely results obtained through sampling (*see notes.)

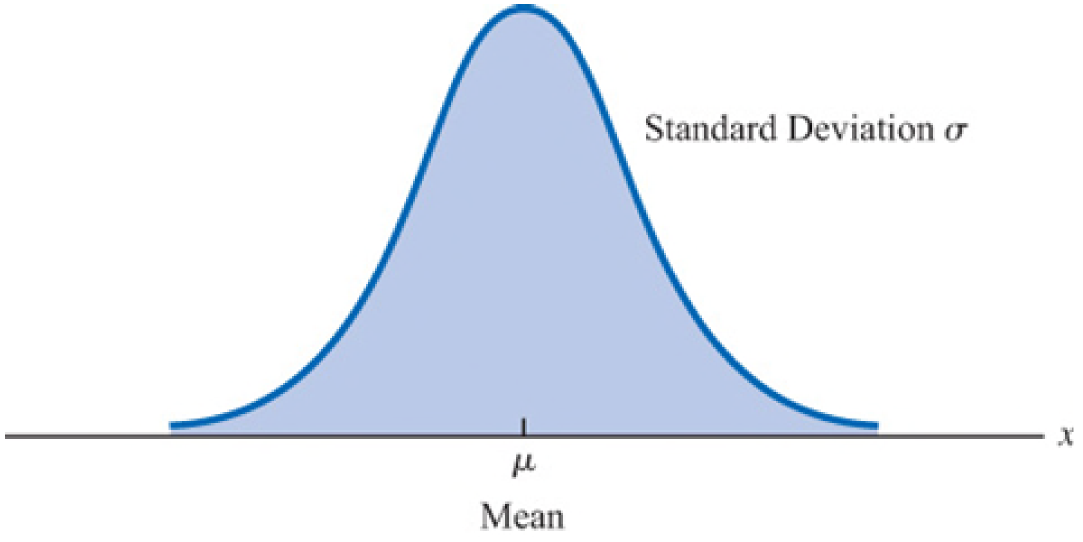
The *normal probability density function* is described by two parameters: the mean μ and the standard deviation σ.

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

Where

$\pi = 3.14159$ pi

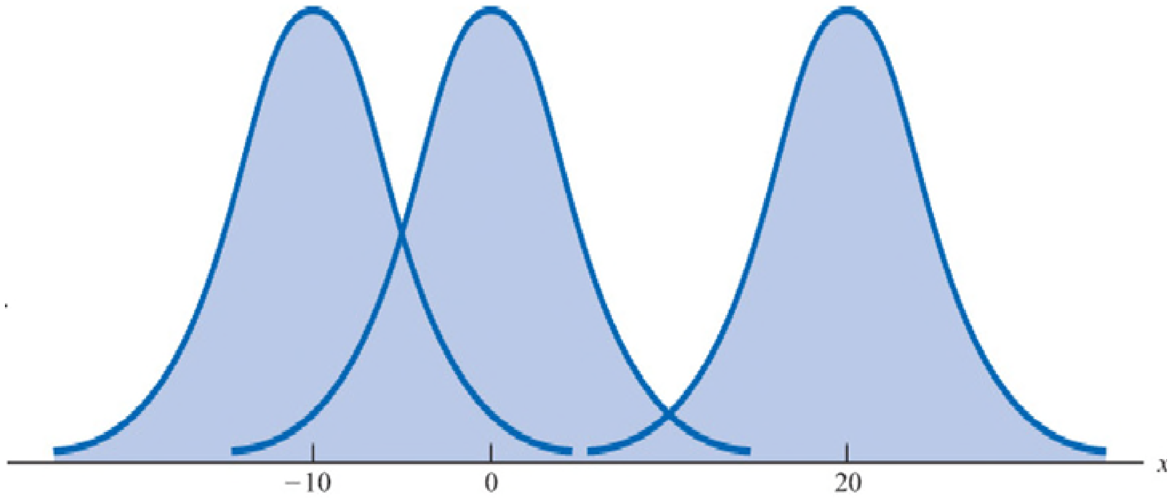
$e = 2.71828$ Euler's number



6.2 The Mean of the Normal Distribution

We make the following observations about the mean μ of the normal distribution:

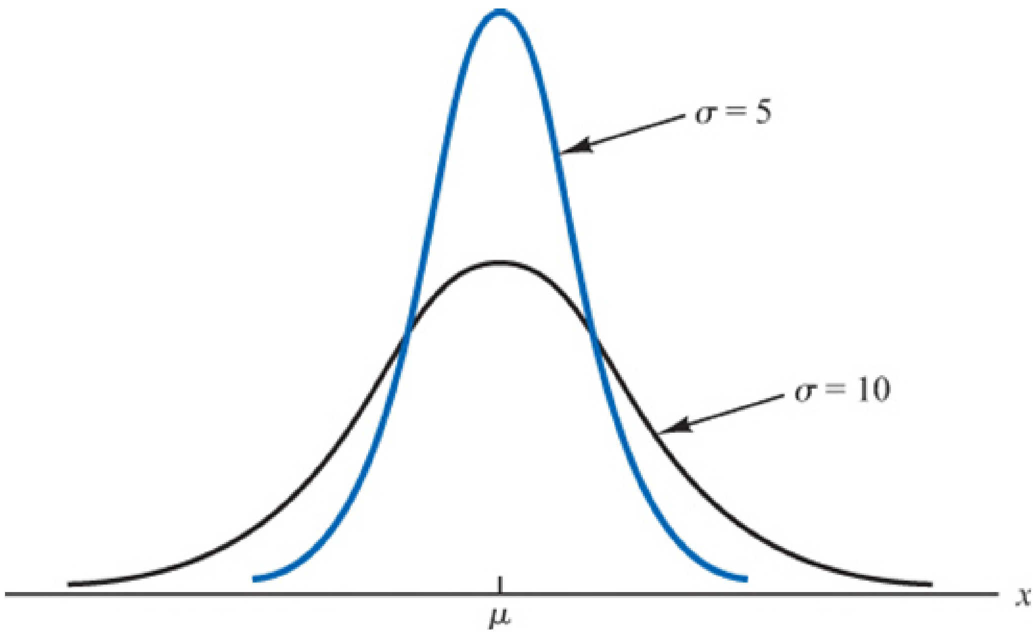
- The mean μ is the highest point on the normal curve.
- The mean μ is also the median and mode of the distribution.
- The mean of the distribution can be any numerical value: negative, zero, or positive. The figure shows three normal distributions with the same σ and $\mu = -10, 0,$ and 20.
- The normal distribution is symmetric.
- The area under the curve to the left of the mean and the area under the curve to the right of the mean are both 0.50.
- The tails of the normal curve extend to infinity in both directions and theoretically never touch the horizontal axis.



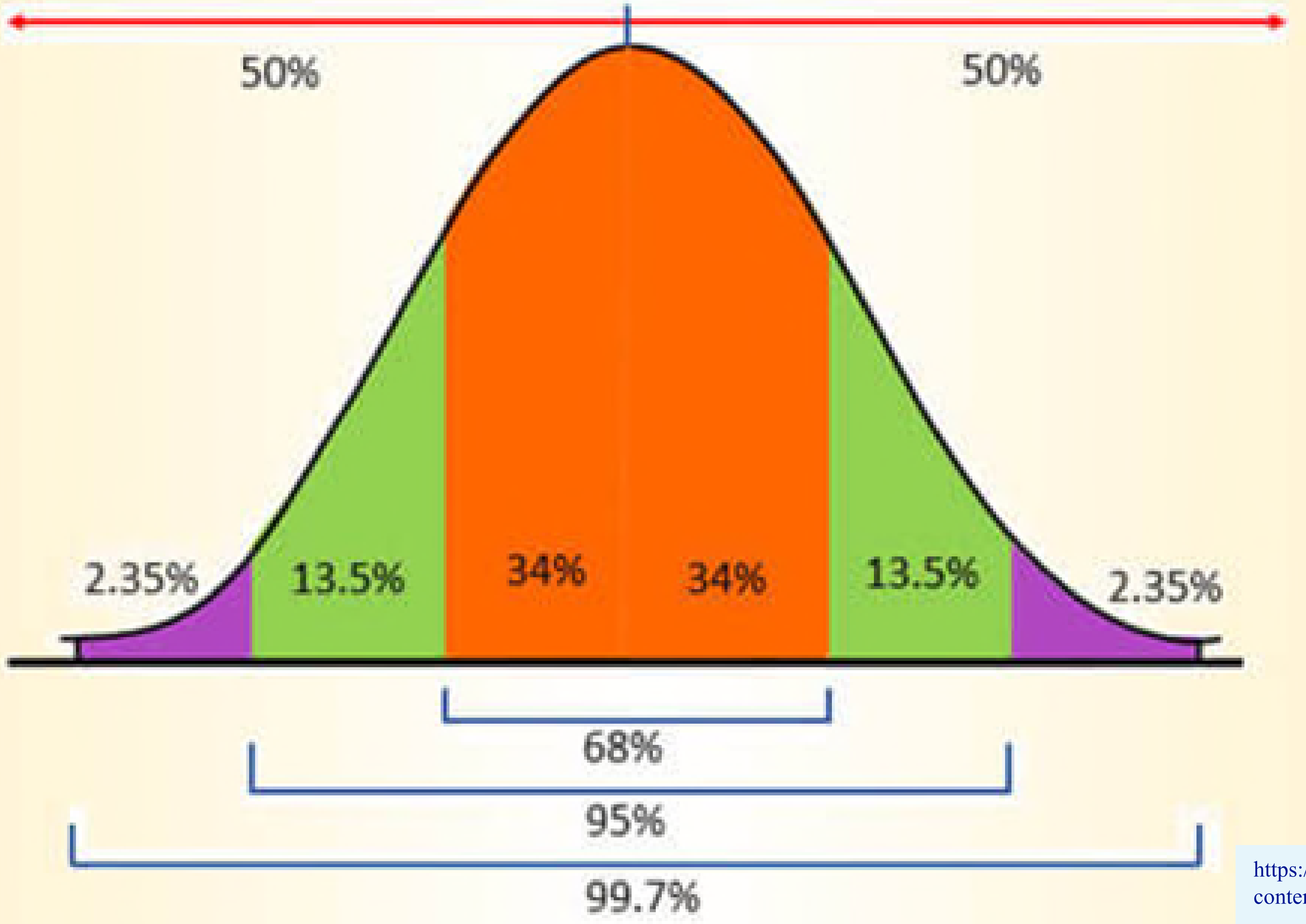
6.2 The Spread of the Normal Distribution

We make the following observations about the standard deviation σ of the normal distribution:

- The standard deviation determines how flat and wide the normal curve is.
  - Larger values of σ result in wider, flatter curves, showing more variability in the data.
  - The figure shows two normal distributions with the same mean but $\sigma = 5$ and $\sigma = 10$.
- The percentage of values in some typical intervals are (see Empirical rule, Chapter 3):
  - 68.3% of the values of a normal random variable are within $\pm 1\ \sigma$ of μ.
  - 95.4% of the values of a normal random variable are within $\pm 2\ \sigma$ of μ.
  - 99.7% of the values of a normal random variable are within $\pm 3\ \sigma$ of μ.



Empirical Rule



https://www.wallstreetmojo.com/wp-content/uploads/2020/06/Empirical-Rule.jpg

6.2 Standard Normal Probability Distribution

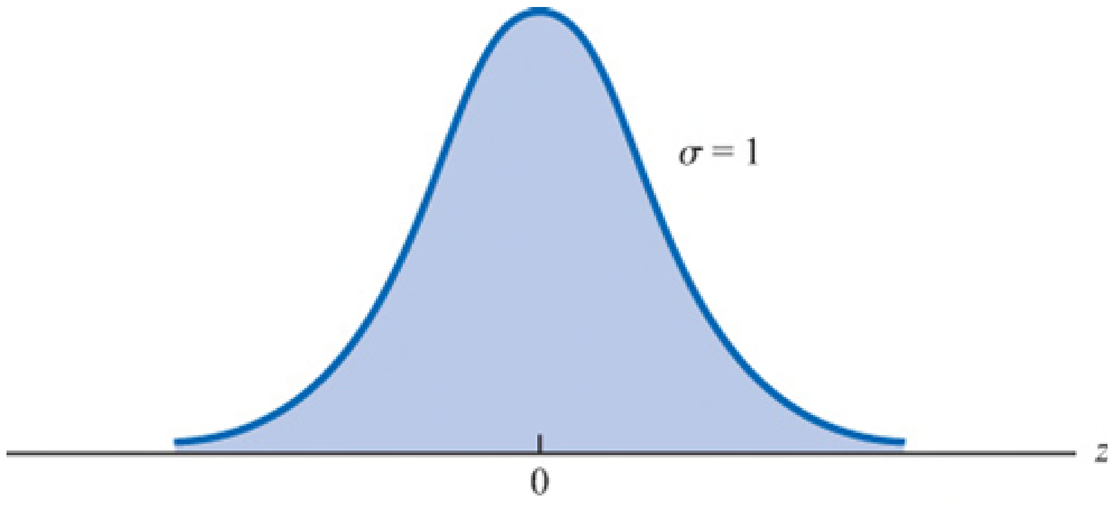
A random variable having a normal distribution with a mean of 0 and a standard deviation of 1 is said to have a **standard normal probability distribution**.

The letter z is commonly used to designate the standard normal random variable.

The standard normal distribution has the same general appearance as other normal distributions, but with the special properties of having $\mu = 0$ and $\sigma = 1$.

Because of $\mu = 0$ and $\sigma = 1$, the formula for the standard normal probability density function is a simpler version of the equation shown earlier for the general normal probability distribution:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$



6.2 Probability Calculations with the Normal Distribution

As with other continuous random variables, probability calculations with a normal distribution are made by computing the area under the graph of the normal curve over a given interval.

Excel can be used to calculate the probabilities associated with the

- normal probability distributions using the function $=\mathrm{NORM.DIST}(x, \mu, \sigma, \mathrm{TRUE})$
- the standard normal probability distribution using the function $=\mathrm{NORM.S.DIST}(z, \mathrm{TRUE})$

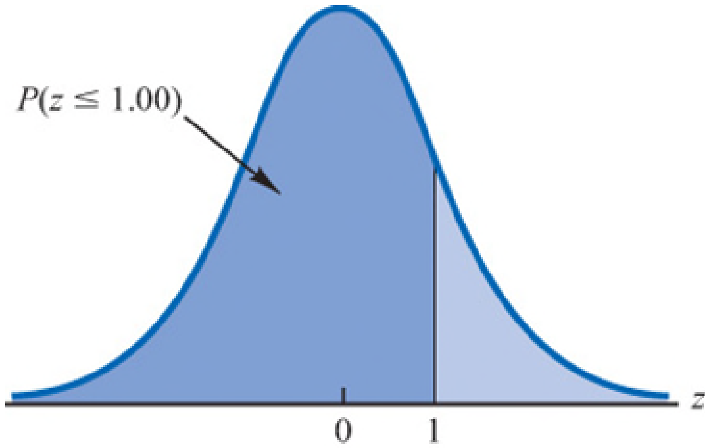
We will now show how to use Table 1 in Appendix B to calculate the following cumulative probabilities associated with a standard normal probability distribution

- the probability that the standard normal random variable z will be less than or equal to a given value Z: $P(z \leq Z)$
- the probability that z will be between two given values, Z_1 and Z_2: $P(Z_1 \leq z \leq Z_2)$
- the probability that z will be greater than or equal to a given value Z: $P(z \geq Z)$

CENGAGE

6.2 Cumulative Probability for $P(z \leq 1.00)$

The cumulative probability that z is less than or equal to 1.00, that is, $P(z \leq 1.00)$, is the area under the standard normal curve to the left of $z = 1.00$.



To find the cumulative probability corresponding to $z \leq 1.00$ in Table 1 of Appendix B, we identify the value located at the intersection of the row labeled 1.0 and the column labeled 0.00.

1. we find 1.0 in the left column of the table
2. we find 0.00 in the top row of the table.

z	0.00	0.01	0.02	...	0.09
⋮	⋮	⋮	⋮		⋮
0.9	0.8159	0.8186	0.8212	...	0.8389
1.0	$\boxed{0.8413}$	0.8438	0.8461	...	0.8621
1.1	0.8643	0.8665	0.8686	...	0.8830
⋮	⋮	⋮	⋮		⋮

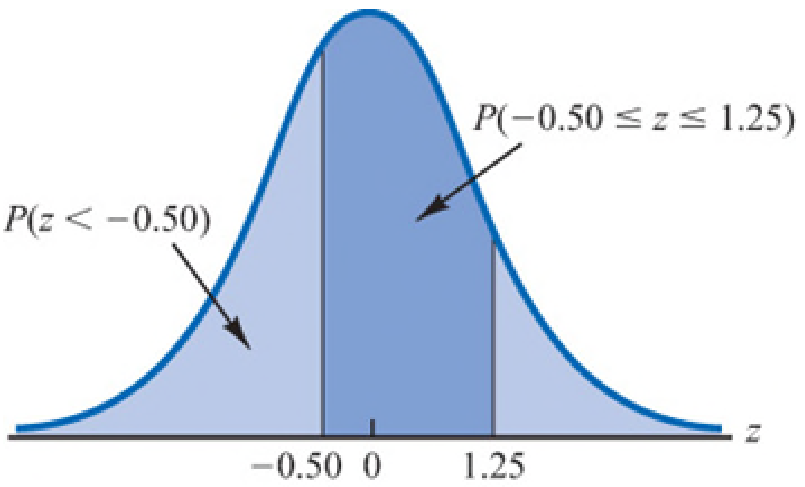
By looking in the body of the table, we find that the 1.0 row and the 0.00 column intersect at the value of 0.8413. Thus

$$P(z \leq 1.00) = 0.8413$$

CENGAGE

6.2 Cumulative Probability for $P(-0.50 \leq z \leq 1.25)$

The cumulative probability that z is in the interval between –0.50 and 1,25, that is, $P(-0.50 \leq z \leq 1.25)$, is the area depicted in the figure to the right.



Three steps are required to compute this probability using Table 1 of Appendix B.

1. we calculate $P(z \leq -0.50)$ as the intersection of the –0.5 row and the 0.00 column:

   $$P(z \leq -0.50) = 0.3085$$

2. we calculate $P(z \leq 1.25)$ as the intersection of the 1.2 row and the 0.05 column:

   $$P(z \leq 1.25) = 0.8944$$

3. we calculate the difference of the cumulative probabilities to the left of $z = 1.25$ and $z = -0.50$ that we calculated in steps 1 and 2, so that:

   $$P(-0.50 \leq z \leq 1.25) = P(z \leq 1.25) - P(z \leq -0.50) = 0.8944 - 0.3085 = 0.5859$$

CENGAGE

6.2 Cumulative Probability for $P(z \geq 1.58)$

The cumulative probability that z is greater than or equal to 1.58, that is, $P(z \geq 1.58)$, is the area under the standard normal curve to the right of $z = 1.58$.

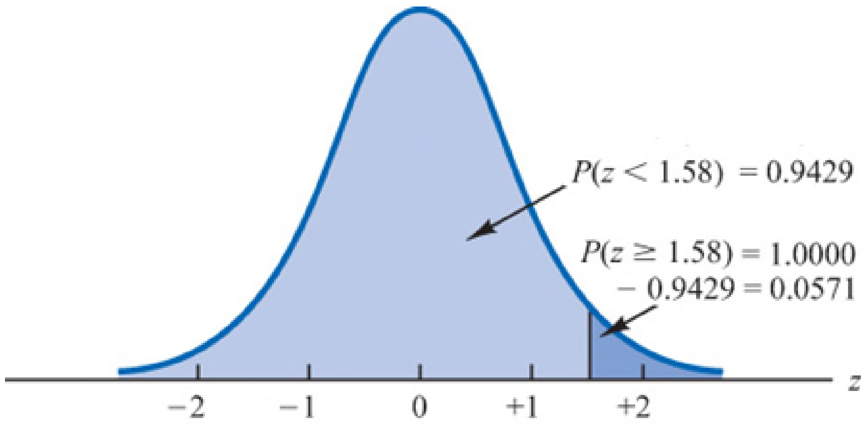
To find the probability to the right of $z = 1.58$, we need to first find the probability corresponding to the left of $z = 1.58$, and then calculate its complement.

To calculate the area to the left of $z = 1.58$, identify the value located at the intersection of the row labeled 1.5 and the column labeled 0.08. Thus

$$P(z \leq 1.58) = 0.9429$$

However, because the total area under the normal curve is 1, we have

$$P(z \geq 1.58) = 1 - 0.9429 = 0.0571$$



z	0.00	0.01	...	0.08	0.09
⋮	⋮	⋮		⋮	⋮
1.4	0.9192	0.9207	...	0.9306	0.9319
1.5	0.9332	0.9345	...	$\boxed{0.9429}$	0.9441
1.6	0.9452	0.9463	...	0.9535	0.9545
⋮	⋮	⋮		⋮	⋮

6.2 Find Z Such That $P(z \geq z^*) = 0.10$

We want to find a z^* value such that the probability of obtaining a larger z value is 0.10. That is

$$P(z \geq z^*) = 0.10$$

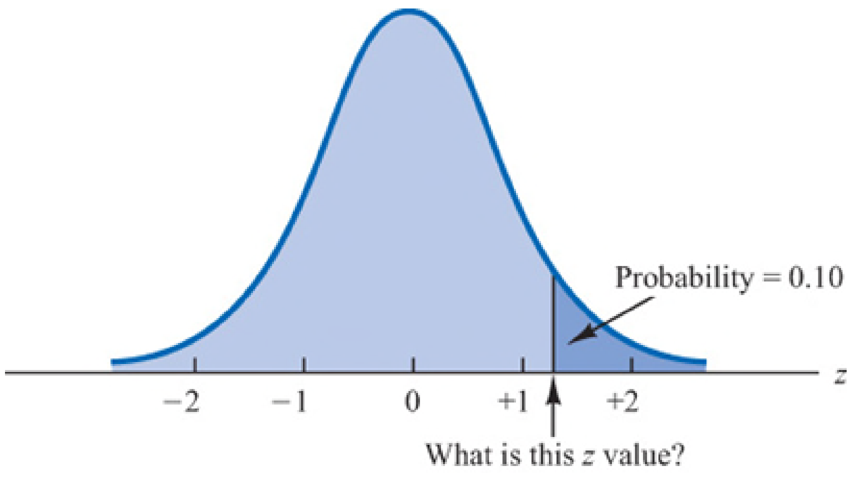
In this inverse problem, we need to find the z value that corresponds to a given probability.

However, the standard normal probability table gives the area under the curve to the *left* of a particular z value.

Thus, we use again the complement rule to find a z^* value such that

$$P(z \leq z^*) = 1 - P(z \geq z^*) = 1 - 0.10 = 0.90$$

Table 1 of Appendix B shows the closest probability to 0.9000 to be 0.8997, corresponding to the 1.2 row and the 0.08 column. That is, $z^* = 1.28$ (*see notes.)



z	0.00	0.01	...	0.08	0.09
⋮	⋮	⋮		⋮	⋮
1.1	0.8643	0.8665	...	0.8810	0.8830
1.2	0.8849	0.8869	...	$\boxed{0.8997}$	0.9015
1.3	0.9032	0.9049	...	0.9162	0.9177
⋮	⋮	⋮		⋮	⋮

CENGAGE

6.2 Computing Probabilities for Any Normal Curve

We use the following formula to convert any normal random variable x with mean μ and standard deviation σ to the standard normal random variable z (*see notes)

$$\boldsymbol{z = \frac{x - \mu}{\sigma}}$$

We see that a value of x equal to its mean μ corresponds to $z = 0$, and that a value of x that is one standard deviation above its mean μ, that is, $x = \mu + \sigma$, has z value

$$z = [(\mu + \sigma) - \mu]/\sigma = \sigma/\sigma = 1.$$

Thus, we can interpret z as the number of standard deviations that x is from its mean μ.

Example: find $P(10 \leq x \leq 14)$ for a normal random variable x described by $\mu = 10$ and $\sigma = 2$.

$$P(10 \leq x \leq 14) = P\left(\frac{10 - \mu}{\sigma} \leq z \leq \frac{14 - \mu}{\sigma}\right) = P\left(\frac{10 - 10}{2} \leq z \leq \frac{14 - 10}{2}\right)$$

$$= P(0 \leq z \leq 2) = P(z \leq 2) - P(z \leq 0) = 0.9772 - 0.5000 = 0.4772$$

CENGAGE

6.2 Grear Tire Company Problem

Grear Tire Company has developed a new steel-belted radial tire to be sold through a national chain of discount stores. From tire road tests, Grear's engineering group estimated

mean tire mileage: $\mu = 36{,}500$ miles

standard deviation: $\sigma = 5{,}000$ miles

In addition, the data collected indicate that a normal distribution is a reasonable assumption.

Thus, we define a normal random variable as

x = number of miles the tire will last

Before finalizing the tire mileage guarantee policy, Grear's managers want an answer to the following two questions:

1. What is the probability that the tire mileage will exceed 40,000 miles?
2. What should the guaranteed mileage be if Grear wants no more than 10% of the tires to be eligible for a discount?

6.2 Probability Tire Mileage Exceeds 40,000 Miles

As depicted by the darker area in the figure, we need to find $P(x \geq 40{,}000)$.

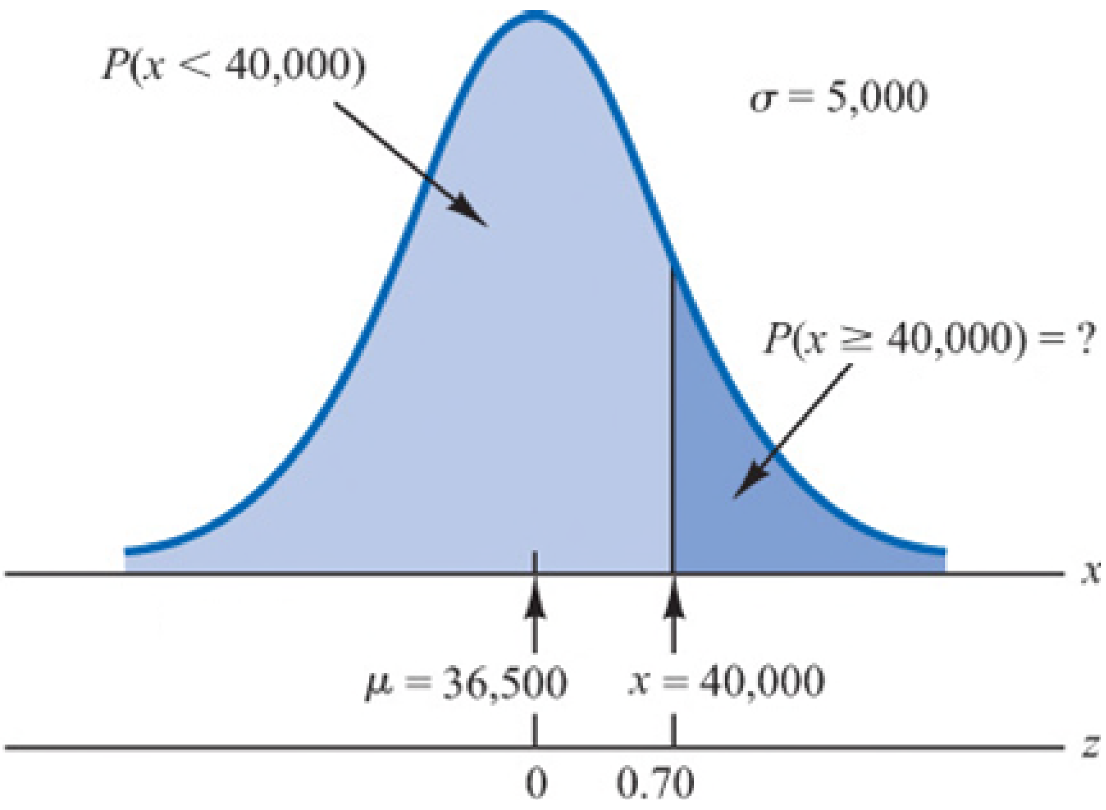
We can convert the normal random variable x to the standard normal random variable z, as

$$z = \frac{x - \mu}{\sigma} = \frac{40{,}000 - 36{,}500}{5{,}000} = \frac{3{,}500}{5{,}000} = 0.70$$

The probability that z is greater than or equal to 0.70 is the complement to the probability that z is less than or equal to 0.70. Thus,

$$P(x \geq 40{,}000) = P(z \geq 0.70)$$
$$= 1 - \mathrm{P}(\mathrm{z} < 0.70) = 1 - 0.7580 = 0.2420$$

We can conclude that about 24.2% of the tires will exceed 40,000 in mileage.



CENGAGE

6.2 Great Tire Guaranteed Mileage

In this inverse problem, we need to find a value of x^* such that: $P(x \leq x^*) = 0.10$

Table 1 of Appendix B shows the closest probability to 0.1000 to be 0.1003, corresponding to the –1.2 row and the 0.08 column. That is, $z^* = -1.28$.

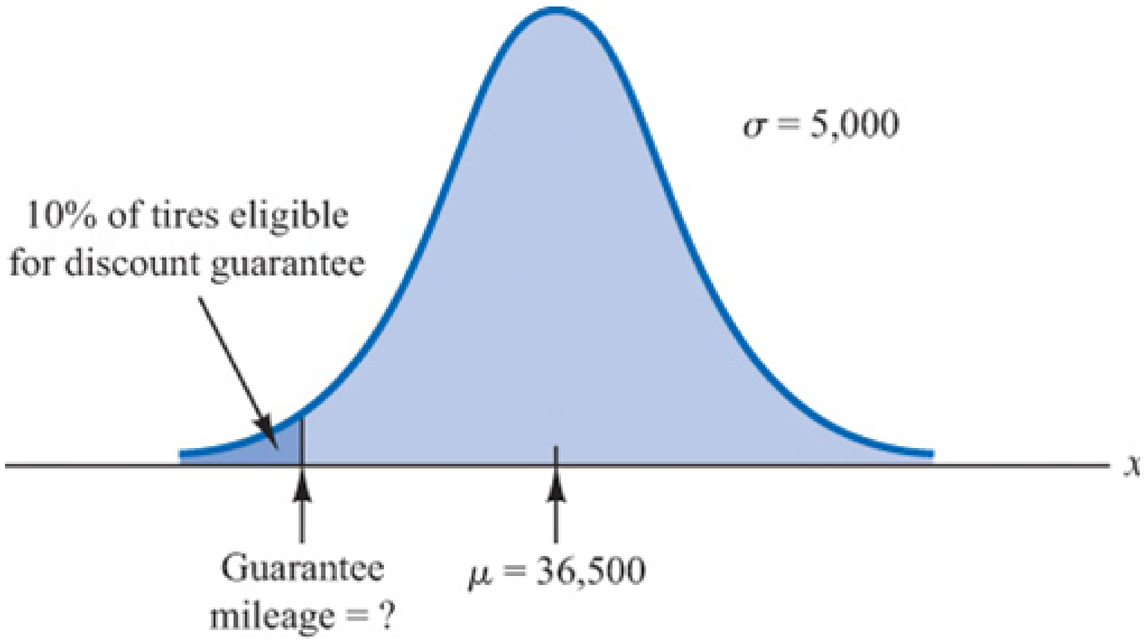
Thus, we have

$$z^* = \frac{x^* - \mu}{\sigma} = -1.28$$

Solving for x^*, we find

$$x^* = \mu - 1.28\sigma = 36{,}500 - 1.28(5{,}000) = 30{,}100 \approx 30{,}000 \text{ miles}$$

Thus, a guarantee of 30,000 miles will meet the requirement that no more than 10% of the tires will be eligible for the discount.



6.3 Normal Approximation of Binomial Probabilities

When the number of trials, n, in a binomial experiment becomes large, evaluating the binomial probability function by hand or with a calculator can become cumbersome.

In cases where $np \geq 5$, and $n(1-p) \geq 5$, where p represents the probability of success in a trial, the normal distribution provides an easy-to-use approximation of binomial probabilities.

When using the normal approximation to the binomial, we set a normal curve defined as

$$\boldsymbol{\mu = np}$$

$$\boldsymbol{\sigma = \sqrt{np(1-p)}}$$

Let us illustrate the normal approximation to the binomial distribution by supposing that a particular company has a history of making errors in 10% of its invoices ($p = 0.10$.)

Using a sample of $n = 100$ invoices, we want to compute the following probabilities:

1. The probability that $x = 12$ of the 100 invoices contain errors.
2. The probability that 13 or fewer ($x \leq 13$) of the 100 invoices contain errors.

6.3 Probability that 12 of 100 Invoices Contain Errors

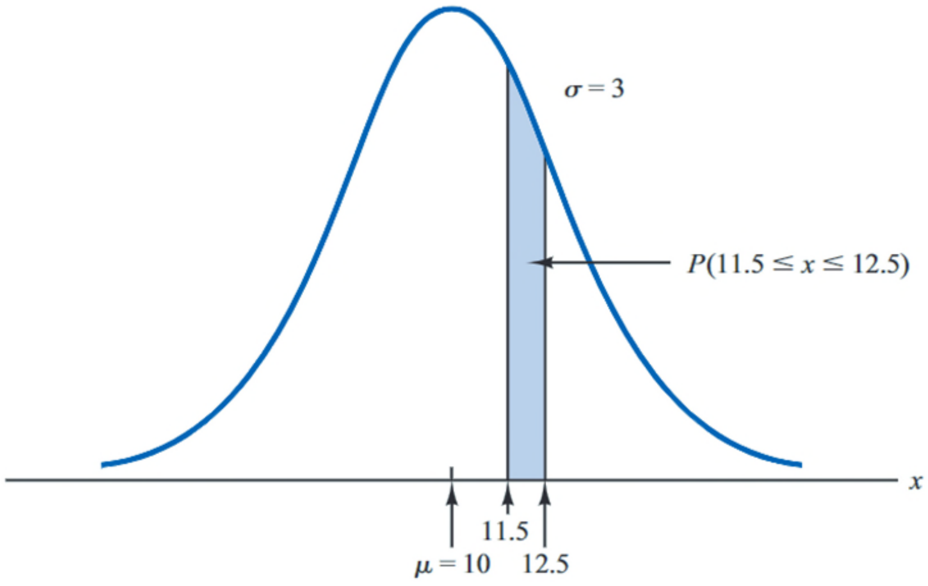
To find the binomial probability of $x = 12$ successes in $n = 100$ trials, we can apply the normal approximation of the binomial distribution because $np = 100(0.10) = 10 \geq 5$, and $n(1-p) = 100(1-0.10) = 90 \geq 5$.

Thus, mean and standard deviation are

$$\mu = np = 100(0.10) = 10$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{100(0.10)(0.90)} = \sqrt{9} = 3$$

To approximate the binomial probability of $x = 12$, we apply the **continuity correction factor** and compute the area under the corresponding normal curve between 11.5 and 12.5 (*see notes.)



$$P(11.5 \leq x \leq 12.5) = P[(11.5 - 10)/3 \leq z \leq (12.5 - 10)/3]$$

$$= P(0.5 \leq z \leq 0.83) = P(z \leq 0.83) - P(z \leq 0.50) = 0.7967 - 0.6915 = 0.1052$$

CENGAGE

6.3 Probability that 13 or Fewer Invoices Contain Errors

The figure shows the area under the normal curve that approximates the probability that 13 of fewer invoices out of a sample of 100 contain errors.

Because of the continuity correction factor introduced in the previous slide, we set $x = 13.5$.

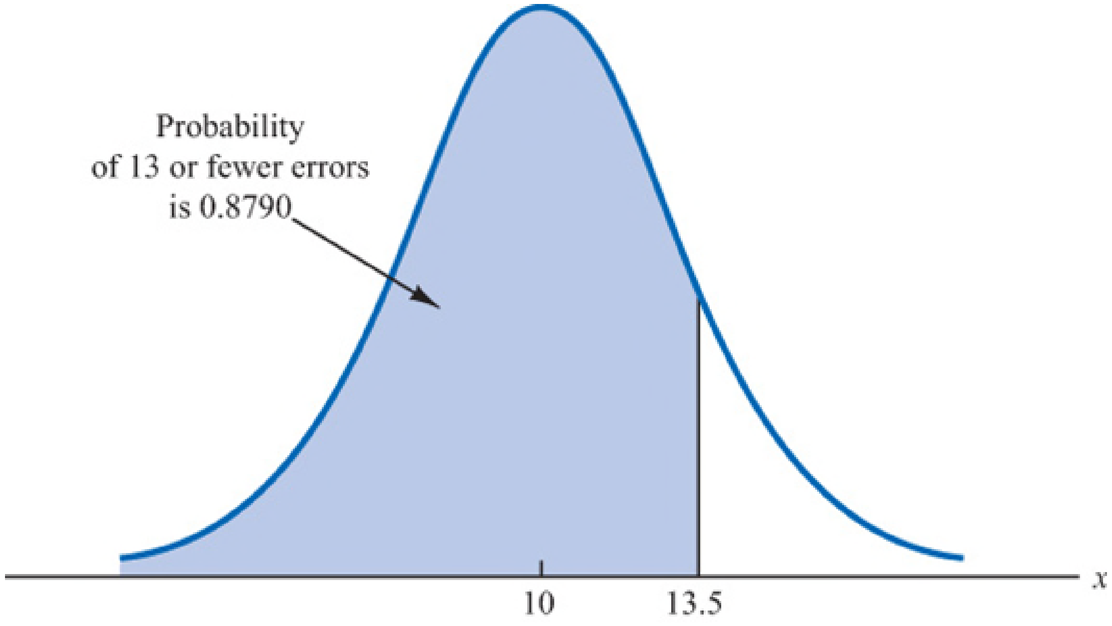
The z-score can be calculated as:

$$z = \frac{x - \mu}{\sigma} = \frac{13.5 - 10}{3} = \frac{3.5}{3} = 1.17$$

Thus, the probability to have 13 or fewer invoices is

$$P(x \leq 13.5) = P(z \leq 1.17) = 0.8790$$

We can conclude that about 88% of the invoices contain 13 or fewer errors.



CENGAGE

6.4 Exponential Probability Distribution

The **exponential probability distribution** may be used for random variables that describe the length of the interval between occurrences, such as the time between arrivals at a hospital emergency room or the distance between major defects in a highway.

The exponential probability density function is

$$\boldsymbol{f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}} \qquad \text{for } x \geq 0$$

Where μ is the expected value or mean.

For example, if the loading time x for a truck at the Schips loading dock has mean $\mu = 15$ minutes, and follows an exponential distribution, the probability density function is

$$f(x) = \frac{1}{15} e^{-\frac{x}{15}}$$

Let us calculate the probability that the loading time x is between 6 and 18 minutes.

CENGAGE

6.4 Exponential Distribution: Cumulative Probability

To compute exponential probabilities, we use the formula for the cumulative probability of obtaining a value for the exponential random variable that is less than or equal to some specific value denoted by x_0.

$$\boldsymbol{P(x \le x_0) = 1 - e^{-\frac{x_0}{\mu}}}$$

For the Schips loading dock example, with x = loading time in minutes and $\mu = 15$ minutes, we have

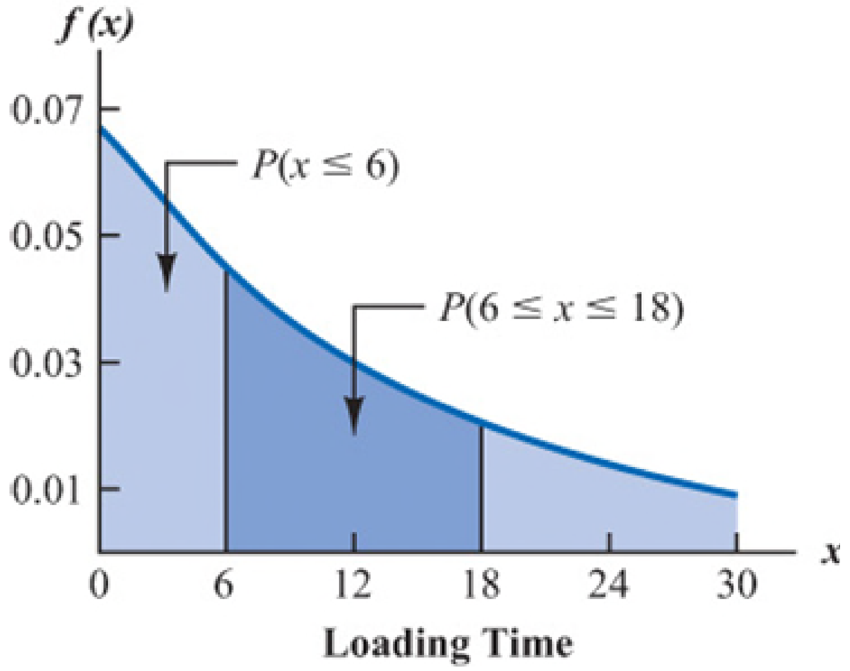
$$P(x \le 6) = 1 - e^{-\frac{6}{15}} = 0.3297$$

$$P(x \le 18) = 1 - e^{-\frac{18}{15}} = 0.6988$$

Thus,

$$P(6 \le x \le 18) = P(x \le 18) - P(x \le 6)$$

$$= 0.6988 - 0.3297 = 0.3691$$



CENGAGE

6.4 Poisson vs. Exponential Distributions

In Chapter 5, we introduced the Poisson distribution as a discrete probability distribution that describes the *number of occurrences per interval*.

Whereas the exponential distribution describes the *length of the interval between occurrences*.

To see how these two distributions are related, consider the arrival of patients at a hospital emergency room with a mean of 10 patients per hour.

The Poisson distribution that describes the number of x arrivals per hour is

$$f(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{10^x e^{-10}}{x!}$$

If, on average 10 patients arrive per hour, it implies that the interval between arrivals has a mean of $1/10 = 0.1$ hour/patient.

Thus, the exponential distribution that describes the time between arrivals is

$$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}} = \frac{1}{0.1} e^{-\frac{x}{0.1}} = 10e^{-10x}$$

CENGAGE

Summary

- In this chapter, we discussed probability distributions for continuous random variables.
- The major conceptual difference between discrete and continuous probability distributions involves the method of computing probabilities.
  - With discrete distributions, the probability function $f(x)$ provides the probability that the random variable x assumes a given value.
  - With continuous distributions, probabilities are given by areas under the curve or graph of the probability density function $f(x)$.
- Because the area under the curve above a single point is zero, we observe that the probability of any particular value is zero for a continuous random variable.
- Three continuous probability distributions—the uniform, normal, and exponential distributions—were considered in detail.
- The normal distribution is used widely in statistical inference and will be used extensively throughout the remainder of the text.

CENGAGE