4-6 Normal Distribution

Definition

A random variable X with probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty \tag{4-8}$$

is a **normal random variable** with parameters μ, where $-\infty < \mu < \infty$, and $\sigma > 0$. Also,

$$E(X) = \mu \quad \text{and} \quad V(X) = \sigma^2 \tag{4-9}$$

and the notation $N(\mu, \sigma^2)$ is used to denote the distribution. The mean and variance of X are shown to equal μ and σ^2, respectively, at the end of this Section 5-6.

4-6 Normal Distribution

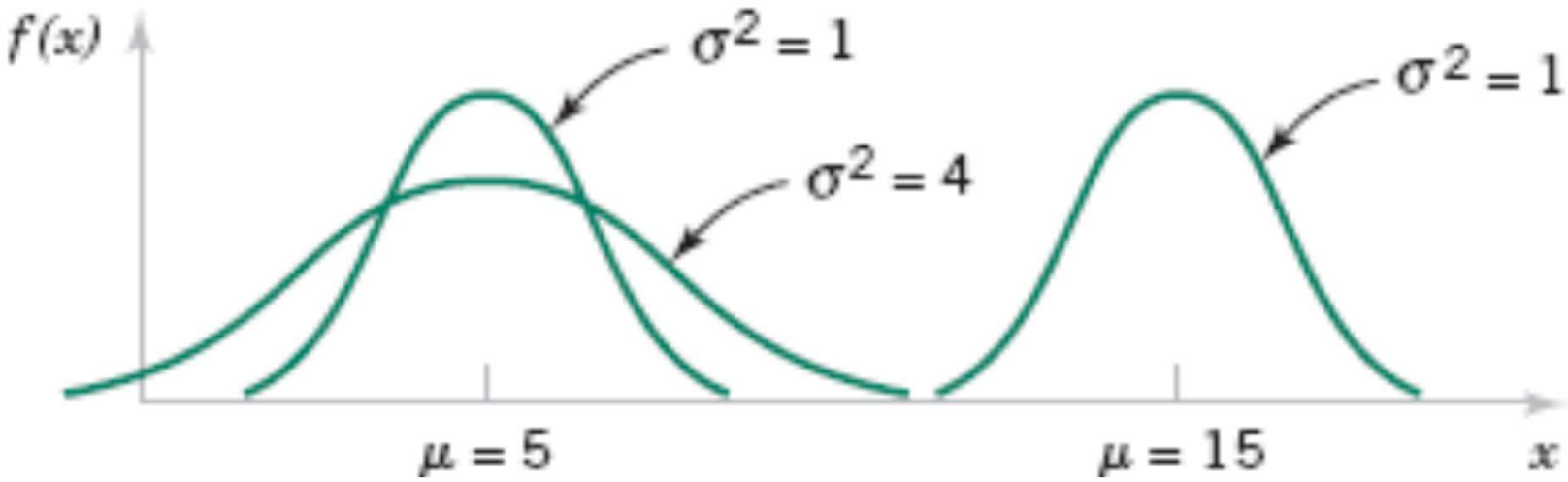


Figure 4-10 Normal probability density functions for selected values of the parameters μ and σ^2.

4-6 Normal Distribution

Some useful results concerning the normal distribution

For any normal random variable,

$$P(\mu - \sigma < X < \mu + \sigma) = 0.6827$$
$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9545$$
$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9973$$

4-6 Normal Distribution

Definition : Standard Normal

A normal random variable with

$$\mu = 0 \quad \text{and} \quad \sigma^2 = 1$$

is called a **standard normal random variable** and is denoted as Z.

The cumulative distribution function of a standard normal random variable is denoted as

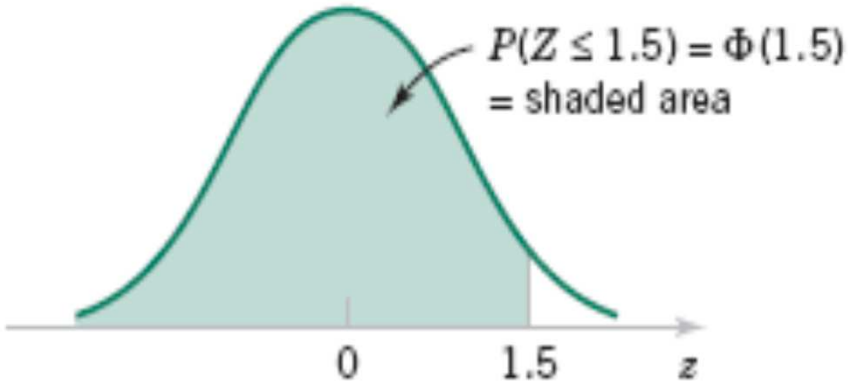
$$\Phi(z) = P(Z \leq z)$$

4-6 Normal Distribution

Example 4-11

Assume Z is a standard normal random variable. Appendix Table II provides probabilities of the form $P(Z \leq z)$. The use of Table II to find $P(Z \leq 1.5)$ is illustrated in Fig. 4-13. Read down the z column to the row that equals 1.5. The probability is read from the adjacent column, labeled 0.00, to be 0.93319.

The column headings refer to the hundredth's digit of the value of z in $P(Z \leq z)$. For example, $P(Z \leq 1.53)$ is found by reading down the z column to the row 1.5 and then selecting the probability from the column labeled 0.03 to be 0.93699.



z	0.00	0.01	0.02	0.03
0	0.50000	0.50399	0.50398	0.51197
⋮		⋮		
1.5	0.93319	0.93448	0.93574	0.93699

Figure 4-13 Standard normal probability density function.

4-6 Normal Distribution

Standardizing

If X is a normal random variable with $E(X) = \mu$ and $V(X) = \sigma^2$, the random variable

$$Z = \frac{X - \mu}{\sigma} \qquad (4\text{-}10)$$

is a normal random variable with $E(Z) = 0$ and $V(Z) = 1$. That is, Z is a standard normal random variable.

4-6 Normal Distribution

Example 4-13

Suppose the current measurements in a strip of wire are assumed to follow a normal distribution with a mean of 10 milliamperes and a variance of 4 (milliamperes)^2. What is the probability that a measurement will exceed 13 milliamperes?

Let X denote the current in milliamperes. The requested probability can be represented as $P(X > 13)$. Let $Z = (X - 10)/2$. The relationship between the several values of X and the transformed values of Z are shown in Fig. 4-15. We note that $X > 13$ corresponds to $Z > 1.5$. Therefore, from Appendix Table II,

$$P(X > 13) = P(Z > 1.5) = 1 - P(Z \leq 1.5) = 1 - 0.93319 = 0.06681$$

Rather than using Fig. 4-15, the probability can be found from the inequality $X > 13$. That is,

$$P(X > 13) = P\left(\frac{(X - 10)}{2} > \frac{(13 - 10)}{2}\right) = P(Z > 1.5) = 0.06681$$

4-6 Normal Distribution

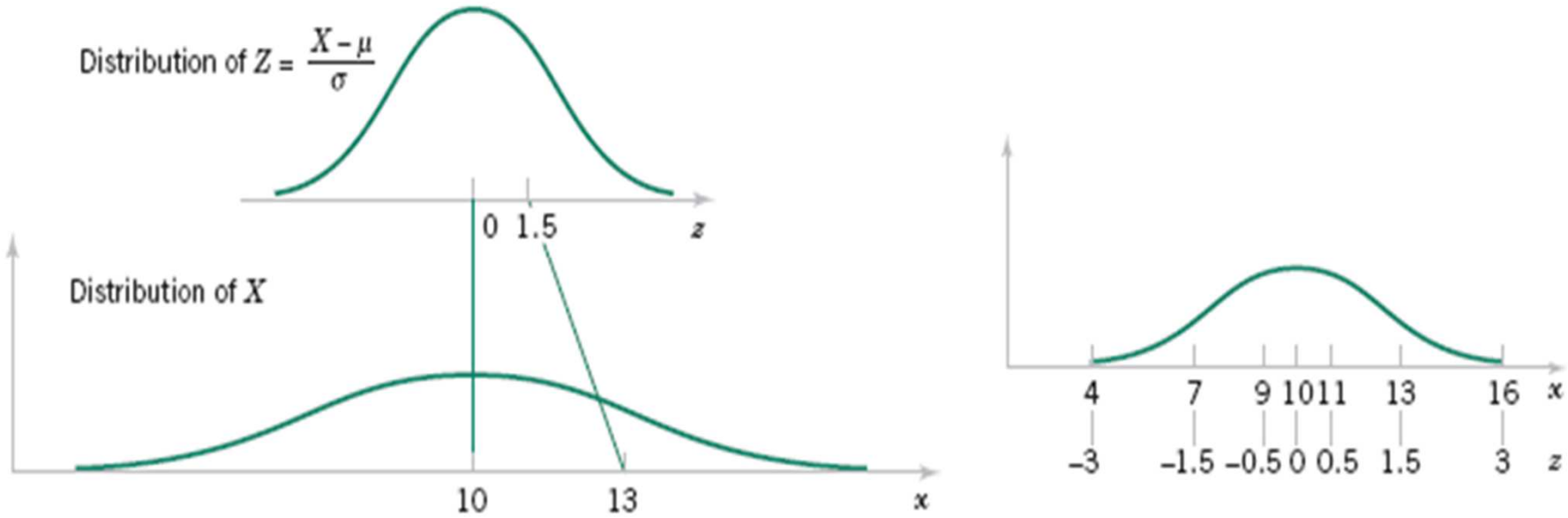


Figure 4-15 Standardizing a normal random variable.

4-6 Normal Distribution

To Calculate Probability

Suppose X is a normal random variable with mean μ and variance σ^2. Then,

$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P(Z \leq z) \qquad (4\text{-}11)$$

where Z is a **standard normal random variable,** and $z = \dfrac{(x-\mu)}{\sigma}$ is the **z-value** obtained by **standardizing** X.

The probability is obtained by entering Appendix Table II with $z = (x-\mu)/\sigma$.

4-6 Normal Distribution

Example 4-14

Continuing the previous example, what is the probability that a current measurement is between 9 and 11 milliamperes? From Fig. 4-15, or by proceeding algebraically, we have

$$\begin{aligned} P(9 < X < 11) &= P((9 - 10)/2 < (X - 10)/2 < (11 - 10)/2) \\ &= P(-0.5 < Z < 0.5) = P(Z < 0.5) - P(Z < -0.5) \\ &= 0.69146 - 0.30854 = 0.38292 \end{aligned}$$

4-6 Normal Distribution

Example 4-14 (continued)

Determine the value for which the probability that a current measurement is below this value is 0.98. The requested value is shown graphically in Fig. 4-16. We need the value of x such that $P(X < x) = 0.98$. By standardizing, this probability expression can be written as

$$\begin{aligned} P(X < x) &= P((X - 10)/2 < (x - 10)/2) \\ &= P(Z < (x - 10)/2) \\ &= 0.98 \end{aligned}$$

Appendix Table II is used to find the z-value such that $P(Z < z) = 0.98$. The nearest probability from Table II results in

$$P(Z < 2.05) = 0.97982$$

Therefore, $(x - 10)/2 = 2.05$, and the standardizing transformation is used in reverse to solve for x. The result is

$$x = 2(2.05) + 10 = 14.1 \text{ milliamperes}$$

4-6 Normal Distribution

Example 4-14 (continued)

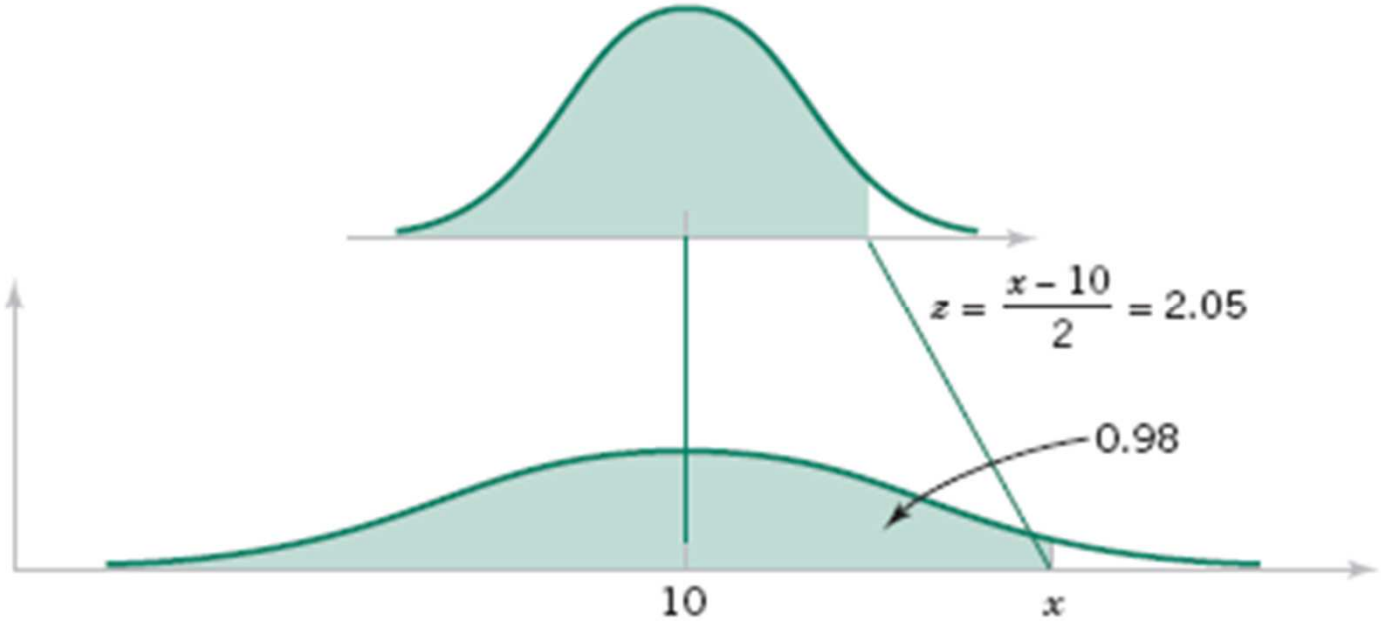


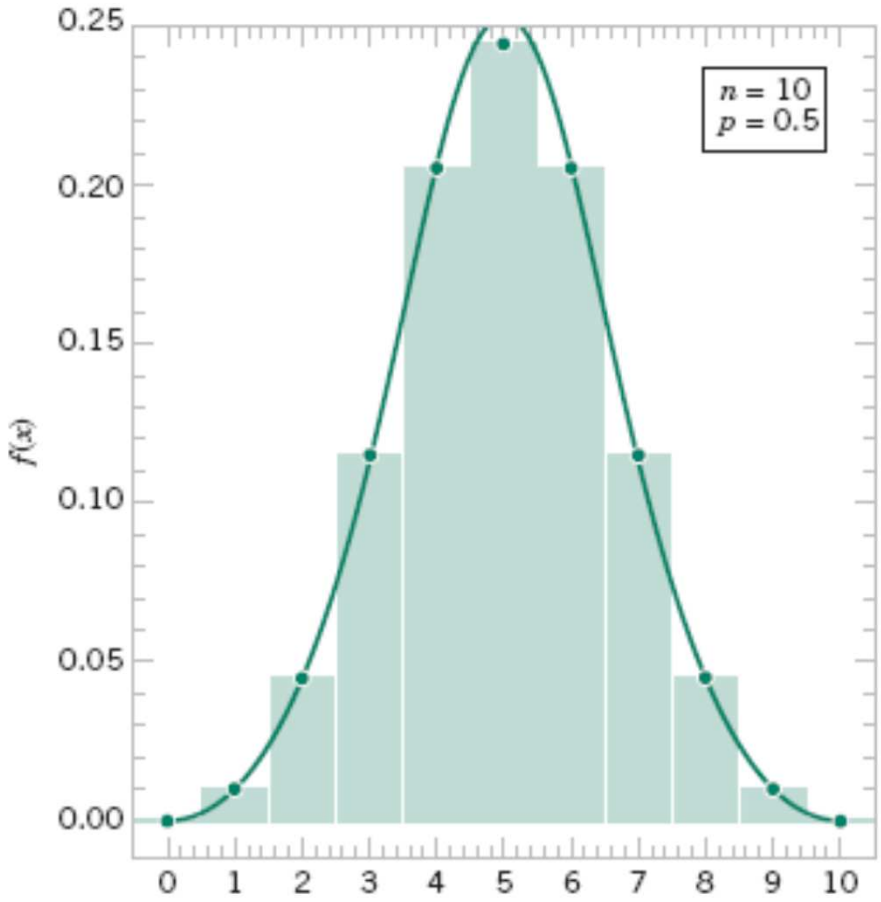
Figure 4-16 Determining the value of x to meet a specified probability.

4-7 Normal Approximation to the Binomial and Poisson Distributions

• Under certain conditions, the normal distribution can be used to approximate the binomial distribution and the Poisson distribution.

4-7 Normal Approximation to the Binomial and Poisson Distributions

Figure 4-19 Normal approximation to the binomial.



4-7 Normal Approximation to the Binomial and Poisson Distributions

Example 4-17

In a digital communication channel, assume that the number of bits received in error can be modeled by a binomial random variable, and assume that the probability that a bit is received in error is 1×10^{-5}. If 16 million bits are transmitted, what is the probability that more than 150 errors occur?

Let the random variable X denote the number of errors. Then X is a binomial random variable and

$$P(X > 150) = 1 - P(x \leq 150) = 1 - \sum_{x=0}^{150} \binom{16{,}000{,}000}{x} (10^{-5})^x (1 - 10^{-5})^{16{,}000{,}000 - x}$$

Clearly, the probability in Example 4-17 is difficult to compute. Fortunately, the normal distribution can be used to provide an excellent approximation in this example.

4-7 Normal Approximation to the Binomial and Poisson Distributions

Normal Approximation to the Binomial Distribution

If X is a binomial random variable, with parameters n and p

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \qquad (4\text{-}12)$$

is approximately a standard normal random variable. To approximate a binomial probability with a normal distribution a **continuity correction** is applied as follows

$$P(X \leq x) = P(X \leq x + 0.5) \cong P\left(Z \leq \frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

and

$$P(x \leq X) = P(x - 0.5 \leq X) \cong P\left(\frac{x - 0.5 - np}{\sqrt{np(1-p)}} \leq Z\right)$$

The approximation is good for $np > 5$ and $n(1-p) > 5$.

4-7 Normal Approximation to the Binomial and Poisson Distributions

Example 4-18

The digital communication problem in the previous example is solved as follows:

$$P(X > 150) = P\left(\frac{X - 160}{\sqrt{160(1 - 10^{-5})}} > \frac{150 - 160}{\sqrt{160(1 - 10^{-5})}}\right)$$

$$= P(Z > -0.79) = P(Z < 0.79) = 0.785$$

Because $np = (16 \times 10^6)(1 \times 10^{-5}) = 160$ and $n(1 - p)$ is much larger, the approximation is expected to work well in this case.

4-7 Normal Approximation to the Binomial and Poisson Distributions

hypergometric distribution $\approx$ ($\frac{n}{N} < 0.1$) binomial distribution $\approx$ ($np > 5$, $n(1-p) > 5$) normal distribution

Figure 4-21 Conditions for approximating hypergeometric and binomial probabilities.

4-7 Normal Approximation to the Binomial and Poisson Distributions

Normal Approximation to the Poisson Distribution

If X is a Poisson random variable with $E(X) = \lambda$ and $V(X) = \lambda$,

$$Z = \frac{X - \lambda}{\sqrt{\lambda}} \qquad (4\text{-}13)$$

is approximately a standard normal random variable. The approximation is good for

$$\lambda > 5$$

4-7 Normal Approximation to the Binomial and Poisson Distributions

Example 4-20

Assume that the number of asbestos particles in a squared meter of dust on a surface follows a Poisson distribution with a mean of 1000. If a squared meter of dust is analyzed, what is the probability that less than 950 particles are found?

This probability can be expressed exactly as

$$P(X \leq 950) = \sum_{x=0}^{950} \frac{e^{-1000} x^{1000}}{x!}$$

The computational difficulty is clear. The probability can be approximated as

$$P(X \leq x) = P\left(Z \leq \frac{950 - 1000}{\sqrt{1000}}\right) = P(Z \leq -1.58) = 0.057$$

4-8 Exponential Distribution

Definition

The random variable X that equals the distance between successive events of a Poisson process with mean $\lambda > 0$ is an **exponential random variable** with parameter λ. The probability density function of X is

$$f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad 0 \leq x < \infty \qquad (4\text{-}14)$$

4-8 Exponential Distribution

Mean and Variance

If the random variable X has an exponential distribution with parameter λ,

$$\mu = E(X) = \frac{1}{\lambda} \quad \text{and} \quad \sigma^2 = V(X) = \frac{1}{\lambda^2} \tag{4-15}$$

It is important to **use consistent units** in the calculation of probabilities, means, and variances involving exponential random variables. The following example illustrates unit conversions.

4-8 Exponential Distribution

Example 4-21

In a large corporate computer network, user log-ons to the system can be modeled as a Poisson process with a mean of 25 log-ons per hour. What is the probability that there are no log-ons in an interval of 6 minutes?

Let X denote the time in hours from the start of the interval until the first log-on. Then, X has an exponential distribution with $\lambda = 25$ log-ons per hour. We are interested in the probability that X exceeds 6 minutes. Because λ is given in log-ons per hour, we express all time units in hours. That is, 6 minutes $= 0.1$ hour. The probability requested is shown as the shaded area under the probability density function in Fig. 4-23. Therefore,

$$P(X > 0.1) = \int_{0.1}^{\infty} 25e^{-25x}\,dx = e^{-25(0.1)} = 0.082$$

4-8 Exponential Distribution

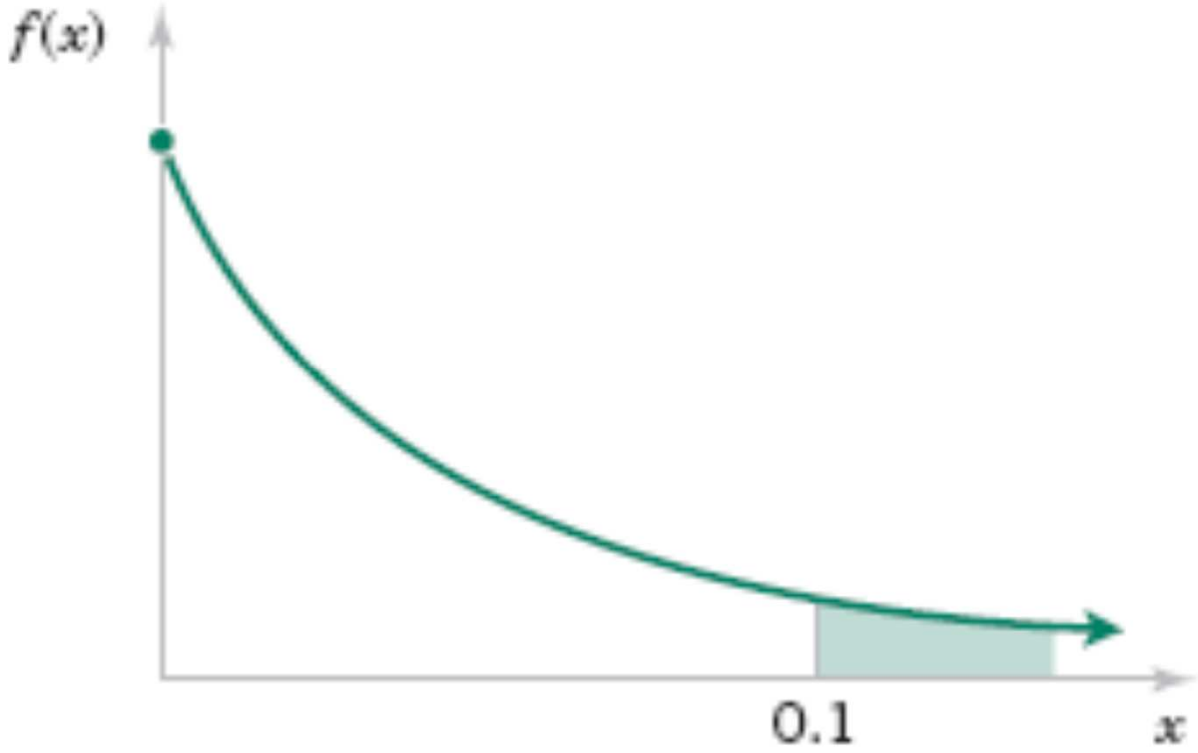


Figure 4-23 Probability for the exponential distribution in Example 4-21.

4-8 Exponential Distribution

Example 4-21 (continued)

Also, the cumulative distribution function can be used to obtain the same result as follows:

$$P(X > 0.1) = 1 - F(0.1) = e^{-25(0.1)}$$

An identical answer is obtained by expressing the mean number of log-ons as 0.417 log-ons per minute and computing the probability that the time until the next log-on exceeds 6 minutes. Try it.

What is the probability that the time until the next log-on is between 2 and 3 minutes? Upon converting all units to hours,

$$P(0.033 < X < 0.05) = \int_{0.033}^{0.05} 25e^{-25x}\,dx = -e^{-25x}\Big|_{0.033}^{0.05} = 0.152$$

4-8 Exponential Distribution

Example 4-21 (continued)

An alternative solution is

$$P(0.033 < X < 0.05) = F(0.05) - F(0.033) = 0.152$$

Determine the interval of time such that the probability that no log-on occurs in the i
val is 0.90. The question asks for the length of time x such that $P(X > x) = 0.90$. Now,

$$P(X > x) = e^{-25x} = 0.90$$

Take the (natural) log of both sides to obtain $-25x = \ln(0.90) = -0.1054$. Therefore,

$$x = 0.00421 \text{ hour} = 0.25 \text{ minute}$$

4-8 Exponential Distribution

Example 4-21 (continued)

Furthermore, the mean time until the next log-on is

$$\mu = 1/25 = 0.04 \text{ hour} = 2.4 \text{ minutes}$$

The standard deviation of the time until the next log-on is

$$\sigma = 1/25 \text{ hours} = 2.4 \text{ minutes}$$

4-8 Exponential Distribution

Our starting point for observing the system does not matter.

•An even more interesting property of an exponential random variable is the **lack of memory property**.

In Example 4-21, suppose that there are no log-ons from 12:00 to 12:15; the probability that there are no log-ons from 12:15 to 12:21 is still 0.082. Because we have already been waiting for 15 minutes, we feel that we are "due." That is, the probability of a log-on in the next 6 minutes should be greater than 0.082. ***However, for an exponential distribution this is not true.***

4-8 Exponential Distribution

Example 4-22

Let X denote the time between detections of a particle with a geiger counter and assume that X has an exponential distribution with $\lambda = 1.4$ minutes. The probability that we detect a particle within 30 seconds of starting the counter is

$$P(X < 0.5 \text{ minute}) = F(0.5) = 1 - e^{-0.5/1.4} = 0.30$$

In this calculation, all units are converted to minutes. Now, suppose we turn on the geiger counter and wait 3 minutes without detecting a particle. What is the probability that a particle is detected in the next 30 seconds?

4-8 Exponential Distribution

Example 4-22 (continued)

Because we have already been waiting for 3 minutes, we feel that we are "due." is, the probability of a detection in the next 30 seconds should be greater than 0.3. How for an exponential distribution, this is not true. The requested probability can be expre as the conditional probability that $P(X < 3.5 | X > 3)$. From the definition of conditi probability,

$$P(X < 3.5 | X > 3) = P(3 < X < 3.5)/P(X > 3)$$

where

$$P(3 < X < 3.5) = F(3.5) - F(3) = [1 - e^{-3.5/1.4}] - [1 - e^{-3/1.4}] = 0.0035$$

and

$$P(X > 3) = 1 - F(3) = e^{-3/1.4} = 0.117$$

4-8 Exponential Distribution

Example 4-22 (continued)

Therefore,

$$P(X < 3.5 \mid X > 3) = 0.035/0.117 = 0.30$$

After waiting for 3 minutes without a detection, the probability of a detection in the next 30 seconds is the same as the probability of a detection in the 30 seconds immediately after starting the counter. The fact that you have waited 3 minutes without a detection does not change the probability of a detection in the next 30 seconds.

4-8 Exponential Distribution

Example 4-22 illustrates the **lack of memory property** of an exponential random variable and a general statement of the property follows. In fact, the exponential distribution is the only continuous distribution with this property.

Lack of Memory Property

For an exponential random variable X,

$$P(X < t_1 + t_2 | X > t_1) = P(X < t_2) \qquad (4\text{-}16)$$

4-8 Exponential Distribution

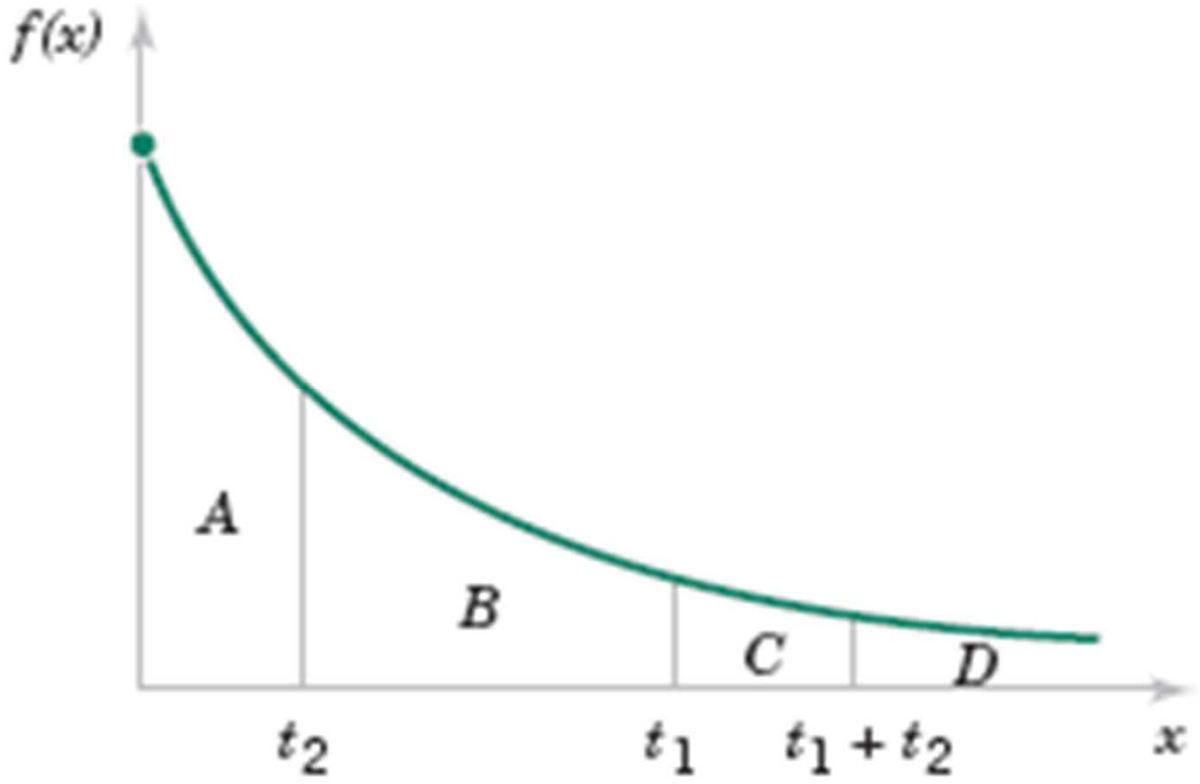


Figure 4-24 Lack of memory property of an Exponential distribution.