3

Continuous Random Variables

Continuous Sample Space

Until now, we have studied discrete random variables. By definition the range of a discrete random variable is a countable set of numbers. This chapter analyzes random variables that range over continuous sets of numbers. A continuous set of numbers, sometimes referred to as an *interval*, contains all of the real numbers between two limits. For the limits x_1 and x_2 with $x_1 < x_2$, there are four different intervals distinguished by which of the limits are contained in the interval. Thus we have definitions and notation for the four continuous sets bounded by the lower limit x_1 and upper limit x_2.

- (x_1, x_2) is the open interval defined as all numbers between x_1 and x_2 but not including either x_1 or x_2. Formally, $(x_1, x_2) = \{x|x_1 < x < x_2\}$.
- $[x_1, x_2]$ is the closed interval defined as all numbers between x_1 and x_2 including both x_1 and x_2. Formally $[x_1, x_2] = \{x|x_1 \le x \le x_2\}$.
- $[x_1, x_2)$ is the interval defined as all numbers between x_1 and x_2 including x_1 but not including x_2. Formally, $[x_1, x_2) = \{x|x_1 \le x < x_2\}$.
- $(x_1, x_2]$ is the interval defined as all numbers between x_1 and x_2 including x_2 but not including x_1. Formally, $(x_1, x_2] = \{x|x_1 < x \le x_2\}$.

Many experiments lead to random variables with a range that is a continuous interval. Examples include measuring T, the arrival time of a particle ($S_T = \{t|0 \le t < \infty\}$); measuring V, the voltage across a resistor ($S_V = \{v| -\infty < v < \infty\}$); and measuring the phase angle A of a sinusoidal radio wave ($S_A = \{a|0 \le a < 2\pi\}$). We will call T, V, and A *continuous random variables* although we will defer a formal definition until Section 3.1.

Consistent with the axioms of probability, we assign numbers between zero and one to all events (sets of elements) in the sample space. A distinguishing feature of the models of continuous random variables is that the probability of each individual outcome is zero! To understand this intuitively, consider an experiment in which the observation is the arrival time of the professor at a class. Assume this professor always arrives between 8:55 and 9:05. We model the arrival time as a random variable T minutes relative to 9:00 o'clock. Therefore, $S_T = \{t| -5 \le t \le 5\}$. Think about predicting the professor's arrival time.

The more precise the prediction, the lower the chance it will be correct. For example, you might guess the interval $-1 \leq T \leq 1$ minute (8:59 to 9:01). Your probability of being correct is higher than if you guess $-0.5 \leq T \leq 0.5$ minute (8:59:30 to 9:00:30). As your prediction becomes more and more precise, the probability that it will be correct gets smaller and smaller. The chance that the professor will arrive within a femtosecond of 9:00 is microscopically small (on the order of 10^{-15}), and the probability of a precise 9:00 arrival is zero.

One way to think about continuous random variables is that the *amount of probability* in an interval gets smaller and smaller as the interval shrinks. This is like the mass in a continuous volume. Even though any finite volume has some mass, there is no mass at a single point. In physics, we analyze this situation by referring to densities of matter. Similarly, we refer to *probability density functions* to describe probabilities related to continuous random variables. The next section introduces these ideas formally by describing an experiment in which the sample space contains all numbers between zero and one.

In many practical applications of probability, we encounter uniform random variables. The sample space of a uniform random variable is an interval with finite limits. The probability model of a uniform random variable states that any two intervals of equal size within the sample space have equal probability. To introduce many concepts of continuous random variables, we will refer frequently to a uniform random variable with limits 0 and 1. Most computer languages include a random number generator. In MATLAB, this is the `rand` function introduced in Chapter 1. These random number generators produce pseudo-random numbers that approximate sample values of a uniform random variable.

In the following example, we examine this random variable by defining an experiment in which the procedure is to spin a pointer in a circle of circumference one meter. This model is very similar to the model of the phase angle of the signal that arrives at the radio receiver of a cellular telephone. Instead of a pointer with stopping points that can be anywhere between 0 and 1 meter, the phase angle can have any value between 0 and 2π radians. By referring to the spinning pointer in the examples in this chapter, we arrive at mathematical expressions that illustrate the main properties of continuous random variables. The formulas that arise from analyzing phase angles in communications engineering models have factors of 2π that do not appear in the examples in this chapter. Example 3.1 defines the sample space of the pointer experiment and demonstrates that all outcomes have probability zero.

Example 3.1 Suppose we have a wheel of circumference one meter and we mark a point on the perimeter at the top of the wheel. In the center of the wheel is a radial pointer that we spin. After spinning the pointer, we measure the distance, X meters, around the circumference of the wheel going clockwise from the marked point to the pointer position as shown in Figure 3.1. Clearly, $0 \leq X < 1$. Also, it is reasonable to believe that if the spin is hard enough, the pointer is just as likely to arrive at any part of the circle as at any other. For a given x, what is the probability $P[X = x]$?

....

This problem is surprisingly difficult. However, given that we have developed methods for discrete random variables in Chapter 2, a reasonable approach is to find a discrete approximation to X. As shown on the right side of Figure 3.1, we can mark the perimeter with n equal-length arcs numbered 1 to n and let Y denote the number of the arc in which the pointer stops. Y is a discrete random variable with range $S_Y = \{1, 2, \ldots, n\}$. Since all parts of the wheel are equally likely, all arcs have the

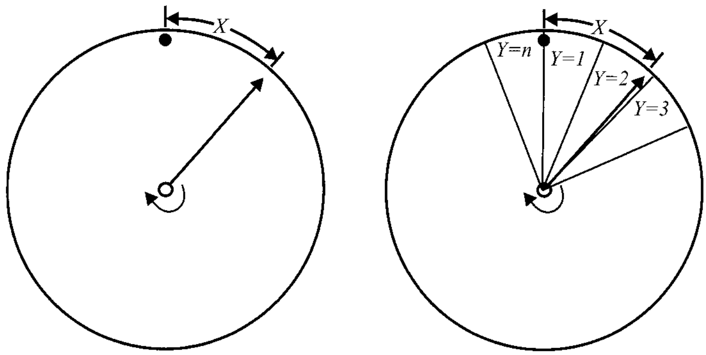


Figure 3.1 The random pointer on disk of circumference 1.

same probability. Thus the PMF of Y is

$$P_Y(y) = \begin{cases} 1/n & y = 1, 2, \ldots, n, \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

From the wheel on the right side of Figure 3.1, we can deduce that if $X = x$, then $Y = \lceil nx \rceil$, where the notation $\lceil a \rceil$ is defined as the smallest integer greater than or equal to a. Note that the event $\{X = x\} \subset \{Y = \lceil nx \rceil\}$, which implies that

$$P[X = x] \leq P[Y = \lceil nx \rceil] = \frac{1}{n}. \tag{3.2}$$

We observe this is true no matter how finely we divide up the wheel. To find $P[X = x]$, we consider larger and larger values of n. As n increases, the arcs on the circle decrease in size, approaching a single point. The probability of the pointer arriving in any particular arc decreases until we have in the limit,

$$P[X = x] \leq \lim_{n\to\infty} P[Y = \lceil nx \rceil] = \lim_{n\to\infty} \frac{1}{n} = 0. \tag{3.3}$$

This demonstrates that $P[X = x] \leq 0$. The first axiom of probability states that $P[X = x] \geq 0$. Therefore, $P[X = x] = 0$. This is true regardless of the outcome, x. It follows that every outcome has probability zero.

Just as in the discussion of the professor arriving in class, similar reasoning can be applied to other experiments to show that for any continuous random variable, the probability of any individual outcome is zero. This is a fundamentally different situation than the one we encountered in our study of discrete random variables. Clearly a probability mass function defined in terms of probabilities of individual outcomes has no meaning in this context. For a continuous random variable, the interesting probabilities apply to intervals.

3.1 The Cumulative Distribution Function

Example 3.1 shows that when X is a continuous random variable, $P[X = x] = 0$ for $x \in S_X$. This implies that when X is continuous, it is impossible to define a probability mass function $P_X(x)$. On the other hand, we will see that the cumulative distribution function, $F_X(x)$ in Definition 2.11, is a very useful probability model for a continuous random variable. We repeat the definition here.

Definition 3.1 ***Cumulative Distribution Function (CDF)***

The ***cumulative distribution function*** *(CDF) of random variable X is*

$$F_X(x) = P[X \leq x].$$

The key properties of the CDF, described in Theorem 2.2 and Theorem 2.3, apply to *all* random variables. Graphs of all cumulative distribution functions start at zero on the left and end at one on the right. All are nondecreasing, and, most importantly, the probability that the random variable is in an interval is the difference in the CDF evaluated at the ends of the interval.

Theorem 3.1 *For any random variable X,*

(a) $F_X(-\infty) = 0$

(b) $F_X(\infty) = 1$

(c) $P[x_1 < X \leq x_2] = F_X(x_2) - F_X(x_1)$

Although these properties apply to any CDF, there is one important difference between the CDF of a discrete random variable and the CDF of a continuous random variable. Recall that for a discrete random variable X, $F_X(x)$ has zero slope everywhere except at values of x with nonzero probability. At these points, the function has a discontinuity in the form of a jump of magnitude $P_X(x)$. By contrast, the defining property of a continuous random variable X is that $F_X(x)$ is a continuous function of X.

Definition 3.2 ***Continuous Random Variable***

X is a ***continuous random variable*** *if the CDF $F_X(x)$ is a continuous function.*

Example 3.2 In the wheel-spinning experiment of Example 3.1, find the CDF of X.

We begin by observing that any outcome $x \in S_X = [0, 1)$. This implies that $F_X(x) = 0$ for $x < 0$, and $F_X(x) = 1$ for $x \geq 1$. To find the CDF for x between 0 and 1 we consider the event $\{X \leq x\}$ with x growing from 0 to 1. Each event corresponds to an arc on the circle in Figure 3.1. The arc is small when $x \approx 0$ and it includes nearly the whole circle when $x \approx 1$. $F_X(x) = P[X \leq x]$ is the probability that the pointer stops somewhere in the arc. This probability grows from 0 to 1 as the arc increases to include the whole circle. Given our assumption that the pointer has no preferred stopping places, it is reasonable to expect the probability to grow in proportion to the fraction of the circle occupied by the arc $X \le x$. This fraction is simply x. To be more formal, we can refer to Figure 3.1 and note that with the circle divided into n arcs,

$$\{Y \le \lceil nx \rceil - 1\} \subset \{X \le x\} \subset \{Y \le \lceil nx \rceil\}. \tag{3.4}$$

Therefore, the probabilities of the three events satisfy

$$F_Y(\lceil nx \rceil - 1) \le F_X(x) \le F_Y(\lceil nx \rceil). \tag{3.5}$$

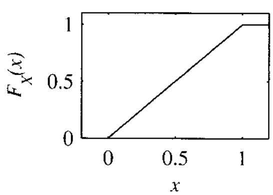
Note that Y is a discrete random variable with CDF

$$F_Y(y) = \begin{cases} 0 & y < 0, \\ k/n & (k-1)/n < y \le k/n, k = 1, 2, \ldots, n, \\ 1 & y > 1. \end{cases} \tag{3.6}$$

Thus for $x \in [0, 1)$ and for all n, we have

$$\frac{\lceil nx \rceil - 1}{n} \le F_X(x) \le \frac{\lceil nx \rceil}{n}. \tag{3.7}$$

In Problem 3.1.4, we ask the reader to verify that $\lim_{n\to\infty} \lceil nx \rceil / n = x$. This implies that as $n \to \infty$, both fractions approach x. The CDF of X is



$$F_X(x) = \begin{cases} 0 & x < 0, \\ x & 0 \le x < 1, \\ 1 & x \ge 1. \end{cases} \tag{3.8}$$

Quiz 3.1 *The cumulative distribution function of the random variable Y is*

$$F_Y(y) = \begin{cases} 0 & y < 0, \\ y/4 & 0 \le y \le 4, \\ 1 & y > 4. \end{cases} \tag{3.9}$$

Sketch the CDF of Y and calculate the following probabilities:

(1) $P[Y \le -1]$ *(2)* $P[Y \le 1]$

(3) $P[2 < Y \le 3]$ *(4)* $P[Y > 1.5]$

3.2 Probability Density Function

The slope of the CDF contains the most interesting information about a continuous random variable. The slope at any point x indicates the probability that X is *near* x. To understand this intuitively, consider the graph of a CDF $F_X(x)$ given in Figure 3.2. Theorem 3.1(c) states that the probability that X is in the interval of width Δ to the right of x_1 is

$$p_1 = P[x_1 < X \le x_1 + \Delta] = F_X(x_1 + \Delta) - F_X(x_1). \tag{3.10}$$

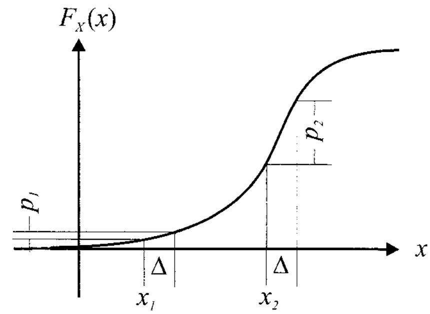


Figure 3.2 The graph of an arbitrary CDF $F_X(x)$.

Note in Figure 3.2 that this is less than the probability of the interval of width Δ to the right of x_2,

$$p_2 = P[x_2 < X \leq x_2 + \Delta] = F_X(x_2 + \Delta) - F_X(x_2). \tag{3.11}$$

The comparison makes sense because both intervals have the same length. If we reduce Δ to focus our attention on outcomes nearer and nearer to x_1 and x_2, both probabilities get smaller. However, their relative values still depend on the average slope of $F_X(x)$ at the two points. This is apparent if we rewrite Equation (3.10) in the form

$$P[x_1 < X \leq x_1 + \Delta] = \frac{F_X(x_1 + \Delta) - F_X(x_1)}{\Delta}\Delta. \tag{3.12}$$

Here the fraction on the right side is the average slope, and Equation (3.12) states that the probability that a random variable is in an interval near x_1 is the average slope over the interval times the length of the interval. By definition, the limit of the average slope as $\Delta \to 0$ is the derivative of $F_X(x)$ evaluated at x_1.

We conclude from the discussion leading to Equation (3.12) that the slope of the CDF in a region near any number x is an indicator of the probability of observing the random variable X near x. Just as the amount of matter in a small volume is the density of the matter times the size of volume, the amount of probability in a small region is the slope of the CDF times the size of the region. This leads to the term *probability density*, defined as the slope of the CDF.

Definition 3.3 ***Probability Density Function (PDF)***

*The **probability density function** (PDF) of a continuous random variable X is*

$$f_X(x) = \frac{dF_X(x)}{dx}.$$

This definition displays the conventional notation for a PDF. The name of the function is a lowercase f with a subscript that is the name of the random variable. As with the PMF and the CDF, the argument is a dummy variable: $f_X(x)$, $f_X(u)$, and $f_X(\cdot)$ are all the same PDF.

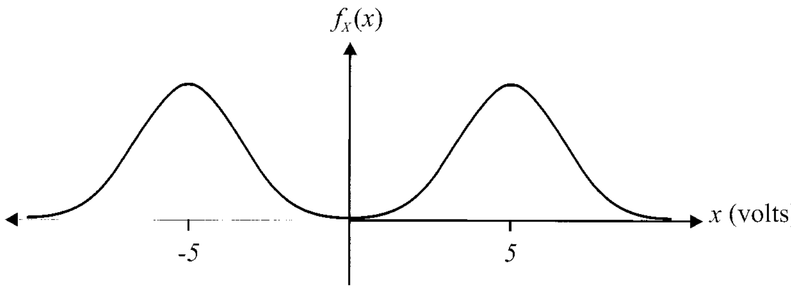


Figure 3.3 The PDF of the modem receiver voltage X.

The PDF is a complete probability model of a continuous random variable. While there are other functions that also provide complete models (the CDF and the moment generating function that we study in Chapter 6), the PDF is the most useful. One reason for this is that the graph of the PDF provides a good indication of the likely values of observations.

Example 3.3 Figure 3.3 depicts the PDF of a random variable X that describes the voltage at the receiver in a modem. What are probable values of X?

Note that there are two places where the PDF has high values and that it is low elsewhere. The PDF indicates that the random variable is likely to be near -5 V (corresponding to the symbol 0 transmitted) and near $+5$ V (corresponding to a 1 transmitted). Values far from ± 5 V (due to strong distortion) are possible but much less likely.

Another reason why the PDF is the most useful probability model is that it plays a key role in calculating the expected value of a random variable, the subject of the next section. Important properties of the PDF follow directly from Definition 3.3 and the properties of the CDF.

Theorem 3.2 *For a continuous random variable X with PDF $f_X(x)$,*

(a) $f_X(x) \geq 0$ *for all* x,

(b) $F_X(x) = \int_{-\infty}^{x} f_X(u)\,du$,

(c) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$.

Proof The first statement is true because $F_X(x)$ is a nondecreasing function of x and therefore its derivative, $f_X(x)$, is nonnegative. The second fact follows directly from the definition of $f_X(x)$ and the fact that $F_X(-\infty) = 0$. The third statement follows from the second one and Theorem 3.1(b).

Given these properties of the PDF, we can prove the next theorem, which relates the PDF to the probabilities of events.

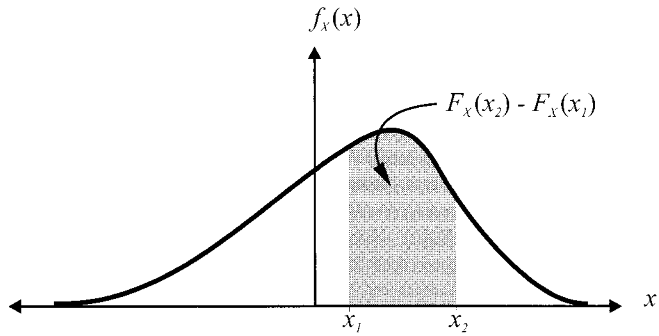


Figure 3.4 The PDF and CDF of X.

Theorem 3.3

$$P[x_1 < X \le x_2] = \int_{x_1}^{x_2} f_X(x)\, dx.$$

Proof From Theorem 3.2(b) and Theorem 3.1,

$$P[x_1 < X \le x_2] = P[X \le x_2] - P[X \le x_1] = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)\, dx. \quad (3.13)$$

Theorem 3.3 states that the probability of observing X in an interval is the area under the PDF graph between the two end points of the interval. This property of the PDF is depicted in Figure 3.4. Theorem 3.2(c) states that the area under the entire PDF graph is one. Note that the value of the PDF can be any nonnegative number. It is not a probability and need not be between zero and one. To gain further insight into the PDF, it is instructive to reconsider Equation (3.12). For very small values of Δ, the right side of Equation (3.12) approximately equals $f_X(x_1)\Delta$. When Δ becomes the infinitesimal dx, we have

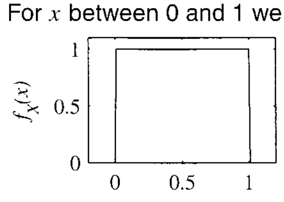
$$P[x < X \le x + dx] = f_X(x)\, dx. \quad (3.14)$$

Equation (3.14) is useful because it permits us to interpret the integral of Theorem 3.3 as the limiting case of a sum of probabilities of events $\{x < X \le x + dx\}$.

Example 3.4 For the experiment in Examples 3.1 and 3.2, find the PDF of X and the probability of the event $\{1/4 < X \le 3/4\}$.

. . . .

Taking the derivative of the CDF in Equation (3.8), $f_X(x) = 0$, when $x < 0$ or $x \ge 1$. For x between 0 and 1 we have $f_X(x) = dF_X(x)/dx = 1$. Thus the PDF of X is



$$f_X(x) = \begin{cases} 1 & 0 \le x < 1, \\ 0 & \text{otherwise.} \end{cases} \quad (3.15)$$

The fact that the PDF is constant over the range of possible values of X reflects the fact that the pointer has no favorite stopping places on the circumference of the circle. To find the probability that X is between $1/4$ and $3/4$, we can use either Theorem 3.1 or Theorem 3.3. Thus

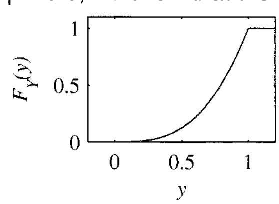
$$P[1/4 < X \leq 3/4] = F_X(3/4) - F_X(1/4) = 1/2, \tag{3.16}$$

and equivalently,

$$P[1/4 < X \leq 3/4] = \int_{1/4}^{3/4} f_X(x)\,dx = \int_{1/4}^{3/4} dx = 1/2. \tag{3.17}$$

When the PDF and CDF are both known it is easier to use the CDF to find the probability of an interval. However, in many cases we begin with the PDF, in which case it is usually easiest to use Theorem 3.3 directly. The alternative is to find the CDF explicitly by means of Theorem 3.2(b) and then to use Theorem 3.1.

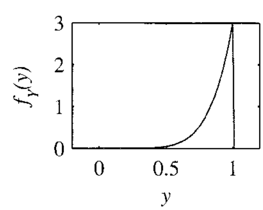
Example 3.5 Consider an experiment that consists of spinning the pointer in Example 3.1 three times and observing Y meters, the maximum value of X in the three spins. In Example 5.8, we show that the CDF of Y is



$$F_Y(y) = \begin{cases} 0 & y < 0, \\ y^3 & 0 \leq y \leq 1, \\ 1 & y > 1. \end{cases} \tag{3.18}$$

Find the PDF of Y and the probability that Y is between $1/4$ and $3/4$.

Applying Definition 3.3,



$$f_Y(y) = \begin{cases} dF_Y(y)/dy = 3y^2 & 0 < y \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.19}$$

Note that the PDF has values between 0 and 3. Its integral between any pair of numbers is less than or equal to 1. The graph of $f_Y(y)$ shows that there is a higher probability of finding Y at the right side of the range of possible values than at the left side. This reflects the fact that the maximum of three spins produces higher numbers than individual spins. Either Theorem 3.1 or Theorem 3.3 can be used to calculate the probability of observing Y between $1/4$ and $3/4$:

$$P[1/4 < Y \leq 3/4] = F_Y(3/4) - F_Y(1/4) = (3/4)^3 - (1/4)^3 = 13/32, \tag{3.20}$$

and equivalently,

$$P[1/4 < Y \leq 3/4] = \int_{1/4}^{3/4} f_Y(y)\,dy = \int_{1/4}^{3/4} 3y^2\,dy = 13/32. \tag{3.21}$$

Note that this probability is less than 1/2, which is the probability of $1/4 < X \leq 3/4$ calculated in Example 3.4 for the uniform random variable.

When we work with continuous random variables, it is usually not necessary to be precise about specifying whether or not a range of numbers includes the endpoints. This is because individual numbers have probability zero. In Example 3.2, there are four different sets of numbers defined by the words *X is between 1/4 and 3/4*:

$$A = (1/4, 3/4), \quad B = (1/4, 3/4], \quad C = [1/4, 3/4), \quad D = [1/4, 3/4]. \tag{3.22}$$

While they are all different events, they all have the same probability because they differ only in whether they include $\{X = 1/4\}$, $\{X = 3/4\}$, or both. Since these two sets have zero probability, their inclusion or exclusion does not affect the probability of the range of numbers. This is quite different from the situation we encounter with discrete random variables. Consider random variable Y with PMF

$$P_Y(y) = \begin{cases} 1/6 & y = 1/4, y = 1/2, \\ 2/3 & y = 3/4, \\ 0 & \text{otherwise.} \end{cases} \tag{3.23}$$

For this random variable Y, the probabilities of the four sets are

$$P[A] = 1/6, \quad P[B] = 5/6, \quad P[C] = 1/3, \quad P[D] = 1. \tag{3.24}$$

So we see that the nature of an inequality in the definition of an event does not affect the probability when we examine continuous random variables. With discrete random variables, it is critically important to examine the inequality carefully.

If we compare other characteristics of discrete and continuous random variables, we find that with discrete random variables, many facts are expressed as sums. With continuous random variables, the corresponding facts are expressed as integrals. For example, when X is discrete,

$$P[B] = \sum_{x \in B} P_X(x). \qquad \text{(Theorem 2.1(c))}$$

When X is continuous and $B = [x_1, x_2]$,

$$P[x_1 < X \leq x_2] = \int_{x_1}^{x_2} f_X(x)\, dx. \qquad \text{(Theorem 3.3)}$$

Quiz 3.2 *Random variable X has probability density function*

$$f_X(x) = \begin{cases} cxe^{-x/2} & x \geq 0, \\ 0 & \textit{otherwise}. \end{cases} \tag{3.25}$$

Sketch the PDF and find the following:

(1) the constant c *(2) the CDF $F_X(x)$*

(3) $P[0 \leq X \leq 4]$ *(4) $P[-2 \leq X \leq 2]$*

3.3 Expected Values

The primary reason that random variables are useful is that they permit us to compute averages. For a discrete random variable Y, the expected value,

$$E[Y] = \sum_{y_i \in S_Y} y_i P_Y(y_i), \tag{3.26}$$

is a sum of the possible values y_i, each multiplied by its probability. For a continuous random variable X, this definition is inadequate because all possible values of X have probability zero. However, we can develop a definition for the expected value of the continuous random variable X by examining a discrete approximation of X. For a small Δ, let

$$Y = \Delta \left\lfloor \frac{X}{\Delta} \right\rfloor, \tag{3.27}$$

where the notation $\lfloor a \rfloor$ denotes the largest integer less than or equal to a. Y is an approximation to X in that $Y = k\Delta$ if and only if $k\Delta \le X < k\Delta + \Delta$. Since the range of Y is $S_Y = \{\ldots, -\Delta, 0, \Delta, 2\Delta, \ldots\}$, the expected value is

$$E[Y] = \sum_{k=-\infty}^{\infty} k\Delta P[Y = k\Delta] = \sum_{k=-\infty}^{\infty} k\Delta P[k\Delta \le X < k\Delta + \Delta]. \tag{3.28}$$

As Δ approaches zero and the intervals under consideration grow smaller, Y more closely approximates X. Furthermore, $P[k\Delta \le X < k\Delta + \Delta]$ approaches $f_X(k\Delta)\Delta$ so that for small Δ,

$$E[X] \approx \sum_{k=-\infty}^{\infty} k\Delta f_X(k\Delta)\Delta. \tag{3.29}$$

In the limit as Δ goes to zero, the sum converges to the integral in Definition 3.4.

Definition 3.4 ***Expected Value***
*The **expected value** of a continuous random variable X is*

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx.$$

When we consider Y, the discrete approximation of X, the intuition developed in Section 2.5 suggests that $E[Y]$ is what we will observe if we add up a very large number n of independent observations of Y and divide by n. This same intuition holds for the continuous random variable X. As $n \to \infty$, the average of n independent samples of X will approach $E[X]$. In probability theory, this observation is known as the *Law of Large Numbers*, Theorem 7.8.

Example 3.6 In Example 3.4, we found that the stopping point X of the spinning wheel experiment was a uniform random variable with PDF

$$f_X(x) = \begin{cases} 1 & 0 \le x < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.30}$$

Find the expected stopping point $E[X]$ of the pointer.

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^1 x\, dx = 1/2 \text{ meter.} \tag{3.31}$$

With no preferred stopping points on the circle, the average stopping point of the pointer is exactly half way around the circle.

Example 3.7 In Example 3.5, find the expected value of the maximum stopping point Y of the three spins:

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \int_0^1 y(3y^2)\, dy = 3/4 \text{ meter.} \tag{3.32}$$

Corresponding to functions of discrete random variables described in Section 2.6, we have functions $g(X)$ of a continuous random variable X. A function of a continuous random variable is also a random variable; however, this random variable is not necessarily continuous!

Example 3.8 Let X be a uniform random variable with PDF

$$f_X(x) = \begin{cases} 1 & 0 \le x < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.33}$$

Let $W = g(X) = 0$ if $X \le 1/2$, and $W = g(X) = 1$ if $X > 1/2$. W is a discrete random variable with range $S_W = \{0, 1\}$.

Regardless of the nature of the random variable $W = g(X)$, its expected value can be calculated by an integral that is analogous to the sum in Theorem 2.10 for discrete random variables.

Theorem 3.4 *The expected value of a function, $g(X)$, of random variable X is*

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx.$$

Many of the properties of expected values of discrete random variables also apply to continuous random variables. Definition 2.16 and Theorems 2.11, 2.12, 2.13, and 2.14 apply to all random variables. All of these relationships are written in terms of expected values. We can summarize these relationships in the following theorem.

Theorem 3.5 *For any random variable X,*

(a) $E[X - \mu_X] = 0$,

(b) $E[aX + b] = aE[X] + b$,

(c) $\text{Var}[X] = E[X^2] - \mu_X^2$,

(d) $\text{Var}[aX + b] = a^2 \text{Var}[X]$.

The method of calculating expected values depends on the type of random variable, discrete or continuous. Theorem 3.4 states that $E[X^2]$, the second moment of X, and $\text{Var}[X]$ are the integrals

$$E\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx, \qquad \text{Var}[X] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\, dx. \tag{3.34}$$

Our interpretation of expected values of discrete random variables carries over to continuous random variables. $E[X]$ represents a typical value of X, and the variance describes the dispersion of outcomes relative to the expected value. Furthermore, if we view the PDF $f_X(x)$ as the density of a mass distributed on a line, then $E[X]$ is the center of mass.

Example 3.9 Find the variance and standard deviation of the pointer position in Example 3.1.

To compute $\text{Var}[X]$, we use Theorem 3.5(c): $\text{Var}[X] = E[X^2] - \mu_X^2$. We calculate $E[X^2]$ directly from Theorem 3.4 with $g(X) = X^2$:

$$E\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_0^1 x^2\, dx = 1/3. \tag{3.35}$$

In Example 3.6, we have $E[X] = 1/2$. Thus $\text{Var}[X] = 1/3 - (1/2)^2 = 1/12$, and the standard deviation is $\sigma_X = \sqrt{\text{Var}[X]} = 1/\sqrt{12} = 0.289$ meters.

Example 3.10 Find the variance and standard deviation of Y, the maximum pointer position after three spins, in Example 3.5.

We proceed as in Example 3.9. We have $f_Y(y)$ from Example 3.5 and $E[Y] = 3/4$ from Example 3.7:

$$E\left[Y^2\right] = \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy = \int_0^1 y^2 \left(3y^2\right) dy = 3/5. \tag{3.36}$$

Thus the variance is

$$\text{Var}[Y] = 3/5 - (3/4)^2 = 3/80 \text{ m}^2, \tag{3.37}$$

and the standard deviation is $\sigma_Y = 0.194$ meters.

Quiz 3.3 *The probability density function of the random variable Y is*

$$f_Y(y) = \begin{cases} 3y^2/2 & -1 \le y \le 1, \\ 0 & \textit{otherwise}. \end{cases} \tag{3.38}$$

Sketch the PDF and find the following:

(1) the expected value $E[Y]$ *(2) the second moment $E[Y^2]$*

(3) the variance $\text{Var}[Y]$ *(4) the standard deviation σ_Y*

3.4 Families of Continuous Random Variables

Section 2.3 introduces several families of discrete random variables that arise in a wide variety of practical applications. In this section, we introduce three important families of continuous random variables: uniform, exponential, and Erlang. We devote all of Section 3.5 to Gaussian random variables. Like the families of discrete random variables, the PDFs of the members of each family all have the same mathematical form. They differ only in the values of one or two parameters. We have already encountered an example of a continuous *uniform random variable* in the wheel-spinning experiment. The general definition is

Definition 3.5 ***Uniform Random Variable***

X is a uniform (a, b) random variable if the PDF of X is

$$f_X(x) = \begin{cases} 1/(b-a) & a \le x < b, \\ 0 & \textit{otherwise}, \end{cases}$$

where the two parameters are $b > a$.

Expressions that are synonymous with X *is a uniform random variable* are X *is uniformly distributed* and X *has a uniform distribution*.

If X is a uniform random variable there is an equal probability of finding an outcome x in any interval of length $\Delta < b - a$ within $S_X = [a, b)$. We can use Theorem 3.2(b), Theorem 3.4, and Theorem 3.5 to derive the following properties of a uniform random variable.

Theorem 3.6 *If X is a uniform (a, b) random variable,*

(a) The CDF of X is

$$F_X(x) = \begin{cases} 0 & x \le a, \\ (x-a)/(b-a) & a < x \le b, \\ 1 & x > b. \end{cases}$$

(b) The expected value of X is $E[X] = (b + a)/2$.

(c) The variance of X is $\text{Var}[X] = (b - a)^2/12$.

Example 3.11 The phase angle, Θ, of the signal at the input to a modem is uniformly distributed between 0 and 2π radians. Find the CDF, the expected value, and the variance of Θ.

From the problem statement, we identify the parameters of the uniform (a, b) random variable as $a = 0$ and $b = 2\pi$. Therefore the PDF of Θ is

$$f_\Theta(\theta) = \begin{cases} 1/(2\pi) & 0 \le \theta < 2\pi, \\ 0 & \text{otherwise}. \end{cases} \tag{3.39}$$

The CDF is

$$F_\Theta(\theta) = \begin{cases} 0 & \theta \le 0, \\ \theta/(2\pi) & 0 < x \le 2\pi, \\ 1 & x > 2\pi. \end{cases} \tag{3.40}$$

The expected value is $E[\Theta] = b/2 = \pi$ radians, and the variance is $\text{Var}[\Theta] = (2\pi)^2/12 = \pi^2/3$ rad^2.

The relationship between the family of discrete uniform random variables and the family of continuous uniform random variables is fairly direct. The following theorem expresses the relationship formally.

Theorem 3.7 *Let X be a uniform (a, b) random variable, where a and b are both integers. Let $K = \lceil X \rceil$. Then K is a discrete uniform $(a + 1, b)$ random variable.*

Proof Recall that for any x, $\lceil x \rceil$ is the smallest integer greater than or equal to x. It follows that the event $\{K = k\} = \{k - 1 < x \leq k\}$. Therefore,

$$P[K = k] = P_K(k) = \int_{k-1}^{k} P_X(x)\, dx = \begin{cases} 1/(b-a) & k = a+1, a+2, \ldots, b, \\ 0 & \text{otherwise.} \end{cases} \tag{3.41}$$

This expression for $P_K(k)$ conforms to Definition 2.9 of a discrete uniform $(a + 1, b)$ PMF.

The continuous relatives of the family of geometric random variables, Definition 2.6, are the members of the family of *exponential random variables*.

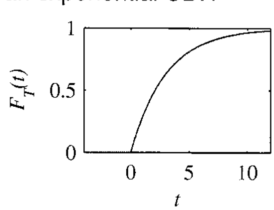
Definition 3.6 ***Exponential Random Variable***

*X is an **exponential** (λ) **random variable** if the PDF of X is*

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & \textit{otherwise,} \end{cases}$$

where the parameter $\lambda > 0$.

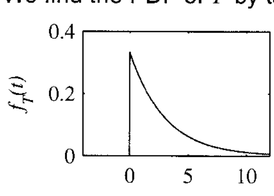
Example 3.12 The probability that a telephone call lasts no more than t minutes is often modeled as an exponential CDF.



$$F_T(t) = \begin{cases} 1 - e^{-t/3} & t \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.42}$$

What is the PDF of the duration in minutes of a telephone conversation? What is the probability that a conversation will last between 2 and 4 minutes?

We find the PDF of T by taking the derivative of the CDF:



$$f_T(t) = \frac{dF_T(t)}{dt} = \begin{cases} (1/3)e^{-t/3} & t \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.43}$$

Therefore, observing Definition 3.6, we recognize that T is an exponential $(\lambda = 1/3)$ random variable. The probability that a call lasts between 2 and 4 minutes is

$$P[2 \le T \le 4] = F_4(4) - F_2(2) = e^{-2/3} - e^{-4/3} = 0.250. \tag{3.44}$$

Example 3.13 In Example 3.12, what is $E[T]$, the expected duration of a telephone call? What are the variance and standard deviation of T? What is the probability that a call duration is within ± 1 standard deviation of the expected call duration?

Using the PDF $f_T(t)$ in Example 3.12, we calculate the expected duration of a call:

$$E[T] = \int_{-\infty}^{\infty} t f_T(t)\, dt = \int_0^{\infty} t \frac{1}{3} e^{-t/3}\, dt. \tag{3.45}$$

Integration by parts (Appendix B, Math Fact B.10) yields

$$E[T] = -te^{-t/3}\Big|_0^{\infty} + \int_0^{\infty} e^{-t/3}\, dt = 3 \text{ minutes.} \tag{3.46}$$

To calculate the variance, we begin with the second moment of T:

$$E\left[T^2\right] = \int_{-\infty}^{\infty} t^2 f_T(t)\, dt = \int_0^{\infty} t^2 \frac{1}{3} e^{-t/3}\, dt. \tag{3.47}$$

Again integrating by parts, we have

$$E\left[T^2\right] = -t^2 e^{-t/3}\Big|_0^{\infty} + \int_0^{\infty} (2t) e^{-t/3}\, dt = 2\int_0^{\infty} te^{-t/3}\, dt. \tag{3.48}$$

With the knowledge that $E[T] = 3$, we observe that $\int_0^{\infty} te^{-t/3}\, dt = 3E[T] = 9$. Thus $E[T^2] = 6E[T] = 18$ and

$$\text{Var}[T] = E\left[T^2\right] - (E[T])^2 = 18 - 3^2 = 9. \tag{3.49}$$

The standard deviation is $\sigma_T = \sqrt{\text{Var}[T]} = 3$ minutes. The probability that the call duration is within 1 standard deviation of the expected value is

$$P[0 \le T \le 6] = F_T(6) - F_T(0) = 1 - e^{-2} = 0.865 \tag{3.50}$$

To derive general expressions for the CDF, the expected value, and the variance of an exponential random variable, we apply Theorem 3.2(b), Theorem 3.4, and Theorem 3.5 to the exponential PDF in Definition 3.6.

Theorem 3.8 *If X is an exponential (λ) random variable,*

(a) $F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$

(b) $E[X] = 1/\lambda$.

(c) $\text{Var}[X] = 1/\lambda^2$.

The following theorem shows the relationship between the family of exponential random variables and the family of geometric random variables.

Theorem 3.9 *If X is an exponential (λ) random variable, then $K = \lceil X \rceil$ is a geometric (p) random variable with $p = 1 - e^{-\lambda}$.*

Proof As in the proof of Theorem 3.7, the definition of K implies $P_K(k) = P[k - 1 < X \leq k]$. Referring to the CDF of X in Theorem 3.8, we observe

$$P_K(k) = F_x(k) - F_x(k-1) = e^{-\lambda(k-1)} - e^{-\lambda k} = (e^{-\lambda})^{k-1}(1 - e^{-\lambda}). \tag{3.51}$$

If we let $p = 1 - e^{-\lambda}$, we have $P_K(k) = p(1-p)^{k-1}$, which conforms to Definition 2.6 of a geometric (p) random variable with $p = 1 - e^{-\lambda}$.

Example 3.14 Phone company A charges $0.15 per minute for telephone calls. For any fraction of a minute at the end of a call, they charge for a full minute. Phone Company B also charges $0.15 per minute. However, Phone Company B calculates its charge based on the exact duration of a call. If T, the duration of a call in minutes, is an exponential ($\lambda = 1/3$) random variable, what are the expected revenues per call $E[R_A]$ and $E[R_B]$ for companies A and B?

. . . .

Because T is an exponential random variable, we have in Theorem 3.8 (and in Example 3.13), $E[T] = 1/\lambda = 3$ minutes per call. Therefore, for phone company B, which charges for the exact duration of a call,

$$E[R_B] = 0.15E[T] = \$0.45 \text{ per call.} \tag{3.52}$$

Company A, by contrast, collects $\$0.15\lceil T \rceil$ for a call of duration T minutes. Theorem 3.9 states that $K = \lceil T \rceil$ is a geometric random variable with parameter $p = 1 - e^{-1/3}$. Therefore, the expected revenue for Company A is

$$E[R_A] = 0.15E[K] = 0.15/p = (0.15)(3.53) = \$0.529 \text{ per call.} \tag{3.53}$$

In Theorem 6.11, we show that the sum of a set of independent identically distributed exponential random variables is an *Erlang* random variable.

Definition 3.7 ***Erlang Random Variable***

*X is an **Erlang** (n, λ) random variable if the PDF of X is*

$$f_X(x) = \begin{cases} \dfrac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

where the parameter $\lambda > 0$, and the parameter $n \geq 1$ is an integer.

The parameter n is often called the *order* of an Erlang random variable. Problem 3.4.10 outlines a procedure to verify that the integral of the Erlang PDF over all x is 1. The Erlang $(n = 1, \lambda)$ random variable is identical to the exponential (λ) random variable. Just as exponential random variables are related to geometric random variables, the family of Erlang continuous random variables is related to the family of Pascal discrete random variables.

Theorem 3.10 *If X is an Erlang (n, λ) random variable, then*

$$E[X] = \frac{n}{\lambda}, \qquad \text{Var}[X] = \frac{n}{\lambda^2}.$$

By comparing Theorem 3.8 and Theorem 3.10, we see for X, an Erlang (n, λ) random variable, and Y, an exponential (λ) random variable, that $E[X] = nE[Y]$ and $\text{Var}[X] = n\,\text{Var}[Y]$. In the following theorem, we can also connect Erlang and Poisson random variables.

Theorem 3.11 *Let K_α denote a Poisson (α) random variable. For any $x > 0$, the CDF of an Erlang (n, λ) random variable X satisfies*

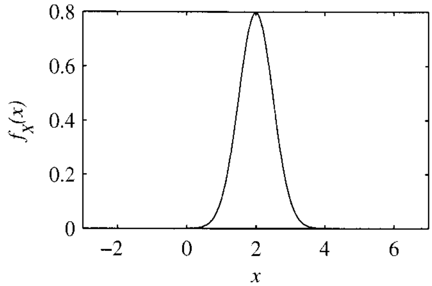
$$F_X(x) = 1 - F_{K_{\lambda x}}(n-1) = 1 - \sum_{k=0}^{n-1} \frac{(\lambda x)^k e^{-\lambda x}}{k!}.$$

Problem 3.4.12 outlines a proof of Theorem 3.11.

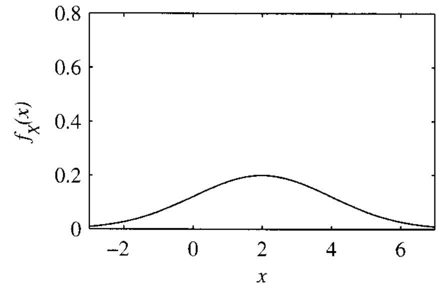
Quiz 3.4 *Continuous random variable X has $E[X] = 3$ and $\text{Var}[X] = 9$. Find the PDF, $f_X(x)$, if*
(1) X has an exponential PDF, *(2) X has a uniform PDF.*

3.5 Gaussian Random Variables

Bell-shaped curves appear in many applications of probability theory. The probability models in these applications are members of the family of *Gaussian random variables.* Chapter 6 contains a mathematical explanation for the prevalence of Gaussian random variables in models of practical phenomena. Because they occur so frequently in practice, Gaussian random variables are sometimes referred to as *normal* random variables.



(a) $\mu = 2, \sigma = 1/2$



(b) $\mu = 2, \sigma = 2$

Figure 3.5 Two examples of a Gaussian random variable X with expected value μ and standard deviation σ.

Definition 3.8 ***Gaussian Random Variable***

X is a Gaussian (μ, σ) random variable if the PDF of X is

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

where the parameter μ can be any real number and the parameter $\sigma > 0$.

Many statistics texts use the notation *X is $N[\mu, \sigma^2]$* as shorthand for *X is a Gaussian (μ, σ) random variable.* In this notation, the N denotes *normal*. The graph of $f_X(x)$ has a bell shape, where the center of the bell is $x = \mu$ and σ reflects the width of the bell. If σ is small, the bell is narrow, with a high, pointy peak. If σ is large, the bell is wide, with a low, flat peak. (The height of the peak is $1/\sigma\sqrt{2\pi}$.) Figure 3.5 contains two examples of Gaussian PDFs with $\mu = 2$. In Figure 3.5(a), $\sigma = 0.5$, and in Figure 3.5(b), $\sigma = 2$. Of course, the area under any Gaussian PDF is $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$. Furthermore, the parameters of the PDF are the expected value and the standard deviation of X.

Theorem 3.12 *If X is a Gaussian (μ, σ) random variable,*

$$E[X] = \mu \qquad \text{Var}[X] = \sigma^2.$$

The proof of Theorem 3.12, as well as the proof that the area under a Gaussian PDF is 1, employs integration by parts and other calculus techniques. We leave them as an exercise for the reader in Problem 3.5.9.

It is impossible to express the integral of a Gaussian PDF between noninfinite limits as a function that appears on most scientific calculators. Instead, we usually find integrals of the Gaussian PDF by referring to tables, such as Table 3.1, that have been obtained by numerical integration. To learn how to use this table, we introduce the following important property of Gaussian random variables.

Theorem 3.13 *If X is Gaussian (μ, σ), $Y = aX + b$ is Gaussian $(a\mu + b, a\sigma)$.*

The theorem states that any linear transformation of a Gaussian random variable produces another Gaussian random variable. This theorem allows us to relate the properties of an arbitrary Gaussian random variable to the properties of a specific random variable.

Definition 3.9 ***Standard Normal Random Variable***
*The **standard normal random variable** Z is the Gaussian $(0, 1)$ random variable.*

Theorem 3.12 indicates that $E[Z] = 0$ and $\text{Var}[Z] = 1$. The tables that we use to find integrals of Gaussian PDFs contain values of $F_Z(z)$, the CDF of Z. We introduce the special notation $\Phi(z)$ for this function.

Definition 3.10 ***Standard Normal CDF***
The CDF of the standard normal random variable Z is

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2}\, du.$$

Given a table of values of $\Phi(z)$, we use the following theorem to find probabilities of a Gaussian random variable with parameters μ and σ.

Theorem 3.14 *If X is a Gaussian (μ, σ) random variable, the CDF of X is*

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

The probability that X is in the interval $(a, b]$ is

$$P[a < X \leq b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

In using this theorem, we transform values of a Gaussian random variable, X, to equivalent values of the standard normal random variable, Z. For a sample value x of the random variable X, the corresponding sample value of Z is

$$z = \frac{x-\mu}{\sigma} \tag{3.54}$$

Note that z is dimensionless. It represents x as a number of standard deviations relative to the expected value of X. Table 3.1 presents $\Phi(z)$ for $0 \leq z \leq 2.99$. People working with probability and statistics spend a lot of time referring to tables like Table 3.1. It seems strange to us that $\Phi(z)$ isn't included in every scientific calculator. For many people, it is far more useful than many of the functions included in ordinary scientific calculators.

Example 3.15 Suppose your score on a test is $x = 46$, a sample value of the Gaussian $(61, 10)$ random variable. Express your test score as a sample value of the standard normal random variable, Z.

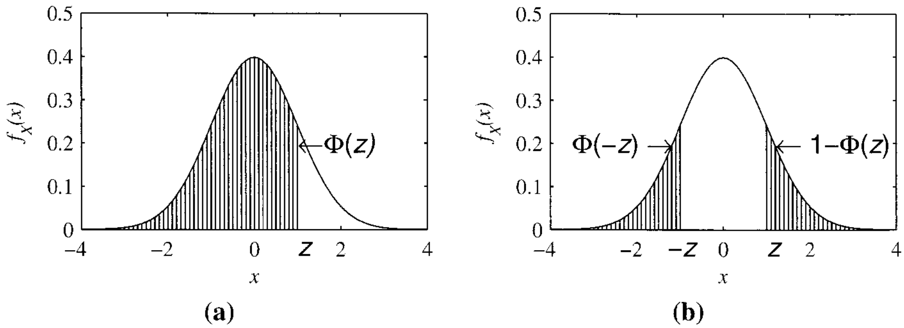


Figure 3.6 Symmetry properties of th Gaussian (0, 1) PDF.

Equation (3.54) indicates that $z = (46 - 61)/10 = -1.5$. Therefore your score is 1.5 *standard deviations less than the expected value.*

To find probabilities of Gaussian random variables, we use the values of $\Phi(z)$ presented in Table 3.1. Note that this table contains entries only for $z \geq 0$. For negative values of z, we apply the following property of $\Phi(z)$.

Theorem 3.15

$$\Phi(-z) = 1 - \Phi(z).$$

Figure 3.6 displays the symmetry properties of $\Phi(z)$. Both graphs contain the standard normal PDF. In Figure 3.6(a), the shaded area under the PDF is $\Phi(z)$. Since the area under the PDF equals 1, the unshaded area under the PDF is $1 - \Phi(z)$. In Figure 3.6(b), the shaded area on the right is $1 - \Phi(z)$ and the shaded area on the left is $\Phi(-z)$. This graph demonstrates that $\Phi(-z) = 1 - \Phi(z)$.

Example 3.16 If X is the Gaussian $(61, 10)$ random variable, what is $P[X \leq 46]$?

Applying Theorem 3.14, Theorem 3.15 and the result of Example 3.15, we have

$$P[X \leq 46] = F_X(46) = \Phi(-1.5) = 1 - \Phi(1.5) = 1 - 0.933 = 0.067. \qquad (3.55)$$

This suggests that if your test score is 1.5 standard deviations below the expected value, you are in the lowest 6.7% of the population of test takers.

Example 3.17 If X is a Gaussian random variable with $\mu = 61$ and $\sigma = 10$, what is $P[51 < X \leq 71]$?

Applying Equation (3.54), $Z = (X - 61)/10$ and the event $\{51 < X \leq 71\}$ corresponds to the event $\{-1 < Z \leq 1\}$. The probability of this event is

$$P[-1 < Z \leq 1] = \Phi(1) - \Phi(-1) = \Phi(1) - [1 - \Phi(1)] = 2\Phi(1) - 1 = 0.683. \qquad (3.56)$$

The solution to Example 3.17 reflects the fact that in an experiment with a Gaussian probability model, 68.3% (about two-thirds) of the outcomes are within ± 1 standard deviation of the expected value. About 95% $(2\Phi(2) - 1)$ of the outcomes are within two standard deviations of the expected value.

Tables of $\Phi(z)$ are useful for obtaining numerical values of integrals of a Gaussian PDF over intervals near the expected value. Regions further than three standard deviations from the expected value (corresponding to $|z| \geq 3$) are in the *tails* of the PDF. When $|z| > 3$, $\Phi(z)$ is very close to one; for example, $\Phi(3) = 0.9987$ and $\Phi(4) = 0.9999768$. The properties of $\Phi(z)$ for extreme values of z are apparent in the *standard normal complementary CDF.*

Definition 3.11 ***Standard Normal Complementary CDF***

The ***standard normal complementary CDF*** *is*

$$Q(z) = P[Z > z] = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-u^2/2}\, du = 1 - \Phi(z).$$

Although we may regard both $\Phi(3) = 0.9987$ and $\Phi(4) = 0.9999768$ as being very close to one, we see in Table 3.2 that $Q(3) = 1.35 \cdot 10^{-3}$ is almost two orders of magnitude larger than $Q(4) = 3.17 \cdot 10^{-5}$.

Example 3.18 In an optical fiber transmission system, the probability of a binary error is $Q(\sqrt{\gamma/2})$, where γ is the signal-to-noise ratio. What is the minimum value of γ that produces a binary error rate not exceeding 10^{-6}?

Referring to Table 3.1, we find that $Q(z) < 10^{-6}$ when $z \geq 4.75$. Therefore, if $\sqrt{\gamma/2} \geq 4.75$, or $\gamma \geq 45$, the probability of error is less than 10^{-6}.

Keep in mind that $Q(z)$ is the probability that a Gaussian random variable exceeds its expected value by more than z standard deviations. We can observe from Table 3.2, $Q(3) = 0.0013$. This means that the probability that a Gaussian random variable is more than three standard deviations above its expected value is approximately one in a thousand. In conversation we refer to the event $\{X - \mu_X > 3\sigma_X\}$ as a *three-sigma event*. It is unlikely to occur. Table 3.2 indicates that the probability of a 5σ event is on the order of 10^{-7}.

Quiz 3.5 *X is the Gaussian $(0, 1)$ random variable and Y is the Gaussian $(0, 2)$ random variable.*

(1) Sketch the PDFs $f_X(x)$ and $f_Y(y)$ on the same axes.

(2) What is $P[-1 < X \leq 1]$?

(3) What is $P[-1 < Y \leq 1]$?

(4) What is $P[X > 3.5]$?

(5) What is $P[Y > 3.5]$?

3.6 Delta Functions, Mixed Random Variables

Thus far, our analysis of continuous random variables parallels our analysis of discrete random variables in Chapter 2. Because of the different nature of discrete and continuous random variables, we represent the probability model of a discrete random variable as a

z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$
0.00	0.5000	0.50	0.6915	1.00	0.8413	1.50	0.9332	2.00	0.97725	2.50	0.99379
0.01	0.5040	0.51	0.6950	1.01	0.8438	1.51	0.9345	2.01	0.97778	2.51	0.99396
0.02	0.5080	0.52	0.6985	1.02	0.8461	1.52	0.9357	2.02	0.97831	2.52	0.99413
0.03	0.5120	0.53	0.7019	1.03	0.8485	1.53	0.9370	2.03	0.97882	2.53	0.99430
0.04	0.5160	0.54	0.7054	1.04	0.8508	1.54	0.9382	2.04	0.97932	2.54	0.99446
0.05	0.5199	0.55	0.7088	1.05	0.8531	1.55	0.9394	2.05	0.97982	2.55	0.99461
0.06	0.5239	0.56	0.7123	1.06	0.8554	1.56	0.9406	2.06	0.98030	2.56	0.99477
0.07	0.5279	0.57	0.7157	1.07	0.8577	1.57	0.9418	2.07	0.98077	2.57	0.99492
0.08	0.5319	0.58	0.7190	1.08	0.8599	1.58	0.9429	2.08	0.98124	2.58	0.99506
0.09	0.5359	0.59	0.7224	1.09	0.8621	1.59	0.9441	2.09	0.98169	2.59	0.99520
0.10	0.5398	0.60	0.7257	1.10	0.8643	1.60	0.9452	2.10	0.98214	2.60	0.99534
0.11	0.5438	0.61	0.7291	1.11	0.8665	1.61	0.9463	2.11	0.98257	2.61	0.99547
0.12	0.5478	0.62	0.7324	1.12	0.8686	1.62	0.9474	2.12	0.98300	2.62	0.99560
0.13	0.5517	0.63	0.7357	1.13	0.8708	1.63	0.9484	2.13	0.98341	2.63	0.99573
0.14	0.5557	0.64	0.7389	1.14	0.8729	1.64	0.9495	2.14	0.98382	2.64	0.99585
0.15	0.5596	0.65	0.7422	1.15	0.8749	1.65	0.9505	2.15	0.98422	2.65	0.99598
0.16	0.5636	0.66	0.7454	1.16	0.8770	1.66	0.9515	2.16	0.98461	2.66	0.99609
0.17	0.5675	0.67	0.7486	1.17	0.8790	1.67	0.9525	2.17	0.98500	2.67	0.99621
0.18	0.5714	0.68	0.7517	1.18	0.8810	1.68	0.9535	2.18	0.98537	2.68	0.99632
0.19	0.5753	0.69	0.7549	1.19	0.8830	1.69	0.9545	2.19	0.98574	2.69	0.99643
0.20	0.5793	0.70	0.7580	1.20	0.8849	1.70	0.9554	2.20	0.98610	2.70	0.99653
0.21	0.5832	0.71	0.7611	1.21	0.8869	1.71	0.9564	2.21	0.98645	2.71	0.99664
0.22	0.5871	0.72	0.7642	1.22	0.8888	1.72	0.9573	2.22	0.98679	2.72	0.99674
0.23	0.5910	0.73	0.7673	1.23	0.8907	1.73	0.9582	2.23	0.98713	2.73	0.99683
0.24	0.5948	0.74	0.7704	1.24	0.8925	1.74	0.9591	2.24	0.98745	2.74	0.99693
0.25	0.5987	0.75	0.7734	1.25	0.8944	1.75	0.9599	2.25	0.98778	2.75	0.99702
0.26	0.6026	0.76	0.7764	1.26	0.8962	1.76	0.9608	2.26	0.98809	2.76	0.99711
0.27	0.6064	0.77	0.7794	1.27	0.8980	1.77	0.9616	2.27	0.98840	2.77	0.99720
0.28	0.6103	0.78	0.7823	1.28	0.8997	1.78	0.9625	2.28	0.98870	2.78	0.99728
0.29	0.6141	0.79	0.7852	1.29	0.9015	1.79	0.9633	2.29	0.98899	2.79	0.99736
0.30	0.6179	0.80	0.7881	1.30	0.9032	1.80	0.9641	2.30	0.98928	2.80	0.99744
0.31	0.6217	0.81	0.7910	1.31	0.9049	1.81	0.9649	2.31	0.98956	2.81	0.99752
0.32	0.6255	0.82	0.7939	1.32	0.9066	1.82	0.9656	2.32	0.98983	2.82	0.99760
0.33	0.6293	0.83	0.7967	1.33	0.9082	1.83	0.9664	2.33	0.99010	2.83	0.99767
0.34	0.6331	0.84	0.7995	1.34	0.9099	1.84	0.9671	2.34	0.99036	2.84	0.99774
0.35	0.6368	0.85	0.8023	1.35	0.9115	1.85	0.9678	2.35	0.99061	2.85	0.99781
0.36	0.6406	0.86	0.8051	1.36	0.9131	1.86	0.9686	2.36	0.99086	2.86	0.99788
0.37	0.6443	0.87	0.8078	1.37	0.9147	1.87	0.9693	2.37	0.99111	2.87	0.99795
0.38	0.6480	0.88	0.8106	1.38	0.9162	1.88	0.9699	2.38	0.99134	2.88	0.99801
0.39	0.6517	0.89	0.8133	1.39	0.9177	1.89	0.9706	2.39	0.99158	2.89	0.99807
0.40	0.6554	0.90	0.8159	1.40	0.9192	1.90	0.9713	2.40	0.99180	2.90	0.99813
0.41	0.6591	0.91	0.8186	1.41	0.9207	1.91	0.9719	2.41	0.99202	2.91	0.99819
0.42	0.6628	0.92	0.8212	1.42	0.9222	1.92	0.9726	2.42	0.99224	2.92	0.99825
0.43	0.6664	0.93	0.8238	1.43	0.9236	1.93	0.9732	2.43	0.99245	2.93	0.99831
0.44	0.6700	0.94	0.8264	1.44	0.9251	1.94	0.9738	2.44	0.99266	2.94	0.99836
0.45	0.6736	0.95	0.8289	1.45	0.9265	1.95	0.9744	2.45	0.99286	2.95	0.99841
0.46	0.6772	0.96	0.8315	1.46	0.9279	1.96	0.9750	2.46	0.99305	2.96	0.99846
0.47	0.6808	0.97	0.8340	1.47	0.9292	1.97	0.9756	2.47	0.99324	2.97	0.99851
0.48	0.6844	0.98	0.8365	1.48	0.9306	1.98	0.9761	2.48	0.99343	2.98	0.99856
0.49	0.6879	0.99	0.8389	1.49	0.9319	1.99	0.9767	2.49	0.99361	2.99	0.99861

Table 3.1 The standard normal CDF $\Phi(y)$.

z	$Q(z)$	z	$Q(z)$	z	$Q(z)$	z	$Q(z)$	z	$Q(z)$
3.00	$1.35\cdot10^{-3}$	3.40	$3.37\cdot10^{-4}$	3.80	$7.23\cdot10^{-5}$	4.20	$1.33\cdot10^{-5}$	4.60	$2.11\cdot10^{-6}$
3.01	$1.31\cdot10^{-3}$	3.41	$3.25\cdot10^{-4}$	3.81	$6.95\cdot10^{-5}$	4.21	$1.28\cdot10^{-5}$	4.61	$2.01\cdot10^{-6}$
3.02	$1.26\cdot10^{-3}$	3.42	$3.13\cdot10^{-4}$	3.82	$6.67\cdot10^{-5}$	4.22	$1.22\cdot10^{-5}$	4.62	$1.92\cdot10^{-6}$
3.03	$1.22\cdot10^{-3}$	3.43	$3.02\cdot10^{-4}$	3.83	$6.41\cdot10^{-5}$	4.23	$1.17\cdot10^{-5}$	4.63	$1.83\cdot10^{-6}$
3.04	$1.18\cdot10^{-3}$	3.44	$2.91\cdot10^{-4}$	3.84	$6.15\cdot10^{-5}$	4.24	$1.12\cdot10^{-5}$	4.64	$1.74\cdot10^{-6}$
3.05	$1.14\cdot10^{-3}$	3.45	$2.80\cdot10^{-4}$	3.85	$5.91\cdot10^{-5}$	4.25	$1.07\cdot10^{-5}$	4.65	$1.66\cdot10^{-6}$
3.06	$1.11\cdot10^{-3}$	3.46	$2.70\cdot10^{-4}$	3.86	$5.67\cdot10^{-5}$	4.26	$1.02\cdot10^{-5}$	4.66	$1.58\cdot10^{-6}$
3.07	$1.07\cdot10^{-3}$	3.47	$2.60\cdot10^{-4}$	3.87	$5.44\cdot10^{-5}$	4.27	$9.77\cdot10^{-6}$	4.67	$1.51\cdot10^{-6}$
3.08	$1.04\cdot10^{-3}$	3.48	$2.51\cdot10^{-4}$	3.88	$5.22\cdot10^{-5}$	4.28	$9.34\cdot10^{-6}$	4.68	$1.43\cdot10^{-6}$
3.09	$1.00\cdot10^{-3}$	3.49	$2.42\cdot10^{-4}$	3.89	$5.01\cdot10^{-5}$	4.29	$8.93\cdot10^{-6}$	4.69	$1.37\cdot10^{-6}$
3.10	$9.68\cdot10^{-4}$	3.50	$2.33\cdot10^{-4}$	3.90	$4.81\cdot10^{-5}$	4.30	$8.54\cdot10^{-6}$	4.70	$1.30\cdot10^{-6}$
3.11	$9.35\cdot10^{-4}$	3.51	$2.24\cdot10^{-4}$	3.91	$4.61\cdot10^{-5}$	4.31	$8.16\cdot10^{-6}$	4.71	$1.24\cdot10^{-6}$
3.12	$9.04\cdot10^{-4}$	3.52	$2.16\cdot10^{-4}$	3.92	$4.43\cdot10^{-5}$	4.32	$7.80\cdot10^{-6}$	4.72	$1.18\cdot10^{-6}$
3.13	$8.74\cdot10^{-4}$	3.53	$2.08\cdot10^{-4}$	3.93	$4.25\cdot10^{-5}$	4.33	$7.46\cdot10^{-6}$	4.73	$1.12\cdot10^{-6}$
3.14	$8.45\cdot10^{-4}$	3.54	$2.00\cdot10^{-4}$	3.94	$4.07\cdot10^{-5}$	4.34	$7.12\cdot10^{-6}$	4.74	$1.07\cdot10^{-6}$
3.15	$8.16\cdot10^{-4}$	3.55	$1.93\cdot10^{-4}$	3.95	$3.91\cdot10^{-5}$	4.35	$6.81\cdot10^{-6}$	4.75	$1.02\cdot10^{-6}$
3.16	$7.89\cdot10^{-4}$	3.56	$1.85\cdot10^{-4}$	3.96	$3.75\cdot10^{-5}$	4.36	$6.50\cdot10^{-6}$	4.76	$9.68\cdot10^{-7}$
3.17	$7.62\cdot10^{-4}$	3.57	$1.78\cdot10^{-4}$	3.97	$3.59\cdot10^{-5}$	4.37	$6.21\cdot10^{-6}$	4.77	$9.21\cdot10^{-7}$
3.18	$7.36\cdot10^{-4}$	3.58	$1.72\cdot10^{-4}$	3.98	$3.45\cdot10^{-5}$	4.38	$5.93\cdot10^{-6}$	4.78	$8.76\cdot10^{-7}$
3.19	$7.11\cdot10^{-4}$	3.59	$1.65\cdot10^{-4}$	3.99	$3.30\cdot10^{-5}$	4.39	$5.67\cdot10^{-6}$	4.79	$8.34\cdot10^{-7}$
3.20	$6.87\cdot10^{-4}$	3.60	$1.59\cdot10^{-4}$	4.00	$3.17\cdot10^{-5}$	4.40	$5.41\cdot10^{-6}$	4.80	$7.93\cdot10^{-7}$
3.21	$6.64\cdot10^{-4}$	3.61	$1.53\cdot10^{-4}$	4.01	$3.04\cdot10^{-5}$	4.41	$5.17\cdot10^{-6}$	4.81	$7.55\cdot10^{-7}$
3.22	$6.41\cdot10^{-4}$	3.62	$1.47\cdot10^{-4}$	4.02	$2.91\cdot10^{-5}$	4.42	$4.94\cdot10^{-6}$	4.82	$7.18\cdot10^{-7}$
3.23	$6.19\cdot10^{-4}$	3.63	$1.42\cdot10^{-4}$	4.03	$2.79\cdot10^{-5}$	4.43	$4.71\cdot10^{-6}$	4.83	$6.83\cdot10^{-7}$
3.24	$5.98\cdot10^{-4}$	3.64	$1.36\cdot10^{-4}$	4.04	$2.67\cdot10^{-5}$	4.44	$4.50\cdot10^{-6}$	4.84	$6.49\cdot10^{-7}$
3.25	$5.77\cdot10^{-4}$	3.65	$1.31\cdot10^{-4}$	4.05	$2.56\cdot10^{-5}$	4.45	$4.29\cdot10^{-6}$	4.85	$6.17\cdot10^{-7}$
3.26	$5.57\cdot10^{-4}$	3.66	$1.26\cdot10^{-4}$	4.06	$2.45\cdot10^{-5}$	4.46	$4.10\cdot10^{-6}$	4.86	$5.87\cdot10^{-7}$
3.27	$5.38\cdot10^{-4}$	3.67	$1.21\cdot10^{-4}$	4.07	$2.35\cdot10^{-5}$	4.47	$3.91\cdot10^{-6}$	4.87	$5.58\cdot10^{-7}$
3.28	$5.19\cdot10^{-4}$	3.68	$1.17\cdot10^{-4}$	4.08	$2.25\cdot10^{-5}$	4.48	$3.73\cdot10^{-6}$	4.88	$5.30\cdot10^{-7}$
3.29	$5.01\cdot10^{-4}$	3.69	$1.12\cdot10^{-4}$	4.09	$2.16\cdot10^{-5}$	4.49	$3.56\cdot10^{-6}$	4.89	$5.04\cdot10^{-7}$
3.30	$4.83\cdot10^{-4}$	3.70	$1.08\cdot10^{-4}$	4.10	$2.07\cdot10^{-5}$	4.50	$3.40\cdot10^{-6}$	4.90	$4.79\cdot10^{-7}$
3.31	$4.66\cdot10^{-4}$	3.71	$1.04\cdot10^{-4}$	4.11	$1.98\cdot10^{-5}$	4.51	$3.24\cdot10^{-6}$	4.91	$4.55\cdot10^{-7}$
3.32	$4.50\cdot10^{-4}$	3.72	$9.96\cdot10^{-5}$	4.12	$1.89\cdot10^{-5}$	4.52	$3.09\cdot10^{-6}$	4.92	$4.33\cdot10^{-7}$
3.33	$4.34\cdot10^{-4}$	3.73	$9.57\cdot10^{-5}$	4.13	$1.81\cdot10^{-5}$	4.53	$2.95\cdot10^{-6}$	4.93	$4.11\cdot10^{-7}$
3.34	$4.19\cdot10^{-4}$	3.74	$9.20\cdot10^{-5}$	4.14	$1.74\cdot10^{-5}$	4.54	$2.81\cdot10^{-6}$	4.94	$3.91\cdot10^{-7}$
3.35	$4.04\cdot10^{-4}$	3.75	$8.84\cdot10^{-5}$	4.15	$1.66\cdot10^{-5}$	4.55	$2.68\cdot10^{-6}$	4.95	$3.71\cdot10^{-7}$
3.36	$3.90\cdot10^{-4}$	3.76	$8.50\cdot10^{-5}$	4.16	$1.59\cdot10^{-5}$	4.56	$2.56\cdot10^{-6}$	4.96	$3.52\cdot10^{-7}$
3.37	$3.76\cdot10^{-4}$	3.77	$8.16\cdot10^{-5}$	4.17	$1.52\cdot10^{-5}$	4.57	$2.44\cdot10^{-6}$	4.97	$3.35\cdot10^{-7}$
3.38	$3.62\cdot10^{-4}$	3.78	$7.84\cdot10^{-5}$	4.18	$1.46\cdot10^{-5}$	4.58	$2.32\cdot10^{-6}$	4.98	$3.18\cdot10^{-7}$
3.39	$3.49\cdot10^{-4}$	3.79	$7.53\cdot10^{-5}$	4.19	$1.39\cdot10^{-5}$	4.59	$2.22\cdot10^{-6}$	4.99	$3.02\cdot10^{-7}$

Table 3.2 The standard normal complementary CDF $Q(z)$.

PMF and we represent the probability model of a continuous random variable as a PDF. These functions are important because they enable us to calculate conveniently parameters of probability models (such as the expected value and the variance) and probabilities of events. Calculations containing a PMF involve sums. The corresponding calculations for a PDF contain integrals.

In this section, we introduce the unit impulse function $\delta(x)$ as a mathematical tool that unites the analyses of discrete and continuous random variables. The unit impulse, often called the *delta function*, allows us to use the same formulas to describe calculations with both types of random variables. It does not alter the calculations, it just provides a new notation for describing them. This is especially convenient when we refer to a *mixed random variable*, which has properties of both continuous and discrete random variables.

The delta function is not completely respectable mathematically because it is zero everywhere except at one point, and there it is infinite. Thus at its most interesting point it has no numerical value at all. While $\delta(x)$ is somewhat disreputable, it is extremely useful. There are various definitions of the delta function. All of them share the key property presented in Theorem 3.16. Here is the definition adopted in this book.

Definition 3.12 ***Unit Impulse (Delta) Function***

Let

$$d_\epsilon(x) = \begin{cases} 1/\epsilon & -\epsilon/2 \le x \le \epsilon/2, \\ 0 & \textit{otherwise}. \end{cases}$$

*The **unit impulse function** is*

$$\delta(x) = \lim_{\epsilon \to 0} d_\epsilon(x).$$

The mathematical problem with Definition 3.12 is that $d_\epsilon(x)$ has no limit at $x = 0$. As indicated in Figure 3.7, $d_\epsilon(0)$ just gets bigger and bigger as $\epsilon \to 0$. Although this makes Definition 3.12 somewhat unsatisfactory, the useful properties of the delta function are readily demonstrated when $\delta(x)$ is approximated by $d_\epsilon(x)$ for very small ϵ. We now present some properties of the delta function. We state these properties as theorems even though they are not theorems in the usual sense of this text because we cannot prove them. Instead of theorem proofs, we refer to $d_\epsilon(x)$ for small values of ϵ to indicate why the properties hold.

Although, $d_\epsilon(0)$ blows up as $\epsilon \to 0$, the area under $d_\epsilon(x)$ is the integral

$$\int_{-\infty}^{\infty} d_\epsilon(x)\, dx = \int_{-\epsilon/2}^{\epsilon/2} \frac{1}{\epsilon}\, dx = 1. \tag{3.57}$$

That is, the area under $d_\epsilon(x)$ is always 1, no matter how small the value of ϵ. We conclude that the area under $\delta(x)$ is also 1:

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1. \tag{3.58}$$

This result is a special case of the following property of the delta function.

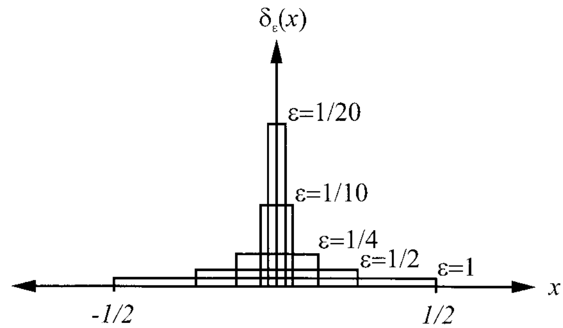


Figure 3.7 As $\epsilon \to 0$, $d_\epsilon(x)$ approaches the delta function $\delta(x)$. For each ϵ, the area under the curve of $d_\epsilon(x)$ equals 1.

Theorem 3.16 *For any continuous function $g(x)$,*

$$\int_{-\infty}^{\infty} g(x)\delta(x - x_0)\,dx = g(x_0).$$

Theorem 3.16 is often called the *sifting property* of the delta function. We can see that Equation (3.58) is a special case of the sifting property for $g(x) = 1$ and $x_0 = 0$. To understand Theorem 3.16, consider the integral

$$\int_{-\infty}^{\infty} g(x)d_\epsilon(x - x_0)\,dx = \frac{1}{\epsilon}\int_{x_0-\epsilon/2}^{x_0+\epsilon/2} g(x)\,dx. \tag{3.59}$$

On the right side, we have the average value of $g(x)$ over the interval $[x_0 - \epsilon/2, x_0 + \epsilon/2]$. As $\epsilon \to 0$, this average value must converge to $g(x_0)$.

The delta function has a close connection to the unit step function.

Definition 3.13 ***Unit Step Function***
*The **unit step function** is*

$$u(x) = \begin{cases} 0 & x < 0, \\ 1 & x \geq 0. \end{cases}$$

Theorem 3.17

$$\int_{-\infty}^{x} \delta(v)\,dv = u(x).$$

To understand Theorem 3.17, we observe that for any $x > 0$, we can choose $\epsilon \leq 2x$ so that

$$\int_{-\infty}^{-x} d_\epsilon(v)\,dv = 0, \qquad \int_{-\infty}^{x} d_\epsilon(v)\,dv = 1. \tag{3.60}$$

Thus for any $x \neq 0$, in the limit as $\epsilon \to 0$, $\int_{-\infty}^{x} d_\epsilon(v)\,dv = u(x)$. Note that we have not yet considered $x = 0$. In fact, it is not completely clear what the value of $\int_{-\infty}^{0} \delta(v)\,dv$ should be. Reasonable arguments can be made for 0, 1/2, or 1. We have adopted the convention that $\int_{-\infty}^{0} \delta(x)\,dx = 1$. We will see that this is a particularly convenient choice when we reexamine discrete random variables.

Theorem 3.17 allows us to write

$$\delta(x) = \frac{du(x)}{dx}. \tag{3.61}$$

Equation (3.61) embodies a certain kind of consistency in its inconsistency. That is, $\delta(x)$ does not really exist at $x = 0$. Similarly, the derivative of $u(x)$ does not really exist at $x = 0$. However, Equation (3.61) allows us to use $\delta(x)$ to define a generalized PDF that applies to discrete random variables as well as to continuous random variables.

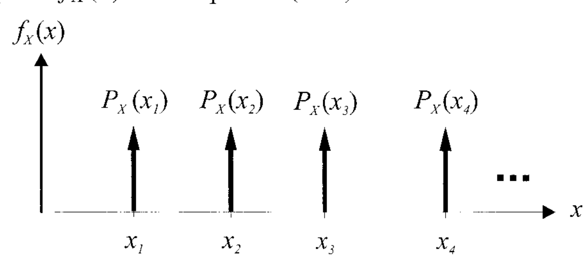
Consider the CDF of a discrete random variable, X. Recall that it is constant everywhere except at points $x_i \in S_X$, where it has jumps of height $P_X(x_i)$. Using the definition of the unit step function, we can write the CDF of X as

$$F_X(x) = \sum_{x_i \in S_X} P_X(x_i)\, u(x - x_i). \tag{3.62}$$

From Definition 3.3, we take the derivative of $F_X(x)$ to find the PDF $f_X(x)$. Referring to Equation (3.61), the PDF of the discrete random variable X is

$$f_X(x) = \sum_{x_i \in S_X} P_X(x_i)\, \delta(x - x_i). \tag{3.63}$$

When the PDF includes delta functions of the form $\delta(x - x_i)$, we say there is an impulse at x_i. When we graph a PDF $f_X(x)$ that contains an impulse at x_i, we draw a vertical arrow labeled by the constant that multiplies the impulse. We draw each arrow representing an impulse at the same height because the PDF is always infinite at each such point. For example, the graph of $f_X(x)$ from Equation (3.63) is



Using delta functions in the PDF, we can apply the formulas in this chapter to all random variables. In the case of discrete random variables, these formulas will be equivalent to the ones presented in Chapter 2. For example, if X is a discrete random variable, Definition 3.4 becomes

$$E[X] = \int_{-\infty}^{\infty} x \sum_{x_i \in S_X} P_X(x_i)\, \delta(x - x_i)\, dx. \tag{3.64}$$

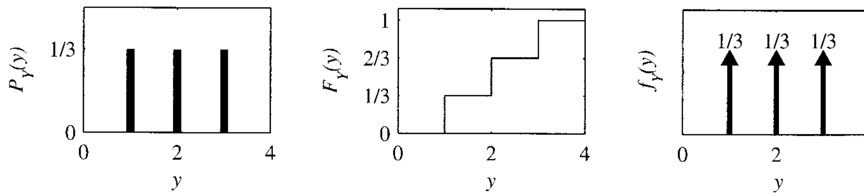


Figure 3.8 The PMF, CDF, and PDF of the mixed random variable Y.

By writing the integral of the sum as a sum of integrals, and using the sifting property of the delta function,

$$E[X] = \sum_{x_i \in S_X} \int_{-\infty}^{\infty} x P_X(x_i)\, \delta(x - x_i)\, dx = \sum_{x_i \in S_X} x_i P_X(x_i), \tag{3.65}$$

which is Definition 2.14.

Example 3.19 Suppose Y takes on the values $1, 2, 3$ with equal probability. The PMF and the corresponding CDF of Y are

$$P_Y(y) = \begin{cases} 1/3 & y = 1, 2, 3, \\ 0 & \text{otherwise}, \end{cases} \qquad F_Y(y) = \begin{cases} 0 & y < 1, \\ 1/3 & 1 \le y < 2, \\ 2/3 & 2 \le y < 3, \\ 1 & y \ge 3. \end{cases} \tag{3.66}$$

Using the unit step function $u(y)$, we can write $F_Y(y)$ more compactly as

$$F_Y(y) = \frac{1}{3}u(y-1) + \frac{1}{3}u(y-2) + \frac{1}{3}u(y-3). \tag{3.67}$$

The PDF of Y is

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{3}\delta(y-1) + \frac{1}{3}\delta(y-2) + \frac{1}{3}\delta(y-3). \tag{3.68}$$

We see that the discrete random variable Y can be represented graphically either by a PMF $P_Y(y)$ with bars at $y = 1, 2, 3$, by a CDF with jumps at $y = 1, 2, 3$, or by a PDF $f_Y(y)$ with impulses at $y = 1, 2, 3$. These three representations are shown in Figure 3.8. The expected value of Y can be calculated either by summing over the PMF $P_Y(y)$ or integrating over the PDF $f_Y(y)$. Using the PDF, we have

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy \tag{3.69}$$

$$= \int_{-\infty}^{\infty} \frac{y}{3}\delta(y-1)\, dy + \int_{-\infty}^{\infty} \frac{y}{3}\delta(y-2)\, dy + \int_{-\infty}^{\infty} \frac{y}{3}\delta(y-3)\, dy \tag{3.70}$$

$$= 1/3 + 2/3 + 1 = 2. \tag{3.71}$$

When $F_X(x)$ has a discontinuity at x, we will use $F_X(x^+)$ and $F_X(x^-)$ to denote the upper and lower limits at x. That is,

$$F_X\left(x^-\right) = \lim_{h\to 0^+} F_X\left(x-h\right), \qquad F_X\left(x^+\right) = \lim_{h\to 0^+} F_X\left(x+h\right). \tag{3.72}$$

Using this notation, we can say that if the CDF $F_X(x)$ has a jump at x_0, then $f_X(x)$ has an impulse at x_0 weighted by the height of the discontinuity $F_X(x_0^+) - F_X(x_0^-)$.

Example 3.20 For the random variable Y of Example 3.19,

$$F_Y\left(2^-\right) = 1/3, \qquad F_Y\left(2^+\right) = 2/3. \tag{3.73}$$

Theorem 3.18 *For a random variable X, we have the following equivalent statements:*

(a) $P[X = x_0] = q$

(b) $P_X(x_0) = q$

(c) $F_X(x_0^+) - F_X(x_0^-) = q$

(d) $f_X(x_0) = q\delta(0)$

In Example 3.19, we saw that $f_Y(y)$ consists of a series of impulses. The value of $f_Y(y)$ is either 0 or ∞. By contrast, the PDF of a continuous random variable has nonzero, finite values over intervals of x. In the next example, we encounter a random variable that has continuous parts and impulses.

Definition 3.14 ***Mixed Random Variable***

*X is a **mixed** random variable if and only if $f_X(x)$ contains both impulses and nonzero, finite values.*

Example 3.21 Observe someone dialing a telephone and record the duration of the call. In a simple model of the experiment, $1/3$ of the calls never begin either because no one answers or the line is busy. The duration of these calls is 0 minutes. Otherwise, with probability $2/3$, a call duration is uniformly distributed between 0 and 3 minutes. Let Y denote the call duration. Find the CDF $F_Y(y)$, the PDF $f_Y(y)$, and the expected value $E[Y]$.

....

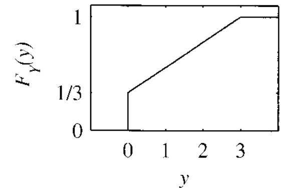
Let A denote the event that the phone was answered. Since $Y \geq 0$, we know that for $y < 0$, $F_Y(y) = 0$. Similarly, we know that for $y > 3$, $F_Y(y) = 1$. For $0 \leq y \leq 3$, we apply the law of total probability to write

$$F_Y\left(y\right) = P\left[Y \leq y\right] = P\left[Y \leq y|A^c\right] P\left[A^c\right] + P\left[Y \leq y|A\right] P\left[A\right]. \tag{3.74}$$

When A^c occurs, $Y = 0$, so that for $0 \leq y \leq 3$, $P[Y \leq y|A^c] = 1$. When A occurs, the call duration is uniformly distributed over $[0, 3]$, so that for $0 \leq y \leq 3$, $P[Y \leq y|A] = y/3$. So, for $0 \leq y \leq 3$,

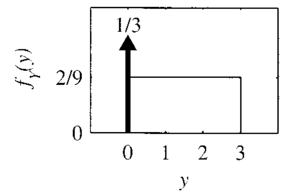
$$F_Y\left(y\right) = (1/3)(1) + (2/3)(y/3) = 1/3 + 2y/9. \tag{3.75}$$

Finally, the complete CDF of Y is



$$F_Y(y) = \begin{cases} 0 & y < 0, \\ 1/3 + 2y/9 & 0 \le y < 3, \\ 1 & y \ge 3. \end{cases} \tag{3.76}$$

Consequently, the corresponding PDF $f_Y(y)$ is



$$f_Y(y) = \begin{cases} \delta(y)/3 + 2/9 & 0 \le y \le 3, \\ 0 & \text{otherwise.} \end{cases} \tag{3.77}$$

For the mixed random variable Y, it is easiest to calculate $E[Y]$ using the PDF:

$$E[Y] = \int_{-\infty}^{\infty} y \frac{1}{3}\delta(y)\,dy + \int_0^3 \frac{2}{9} y\,dy = 0 + \frac{2}{9}\frac{y^2}{2}\bigg|_0^3 = 1. \tag{3.78}$$

In Example 3.21, we see that with probability $1/3$, Y resembles a discrete random variable; otherwise, Y behaves like a continuous random variable. This behavior is reflected in the impulse in the PDF of Y. In many practical applications of probability, mixed random variables arise as functions of continuous random variables. Electronic circuits perform many of these functions. Example 3.25 in Section 3.7 gives one example.

Before going any further, we review what we have learned about random variables. For any random variable X,

- X always has a CDF $F_X(x) = P[X \le x]$.
- If $F_X(x)$ is piecewise flat with discontinuous jumps, then X is discrete.
- If $F_X(x)$ is a continuous function, then X is continuous.
- If $F_X(x)$ is a piecewise continuous function with discontinuities, then X is mixed.
- When X is discrete or mixed, the PDF $f_X(x)$ contains one or more delta functions.

Quiz 3.6 *The cumulative distribution function of random variable X is*

$$F_X(x) = \begin{cases} 0 & x < -1, \\ (x+1)/4 & -1 \le x < 1, \\ 1 & x \ge 1. \end{cases} \tag{3.79}$$

Sketch the CDF and find the following:

(1) $P[X \le 1]$ *(2)* $P[X < 1]$

(3) $P[X = 1]$ *(4) the PDF* $f_X(x)$

3.7 Probability Models of Derived Random Variables

Here we return to derived random variables. If $Y = g(X)$, we discuss methods of determining $f_Y(y)$ from $g(X)$ and $f_X(x)$. The approach is considerably different from the task of determining a derived PMF of a discrete random variable. In the discrete case we derive the new PMF directly from the original one. For continuous random variables we follow a two-step procedure.

1. Find the CDF $F_Y(y) = P[Y \leq y]$.
2. Compute the PDF by calculating the derivative $f_Y(y) = dF_Y(y)/dy$.

This procedure *always* works and is very easy to remember. The method is best demonstrated by examples. However, as we shall see in the examples, following the procedure, in particular finding $F_Y(y)$, can be tricky. Before proceeding to the examples, we add one reminder. If you have to find $E[g(X)]$, it is easier to calculate $E[g(X)]$ directly using Theorem 3.4 than it is to derive the PDF of $Y = g(X)$ and then use the definition of expected value, Definition 3.4. The material in this section applies to situations in which it is necessary to find a complete probability model of $Y = g(X)$.

Example 3.22 In Example 3.2, Y centimeters is the location of the pointer on the 1-meter circumference of the circle. Use the solution of Example 3.2 to derive $f_Y(y)$.

The function $Y = 100X$, where X in Example 3.2 is the location of the pointer measured in meters. To find the PDF of Y, we first find the CDF $F_Y(y)$. Example 3.2 derives the CDF of X,

$$F_X(x) = \begin{cases} 0 & x < 0, \\ x & 0 \leq x < 1, \\ 1 & x \geq 1. \end{cases} \tag{3.80}$$

We use this result to find the CDF $F_Y(y) = P[100X \leq y]$. Equivalently,

$$F_Y(y) = P[X \leq y/100] = F_X(y/100) = \begin{cases} 0 & y/100 < 0, \\ y/100 & 0 \leq y/100 < 1, \\ 1 & y/100 \geq 1. \end{cases} \tag{3.81}$$

We take the derivative of the CDF of Y over each of the three intervals to find the PDF:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} 1/100 & 0 \leq y < 100, \\ 0 & \text{otherwise.} \end{cases} \tag{3.82}$$

We see that Y is the uniform $(0, 100)$ random variable.

We use this two-step procedure in the following theorem to generalize Example 3.22 by deriving the CDF and PDF for any scale change and any continuous random variable.

Theorem 3.19 *If $Y = aX$, where $a > 0$, then Y has CDF and PDF*

$$F_Y(y) = F_X(y/a), \qquad f_Y(y) = \frac{1}{a} f_X(y/a).$$

Proof First, we find the CDF of Y,

$$F_Y(y) = P[aX \leq y] = P[X \leq y/a] = F_X(y/a). \tag{3.83}$$

We take the derivative of $F_Y(y)$ to find the PDF:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{a} f_X(y/a). \tag{3.84}$$

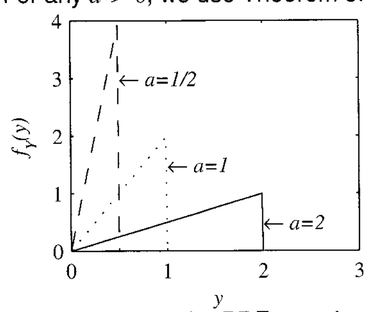
Theorem 3.19 states that multiplying a random variable by a positive constant stretches ($a > 1$) or shrinks ($a < 1$) the original PDF.

Example 3.23 Let X have the triangular PDF

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.85}$$

Find the PDF of $Y = aX$. Sketch the PDF of Y for $a = 1/2, 1, 2$.

For any $a > 0$, we use Theorem 3.19 to find the PDF:



$$f_Y(y) = \frac{1}{a} f_X(y/a) \tag{3.86}$$

$$= \begin{cases} 2y/a^2 & 0 \leq y \leq a, \\ 0 & \text{otherwise.} \end{cases} \tag{3.87}$$

As a increases, the PDF stretches horizontally.

For the families of continuous random variables in Sections 3.4 and 3.5, we can use Theorem 3.19 to show that multiplying a random variable by a constant produces a new family member with transformed parameters.

Theorem 3.20 *$Y = aX$, where $a > 0$.*

(a) If X is uniform (b, c), then Y is uniform (ab, ac).

(b) If X is exponential (λ), then Y is exponential (λ/a).

(c) If X is Erlang (n, λ), then Y is Erlang $(n, \lambda/a)$.

(d) If X is Gaussian (μ, σ), then Y is Gaussian $(a\mu, a\sigma)$.

The next theorem shows that adding a constant to a random variable simply shifts the CDF and the PDF by that constant.

Theorem 3.21 *If $Y = X + b$,*

$$F_Y(y) = F_X(y - b), \qquad f_Y(y) = f_X(y - b).$$

Proof First, we find the CDF of V,

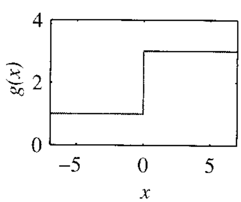
$$F_Y(y) = P[X + b \leq y] = P[X \leq y - b] = F_X(y - b). \tag{3.88}$$

We take the derivative of $F_Y(y)$ to find the PDF:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = f_X(y - b). \tag{3.89}$$

Thus far, the examples and theorems in this section relate to a continuous random variable derived from another continuous random variable. By contrast, in the following example, the function $g(x)$ transforms a continuous random variable to a discrete random variable.

Example 3.24 Let X be a random variable with CDF $F_X(x)$. Let Y be the output of a clipping circuit with the characteristic $Y = g(X)$ where



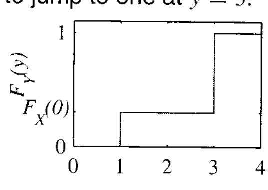
$$g(x) = \begin{cases} 1 & x \leq 0, \\ 3 & x > 0. \end{cases} \tag{3.90}$$

Express $F_Y(y)$ and $f_Y(y)$ in terms of $F_X(x)$ and $f_X(x)$.

Before going deeply into the math, it is helpful to think about the nature of the derived random variable Y. The definition of $g(x)$ tells us that Y has only two possible values, $Y = 1$ and $Y = 3$. Thus Y is a discrete random variable. Furthermore, the CDF, $F_Y(y)$, has jumps at $y = 1$ and $y = 3$; it is zero for $y < 1$ and it is one for $y \geq 3$. Our job is to find the heights of the jumps at $y = 1$ and $y = 3$. In particular,

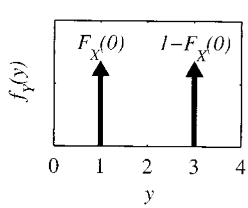
$$F_Y(1) = P[Y \leq 1] = P[X \leq 0] = F_X(0). \tag{3.91}$$

This tells us that the CDF jumps by $F_X(0)$ at $y = 1$. We also know that the CDF has to jump to one at $y = 3$. Therefore, the entire story is



$$F_Y(y) = \begin{cases} 0 & y < 1, \\ F_X(0) & 1 \leq y < 3, \\ 1 & y \geq 3. \end{cases} \tag{3.92}$$

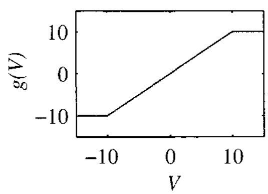
The PDF consists of impulses at $y = 1$ and $y = 3$. The weights of the impulses are the sizes of the two jumps in the CDF: $F_X(0)$ and $1 - F_X(0)$, respectively.



$$f_Y(y) = F_X(0)\,\delta(y - 1) + [1 - F_X(0)]\delta(y - 3). \tag{3.93}$$

The next two examples contain functions that transform continuous random variables to mixed random variables.

Example 3.25 The output voltage of a microphone is a Gaussian random variable V with expected value $\mu_V = 0$ and standard deviation $\sigma_V = 5$ V. The microphone signal is the input to a limiter circuit with cutoff value ± 10 V. The random variable W is the output of the limiter:



$$W = g(V) = \begin{cases} -10 & V < -10, \\ V & -10 \le V \le 10, \\ 10 & V > 10. \end{cases} \tag{3.94}$$

What are the CDF and PDF of W?

To find the CDF, we first observe that the minimum value of W is -10 and the maximum value is 10. Therefore,

$$F_W(w) = P[W \le w] = \begin{cases} 0 & w < -10, \\ 1 & w > 10. \end{cases} \tag{3.95}$$

For $-10 \le v \le 10$, $W = V$ and

$$F_W(w) = P[W \le w] = P[V \le w] = F_V(w). \tag{3.96}$$

Because V is Gaussian $(0, 5)$, Theorem 3.14 states that $F_V(v) = \Phi(v/5)$. Therefore,

$$F_W(w) = \begin{cases} 0 & w < -10, \\ \Phi(w/5) & -10 \le w \le 10, \\ 1 & w > 10. \end{cases} \tag{3.97}$$

Note that the CDF jumps from 0 to $\Phi(-10/5) = 0.023$ at $w = -10$ and that it jumps from $\Phi(10/5) = 0.977$ to 1 at $w = 10$. Therefore,

$$f_W(w) = \frac{dF_W(w)}{dw} = \begin{cases} 0.023\delta(w+10) & w = -10, \\ \frac{1}{5\sqrt{2\pi}} e^{-w^2/50} & -10 < w < 10, \\ 0.023\delta(w-10) & w = 10, \\ 0 & \text{otherwise.} \end{cases} \tag{3.98}$$

Derived density problems like the ones in the previous three examples are difficult because there are no simple cookbook procedures for finding the CDF. The following example is tricky because $g(X)$ transforms more than one value of X to the same Y.

Example 3.26 Suppose X is uniformly distributed over $[-1, 3]$ and $Y = X^2$. Find the CDF $F_Y(y)$ and the PDF $f_Y(y)$.

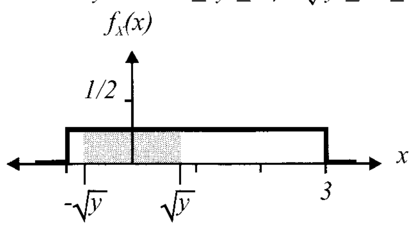
From the problem statement and Definition 3.5, the PDF of X is

$$f_X(x) = \begin{cases} 1/4 & -1 \le x \le 3, \\ 0 & \text{otherwise.} \end{cases} \tag{3.99}$$

Following the two-step procedure, we first observe that $0 \le Y \le 9$, so $F_Y(y) = 0$ for $y < 0$, and $F_Y(y) = 1$ for $y > 9$. To find the entire CDF,

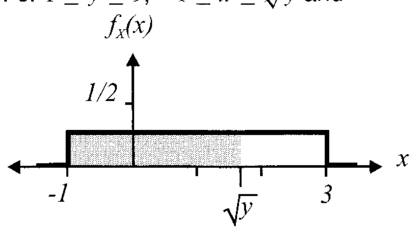
$$F_Y(y) = P\left[X^2 \le y\right] = P\left[-\sqrt{y} \le X \le \sqrt{y}\right] = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\, dx. \tag{3.100}$$

This is somewhat tricky because the calculation of the integral depends on the exact value of y. For $0 \le y \le 1$, $-\sqrt{y} \le x \le \sqrt{y}$ and



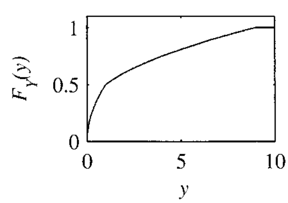
$$F_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{4}\, dx = \frac{\sqrt{y}}{2}. \tag{3.101}$$

For $1 \le y \le 9$, $-1 \le x \le \sqrt{y}$ and

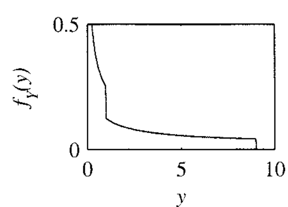


$$F_Y(y) = \int_{-1}^{\sqrt{y}} \frac{1}{4}\, dx = \frac{\sqrt{y}+1}{4}. \tag{3.102}$$

By combining the separate pieces, we can write a complete expression for $F_Y(y)$. To find $f_Y(y)$, we take the derivative of $F_Y(y)$ over each interval.



$$F_Y(y) = \begin{cases} 0 & y < 0, \\ \sqrt{y}/2 & 0 \le y \le 1, \\ (\sqrt{y}+1)/4 & 1 \le y \le 9, \\ 1 & y \ge 9. \end{cases} \tag{3.103}$$



$$f_Y(y) = \begin{cases} 1/4\sqrt{y} & 0 \le y \le 1, \\ 1/8\sqrt{y} & 1 \le y \le 9, \\ 0 & \text{otherwise.} \end{cases} \tag{3.104}$$

We end this section with a useful application of derived random variables. The following theorem shows how to derive various types of random variables from the transformation $X = g(U)$ where U is a uniform $(0, 1)$ random variable. In Section 3.9, we use this technique with the MATLAB `rand` function approximating U to generate sample values of a random variable X.

Theorem 3.22 *Let U be a uniform $(0, 1)$ random variable and let $F(x)$ denote a cumulative distribution function with an inverse $F^{-1}(u)$ defined for $0 < u < 1$. The random variable $X = F^{-1}(U)$ has CDF $F_X(x) = F(x)$.*

Proof First, we verify that $F^{-1}(u)$ is a nondecreasing function. To show this, suppose that for $u \geq u'$, $x = F^{-1}(u)$ and $x' = F^{-1}(u')$. In this case, $u = F(x)$ and $u' = F(x')$. Since $F(x)$ is nondecreasing, $F(x) \geq F(x')$ implies that $x \geq x'$. Hence, for the random variable $X = F^{-1}(U)$, we can write

$$F_X(x) = P\left[F^{-1}(U) \leq x\right] = P\left[U \leq F(x)\right] = F(x). \qquad (3.105)$$

We observe that the requirement that $F_X(u)$ have an inverse for $0 < u < 1$ is quite strict. For example, this requirement is not met by the mixed random variables of Section 3.6. A generalizaton of the theorem that does hold for mixed random variables is given in Problem 3.7.18. The following examples demonstrate the utility of Theorem 3.22.

Example 3.27 U is the uniform $(0, 1)$ random variable and $X = g(U)$. Derive $g(U)$ such that X is the exponential (1) random variable.

The CDF of X is simply

$$F_X(x) = \begin{cases} 0 & x < 0, \\ 1 - e^{-x} & x \geq 0. \end{cases} \qquad (3.106)$$

Note that if $u = F_X(x) = 1 - e^{-x}$, then $x = -\ln(1-u)$. That is, for any $u \geq 0$, $F_X^{-1}(u) = -\ln(1-u)$. Thus, by Theorem 3.22,

$$X = g(U) = -\ln(1-U) \qquad (3.107)$$

is an exponential random variable with parameter $\lambda = 1$. Problem 3.7.5 asks the reader to derive the PDF of $X = -\ln(1-U)$ directly from first principles.

Example 3.28 For a uniform $(0, 1)$ random variable U, find a function $g(\cdot)$ such that $X = g(U)$ has a uniform (a, b) distribution.

The CDF of X is

$$F_X(x) = \begin{cases} 0 & x < a, \\ (x-a)/(b-a) & a \leq x \leq b, \\ 1 & x > b. \end{cases} \qquad (3.108)$$

For any u satisfying $0 \leq u \leq 1$, $u = F_X(x) = (x-a)/(b-a)$ if and only if

$$x = F_X^{-1}(u) = a + (b-a)u. \qquad (3.109)$$

Thus by Theorem 3.22, $X = a + (b-a)U$ is a uniform (a, b) random variable. Note that we could have reached the same conclusion by observing that Theorem 3.20 implies $(b-a)U$ has a uniform $(0, b-a)$ distribution and that Theorem 3.21 implies $a + (b-a)U$ has a uniform $(a, (b-a)+a)$ distribution. Another approach, as taken in Problem 3.7.13, is to derive the CDF and PDF of $a + (b-a)U$.

The technique of Theorem 3.22 is particularly useful when the CDF is an easily invertible function. Unfortunately, there are many cases, including Gaussian and Erlang random variables, when the CDF is difficult to compute much less to invert. In these cases, we will need to develop other methods.

Quiz 3.7 *Random variable X has probability density function*

$$f_X(x) = \begin{cases} 1 - x/2 & 0 \le x \le 2, \\ 0 & \text{otherwise.} \end{cases} \tag{3.110}$$

A hard limiter produces

$$Y = \begin{cases} X & X \le 1, \\ 1 & X > 1. \end{cases} \tag{3.111}$$

(1) What is the CDF $F_X(x)$?

(2) What is $P[Y = 1]$?

(3) What is $F_Y(y)$?

(4) What is $f_Y(y)$?

3.8 Conditioning a Continuous Random Variable

In an experiment that produces a random variable X, there are occasions in which we cannot observe X. Instead, we obtain information about X without learning its precise value.

Example 3.29 Recall the experiment in which you wait for the professor to arrive for the probability lecture. Let X denote the arrival time in minutes either before ($X < 0$) or after ($X > 0$) the scheduled lecture time. When you observe that the professor is already two minutes late but has not yet arrived, you have learned that $X > 2$ but you have not learned the precise value of X.

In general, we learn that an event B has occurred, where B is defined in terms of the random variable X. For example, B could be the event $\{X \le 33\}$ or $\{|X| > 1\}$. Given the occurrence of the conditioning event B, we define a conditional probability model for the random variable X.

Definition 3.15 ***Conditional PDF given an Event***
For a random variable X with PDF $f_X(x)$ and an event $B \subset S_X$ with $P[B] > 0$, the ***conditional PDF of X given B*** *is*

$$f_{X|B}(x) = \begin{cases} \dfrac{f_X(x)}{P[B]} & x \in B, \\ 0 & \text{otherwise.} \end{cases}$$

The function $f_{X|B}(x)$ is a probability model for a new random variable related to X. Thus it has the same properties as any PDF $f_X(x)$. For example, the integral of the conditional PDF over all x is 1 (Theorem 3.2(c)) and the conditional probability of any interval is the integral of the conditional PDF over the interval (Theorem 3.3).

The definition of the conditional PDF follows naturally from the formula for conditional probability $P[A|B] = P[AB]/P[B]$ for the infinitesimal event $A = \{x < X \le x + dx\}$.

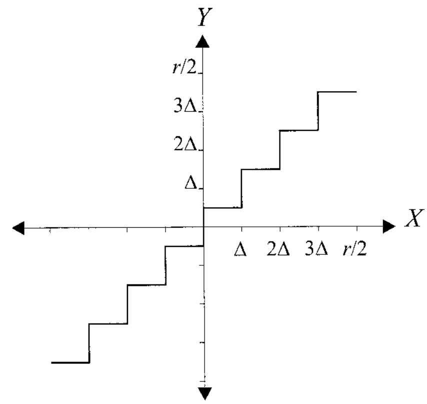


Figure 3.9 The b-bit uniform quantizer shown for $b = 3$ bits.

Since $f_{X|B}(x)$ is a probability density function, the conditional probability formula yields

$$f_{X|B}(x)\,dx = P[x < X \le x + dx|B] = \frac{P[x < X \le x + dx, B]}{P[B]}. \tag{3.112}$$

Example 3.30 For the wheel-spinning experiment of Example 3.1, find the conditional PDF of the pointer position for spins in which the pointer stops on the left side of the circle. What are the conditional expected value and the conditional standard deviation?

Let L denote the left side of the circle. In terms of the stopping position, $L = [1/2, 1)$. Recalling from Example 3.4 that the pointer position X has a uniform PDF over $[0, 1)$,

$$P[L] = \int_{1/2}^{1} f_X(x)\,dx = \int_{1/2}^{1} dx = 1/2. \tag{3.113}$$

Therefore,

$$f_{X|L}(x) = \begin{cases} 2 & 1/2 \le x < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.114}$$

Example 3.31 The uniform $(-r/2, r/2)$ random variable X is processed by a b-bit uniform quantizer to produce the quantized output Y. Random variable X is rounded to the nearest quantizer level. With a b-bit quantizer, there are $n = 2^b$ quantization levels. The quantization step size is $\Delta = r/n$, and Y takes on values in the set

$$Q_Y = \{y_i = \Delta/2 + i\Delta | i = -n/2, -n/2+1, \ldots, n/2-1\}. \tag{3.115}$$

This relationship is shown for $b = 3$ in Figure 3.9. Given the event B_i that $Y = y_i$, find the conditional PDF of X given B_i.

In terms of X, we observe that $B_i = \{i\Delta \le X < (i+1)\Delta\}$. Thus,

$$P[B_i] = \int_{i\Delta}^{(i+1)\Delta} f_X(x)\,dx = \frac{\Delta}{r} = \frac{1}{n}. \tag{3.116}$$

By Definition 3.15,

$$f_{X|B_i}(x) = \begin{cases} \dfrac{f_X(x)}{P[B_i]} & x \in B_i, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} 1/\Delta & i\Delta \le x < (i+1)\Delta, \\ 0 & \text{otherwise.} \end{cases} \tag{3.117}$$

Given B_i, the conditional PDF of X is uniform over the ith quantization interval.

We observe in Example 3.31 that $\{B_i\}$ is an event space. The following theorem shows how we can can reconstruct the PDF of X given the conditional PDFs $f_{X|B_i}(x)$.

Theorem 3.23 *Given an event space $\{B_i\}$ and the conditional PDFs $f_{X|B_i}(x)$,*

$$f_X(x) = \sum_i f_{X|B_i}(x) P[B_i].$$

Although we initially defined the event B_i as a subset of S_X, Theorem 3.23 extends naturally to arbitrary event spaces $\{B_i\}$ for which we know the conditional PDFs $f_{X|B_i}(x)$.

Example 3.32 Continuing Example 3.3, when symbol "0" is transmitted (event B_0), X is the Gaussian $(-5, 2)$ random variable. When symbol "1" is transmitted (event B_1), X is the Gaussian $(5, 2)$ random variable. Given that symbols "0" and "1" are equally likely to be sent, what is the PDF of X?

The problem statement implies that $P[B_0] = P[B_1] = 1/2$ and

$$f_{X|B_0}(x) = \frac{1}{2\sqrt{2\pi}} e^{-(x+5)^2/8}, \qquad f_{X|B_1}(x) = \frac{1}{2\sqrt{2\pi}} e^{-(x-5)^2/8}. \tag{3.118}$$

By Theorem 3.23,

$$f_X(x) = f_{X|B_0}(x) P[B_0] + f_{X|B_1}(x) P[B_1] \tag{3.119}$$

$$= \frac{1}{4\sqrt{2\pi}} \left(e^{-(x+5)^2/8} + e^{-(x-5)^2/8}\right). \tag{3.120}$$

Problem 3.9.2 asks the reader to graph $f_X(x)$ to show its similarity to Figure 3.3.

Conditional probability models have parameters corresponding to the parameters of unconditional probability models.

Definition 3.16 ***Conditional Expected Value Given an Event***

If $\{x \in B\}$, the conditional expected value of X is

$$E[X|B] = \int_{-\infty}^{\infty} x f_{X|B}(x)\, dx.$$

The conditional expected value of $g(X)$ is

$$E[g(X)|B] = \int_{-\infty}^{\infty} g(x) f_{X|B}(x)\, dx. \tag{3.121}$$

The conditional variance is

$$\operatorname{Var}[X|B] = E\left[\left(X - \mu_{X|B}\right)^2 |B\right] = E\left[X^2|B\right] - \mu_{X|B}^2. \tag{3.122}$$

The conditional standard deviation is $\sigma_{X|B} = \sqrt{\operatorname{Var}[X|B]}$. The conditional variance and conditional standard deviation are useful because they measure the spread of the random variable after we learn the conditioning information B. If the conditional standard deviation $\sigma_{X|B}$ is much smaller than σ_X, then we can say that learning the occurrence of B reduces our uncertainty about X because it shrinks the range of typical values of X.

Example 3.33 Continuing the wheel spinning of Example 3.30, find the conditional expected value and the conditional standard deviation of the pointer position X given the event L that the pointer stops on the left side of the circle.

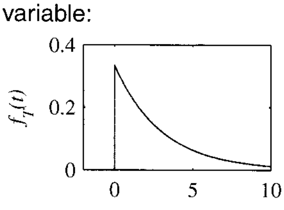
The conditional expected value and the conditional variance are

$$E[X|L] = \int_{-\infty}^{\infty} x f_{X|L}(x)\, dx = \int_{1/2}^{1} 2x\, dx = 3/4 \text{ meters}. \tag{3.123}$$

$$\operatorname{Var}[X|L] = E\left[X^2|L\right] - (E[X|L])^2 = \frac{7}{12} - \left(\frac{3}{4}\right)^2 = 1/48 \text{ m}^2. \tag{3.124}$$

The conditional standard deviation is $\sigma_{X|L} = \sqrt{\operatorname{Var}[X|L]} = 0.144$ meters. Example 3.9 derives $\sigma_X = 0.289$ meters. That is, $\sigma_X = 2\sigma_{X|L}$. It follows that learning that the pointer is on the left side of the circle leads to a set of typical values that are within 0.144 meters of 0.75 meters. Prior to learning which half of the circle the pointer is in, we had a set of typical values within 0.289 of 0.5 meters.

Example 3.34 Suppose the duration T (in minutes) of a telephone call is an exponential $(1/3)$ random variable:



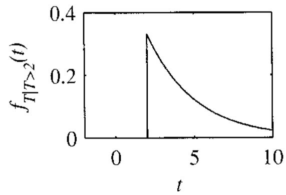
$$f_T(t) = \begin{cases} (1/3)e^{-t/3} & t \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.125}$$

For calls that last at least 2 minutes, what is the conditional PDF of the call duration?

In this case, the conditioning event is $T > 2$. The probability of the event is

$$P[T > 2] = \int_{2}^{\infty} f_T(t)\, dt = e^{-2/3}. \tag{3.126}$$

The conditional PDF of T given $T > 2$ is



$$f_{T|T>2}(t) = \begin{cases} \frac{f_T(t)}{P[T>2]} & t > 2, \\ 0 & \text{otherwise}, \end{cases} \tag{3.127}$$

$$= \begin{cases} \frac{1}{3}e^{-(t-2)/3} & t > 2, \\ 0 & \text{otherwise}. \end{cases} \tag{3.128}$$

Note that $f_{T|T>2}(t)$ is a time-shifted version of $f_T(t)$. In particular, $f_{T|T>2}(t) = f_T(t-2)$. An interpretation of this result is that if the call is in progress after 2 minutes, the duration of the call is 2 minutes plus an exponential time equal to the duration of a new call.

The conditional expected value is

$$E[T|T > 2] = \int_2^\infty t\frac{1}{3}e^{-(t-2)/3}\,dt. \tag{3.129}$$

Integration by parts (Appendix B, Math Fact B.10) yields

$$E[T|T > 2] = -te^{-(t-2)/3}\Big|_2^\infty + \int_2^\infty e^{-(t-2)/3}\,dt = 2 + 3 = 5 \text{ minutes.} \tag{3.130}$$

Recall in Example 3.13 that the expected duration of the call is $E[T] = 3$ minutes. We interpret $E[T|T > 2]$ by saying that if the call is still in progress after 2 minutes, the additional duration is 3 minutes (the same as the expected time of a new call) and the expected total time is 5 minutes.

Quiz 3.8 *The probability density function of random variable Y is*

$$f_Y(y) = \begin{cases} 1/10 & 0 \le y < 10, \\ 0 & \textit{otherwise}. \end{cases} \tag{3.131}$$

Find the following:

(1) $P[Y \le 6]$

(2) the conditional PDF $f_{Y|Y\le 6}(y)$

(3) $P[Y > 8]$

(4) the conditional PDF $f_{Y|Y>8}(y)$

(5) $E[Y|Y \le 6]$

(6) $E[Y|Y > 8]$

3.9 MATLAB

Probability Functions

Now that we have introduced continuous random variables, we can say that the built-in function `y=rand(m,n)` is MATLAB's approximation to a uniform (0, 1) random variable. It is an approximation for two reasons. First, `rand` produces pseudorandom numbers; the numbers seem random but are actually the output of a deterministic algorithm. Second, `rand` produces a double precision floating point number, represented in the computer by 64 bits. Thus MATLAB distinguishes no more than 2^{64} unique double precision floating point numbers. By comparision, there are uncountably infinite real numbers in $[0, 1)$. Even though `rand` is not random and does not have a continuous range, we can for all practical purposes use it as a source of independent sample values of the uniform $(0, 1)$ random variable.

Table 3.3 describes MATLAB functions related to four families of continuous random variables introduced in this chapter: uniform, exponential, Erlang, and Gaussian. The functions calculate directly the CDFs and PDFs of uniform and exponential random variables. The corresponding `pdf` and `cdf` functions are simply defined for our convenience. For Erlang and Gaussian random variables, the PDFs can be calculated directly but the CDFs require numerical integration. For Erlang random variables, we can use Theorem 3.11 in MATLAB:

```
function F=erlangcdf(n,lambda,x)
F=1.0-poissoncdf(lambda*x,n-1);
```

For the Gaussian CDF, we use the standard MATLAB error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\, du. \tag{3.132}$$

It is related to the Gaussian CDF by

$$\Phi(x) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right), \tag{3.133}$$

which is how we implement the MATLAB function `phi(x)`. In each function description in Table 3.3, `x` denotes a vector $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}'$. The `pdf` function output is a vector $\mathbf{y}$ such that $y_i = f_X(x_i)$. The `cdf` function output is a vector $\mathbf{y}$ such that $y_i = F_X(x_i)$. The `rv` function output is a vector $\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_m \end{bmatrix}'$ such that each X_i is a sample value of the random variable X. If $m = 1$, then the output is a single sample value of random variable X.

Random Samples

We have already employed the `rand` function to generate random samples of uniform $(0, 1)$ random variables. Conveniently, MATLAB also includes the built-in function `randn` to generate random samples of standard normal random variables. Thus we generate Gaussian (μ, σ) random variables by stretching and shifting standard normal random variables

```
function x=gaussrv(mu,sigma,m)
x=mu +(sigma*randn(m,1));
```

For other continuous random variables, we use Theorem 3.22 to transform a uniform $(0, 1)$ random variable U into other types of random variables.

Example 3.35 Use Example 3.27 to write a MATLAB program that generates m samples of an exponential (λ) random variable.

Random Variable	Matlab Function	Function Output
X Uniform (a, b)	`y=uniformpdf(a,b,x)`	$y_i = f_X(x_i)$
	`y=uniformcdf(a,b,x)`	$y_i = F_X(x_i)$
	`x=uniformrv(a,b,m)`	$\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_m \end{bmatrix}'$
X Exponential (λ)	`y=exponentialpdf(lambda,x)`	$y_i = f_X(x_i)$
	`y=exponentialcdf(lambda,x)`	$y_i = F_X(x_i)$
	`x=exponentialrv(lambda,m)`	$\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_m \end{bmatrix}'$
X Erlang (n, λ)	`y=erlangpdf(n,lambda,x)`	$y_i = f_X(x_i)$
	`y=erlangcdf(n,lambda,x)`	$y_i = F_X(x_i)$
	`x=erlangrv(n,lambda,m)`	$\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_m \end{bmatrix}'$
X Gaussian (μ, σ^2)	`y=gausspdf(mu,sigma,x)`	$y_i = f_X(x_i)$
	`y=gausscdf(mu,sigma,x)`	$y_i = F_X(x_i)$
	`x=gaussrv(mu,sigma,m)`	$\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_m \end{bmatrix}'$

Table 3.3 MATLAB functions for continuous random variables.

In Example 3.27, we found that if U is a uniform $(0, 1)$ random variable, then

$$Y = -\ln(1 - U) \tag{3.134}$$

is an exponential $(\lambda = 1)$ random variable. By Theorem 3.20(b), $X = Y/\lambda$ is an exponential (λ) random variable. Using `rand` to approximate U, we have the following MATLAB code:

```
function x=exponentialrv(lambda,m)
x=-(1/lambda)*log(1-rand(m,1));
```

Example 3.36 Use Example 3.28 to write a MATLAB function that generates m samples of a uniform (a, b) random variable.

Example 3.28 says that $Y = a + (b - a)U$ is a uniform (a, b) random variable. Thus we use the following code:

```
function x=uniformrv(a,b,m)
x=a+(b-a)*rand(m,1);
```

Theorem 6.11 will demonstrate that the sum of n independent exponential (λ) random variables is an Erlang random variable. The following code generates m sample values of the Erlang (n, λ) random variable.

```
function x=erlangrv(n,lambda,m)
y=exponentialrv(lambda,m*n);
x=sum(reshape(y,m,n),2);
```

Note that we first generate nm exponential random variables. The `reshape` function arranges these samples in an $m \times n$ array. Summing across the rows yields m Erlang samples.

Finally, for a random variable X with an arbitrary CDF $F_X(x)$, we implement the function `icdfrv.m` which uses Theorem 3.22 for generating random samples. The key is that we need to define a MATLAB function `x=icdfx(u)` that calculates $x = F_X^{-1}(u)$. The function `icdfx(u)` is then passed as an argument to `icdfrv.m` which generates samples of X. Note that MATLAB passes a function as an argument to another function using a function *handle*, which is a kind of pointer. Here is the code for `icdfrv.m`:

```
function x=icdfrv(icdfhandle,m)
%Usage: x=icdfrv(@icdf,m)
%returns m samples of rv X
%with inverse CDF icdf.m
u=rand(m,1);
x=feval(icdfhandle,u);
```

The following example shows how to use `icdfrv.m`.

Example 3.37 Write a MATLAB function that uses `icdfrv.m` to generate samples of Y, the maximum of three pointer spins, in Example 3.5.

From Equation (3.18), we see that for $0 \leq y \leq 1$, $F_Y(y) = y^3$. If $u = F_Y(y) = y^3$, then $y = F_Y^{-1}(u) = u^{1/3}$. So we define (and save to disk) `icdf3spin.m`:

```
function y = icdf3spin(u);
y=u.^(1/3);
```

Now, `y=icdfrv(@icdf3spin,1000)` generates a vector holding 1000 samples of random variable Y. The notation `@icdf3spin` is the function handle for the function `icdf3spin.m`.

Keep in mind that for the MATLAB code to run quickly, it is best for the inverse CDF function, `icdf3spin.m` in the case of the last example, to process the vector `u` without using a `for` loop to find the inverse CDF for each element `u(i)`. We also note that this same technique can be extended to cases where the inverse CDF $F_X^{-1}(u)$ does not exist for all $0 \leq u \leq 1$. For example, the inverse CDF does not exist if X is a mixed random variable or if $f_X(x)$ is constant over an interval (a, b). How to use `icdfrv.m` in these cases is addressed in Problems 3.7.18 and 3.9.9.

Quiz 3.9 *Write a* MATLAB *function* `t=t2rv(m)` *that generates m samples of a random variable with the PDF $f_{T|T>2}(t)$ as given in Example 3.34.*

Chapter Summary

This chapter introduces continuous random variables. Most of the chapter parallels Chapter 2 on discrete random variables. In the case of a continuous random variable, probabilities and expected values are integrals. For a discrete random variable, they are sums.

- *A random variable X is continuous* if the range S_X consists of one or more intervals. Each possible value of X has probability zero.
- *The PDF $f_X(x)$* is a probability model for a continuous random variable X. The PDF $f_X(x)$ is proportional to the probability that X is close to x.

- *The expected value* $E[X]$ of a continuous random variable has the same interpretation as the expected value of a discrete random variable. $E[X]$ is a typical value of X.
- *A random variable X is mixed* if it has at least one sample value with nonzero probability (like a discrete random variable) but also has sample values that cover an interval (like a continuous random variable.) The PDF of a mixed random variable contains finite nonzero values and delta functions.
- *A function of a random variable* transforms a random variable X into a new random variable $Y = g(X)$. If X is continuous, we find the probability model of Y by deriving the CDF, $F_Y(y)$, from $F_X(x)$ and $g(x)$.
- *The conditional PDF* $f_{X|B}(x)$ is a probability model of X that uses the information that $X \in B$.

Problems

Difficulty: Easy Moderate Difficult Experts Only

3.1.1 The cumulative distribution function of random variable X is

$$F_X(x) = \begin{cases} 0 & x < -1, \\ (x+1)/2 & -1 \le x < 1, \\ 1 & x \ge 1. \end{cases}$$

(a) What is $P[X > 1/2]$?

(b) What is $P[-1/2 < X \le 3/4]$?

(c) What is $P[|X| \le 1/2]$?

(d) What is the value of a such that $P[X \le a] = 0.8$?

3.1.2 The cumulative distribution function of the continuous random variable V is

$$F_V(v) = \begin{cases} 0 & v < -5, \\ c(v+5)^2 & -5 \le v < 7, \\ 1 & v \ge 7. \end{cases}$$

(a) What is c?

(b) What is $P[V > 4]$?

(c) $P[-3 < V \le 0]$?

(d) What is the value of a such that $P[V > a] = 2/3$?

3.1.3 The CDF of random variable W is

$$F_W(w) = \begin{cases} 0 & w < -5, \\ (w+5)/8 & -5 \le w < -3, \\ 1/4 & -3 \le w < 3, \\ 1/4 + 3(w-3)/8 & 3 \le w < 5, \\ 1 & w \ge 5. \end{cases}$$

(a) What is $P[W \le 4]$?

(b) What is $P[-2 < W \le 2]$?

(c) What is $P[W > 0]$?

(d) What is the value of a such that $P[W \le a] = 1/2$?

3.1.4 In this problem, we verify that $\lim_{n\to\infty} \lceil nx \rceil / n = x$.

(a) Verify that $nx \le \lceil nx \rceil \le nx + 1$.

(b) Use part (a) to show that $\lim_{n\to\infty} \lceil nx \rceil / n = x$.

(c) Use a similar argument to show that $\lim_{n\to\infty} \lfloor nx \rfloor / n = x$.

3.2.1 The random variable X has probability density function

$$f_X(x) = \begin{cases} cx & 0 \le x \le 2, \\ 0 & \text{otherwise.} \end{cases}$$

Use the PDF to find

(a) the constant c,

(b) $P[0 \le X \le 1]$,

(c) $P[-1/2 \le X \le 1/2]$,

(d) the CDF $F_X(x)$.

3.2.2 The cumulative distribution function of random variable X is

$$F_X(x) = \begin{cases} 0 & x < -1, \\ (x+1)/2 & -1 \le x < 1, \\ 1 & x \ge 1. \end{cases}$$

Find the PDF $f_X(x)$ of X.

3.2.3 Find the PDF $f_U(u)$ of the random variable U in Problem 3.1.3.

3.2.4 For a constant parameter $a > 0$, a Rayleigh random variable X has PDF

$$f_X(x) = \begin{cases} a^2 x e^{-a^2x^2/2} & x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

What is the CDF of X?

3.2.5 For constants a and b, random variable X has PDF

$$f_X(x) = \begin{cases} ax^2 + bx & 0 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

What conditions on a and b are necessary and sufficient to guarantee that $f_X(x)$ is a valid PDF?

3.3.1 Continuous random variable X has PDF

$$f_X(x) = \begin{cases} 1/4 & -1 \le x \le 3, \\ 0 & \text{otherwise.} \end{cases}$$

Define the random variable Y by $Y = h(X) = X^2$.

(a) Find $E[X]$ and $\text{Var}[X]$.

(b) Find $h(E[X])$ and $E[h(X)]$.

(c) Find $E[Y]$ and $\text{Var}[Y]$.

3.3.2 Let X be a continuous random variable with PDF

$$f_X(x) = \begin{cases} 1/8 & 1 \le x \le 9, \\ 0 & \text{otherwise.} \end{cases}$$

Let $Y = h(X) = 1/\sqrt{X}$.

(a) Find $E[X]$ and $\text{Var}[X]$.

(b) Find $h(E[X])$ and $E[h(X)]$.

(c) Find $E[Y]$ and $\text{Var}[Y]$.

3.3.3 Random variable X has CDF

$$F_X(x) = \begin{cases} 0 & x < 0, \\ x/2 & 0 \le x \le 2, \\ 1 & x > 2. \end{cases}$$

(a) What is $E[X]$?

(b) What is $\text{Var}[X]$?

3.3.4 The probability density function of random variable Y is

$$f_Y(y) = \begin{cases} y/2 & 0 \le y < 2, \\ 0 & \text{otherwise.} \end{cases}$$

What are $E[Y]$ and $\text{Var}[Y]$?

3.3.5 The cumulative distribution function of the random variable Y is

$$F_Y(y) = \begin{cases} 0 & y < -1, \\ (y+1)/2 & -1 \le y \le 1, \\ 1 & y > 1. \end{cases}$$

(a) What is $E[Y]$?

(b) What is $\text{Var}[Y]$?

3.3.6 The cumulative distribution function of random variable V is

$$F_V(v) = \begin{cases} 0 & v < -5, \\ (v+5)^2/144 & -5 \le v < 7, \\ 1 & v \ge 7. \end{cases}$$

(a) What is $E[V]$?

(b) What is $\text{Var}[V]$?

(c) What is $E[V^3]$?

3.3.7 The cumulative distribution function of random variable U is

$$F_U(u) = \begin{cases} 0 & u < -5, \\ (u+5)/8 & -5 \le u < -3m, \\ 1/4 & -3 \le u < 3, \\ 1/4 + 3(u-3)/8 & 3 \le u < 5, \\ 1 & u \ge 5. \end{cases}$$

(a) What is $E[U]$?

(b) What is $\text{Var}[U]$?

(c) What is $E[2^U]$?

3.3.8 X is a Pareto (α, μ) random variable, as defined in Appendix A. What is the largest value of n for which the nth moment $E[X^n]$ exists? For all feasible values of n, find $E[X^n]$.

3.4.1 Radars detect flying objects by measuring the power reflected from them. The reflected power of an aircraft can be modeled as a random variable Y with PDF

$$f_Y(y) = \begin{cases} \frac{1}{P_0} e^{-y/P_0} & y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

where $P_0 > 0$ is some constant. The aircraft is correctly identified by the radar if the reflected power of the aircraft is larger than its average value. What is the probability $P[C]$ that an aircraft is correctly identified?

3.4.2 Y is an exponential random variable with variance $\text{Var}[Y] = 25$.

(a) What is the PDF of Y?

(b) What is $E[Y^2]$?

(c) What is $P[Y > 5]$?

3.4.3 X is an Erlang (n, λ) random variable with parameter $\lambda = 1/3$ and expected value $E[X] = 15$.

(a) What is the value of the parameter n?

(b) What is the PDF of X?

(c) What is $\mathrm{Var}[X]$?

3.4.4 Y is an Erlang $(n = 2, \lambda = 2)$ random variable.

(a) What is $E[Y]$?

(b) What is $\mathrm{Var}[Y]$?

(c) What is $P[0.5 \leq Y < 1.5]$?

3.4.5 X is a continuous uniform $(-5, 5)$ random variable.

(a) What is the PDF $f_X(x)$?

(b) What is the CDF $F_X(x)$?

(c) What is $E[X]$?

(d) What is $E[X^5]$?

(e) What is $E[e^X]$?

3.4.6 X is a uniform random variable with expected value $\mu_X = 7$ and variance $\mathrm{Var}[X] = 3$. What is the PDF of X?

3.4.7 The probability density function of random variable X is

$$f_X(x) = \begin{cases} (1/2)e^{-x/2} & x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

(a) What is $P[1 \leq X \leq 2]$?

(b) What is $F_X(x)$, the cumulative distribution function of X?

(c) What is $E[X]$, the expected value of X?

(d) What is $\mathrm{Var}[X]$, the variance of X?

3.4.8 Verify parts (b) and (c) of Theorem 3.6 by directly calculating the expected value and variance of a uniform random variable with parameters $a < b$.

3.4.9 Long-distance calling plan A offers flat rate service at 10 cents per minute. Calling plan B charges 99 cents for every call under 20 minutes; for calls over 20 minutes, the charge is 99 cents for the first 20 minutes plus 10 cents for every additional minute. (Note that these plans measure your call duration exactly, without rounding to the next minute or even second.) If your long-distance calls have exponential distribution with expected value τ minutes, which plan offers a lower expected cost per call?

3.4.10 In this problem we verify that an Erlang (n, λ) PDF integrates to 1. Let the integral of the nth order Erlang PDF be denoted by

$$I_n = \int_0^\infty \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}\, dx.$$

First, show directly that the Erlang PDF with $n = 1$ integrates to 1 by verifying that $I_1 = 1$. Second, use integration by parts (Appendix B, Math Fact B.10) to show that $I_n = I_{n-1}$.

3.4.11 Calculate the kth moment $E[X^k]$ of an Erlang (n, λ) random variable X. Use your result to verify Theorem 3.10. Hint: Remember that the Erlang $(n + k, \lambda)$ PDF integrates to 1.

3.4.12 In this problem, we outline the proof of Theorem 3.11.

(a) Let X_n denote an Erlang (n, λ) random variable. Use the definition of the Erlang PDF to show that for any $x \geq 0$,

$$F_{X_n}(x) = \int_0^x \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}\, dt.$$

(b) Apply integration by parts (Appendix B, Math Fact B.10) to this integral to show that for $x \geq 0$,

$$F_{X_n}(x) = F_{X_{n-1}}(x) - \frac{(\lambda x)^{n-1} e^{-\lambda x}}{(n-1)!}.$$

(c) Use the fact that $F_{X_1}(x) = 1 - e^{-\lambda x}$ for $x \geq 0$ to verify the claim of Theorem 3.11.

3.4.13 Prove by induction that an exponential random variable X with expected value $1/\lambda$ has nth moment

$$E\left[X^n\right] = \frac{n!}{\lambda^n}.$$

Hint: Use integration by parts (Appendix B, Math Fact B.10).

3.4.14 This problem outlines the steps needed to show that a nonnegative continuous random variable X has expected value

$$E[X] = \int_0^\infty x f_X(x)\, dx = \int_0^\infty [1 - F_X(x)]\, dx.$$

(a) For any $r \geq 0$, show that

$$rP[X > r] \leq \int_r^\infty x f_X(x)\, dx.$$

(b) Use part (a) to argue that if $E[X] < \infty$, then

$$\lim_{r\to\infty} rP[X > r] = 0.$$

(c) Now use integration by parts (Appendix B, Math Fact B.10) to evaluate

$$\int_0^\infty [1 - F_X(x)]\, dx.$$

3.5.1 The peak temperature T, as measured in degrees Fahrenheit, on a July day in New Jersey is the Gaussian (85, 10) random variable. What is $P[T > 100]$, $P[T < 60]$, and $P[70 \le T \le 100]$?

3.5.2 What is the PDF of Z, the standard normal random variable?

3.5.3 X is a Gaussian random variable with $E[X] = 0$ and $P[|X| \le 10] = 0.1$. What is the standard deviation σ_X?

3.5.4 A function commonly used in communications textbooks for the tail probabilities of Gaussian random variables is the complementary error function, defined as

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-x^2}\, dx.$$

Show that

$$Q(z) = \frac{1}{2}\operatorname{erfc}\left(\frac{z}{\sqrt{2}}\right).$$

3.5.5 The peak temperature T, in degrees Fahrenheit, on a July day in Antarctica is a Gaussian random variable with a variance of 225. With probability 1/2, the temperature T exceeds 10 degrees. What is $P[T > 32]$, the probability the temperature is above freezing? What is $P[T < 0]$? What is $P[T > 60]$?

3.5.6 A professor pays 25 cents for each blackboard error made in lecture to the student who points out the error. In a career of n years filled with blackboard errors, the total amount in dollars paid can be approximated by a Gaussian random variable Y_n with expected value $40n$ and variance $100n$. What is the probability that Y_{20} exceeds 1000? How many years n must the professor teach in order that $P[Y_n > 1000] > 0.99$?

3.5.7 Suppose that out of 100 million men in the United States, 23,000 are at least 7 feet tall. Suppose that the heights of U.S. men are independent Gaussian random variables with a expected value of 5′10″. Let N equal the number of men who are at least 7′6″ tall.

(a) Calculate σ_X, the standard deviation of the height of men in the United States.

(b) In terms of the $\Phi(\cdot)$ function, what is the probability that a randomly chosen man is at least 8 feet tall?

(c) What is the probability that there is no man alive in the U.S. today that is at least 7′6″ tall?

(d) What is $E[N]$?

3.5.8 In this problem, we verify that for $x \ge 0$,

$$\Phi(x) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right).$$

(a) Let Y have a Gaussian $(0, 1/\sqrt{2})$ distribution and show that

$$F_Y(y) = \int_{-\infty}^{y} f_Y(u)\, du = \frac{1}{2} + \operatorname{erf}(y).$$

(b) Observe that $Z = \sqrt{2}Y$ is Gaussian (0, 1) and show that

$$\Phi(z) = F_Z(z) = P\left[Y \le \frac{z}{\sqrt{2}}\right] = F_Y\left(\frac{z}{\sqrt{2}}\right).$$

3.5.9 This problem outlines the steps needed to show that the Gaussian PDF integrates to unity. For a Gaussian (μ, σ) random variable W, we will show that

$$I = \int_{-\infty}^{\infty} f_W(w)\, dw = 1.$$

(a) Use the substitution $x = (w - \mu)/\sigma$ to show that

$$I = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\, dx.$$

(b) Show that

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\, dx\, dy.$$

(c) Change the integral for I^2 to polar coordinates to show that it integrates to 1.

3.5.10 In mobile radio communications, the radio channel can vary randomly. In particular, in communicating with a fixed transmitter power over a "Rayleigh fading" channel, the receiver signal-to-noise ratio Y is an exponential random variable with expected value γ. Moreover, when $Y = y$, the probability of an error in decoding a transmitted bit is $P_e(y) = Q(\sqrt{2y})$ where $Q(\cdot)$ is the standard normal complementary CDF. The average probability of bit error, also known as the bit error rate or BER, is

$$\overline{P}_e = E\left[P_e(Y)\right] = \int_{-\infty}^{\infty} Q(\sqrt{2y}) f_Y(y)\, dy.$$

Find a simple formula for the BER $\overline{P}_e$ as a function of the average SNR γ.

3.6.1 Let X be a random variable with CDF

$$F_X(x) = \begin{cases} 0 & x < -1, \\ x/3 + 1/3 & -1 \le x < 0, \\ x/3 + 2/3 & 0 \le x < 1, \\ 1 & 1 \le x. \end{cases}$$

Sketch the CDF and find

(a) $P[X < -1]$ and $P[X \le -1]$,

(b) $P[X < 0]$ and $P[X \le 0]$,

(c) $P[0 < X \le 1]$ and $P[0 \le X \le 1]$.

3.6.2 Let X be a random variable with CDF

$$F_X(x) = \begin{cases} 0 & x < -1, \\ x/4 + 1/2 & -1 \le x < 1, \\ 1 & 1 \le x. \end{cases}$$

Sketch the CDF and find

(a) $P[X < -1]$ and $P[X \le -1]$,

(b) $P[X < 0]$ and $P[X \le 0]$,

(c) $P[X > 1]$ and $P[X \ge 1]$.

3.6.3 For random variable X of Problem 3.6.2, find

(a) $f_X(x)$

(b) $E[X]$

(c) $\text{Var}[X]$

3.6.4 X is Bernoulli random variable with expected value p. What is the PDF $f_X(x)$?

3.6.5 X is a geometric random variable with expected value $1/p$. What is the PDF $f_X(x)$?

3.6.6 When you make a phone call, the line is busy with probability 0.2 and no one answers with probability 0.3. The random variable X describes the conversation time (in minutes) of a phone call that is answered. X is an exponential random variable with $E[X] = 3$ minutes. Let the random variable W denote the conversation time (in seconds) of all calls ($W = 0$ when the line is busy or there is no answer.)

(a) What is $F_W(w)$?

(b) What is $f_W(w)$?

(c) What are $E[W]$ and $\text{Var}[W]$?

3.6.7 For 80% of lectures, Professor X arrives on time and starts lecturing with delay $T = 0$. When Professor X is late, the starting time delay T is uniformly distributed between 0 and 300 seconds. Find the CDF and PDF of T.

3.6.8 With probability 0.7, the toss of an Olympic shot-putter travels $D = 60 + X$ feet, where X is an exponential random variable with expected value $\mu = 10$. Otherwise, with probability 0.3, a foul is committed by stepping outside of the shot-put circle and we say $D = 0$. What are the CDF and PDF of random variable D?

3.6.9 For 70% of lectures, Professor Y arrives on time. When Professor Y is late, the arrival time delay is a continuous random variable uniformly distributed from 0 to 10 minutes. Yet, as soon as Professor Y is 5 minutes late, all the students get up and leave. (It is unknown if Professor Y still conducts the lecture.) If a lecture starts when Professor Y arrives and always ends 80 minutes after the scheduled starting time, what is the PDF of T, the length of time that the students observe a lecture.

3.7.1 The voltage X across a 1 Ω resistor is a uniform random variable with parameters 0 and 1. The instantaneous power is $Y = X^2$. Find the CDF $F_Y(y)$ and the PDF $f_Y(y)$ of Y.

3.7.2 Let X have an exponential (λ) PDF. Find the CDF and PDF of $Y = \sqrt{X}$. Show that Y is a Rayleigh random variable (see Appendix A.2). Express the Rayleigh parameter a in terms of the exponential parameter λ.

3.7.3 If X has an exponential (λ) PDF, what is the PDF of $W = X^2$?

3.7.4 X is the random variable in Problem 3.6.1. $Y = g(X)$ where

$$g(X) = \begin{cases} 0 & X < 0, \\ 100 & X \ge 0. \end{cases}$$

(a) What is $F_Y(y)$?

(b) What is $f_Y(y)$?

(c) What is $E[Y]$?

3.7.5 U is a uniform (0, 1) random variable and $X = -\ln(1 - U)$.

(a) What is $F_X(x)$?

(b) What is $f_X(x)$?

(c) What is $E[X]$?

3.7.6 X is uniform random variable with parameters 0 and 1. Find a function $g(x)$ such that the PDF of $Y = g(X)$ is

$$f_Y(y) = \begin{cases} 3y^2 & 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

3.7.7 The voltage V at the output of a microphone is a uniform random variable with limits -1 volt and 1 volt. The microphone voltage is processed by a hard limiter with cutoff points -0.5 volt and 0.5 volt. The magnitude of the limiter output L is a random variable such that

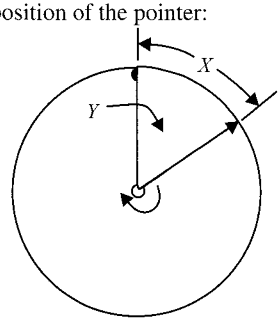
$$L = \begin{cases} V & |V| \le 0.5, \\ 0.5 & \text{otherwise.} \end{cases}$$

(a) What is $P[L = 0.5]$?

(b) What is $F_L(l)$?

(c) What is $E[L]$?

3.7.8 Let X denote the position of the pointer after a spin on a wheel of circumference 1. For that same spin, let Y denote the area within the arc defined by the stopping position of the pointer:



(a) What is the relationship between X and Y?

(b) What is $F_Y(y)$?

(c) What is $f_Y(y)$?

(d) What is $E[Y]$?

3.7.9 U is a uniform random variable with parameters 0 and 2. The random variable W is the output of the clipper:

$$W = g(U) = \begin{cases} U & U \le 1, \\ 1 & U > 1. \end{cases}$$

Find the CDF $F_W(w)$, the PDF $f_W(w)$, and the expected value $E[W]$.

3.7.10 X is a random variable with CDF $F_X(x)$. Let $Y = g(X)$ where

$$g(x) = \begin{cases} 10 & x < 0, \\ -10 & x \ge 0. \end{cases}$$

Express $F_Y(y)$ in terms of $F_X(x)$.

3.7.11 The input voltage to a rectifier is a random variable U with a uniform distribution on $[-1, 1]$. The rectifier output is a random variable W defined by

$$W = g(U) = \begin{cases} 0 & U < 0, \\ U & U \ge 0. \end{cases}$$

Find the CDF $F_W(w)$ and the expected value $E[W]$.

3.7.12 Use Theorem 3.19 to prove Theorem 3.20.

3.7.13 For a uniform $(0, 1)$ random variable U, find the CDF and PDF of $Y = a + (b - a)U$ with $a < b$. Show that Y is a uniform (a, b) random variable.

3.7.14 Theorem 3.22 required the inverse CDF $F^{-1}(u)$ to exist for $0 < u < 1$. Why was it *not* necessary that $F^{-1}(u)$ exist at either $u = 0$ or $u = 1$.

3.7.15 Random variable X has PDF

$$f_X(x) = \begin{cases} x/2 & 0 \le x \le 2, \\ 0 & \text{otherwise.} \end{cases}$$

X is processed by a clipping circuit with output Y. The circuit is defined by:

$$Y = \begin{cases} 0.5 & 0 \le X \le 1, \\ X & X > 1. \end{cases}$$

(a) What is $P[Y = 0.5]$?

(b) Find the CDF $F_Y(y)$.

3.7.16 X is a continuous random variable. $Y = aX + b$, where $a, b \ne 0$. Prove that

$$f_Y(y) = \frac{f_X((y - b)/a)}{|a|}.$$

Hint: Consider the cases $a < 0$ and $a > 0$ separately.

3.7.17 Let continuous random variable X have a CDF $F(x)$ such that $F^{-1}(u)$ exists for all u in $[0, 1]$. Show that $U = F(X)$ is uniformly distributed over $[0, 1]$. Hint: U is a random variable such that when

$X = x'$, $U = F(x')$. That is, we evaluate the CDF of X at the observed value of X.

3.7.18 In this problem we prove a generalization of Theorem 3.22. Given a random variable X with CDF $F_X(x)$, define

$$\tilde{F}(u) = \min\{x | F_X(x) \geq u\}.$$

This problem proves that for a continuous uniform $(0, 1)$ random variable U, $\hat{X} = \tilde{F}(U)$ has CDF $F_{\hat{X}}(x) = F_X(x)$.

(a) Show that when $F_X(x)$ is a continuous, strictly increasing function (i.e., X is not mixed, $F_X(x)$ has no jump discontinuities, and $F_X(x)$ has no "flat" intervals (a, b) where $F_X(x) = c$ for $a \leq x \leq b$), then $\tilde{F}(u) = F_X^{-1}(u)$ for $0 < u < 1$.

(b) Show that if $F_X(x)$ has a jump at $x = x_0$, then $\tilde{F}(u) = x_0$ for all u in the interval

$$F_X\left(x_0^-\right) \leq u \leq F_X\left(x_0^+\right).$$

(c) Prove that $\hat{X} = \tilde{F}(U)$ has CDF $F_{\hat{X}}(x) = F_X(x)$.

3.8.1 X is a uniform random variable with parameters -5 and 5. Given the event $B = \{|X| \leq 3\}$,

(a) Find the conditional PDF, $f_{X|B}(x)$.

(b) Find the conditional expected value, $E[X|B]$.

(c) What is the conditional variance, $\text{Var}[X|B]$?

3.8.2 Y is an exponential random variable with parameter $\lambda = 0.2$. Given the event $A = \{Y < 2\}$,

(a) What is the conditional PDF, $f_{Y|A}(y)$?

(b) Find the conditional expected value, $E[Y|A]$.

3.8.3 For the experiment of spinning the pointer three times and observing the maximum pointer position, Example 3.5, find the conditional PDF given the event R that the maximum position is on the right side of the circle. What are the conditional expected value and the conditional variance?

3.8.4 W is a Gaussian random variable with expected value $\mu = 0$, and variance $\sigma^2 = 16$. Given the event $C = \{W > 0\}$,

(a) What is the conditional PDF, $f_{W|C}(w)$?

(b) Find the conditional expected value, $E[W|C]$.

(c) Find the conditional variance, $\text{Var}[W|C]$.

3.8.5 The time between telephone calls at a telephone switch is an exponential random variable T with expected value 0.01. Given $T > 0.02$,

(a) What is $E[T|T > 0.02]$, the conditional expected value of T?

(b) What is $\text{Var}[T|T > 0.02]$, the conditional variance of T?

3.8.6 For the distance D of a shot-put toss in Problem 3.6.8, find

(a) the conditional PDF of D given that $D > 0$,

(b) the conditional PDF of D given $D \leq 70$.

3.8.7 A test for diabetes is a measurement X of a person's blood sugar level following an overnight fast. For a healthy person, a blood sugar level X in the range of $70 - 110$ mg/dl is considered normal. When a measurement X is used as a test for diabetes, the result is called positive (event T^+) if $X \geq 140$; the test is negative (event T^-) if $X \leq 110$, and the test is ambiguous (event T^0) if $110 < X < 140$.

Given that a person is healthy (event H), a blood sugar measurement X is a Gaussian ($\mu = 90, \sigma = 20$) random variable. Given that a person has diabetes, (event D), X is a Gaussian ($\mu = 160, \sigma = 40$) random variable. A randomly chosen person is healthy with probability $P[H] = 0.9$ or has diabetes with probability $P[D] = 0.1$.

(a) What is the conditional PDF $f_{X|H}(x)$?

(b) In terms of the $\Phi(\cdot)$ function, find the conditional probabilities $P[T^+|H]$, and $P[T^-|H]$.

(c) Find the conditional conditional probability $P[H|T^-]$ that a person is healthy given the event of a negative test.

(d) When a person has an ambiguous test result, (T^0) the test is repeated, possibly many times, until either a positive T^+ or negative T^- result is obtained. Let N denote the number of times the test is given. Assuming that for a given person, the result of each test is independent of the result of all other tests, find the condtional PMF of N given event H that a person is healthy. Note that $N = 1$ if the person has a positive T^+ or negative result T^- on the first test.

3.8.8 For the quantizer of Example 3.31, the difference $Z = X - Y$ is the quantization error or quantization "noise." As in Example 3.31, assume that X has a uniform $(-r/2, r/2)$ PDF.

(a) Given event B_i that $Y = y_i = \Delta/2 + i\Delta$ and X is in the ith quantization interval, find the conditional PDF of Z.

(b) Show that Z is a uniform random variable. Find the PDF, the expected value, and the variance of Z.

3.8.9 For the quantizer of Example 3.31, we showed in Problem 3.8.8 that the quantization noise Z is a uniform random variable. If X is not uniform, show that Z is nonuniform by calculating the PDF of Z for a simple example.

3.9.1 Write a MATLAB function `y=quiz31rv(m)` that produces m samples of random variable Y defined in Quiz 3.1.

3.9.2 For the modem receiver voltage X with PDF given in Example 3.32, use MATLAB to plot the PDF and CDF of random variable X. Write a MATLAB function `x=modemrv(m)` that produces `m` samples of the modem voltage X.

3.9.3 For the Gaussian (0, 1) complementary CDF $Q(z)$, a useful numerical approximation for $z \geq 0$ is

$$\hat{Q}(z) = (a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5) e^{-z^2/2},$$

where

$$t = \frac{1}{1 + 0.231641888z} \qquad a_1 = 0.127414796$$
$$a_2 = -0.142248368 \qquad a_3 = 0.7107068705$$
$$a_4 = -0.7265760135 \qquad a_5 = 0.5307027145$$

To compare this approximation to $Q(z)$, use MATLAB to graph

$$e(z) = \frac{Q(z) - \hat{Q}(z)}{Q(z)}.$$

3.9.4 Use Theorem 3.9 and `exponentialrv.m` to write a MATLAB function `k=georv(p,m)` that generates m samples of a geometric (p) random variable K. Compare the resulting algorithm to the technique employed in Problem 2.10.7 for `geometricrv(p,m)`.

3.9.5 Use `icdfrv.m` to write a function `w=wrv1(m)` that generates m samples of random variable W from Problem 3.1.3. Note that $F_W^{-1}(u)$ does not exist for $u = 1/4$; however, you must define a function `icdfw(u)` that returns a value for `icdfw(0.25)`. Does it matter what value you return for `u=0.25`?

3.9.6 Applying Equation (3.14) with x replaced by $i\Delta$ and dx replaced by Δ, we obtain

$$P[i\Delta < X \leq i\Delta + \Delta] = f_X(i\Delta)\,\Delta.$$

If we generate a large number n of samples of random variable X, let n_i denote the number of occurrences of the event

$$\{i\Delta < X \leq (i+1)\Delta\}.$$

We would expect that $\lim_{n\to\infty} \frac{n_i}{n} = f_X(i\Delta)\Delta$, or equivalently,

$$\lim_{n\to\infty} \frac{n_i}{n\Delta} = f_X(i\Delta).$$

Use MATLAB to confirm this with $\Delta = 0.01$ for

(a) an exponential ($\lambda = 1$) random variable X and for $i = 0, \ldots, 500$,

(b) a Gaussian (3, 1) random variable X and for $i = 0, \ldots, 600$.

3.9.7 For the quantizer of Example 3.31, we showed in Problem 3.8.9 that the quantization noise Z is nonuniform if X is nonuniform. In this problem, we examine whether it is a reasonable approximation to model the quantization noise as uniform. Consider the special case of a Gaussian (0, 1) random variable X passed through a uniform b-bit quantizer over the interval $(-r/2, r/2)$ with $r = 6$. Does a uniform approximation get better or worse as b increases? Write a MATLAB program to generate histograms for Z to answer this question.

3.9.8 Write a MATLAB function `u=urv(m)` that generates m samples of random variable U defined in Problem 3.3.7.

3.9.9 Write a MATLAB function `y=quiz36rv(m)` that returns m samples of the random variable X defined in Quiz 3.6. Since $F_X^{-1}(u)$ is not defined for $1/2 \leq u < 1$, you will need to use the result of Problem 3.7.18.