Chapter 10

Continuous Random Variables

10.1 Introduction

In Chapters 5–9 we discussed discrete random variables and the methods employed to describe them probabilistically. The principal assumption necessary in order to do so is that the sample space, which is the set of all possible outcomes, is finite or at most countably infinite. It followed then that a probability mass function (PMF) could be defined as the probability of each sample point and used to calculate the probability of all possible events (which are subsets of the sample space). Most physical measurements, however, do not produce a discrete set of values but rather a *continuum* of values such as the rainfall measurement data previously shown in Figures 1.1 and 1.2. Another example is the maximum temperature measured during the day, which might be anywhere between 20°F and 60°F. The number of possible temperatures in the interval $[20, 60]$ is infinite and *uncountable*. Therefore, we cannot assign a valid PMF to the temperature random variable. Of course, we could always choose to "round off" the measurement to the nearest degree so that the possible outcomes would then become $\{20, 21, \ldots, 60\}$. Then, many valid PMFs could be assigned. But this approach compromises the measurement precision and so is to be avoided if possible. What we are ultimately interested in is the probability of *any interval*, such as the probability of the temperature being in the interval $[20, 25]$ or $[55, 60]$ or the union of intervals $[20, 25] \cup [55, 60]$. To do so we must extend our previous approaches to be able to handle this new case. And if we later decide that less precision is warranted, such that the rounding of 20.6° to 21° is acceptable, we will still be able to determine the probability of observing 21°. To do so we can regard the rounded temperature of 21° as having arisen from all temperatures in the interval $A = [20.5, 21.5)$. Then, $P[\text{rounded temperature } = 21] = P[A]$, so that we have lost nothing by considering a continuum of outcomes (see Problem 10.2).

Chapters 10–14 discuss continuous random variables in a manner similar to Chapters 5–9 for discrete random variables. Since many of the concepts are the same, we will not belabor the discussion but will concentrate our efforts on the algebraic manipulations required to analyze continuous random variables. It may be of interest to note that discrete and continuous random variables can be subsumed under the topic of a general random variable. There exists the mathematical machinery to analyze both types of random variables simultaneously. This theory is called *measure theory* [Capinski, Kopp 2004]. It requires an advanced mathematical background and does not easily lend itself to intuitive interpretations. An alternative means of describing the general random variable that appeals more to engineers and scientists makes use of the Dirac delta function. This approach is discussed later in this chapter under the topic of *mixed random variables.*

In the course of our discussions we will revisit some of the concepts alluded to in Chapters 1 and 2. With the appropriate mathematical tools we will now be able to define these concepts. Hence, the reader may wish to review the relevant sections in those chapters.

10.2 Summary

The definition of a continuous random variable is given in Section 10.3 and illustrated in Figure 10.1. The probabilistic description of a continuous random variable is the probability density function (PDF) $p_X(x)$ with its interpretation as the probability per unit length. As such the probability of an interval is given by the area under the PDF (10.4). The properties of a PDF are that it is nonnegative and integrates to one, as summarized by Properties 10.1 and 10.2 in Section 10.4. Some important PDFs are given in Section 10.5, such as the uniform (10.6), the exponential (10.5), the Gaussian or normal (10.7), the Laplacian (10.8), the Cauchy (10.9), the Gamma (10.10), and the Rayleigh (10.14). Special cases of the Gamma PDF are the exponential, the chi-squared (10.12), and the Erlang (10.13). The cumulative distribution function (CDF) for a continuous random variable is defined the same as for the discrete random variable and is given by (10.16). The corresponding CDFs for the PDFs of Section 10.5 are given in Section 10.6. In particular, the CDF for the standard normal is denoted by $\Phi(x)$ and is related to the Q function by (10.17). The latter function cannot be evaluated in closed form but may be found numerically using the MATLAB subprogram `Q.m` listed in Appendix 10B. An approximation to the Q function is given by (10.23). The CDF is useful in that probabilities of intervals are easily found via (10.25) once the CDF is known. The transformation of a continuous random variable by a one-to-one function produces the PDF of (10.30). If the transformation is many-to-one, then (10.33) can be used to determine the PDF of the transformed random variable. Mixed random variables, ones that exhibit nonzero probabilities for some points but are continuous otherwise, are described in Section 10.8. They can be described by a PDF if we allow the use of the Dirac delta function or impulse. For a general mixed random variable the PDF is given by (10.36). To generate realizations of a continuous random variable on a digital computer one can use a transformation of a uniform random variable as summarized in Theorem 10.9.1. Examples are given in Section 10.9. Estimation of the PDF and CDF can be accomplished by using (10.38) and (10.39). Finally, an example of the application of the theory to the problem of speech clipping is given in Section 10.10.

10.3 Definition of a Continuous Random Variable

A *continuous random variable* X is defined as a mapping from the experimental sample space $\mathcal{S}$ to a numerical (or measurement) sample space $\mathcal{S}_X$, which is a subset of the real line R^1. In contrast to the sample space of a discrete random variable, $\mathcal{S}_X$ consists of an infinite *and uncountable* number of outcomes. As an example, consider an experiment in which a dart is thrown at the circular dartboard shown in Figure 10.1. The outcome of the dart-throwing experiment is a *point* s_1 in the circle

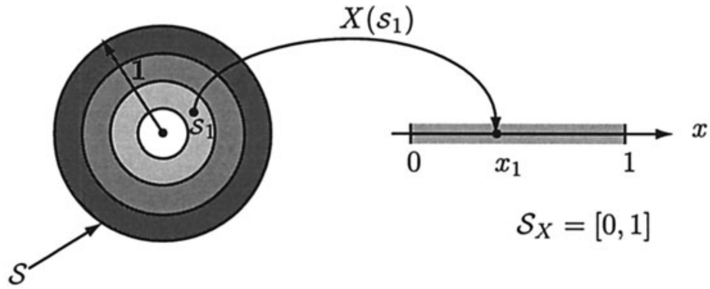


Figure 10.1: Mapping of the outcome of a thrown dart to the real line (example of continuous random variable).

of radius one. The distance from the bullseye (center of the dartboard) is measured and that value is assigned to the random variable as $X(s_1) = x_1$. Clearly then, the possible outcomes of the random variable are in the interval $[0,1]$, which is an uncountably infinite set. We cannot assign a nonzero probability to each value of X and expect the sum of the probabilities to be one. One way out of this dilemma is to assign probabilities to intervals, as was done in Section 3.6. There we had a one-dimensional dartboard and we assigned a probability of the dart landing in an interval to be the *length* of the interval. Similarly, for our problem if each value of X is equally likely so that intervals of the same length are equally likely, we could assign

$$P[a \leq X \leq b] = b - a \qquad 0 \leq a \leq b \leq 1 \tag{10.1}$$

for the probability of the dart landing in the interval $[a,b]$. This probability assignment satisfies the probability axioms given in Section 3.6 and so would suffice to calculate the probability of any interval or union of disjoint intervals (use Axiom 3 for disjoint intervals). But what would we do if the probability of all equal length intervals were not the same? For example, a champion dart thrower would be more likely to obtain a value near $x = 0$ than near $x = 1$. We therefore need a more general approach. For discrete random variables it was just as easy to assign PMFs that were not uniform as ones that were uniform. Our goal then is to extend this approach to encompass continuous random variables. We will do so by examining the approximation afforded by using the PMF to calculate interval probabilities for continuous random variables.

Consider first a possible approximation of (10.1) by a uniform PMF as

$$p_X[x_i] = \frac{1}{M} \qquad x_i = i\Delta x \text{ for } i = 1, 2, \ldots, M$$

where $\Delta x = 1/M$, so that $M\Delta x = 1$ as shown in Figure 10.2. Then to approximate

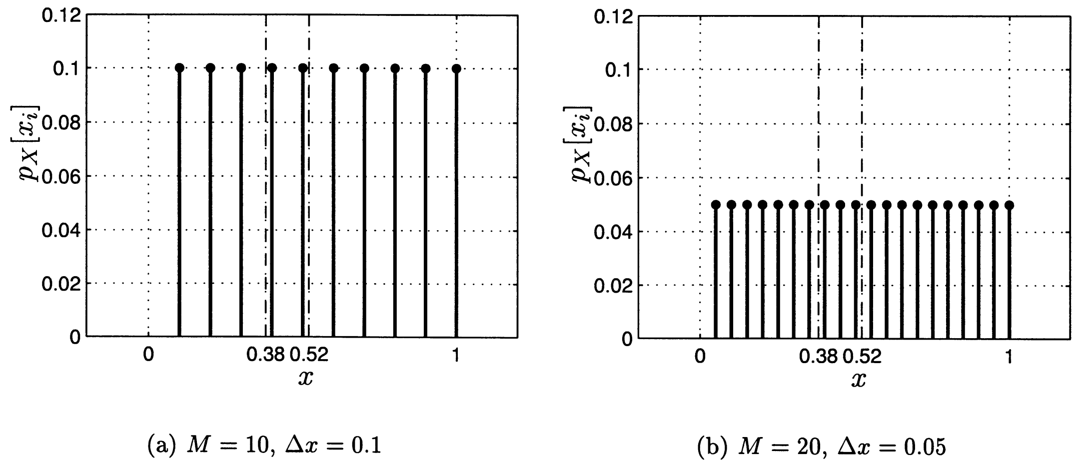


Figure 10.2: Approximating the probability of an interval for a continuous random variable by using a PMF.

the probability of the outcome of X in the interval $[a, b]$ we can use

$$P[a \leq X \leq b] = \sum_{\{i: a \leq x_i \leq b\}} \frac{1}{M}. \tag{10.2}$$

For example, referring to Figure 10.2a, if $a = 0.38$ and $b = 0.52$, then there are two values of x_i that lie in that interval and therefore $P[0.38 \leq X \leq 0.52] = 2/M = 0.2$, even though we know that the true value from (10.1) is 0.14 . To improve the quality of our approximation we increase M to $M = 20$ as shown in Figure 10.2b. Then, we have three values of x_i that lie in the interval and therefore $P[0.38 \leq X \leq 0.52] = 3/M = 0.15$, which is closer to the true value. Clearly, if we let $M \to \infty$ or equivalently let $\Delta x \to 0$, our approximation will become exact. Considering again

(10.2) with $\Delta x = 1/M$, we have

$$P[a \leq X \leq b] = \sum_{\{i:a \leq x_i \leq b\}} 1 \cdot \Delta x$$

and defining $p_X(x) = 1$ for $0 < x < 1$ and zero otherwise, we can write this as

$$P[a \leq X \leq b] = \sum_{\{i:a \leq x_i \leq b\}} p_X(x_i)\Delta x. \qquad (10.3)$$

Finally, letting $\Delta x \to 0$ to yield no error in the approximation, the sum in (10.3) becomes an integral and $p_X(x_i) \to p_X(x)$ so that

$$P[a \leq X \leq b] = \int_a^b p_X(x)dx \qquad (10.4)$$

which gives the same result for the probability of an interval as (10.1). Note that $p_X(x)$ is defined to be 1 for *all* $0 < x < 1$. To interpret this new function $p_X(x)$ we have from (10.3) with $x_0 = k\Delta x$ for k an integer

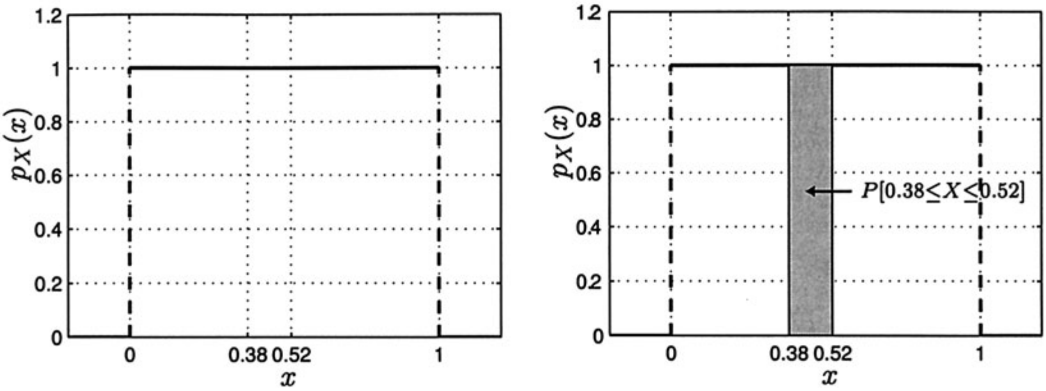
$$P[x_0 - \Delta x/2 \leq X \leq x_0 + \Delta x/2]$$

$$\begin{aligned} &= \sum_{\{i:x_0-\Delta x/2 \leq x_i \leq x_0+\Delta x/2\}} p_X(x_i)\Delta x \\ &= \sum_{\{i:x_i=x_0\}} p_X(x_i)\Delta x \qquad \text{(only one value of } x_i \text{ within interval)} \\ &= p_X(x_0)\Delta x \end{aligned}$$

which yields

$$p_X(x_0) = \frac{P[x_0 - \Delta x/2 \leq X \leq x_0 + \Delta x/2]}{\Delta x}.$$

This is the probability of X being in the interval $[x_0 - \Delta x/2, x_0 + \Delta x/2]$ divided by the interval length Δx. Hence, $p_X(x_0)$ is the *probability per unit length* and is termed the *probability density function* (PDF). It can be used to find the probability of any interval by using (10.4). Equivalently, since the value of an integral may be interpreted as the area under a curve, the probability is found by determining the area under the PDF curve. This is shown in Figure 10.3. The PDF is denoted by $p_X(x)$, where we now use *parentheses* since the argument is no longer discrete but continuous. Also, for the same reason we omit the subscript i, which was used for the PMF argument. Hence, the PDF for a continuous random variable is the extension of the PMF that we sought. Before continuing we examine this example further.



(a) Probability density function (b) Probability shown as shaded area

Figure 10.3: Example of probability density function and how probability is found as the area under it.

Example 10.1 – PDF for a uniform random variable and the MATLAB command rand

The PDF given by

$$p_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

is known as a *uniform* PDF. Equivalently, X is said to be a uniform random variable or we say that X is uniformly distributed on $(0, 1)$. The shorthand notation is $X \sim \mathcal{U}(0, 1)$. Observe that this is the continuous random variable for which MATLAB uses `rand` to produce a realization. Hence, in simulating a coin toss with a probability of heads of $p = 0.75$, we use (10.4) to obtain

$$\begin{aligned} P[a \leq X \leq b] &= \int_a^b p_X(x)dx \\ &= \int_a^b 1\, dx \\ &= b - a = 0.75 \end{aligned}$$

and choose $a = 0$ and $b = 0.75$. The probability of obtaining an outcome in the interval $(0, 0.75]$ for a random variable $X \sim \mathcal{U}(0, 1)$ is now seen to be 0.75. Hence, the code below can be used to generate the outcomes of a repeated coin tossing experiment with $p = 0.75$.

```
for i=1:M
u=rand(1,1);
```

```
      if u<=0.75
        x(i,1)=1; % head mapped into 1
      else
        x(i,1)=0; % tail mapped into 0
      end
    end
```

Could we have used any other values for a and b?

$\diamond$

Now returning to our dart thrower, we can acknowledge her superior dart-throwing ability by assigning a nonuniform PDF as shown in Figure 10.4. The probability of

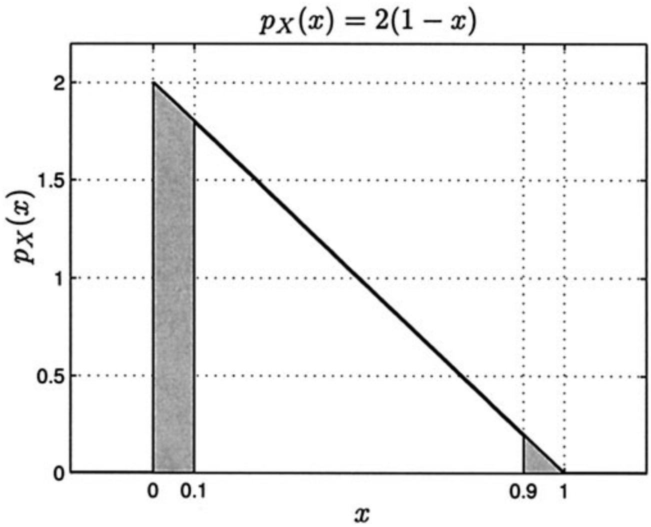


Figure 10.4: Nonuniform PDF.

throwing a dart within a circle of radius 0.1 or $X \in [0, 0.1]$ will be larger than for the region between the circles with radii 0.9 and 1 or $X \in [0.9, 1]$. Specifically, using (10.4)

$$P[0 \leq X \leq 0.1] = \int_0^{0.1} 2(1-x)dx = 2(x - x^2/2)\Big|_0^{0.1} = 0.19$$
$$P[0.9 \leq X \leq 1] = \int_{0.9}^{1} 2(1-x)dx = 2(x - x^2/2)\Big|_{0.9}^{1} = 0.01.$$

Note that in this example $p_X(x) \geq 0$ for all x and also $\int_{-\infty}^{\infty} p_X(x)dx = 1$. These are properties that must be satisfied for a valid PDF. We will say more about these properties in the next section.

It may be helpful to consider a mass analogy to the PDF. An example is shown in Figure 10.5. It can be thought of as a slice of Jarlsberg cheese with length 2

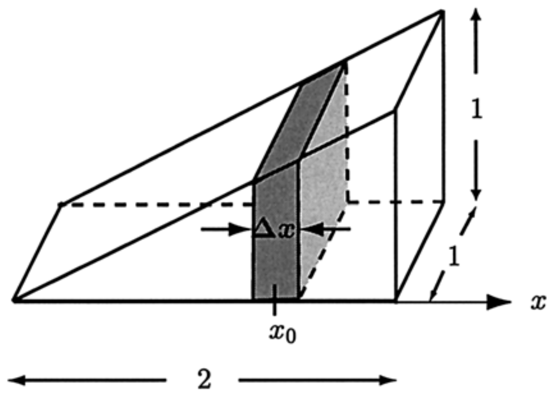


Figure 10.5: Jarlsberg cheese slice used for mass analogy to PDF.

meters, height of 1 meter, and depth of 1 meter, which might be purchased for a New Year's Eve party (with a lot of guests!). If its mass is 1 kilogram (it is a new "lite" cheese), then its overall density D is

$$D = \frac{\text{mass}}{\text{volume}} = \frac{M}{V} = \frac{1 \text{ kg}}{1 \text{ m}^3} = 1 \text{ kg/m}^3.$$

However, its *linear* density or mass per meter which is defined as $\Delta M/\Delta x$ will change with x. If each guest is allowed to cut a wedge of cheese of length Δx as shown in Figure 10.5, then clearly the hungriest guests should choose a wedge near $x = 2$ for the greatest amount of cheese. To determine the linear density we compute $\Delta M/\Delta x$ versus x. To do so first note that $\Delta M = D\Delta V = \Delta V$ and $\Delta V = 1 \cdot (\text{area of face})$, where the face is seen to be trapezoidal. Thus,

$$\Delta V = \frac{1}{2}\Delta x \left(\frac{x_0 - \Delta x/2}{2} + \frac{x_0 + \Delta x/2}{2} \right) = \frac{1}{2} x_0 \Delta x.$$

Hence, $\Delta M/\Delta x = \Delta V/\Delta x = x_0/2$ and this is the same even as $\Delta x \to 0$. Thus,

$$\frac{dM}{dx} = \frac{1}{2}x \qquad 0 \leq x \leq 2$$

and to obtain the mass for any wedge from $x = a$ to $x = b$ we need only integrate dM/dx to obtain the mass as a function of x. This yields

$$M([a,b]) = \int_a^b \frac{1}{2} x dx = \int_a^b m(x) dx$$

where $m(x) = x/2$ is the *linear mass density or the mass per unit length.* It is perfectly analogous to the PDF which is the *probability per unit length.* Can you find the total mass of cheese from $M([a,b])$? See also Problem 10.3.

10.4 The PDF and Its Properties

The PDF must have certain properties so that the probabilities obtained using (10.4) satisfy the axioms given in Section 3.4. Since the probability of an interval is given by

$$P[a \leq X \leq b] = \int_a^b p_X(x)dx$$

the PDF must have the following properties.

Property 10.1 – PDF must be nonnegative.

$$p_X(x) \geq 0 \qquad -\infty < x < \infty.$$

Proof: If $p_X(x) < 0$ on some small interval $[x_0 - \Delta x/2, x_0 + \Delta x/2]$, then

$$P[x_0 - \Delta x/2 \leq X \leq x_0 + \Delta x/2] = \int_{x_0-\Delta x/2}^{x_0+\Delta x/2} p_X(x)dx < 0$$

which violates Axiom 1 that $P[E] \geq 0$ for all events E.

□

Property 10.2 – PDF must integrate to one.

$$\int_{-\infty}^{\infty} p_X(x)dx = 1$$

Proof:

$$1 = P[X \in \mathcal{S}_X] = P[-\infty < X < \infty] = \int_{-\infty}^{\infty} p_X(x)dx$$

□

Hence, any nonnegative function that integrates to one can be considered as a PDF. An example follows.

Example 10.2 – Exponential PDF

Consider the function

$$p_X(x) = \begin{cases} \lambda \exp(-\lambda x) & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{10.5}$$

for $\lambda > 0$. This is called the *exponential PDF* and is shown in Figure 10.6. Note that it is discontinuous at $x = 0$. Hence, a PDF need not be continuous (see also Figure 10.3a for the uniform PDF which also has points of discontinuity). Also, for $\lambda > 1$, we have $p_X(0) = \lambda > 1$. In contrast to a PMF, the PDF can exceed one in

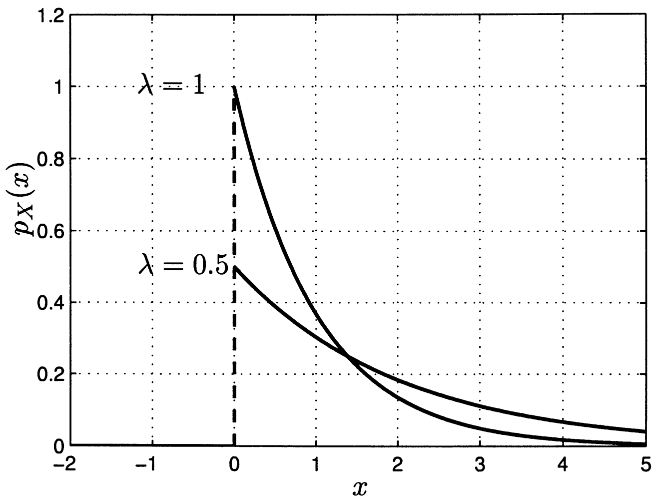


Figure 10.6: Exponential PDF.

value. It is the *area* under the PDF that cannot exceed one. As expected $p_X(x) \geq 0$ for $-\infty < x < \infty$ and

$$\begin{aligned}\int_{-\infty}^{\infty} p_X(x)dx &= \int_{0}^{\infty} \lambda \exp(-\lambda x)dx \\ &= -\exp(-\lambda x)|_0^\infty = 1\end{aligned}$$

for $\lambda > 0$. This PDF is often used as a model for the lifetime of a product. For example, if X is the failure time in days of a lightbulb, then $P[X > 100]$ is the probability that the lightbulb will fail after 100 days or it will last for at least 100 days. This is found to be

$$\begin{aligned}P[X > 100] &= \int_{100}^{\infty} \lambda \exp(-\lambda x)dx \\ &= -\exp(-\lambda x)|_{100}^\infty \\ &= \exp(-100\lambda) \\ &= \begin{cases} 0.367 & \lambda = 0.01 \\ 0.904 & \lambda = 0.001\,. \end{cases}\end{aligned}$$

$\diamond$

⚠ The probability of a sample point is zero.

If X is a continuous random variable, then it was argued in Section 3.6 that the probability of a point is zero. This is consistent with our definition of a PDF. If the width of the interval shrinks to zero, then the area under the PDF also goes to zero. Hence, $P[X = x] = 0$. This is true whether or not $p_X(x)$ is continuous at the point of interest (as long as the discontinuity is a finite jump). In the previous example of an exponential PDF $P[X = 0] = 0$ even though $p_X(0)$ is discontinuous at $x = 0$. This means that we could, if desired, have defined the exponential PDF as

$$p_X(x) = \begin{cases} \lambda \exp(-\lambda x) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

for which $p_X(0)$ is now defined to be 0. It makes no difference in our probability calculations whether we include $x = 0$ in the interval or not. Hence, we see that

$$\int_{0^-}^{b} p_X(x)dx = \int_{0^+}^{b} p_X(x)dx = \int_{0}^{b} p_X(x)dx$$

and in a similar manner if X is a continuous random variable, then

$$P[a \leq X \leq b] = P[a < X \leq b] = P[a \leq X < b] = P[a < X < b].$$

In summary, the value assigned to the PDF at a discontinuity is arbitrary since it does not affect any subsequent probability calculation involving a continuous random variable. However, for discontinuities other than step discontinuities (which are jumps of finite magnitude) we will see in Section 10.8 that we must be more careful.

10.5 Important PDFs

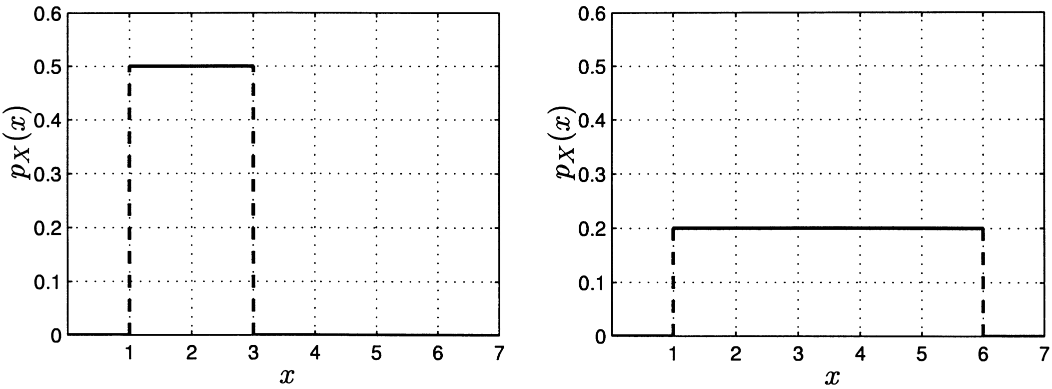
There are a multitude of PDFs in use in various scientific disciplines. The books by [Johnson, Kotz, and Balakrishnan 1994] contain a summary of many of these and should be consulted for further information. We now describe some of the more important PDFs.

10.5.1 Uniform

We have already encountered a special case of the uniform PDF in Figure 10.3. More generally it is defined as

$$p_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases} \tag{10.6}$$

and examples are shown in Figure 10.7. It is given the shorthand notation $X \sim \mathcal{U}(a, b)$. If $a = 0$ and $b = 1$, then an outcome of a $\mathcal{U}(0, 1)$ random variable can be generated in MATLAB using `rand`.



(a) $a = 1, b = 3$ (b) $a = 1, b = 6$

Figure 10.7: Examples of uniform PDF.

10.5.2 Exponential

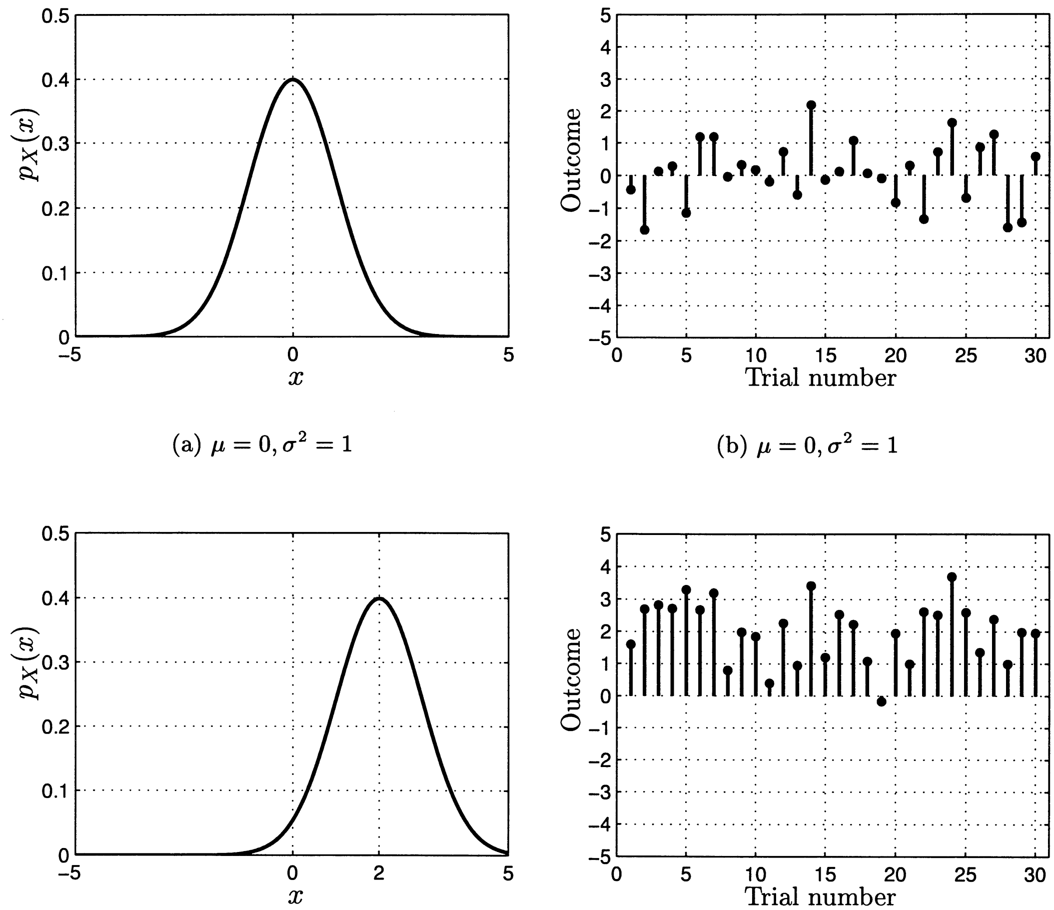
This was previously defined in Example 10.2. The shorthand notation is $X \sim \exp(\lambda)$.

10.5.3 Gaussian or Normal

This is the famous "bell-shaped" curve first introduced in Section 1.3. It is given by

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \qquad -\infty < x < \infty \tag{10.7}$$

where $\sigma^2 > 0$ and $-\infty < \mu < \infty$. Its application in practical problems is ubiquitous. It is shown to integrate to one in Problem 10.9. Some examples of this PDF as well as some outcomes for various values of the parameters (μ, σ^2) are shown in Figures 10.8 and 10.9. It is characterized by the two parameters μ and σ^2. The parameter μ indicates the center of the PDF which is seen in Figures 10.8a and 10.8c. It depicts the "average value" of the random variable as can be observed by examining Figures 10.8b and 10.8d. In Chapter 11 we will show that μ is actually the mean of X. The parameter σ^2 indicates the width of the PDF as is seen in Figures 10.9a and 10.9c. It is related to the variability of the outcomes as seen in Figures 10.9b and 10.9d. In Chapter 11 we will show that σ^2 is actually the variance of X. The PDF is called the *Gaussian* PDF after the famous German mathematician K.F. Gauss and also the *normal* PDF, since "normal" populations tend to exhibit this type of distribution. A *standard* normal PDF is one for which $\mu = 0$ and $\sigma^2 = 1$. The shorthand notation is $X \sim \mathcal{N}(\mu, \sigma^2)$. MATLAB generates a realization of a *standard normal* random variable using `randn`. This was used extensively in Chapter 2.



(a) $\mu = 0, \sigma^2 = 1$

(b) $\mu = 0, \sigma^2 = 1$

(c) $\mu = 2, \sigma^2 = 1$

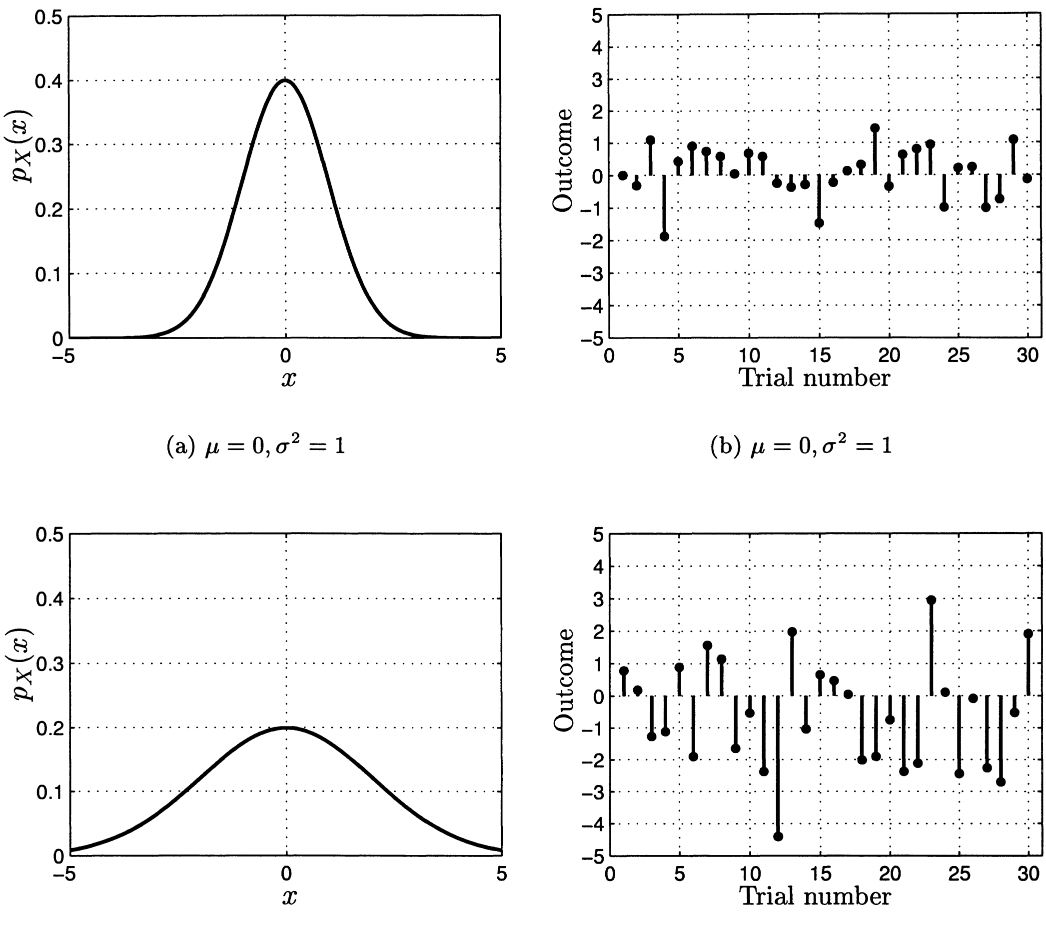
(d) $\mu = 2, \sigma^2 = 1$

Figure 10.8: Examples of Gaussian PDF with different μ's.

To find the probability of the outcome of a Gaussian random variable lying within an interval requires numerical integration (see Problem 1.14) since the integral

$$\int_a^b \frac{1}{\sqrt{2\pi}} \exp(-(1/2)x^2) dx$$

cannot be evaluated analytically. A MATLAB subprogram will be provided and described shortly to do this. The Gaussian PDF is commonly used to model noise in a communication system (see Section 2.6), as well as for numerous other applications. We will see in Chapter 15 that the PDF arises quite naturally as the PDF of a large number of independent random variables that have been added together.



(a) $\mu = 0, \sigma^2 = 1$

(b) $\mu = 0, \sigma^2 = 1$

(c) $\mu = 0, \sigma^2 = 2$

(d) $\mu = 0, \sigma^2 = 2$

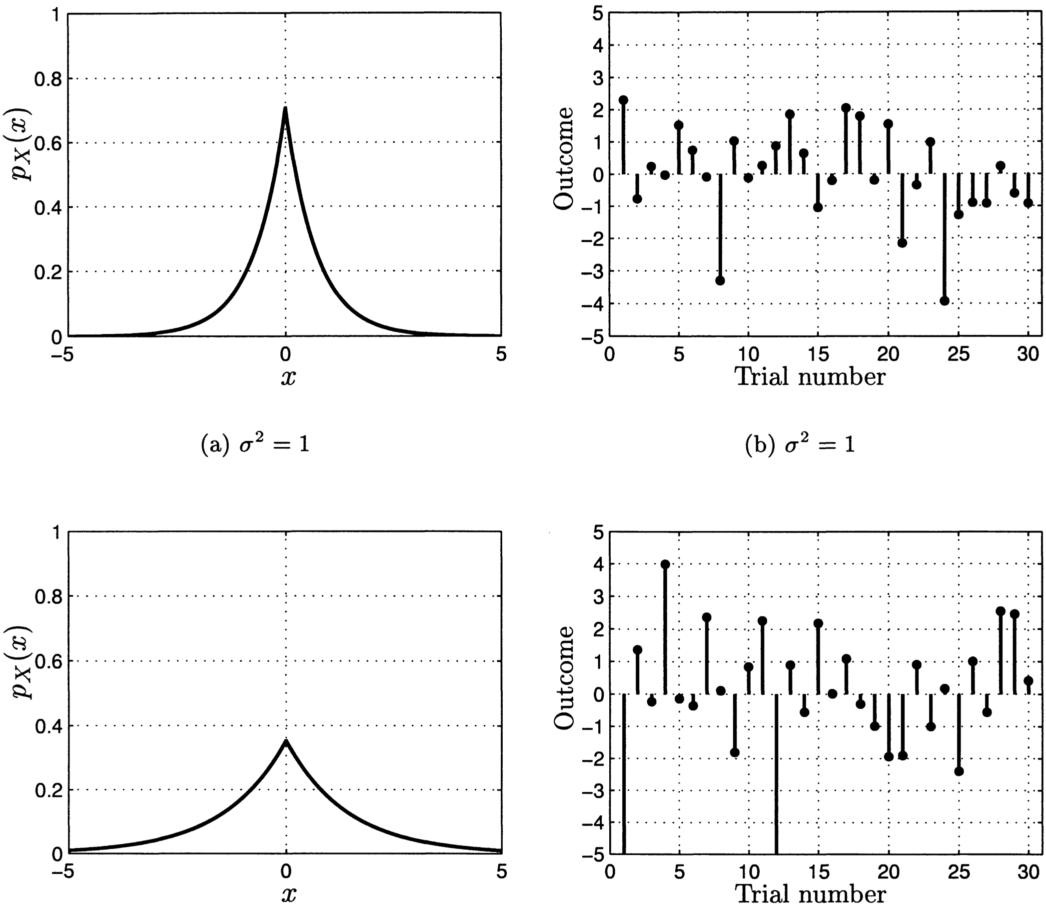
Figure 10.9: Examples of Gaussian PDF with different σ^2's.

10.5.4 Laplacian

This PDF is named after Laplace, the famous French mathematician. It is similar to the Gaussian except that it does not decrease as rapidly from its maximum value. Its PDF is

$$p_X(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left(-\sqrt{\frac{2}{\sigma^2}}|x|\right) \qquad -\infty < x < \infty \tag{10.8}$$

where $\sigma^2 > 0$. Again the parameter σ^2 specifies the width of the PDF, and will be shown in Chapter 11 to be the variance of X. It is seen to be symmetric about $x = 0$. Some examples of the PDF and outcomes are shown in Figure 10.10. Note that for



(a) $\sigma^2 = 1$ (b) $\sigma^2 = 1$

(c) $\sigma^2 = 4$ (d) $\sigma^2 = 4$

Figure 10.10: Examples of Laplacian PDF with different σ^2's.

the same σ^2 as the Gaussian PDF, the outcomes are larger as seen by comparing Figure 10.10b to Figure 10.9b. This is due to the larger probability in the "tails" of the PDF. The "tail" region of the PDF is that for which $|x|$ is large. The Laplacian PDF is easily integrated to find the probability of an interval. This PDF is used as a model for speech amplitudes [Rabiner and Schafer 1978].

10.5.5 Cauchy

The Cauchy PDF is named after another famous French mathematician and is defined as

$$p_X(x) = \frac{1}{\pi(1+x^2)} \qquad -\infty < x < \infty. \tag{10.9}$$

It is shown in Figure 10.11 and is seen to be symmetric about $x = 0$. The Cauchy PDF can easily be integrated to find the probability of any interval. It arises as the PDF of the ratio of two independent $\mathcal{N}(0,1)$ random variables (see Chapter 12).

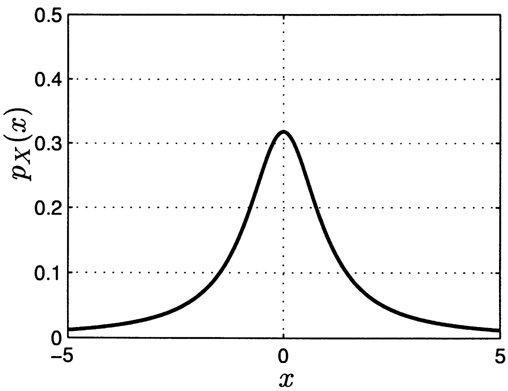


Figure 10.11: Cauchy PDF.

10.5.6 Gamma

The Gamma PDF is a very general PDF that is used for nonnegative random variables. It is given by

$$p_X(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x) & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{10.10}$$

where $\lambda > 0$, $\alpha > 0$, and $\Gamma(z)$ is the *Gamma* function which is defined as

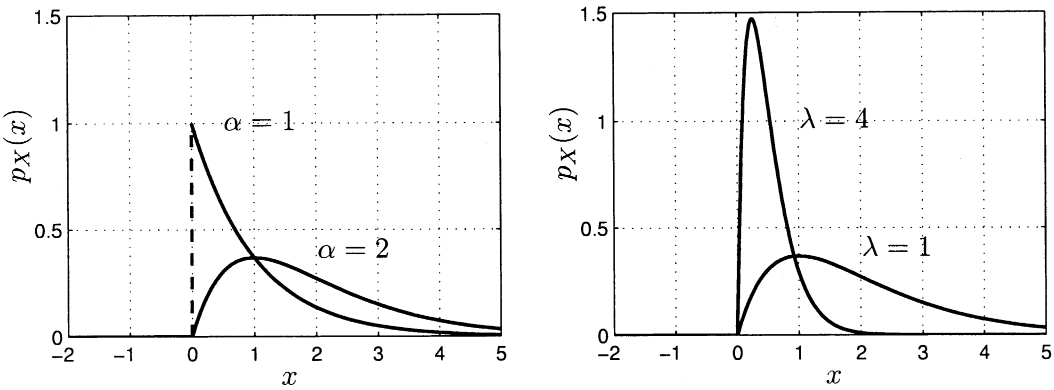
$$\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t) dt. \tag{10.11}$$

Clearly, the $\Gamma(\alpha)$ factor in (10.10) is the normalizing constant needed to ensure that the PDF integrates to one. Some examples of this PDF are shown in Figure 10.12. The shorthand notation is $X \sim \Gamma(\alpha, \lambda)$. Some useful properties of the Gamma function are as follows.

Property 10.3 – $\Gamma(z+1) = z\Gamma(z)$

Proof: See Problem 10.16.

□



(a) $\lambda = 1$ (b) $\alpha = 2$

Figure 10.12: Examples of Gamma PDF.

Property 10.4 – $\Gamma(N) = (N-1)!$

Proof: Follows from Property 10.3 with $z = N-1$ since

$$\begin{aligned}\Gamma(N) &= (N-1)\Gamma(N-1) \\ &= (N-1)(N-2)\Gamma(N-3) \qquad \text{(let } z = N-2 \text{ now)} \\ &= (N-1)(N-2)\ldots 1 = (N-1)!\end{aligned}$$

□

Property 10.5 – $\Gamma(1/2) = \sqrt{\pi}$

Proof:

$$\Gamma(1/2) = \int_0^\infty t^{-1/2} \exp(-t) dt$$

(Note that near $t = 0$ the integrand becomes infinite but $t^{-1/2}\exp(-t) \approx t^{-1/2}$ which is integrable.) Now let $t = u^2/2$ and thus $dt = u du$ which yields

$$\begin{aligned}\Gamma(1/2) &= \int_0^\infty \frac{1}{\sqrt{u^2/2}} \exp(-u^2/2) u du \\ &= \int_0^\infty \sqrt{2} \exp(-u^2/2) du \\ &= \frac{\sqrt{2}}{2} \underbrace{\int_{-\infty}^\infty \exp(-u^2/2) du}_{=\sqrt{2\pi} \text{ why?}} \qquad \text{(integrand is symmetric about } u = 0\text{)} \\ &= \sqrt{\pi}.\end{aligned}$$

□

The Gamma PDF reduces to many well known PDFs for appropriate choices of the parameters α and λ. Some of these are:

1. Exponential for $\alpha = 1$
   From (10.10) we have

$$p_X(x) = \begin{cases} \frac{\lambda}{\Gamma(1)} \exp(-\lambda x) & x \geq 0 \\ 0 & x < 0. \end{cases}$$

   But $\Gamma(1) = 0! = 1$, which results from Property 10.4 so that we have the exponential PDF.

2. Chi-squared PDF with N degrees of freedom for $\alpha = N/2$ and $\lambda = 1/2$
   From (10.10) we have

$$p_X(x) = \begin{cases} \frac{1}{2^{N/2}\Gamma(N/2)} x^{N/2-1} \exp(-x/2) & x \geq 0 \\ 0 & x < 0. \end{cases} \tag{10.12}$$

   This is called the *chi-squared PDF with N degrees of freedom* and is important in statistics. It can be shown to be the PDF for the sum of the squares of N independent random variables all with the same PDF $\mathcal{N}(0, 1)$ (see Problem 12.44). The shorthand notation is $X \sim \chi^2_N$.

3. Erlang for $\alpha = N$
   From (10.10) we have

$$p_X(x) = \begin{cases} \frac{\lambda^N}{\Gamma(N)} x^{N-1} \exp(-\lambda x) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

   and since $\Gamma(N) = (N-1)!$ from Property 10.4, this becomes

$$p_X(x) = \begin{cases} \frac{\lambda^N}{(N-1)!} x^{N-1} \exp(-\lambda x) & x \geq 0 \\ 0 & x < 0. \end{cases} \tag{10.13}$$

   This PDF arises as the PDF of a sum of N independent exponential random variables all with the same λ (see also Problem 10.17).

10.5.7 Rayleigh

The Rayleigh PDF is named after the famous British physicist Lord Rayleigh and is defined as

$$p_X(x) = \begin{cases} \frac{x}{\sigma^2} \exp\left(-\frac{1}{2}\frac{x^2}{\sigma^2}\right) & x \geq 0 \\ 0 & x < 0. \end{cases} \tag{10.14}$$

It is shown in Figure 10.13. The Rayleigh PDF is easily integrated to yield the

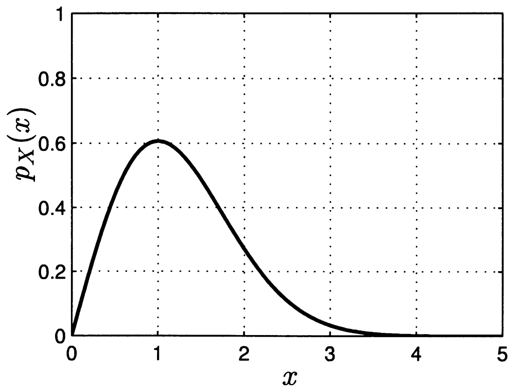


Figure 10.13: Rayleigh PDF with $\sigma^2 = 1$.

probability of any interval. It can be shown to arise as the PDF of the square root of the sum of the squares of two independent $\mathcal{N}(0,1)$ random variables (see Example 12.12).

Finally, note that many of these PDFs arise as the PDFs of transformed Gaussian random variables. Therefore, realizations of the random variable may be obtained by first generating multiple realizations of independent standard normal or $\mathcal{N}(0,1)$ random variables, and then performing the appropriate transformation. An alternative and more general approach to generating realizations of a random variable, once the PDF is known, is via the probability integral transformation to be discussed in Section 10.9.

10.6 Cumulative Distribution Functions

The cumulative distribution function (CDF) for a continuous random variable is defined exactly the same as for a discrete random variable. It is

$$F_X(x) = P[X \leq x] \qquad -\infty < x < \infty \tag{10.15}$$

and is evaluated using the PDF as

$$F_X(x) = \int_{-\infty}^{x} p_X(t)dt \qquad -\infty < x < \infty. \tag{10.16}$$

Avoiding confusion in evaluating CDFs

It is important to note that in evaluating a definite integral such as in (10.16) it is best to replace the variable of integration with another symbol. This is because the upper limit depends on x which would conflict with the dummy variable of integration. We have chosen to use t but of course any other symbol that does not conflict with x can be used.

Some examples of the evaluation of the CDF are given next.

10.6.1 Uniform

Using (10.6) we have

$$F_X(x) = \begin{cases} 0 & x \le a \\ \int_a^x \frac{1}{b-a} dt & a < x < b \\ 1 & x \ge b \end{cases}$$

which is

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

An example is shown in Figure 10.14 for $a = 1$ and $b = 2$.

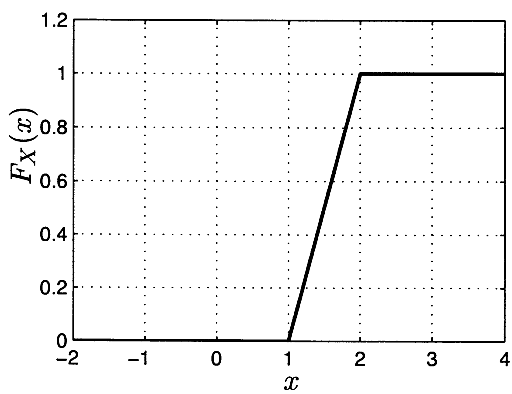


Figure 10.14: CDF for uniform random variable over interval $(1, 2)$.

10.6.2 Exponential

Using (10.5) we have

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \int_0^x \lambda \exp(-\lambda t) dt & x \ge 0\,. \end{cases}$$

But

$$\int_0^x \lambda \exp(-\lambda t) dt = -\exp(-\lambda t)|_0^x = 1 - \exp(-\lambda x)$$

so that

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - \exp(-\lambda x) & x \geq 0\,. \end{cases}$$

An example is shown in Figure 10.15 for $\lambda = 1$.

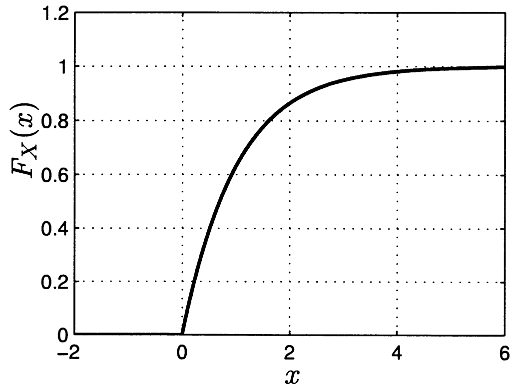


Figure 10.15: CDF for exponential random variable with $\lambda = 1$.

Note that for the uniform and exponential random variables the CDFs are continuous even though the PDFs are discontinuous. This property motivates an alternative definition of a continuous random variable as one for which *the CDF is continuous*. Recall that the CDF of a discrete random variable is always discontinuous, displaying multiple jumps.

10.6.3 Gaussian

Consider a standard normal PDF, which is a Gaussian PDF with $\mu = 0$ and $\sigma^2 = 1$. (If $\mu \neq 0$ and/or $\sigma^2 \neq 1$ the CDF is a simple modification as shown in Problem 10.22.) Then from (10.7) we have

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt \qquad -\infty < x < \infty.$$

This cannot be evaluated further but can be found numerically and is shown in Figure 10.16. The CDF for a standard normal is usually given the special symbol $\Phi(x)$ so that

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt \qquad -\infty < x < \infty.$$

Hence, $\Phi(x)$ represents the area under the PDF to the *left* of the point x as seen in Figure 10.17a. It is sometimes more convenient, however, to have knowledge of the area to the *right* instead. This is called the *right-tail probability* of a standard

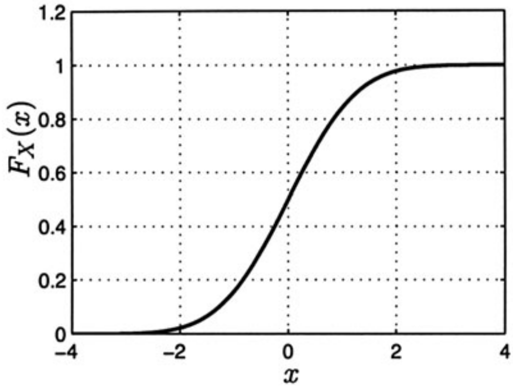
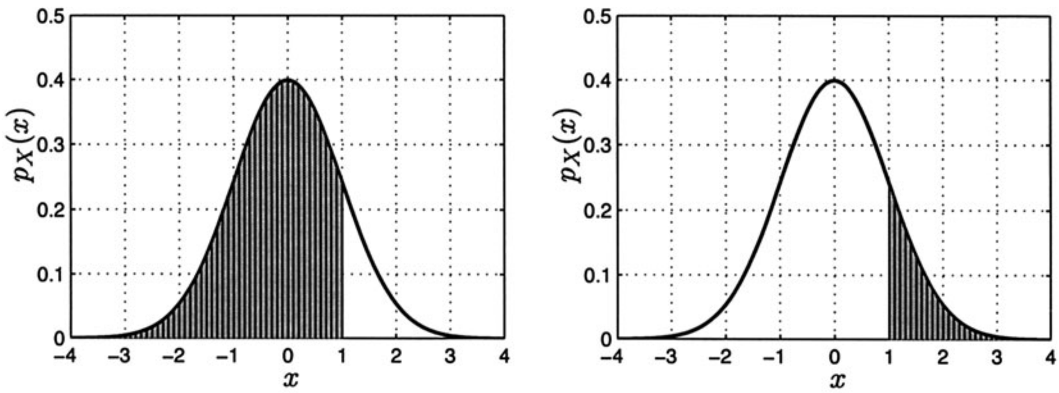


Figure 10.16: CDF for standard normal or Gaussian random variable.



(a) Shaded area = $\Phi(1)$

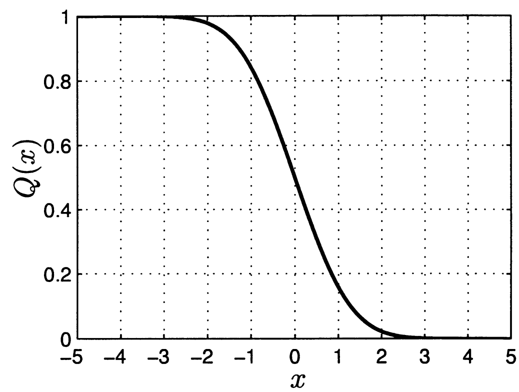
(b) Shaded area = $Q(1)$

Figure 10.17: Definitions of $\Phi(x)$ and $Q(x)$ functions.

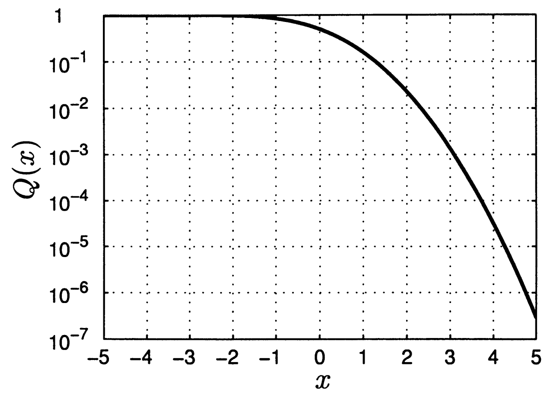
normal and is given the symbol $Q(x)$. It is termed the "Q" function and is defined as the area to the right of x, an example of which is shown in Figure 10.17b. By its definition we have

$$
\begin{aligned}
Q(x) &= 1 - \Phi(x) && (10.17)\\
&= \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt \quad -\infty < x < \infty && (10.18)
\end{aligned}
$$

and is shown in Figure 10.18, plotted on a linear as well as a logarithmic vertical scale. Some of the properties of the Q function that are easily verified are (see



(a) Linear vertical scale



(b) Logarithmic vertical scale – for display of small values of $Q(x)$

Figure 10.18: $Q(x)$ function.

Problem 10.25)

$$\begin{aligned} Q(-\infty) &= 1 & (10.19)\\ Q(\infty) &= 0 & (10.20)\\ Q(0) &= \frac{1}{2} & (10.21)\\ Q(-x) &= 1 - Q(x). & (10.22) \end{aligned}$$

Although the Q function cannot be evaluated analytically, it is related to the well known "error function". Thus, making use of the latter function a MATLAB subprogram `Q.m`, which is listed in Appendix 10B, can be used to evaluate it. An example follows.

Example 10.3 – Probability of error in a communication system

In Section 2.6 we analyzed the probability of error for a PSK digital communication system. The probability of error P_e was given by

$$P_e = P[A/2 + W \leq 0]$$

where $W \sim \mathcal{N}(0,1)$. (In the MATLAB code we used `w=randn(1,1)` and hence the random variable representing the noise was a standard normal random variable.)

To explicitly evaluate P_e we have that

$$\begin{aligned} P_e &= P[A/2 + W \leq 0] \\ &= 1 - P[A/2 + W > 0] \\ &= 1 - P[W > -A/2] \\ &= 1 - Q(-A/2) \qquad \text{(definition)} \\ &= Q(A/2) \qquad \text{(use (10.22))}. \end{aligned}$$

Hence, the true P_e shown in Figure 2.15 as the dashed line can be found by using the MATLAB subprogram `Q.m`, which is listed in Appendix 10B, for the argument $A/2$ (see Problem 10.26). It is also sometimes important to determine A to yield a given P_e. This is found as $A = 2Q^{-1}(P_e)$, where Q^{-1} is the inverse of the Q function. It is defined as the value of x necessary to yield a given value of $Q(x)$. It too cannot be expressed analytically but may be evaluated using the MATLAB subprogram `Qinv.m`, also listed in Appendix 10B.

◊

The Q function can also be approximated for large values of x using [Abramowitz and Stegun 1965]

$$Q(x) \approx \frac{1}{\sqrt{2\pi}x} \exp\left(-\frac{1}{2}x^2\right) \qquad x > 3. \tag{10.23}$$

A comparison of the approximation to the true value is shown in Figure 10.19. If

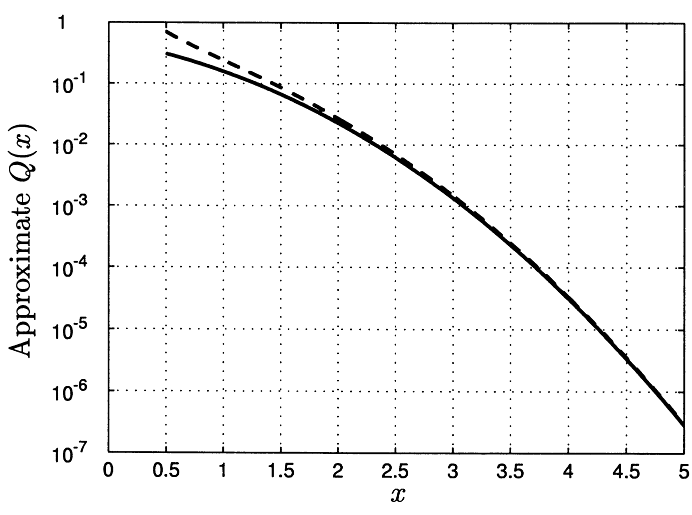


Figure 10.19: Approximation of Q function – true value is shown dashed.

$X \sim \mathcal{N}(\mu, \sigma^2)$, then the right-tail probability becomes

$$P[X > x] = Q\left(\frac{x - \mu}{\sqrt{\sigma^2}}\right) \tag{10.24}$$

(see Problem 10.24). Finally, note that the area under the standard normal Gaussian PDF is mostly contained in the interval $[-3,3]$. As seen in Figure 10.19 $Q(3) \approx 0.001$, which means that the area to the right of $x = 3$ is only 0.001. Since the PDF is symmetric, the total area to the right of $x = 3$ and to the left of $x = -3$ is 0.002 or the area in the $[-3,3]$ interval is 0.998. Hence, 99.8% of the probability lies within this interval. We would not expect to see a value greater than 3 in magnitude very often. This is borne out by an examination of Figure 10.8b. How many realizations would you expect to see in the interval $(1,\infty)$? Is this consistent with Figure 10.8b ?

As we have seen, the CDF for a continuous random variable has certain properties. For the most part they are the same as for a discrete random variable: the CDF is 0 at $x = -\infty$, 1 at $x = \infty$, and is monotonically increasing (or stays the same) between these limits. However, now it is continuous, having no jumps. The most important property for practical purposes is that which allows us to compute probabilities of intervals. This follows from the property

$$P[a \leq X \leq b] = P[a < X \leq b] = F_X(b) - F_X(a) \tag{10.25}$$

which is easily proven (see Problem 10.35). It can be seen to be valid by referring to Figure 10.20. Using the CDF we *no longer have to integrate the PDF* to determine

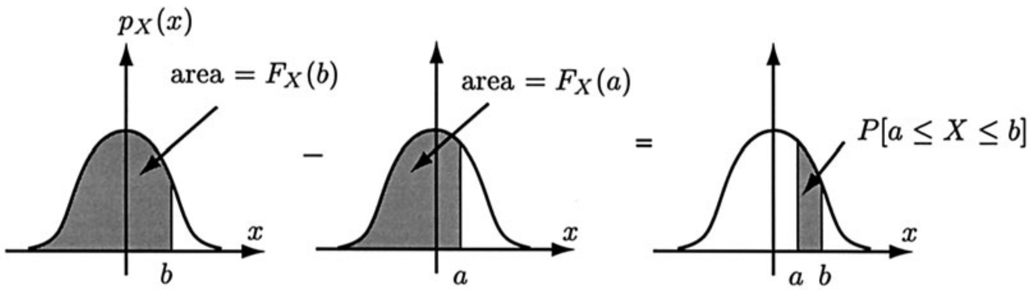


Figure 10.20: Illustration of use of CDF to find probability of interval.

probabilities of intervals. In effect, all the integration has been done for us in finding the CDF. Some examples follow.

Example 10.4 – Probability of interval for exponential PDF

Since $F_X(x) = 1 - \exp(-\lambda x)$ for $x \geq 0$, we have for $a > 0$ and $b > 0$

$$\begin{aligned} P[a \leq X \leq b] &= F_X(b) - F_X(a) \\ &= (1 - \exp(-\lambda b)) - (1 - \exp(-\lambda a)) \\ &= \exp(-\lambda a) - \exp(-\lambda b) \end{aligned}$$

which should be compared to

$$\int_a^b \lambda \exp(-\lambda x) dx.$$

◇

Since we obtained the CDF from the PDF, we might suppose that the PDF could be recovered from the CDF. For a discrete random variable this was the case since $p_X[x_i] = F_X(x_i^+) - F_X(x_i^-)$. For a continuous random variable we consider a small interval $[x_0 - \Delta x/2, x_0 + \Delta x/2]$ and evaluate its probability using (10.25) with

$$F_X(x) = \int_{-\infty}^{x} p_X(t)dt.$$

Then, we have

$$\begin{aligned}
& F_X(x_0 + \Delta x/2) - F_X(x_0 - \Delta x/2) \\
&= \int_{-\infty}^{x_0+\Delta x/2} p_X(t)dt - \int_{-\infty}^{x_0-\Delta x/2} p_X(t)dt \\
&= \int_{x_0-\Delta x/2}^{x_0+\Delta x/2} p_X(t)dt \\
&\approx p_X(x_0) \int_{x_0-\Delta x/2}^{x_0+\Delta x/2} 1\,dt \quad (p_X(t) \approx \text{constant as } \Delta x \to 0) \\
&= p_X(x_0)\Delta x
\end{aligned}$$

so that

$$\begin{aligned}
p_X(x_0) &\approx \frac{F_X(x_0 + \Delta x/2) - F_X(x_0 - \Delta x/2)}{\Delta x} \\
&\to \left.\frac{dF_X(x)}{dx}\right|_{x=x_0} \quad \text{as } \Delta x \to 0.
\end{aligned}$$

Hence, we can obtain the PDF from the CDF by differentiation or

$$p_X(x) = \frac{dF_X(x)}{dx}. \tag{10.26}$$

This relationship is really just the *fundamental theorem of calculus* [Widder 1989]. Note the similarity to the discrete case in which $p_X[x_i] = F_X(x_i^+) - F_X(x_i^-)$. As an example, if $X \sim \exp(\lambda)$, then

$$F_X(x) = \begin{cases} 1 - \exp(-\lambda x) & x \geq 0 \\ 0 & x < 0. \end{cases}$$

For all x except $x = 0$ (at which the CDF does not have a derivative due to the change in slope as seen in Figure 10.15) we have

$$\begin{aligned}
p_X(x) = \frac{dF_X(x)}{dx} &= 0 & x < 0 \\
&= \lambda \exp(-\lambda x) & x > 0
\end{aligned}$$

and as remarked earlier, $p_X(0)$ can be assigned any value.

10.7 Transformations

In discussing transformations for discrete random variables we noted that a transformation can be either one-to-one or many-to-one. For example, the function $g(x) = 2x$ is one-to-one while $g(x) = x^2$ is many-to-one (in this case two-to-one since $-x$ and $+x$ both map into x^2). The determination of the PDF of $Y = g(X)$ will depend upon which type of transformation we have. Initially, we will consider the one-to-one case, which is simpler. For the transformation of a discrete random variable we saw from (5.9) that the PMF of $Y = g(X)$ for any g could be found from the PMF of X using

$$p_Y[y_i] = \sum_{\{j:g(x_j)=y_i\}} p_X[x_j].$$

But if g is one-to-one we have only a single solution for $g(x_j) = y_i$, so that $x_j = g^{-1}(y_i)$ and therefore

$$p_Y[y_i] = p_X[g^{-1}(y_i)] \tag{10.27}$$

and we are done. For example, assume X takes on values $\{1, 2\}$ with a PMF $p_X[1]$ and $p_X[2]$ and we wish to determine the PMF of $Y = g(X) = 2X$, which is shown in Figure 10.21. Then from (10.27)

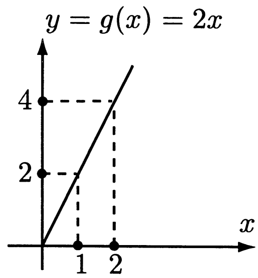


Figure 10.21: Transformation of a discrete random variable.

$$\begin{aligned} p_Y[2] &= p_X[g^{-1}(2)] = p_X[1] \\ p_Y[4] &= p_X[g^{-1}(4)] = p_X[2]. \end{aligned}$$

Because we are now dealing with a PDF, which is a density function, and not a PMF, which is a probability function, the simple relationship of (10.27) is no longer valid. To see what happens instead, consider the problem of determining the PDF of $Y = 2X$, where $X \sim \mathcal{U}(1, 2)$. Clearly, $\mathcal{S}_X = \{x : 1 < x < 2\}$ and therefore $\mathcal{S}_Y = \{y : 2 < y < 4\}$ so that $p_Y(y)$ must be zero outside the interval $(2, 4)$. The results of a MATLAB computer simulation are shown in Figure 10.22. A total of 50 realizations were obtained for X and Y. The generated X outcomes are shown on the x-axis and the resultant Y outcomes obtained from $y = 2x$ are shown on the

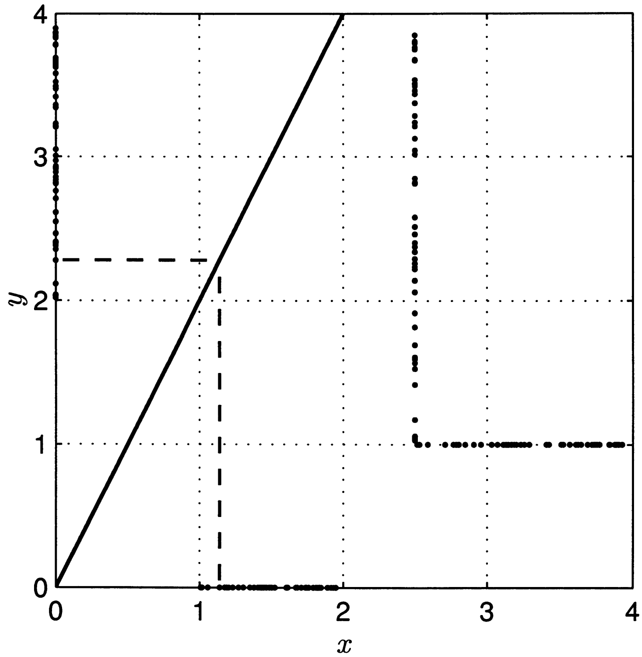


Figure 10.22: Computer generated realizations of X and $Y = 2X$ for $X \sim \mathcal{U}(1,2)$. A 50% expanded version of the realizations is shown to the right.

y-axis. Also, a 50% expanded version of the points is shown to the right. It is seen that the *density* of points on the y-axis is less than that on the x-axis. After some thought the reader will realize that this is the result of the scaling by a factor of 2 due to the transformation. Since the PDF is *probability per unit length*, we should expect $p_Y = p_X/2$ for $2 < y < 4$. To prove that this is so, we note that a small interval on the x-axis, say $[x_0 - \Delta x/2, x_0 + \Delta x/2]$, will map into $[2x_0 - \Delta x, 2x_0 + \Delta x]$ on the y-axis. However, the intervals are equivalent events and so their probabilities must be equal. It follows then that

$$\int_{x_0-\Delta x/2}^{x_0+\Delta x/2} p_X(x)dx = \int_{2x_0-\Delta x}^{2x_0+\Delta x} p_Y(y)dy$$

and as $\Delta x \to 0$, we have that $p_X(x) \to p_X(x_0)$ and $p_Y(y) \to p_Y(2x_0)$ in the small intervals so that

$$p_X(x_0)\Delta x = p_Y(2x_0)2\Delta x$$

or

$$p_Y(2x_0) = p_X(x_0)\frac{1}{2}.$$

As expected, the PDF of Y is scaled by 1/2. If we now let $y_0 = 2x_0$, then this becomes

$$p_Y(y_0) = p_X(y_0/2)\frac{1}{2}$$

or for any arbitrary value of y

$$p_Y(y) = p_X(y/2)\frac{1}{2} \qquad 2 < y < 4. \tag{10.28}$$

This results in the final PDF using $p_X(x) = 1$ for $1 < x < 2$ as

$$p_Y(y) = \begin{cases} \frac{1}{2} & 2 < y < 4 \\ 0 & \text{otherwise} \end{cases} \tag{10.29}$$

and thus if $X \sim \mathcal{U}(1,2)$, then $Y = 2X \sim \mathcal{U}(2,4)$. The general result for the PDF of $Y = g(X)$ is given by

$$p_Y(y) = p_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|. \tag{10.30}$$

For our example, the use of (10.30) with $g(x) = 2x$ and therefore $g^{-1}(y) = y/2$ results in (10.29). The absolute value is needed to allow for the case when g is decreasing and hence g^{-1} is decreasing since otherwise the scaling term would be negative (see Problem 10.57). A formal derivation is given in Appendix 10A. Note the similarity of (10.30) to (10.27). The principal difference is the presence of the derivative or *Jacobian* factor $dg^{-1}(y)/dy$. It is needed to account for the change in scaling due to the mapping of a given length interval into an interval of a different length as illustrated in Figure 10.22. Some examples of the use of (10.30) follow.

Example 10.5 – PDF for linear (actually an affine) transformation

To determine the PDF of $Y = aX + b$, for a and b constants first assume that $\mathcal{S}_X = \{x : -\infty < x < \infty\}$ and hence $\mathcal{S}_Y = \{y : -\infty < y < \infty\}$. Here we have $g(x) = ax + b$ so that the inverse function g^{-1} is found by solving $y = ax + b$ for x. This yields $x = (y - b)/a$ so that

$$g^{-1}(y) = \frac{y-b}{a}$$

and from (10.30) the general result is

$$p_Y(y) = p_X\left(\frac{y-b}{a}\right)\left|\frac{1}{a}\right|. \tag{10.31}$$

As a further example, consider $X \sim \mathcal{N}(0,1)$ and the transformation $Y = \sqrt{\sigma^2}X + \mu$.

Then, letting $\sigma = \sqrt{\sigma^2} > 0$ we have

$$\begin{aligned} p_Y(y) &= p_X\left(\frac{y-\mu}{\sigma}\right)\left|\frac{1}{\sigma}\right| \\ &= p_X\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} \\ &= \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]\frac{1}{\sigma} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] \end{aligned}$$

and therefore $Y \sim \mathcal{N}(\mu, \sigma^2)$. *A linear transformation of a Gaussian random variable results in another Gaussian random variable whose Gaussian PDF has different values of the parameters.* Because of this property we can easily generate a realization of a $\mathcal{N}(\mu, \sigma^2)$ random variable using the MATLAB construction `y=sqrt(sigma2)*randn(1,1)+mu`, since `randn(1,1)` produces a realization of a standard normal random variable (see Problem 10.60).

◇

Example 10.6 – PDF of $Y = \exp(X)$ for $X \sim \mathcal{N}(0,1)$

Here we have that $\mathcal{S}_Y = \{y : y > 0\}$. To find $g^{-1}(y)$ we let $y = \exp(x)$ and solve for x, which is $x = \ln(y)$. Thus, $g^{-1}(y) = \ln(y)$. From (10.30) it follows that

$$p_Y(y) = p_X(\ln(y))\left|\frac{d\ln(y)}{dy}\right| = \begin{cases} p_X(\ln(y))\frac{1}{y} & y > 0 \\ 0 & y \le 0 \end{cases}$$

or

$$p_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}y}\exp\left[-\frac{1}{2}(\ln(y))^2\right] & y > 0 \\ 0 & y \le 0 . \end{cases}$$

This PDF is called the *log-normal* PDF. It is frequently used as a model for a quantity that is measured in decibels (dB) and which has a normal PDF in dB quantities [Members of Technical Staff 1970].

◇

⚠ **Always determine the possible values for Y before using (10.30).**

A common error in determining the PDF of a transformed random variable is to forget that $p_Y(y)$ may be zero over some regions. In the previous example of $y = \exp(x)$, the mapping of $-\infty < x < \infty$ is into $y > 0$. Hence, the PDF of Y must be zero for $y \le 0$ since there are no values of X that produce a zero or negative value of Y. Nonsensical results occur if we attempt to insert values in $p_Y(y)$ for $y \leq 0$. To avoid this potential problem, we should first determine $\mathcal{S}_Y$ and then use (10.30) to find the PDF over the sample space.

When the transformation is not one-to-one, we will have multiple solutions for x in $y = g(x)$. An example is for $y = x^2$ for which the solutions are

$$\begin{aligned} x_1 &= -\sqrt{y} = g_1^{-1}(y) \\ x_2 &= +\sqrt{y} = g_2^{-1}(y). \end{aligned}$$

This is shown in Figure 10.23. In this case we use (10.30) but must add the PDFs

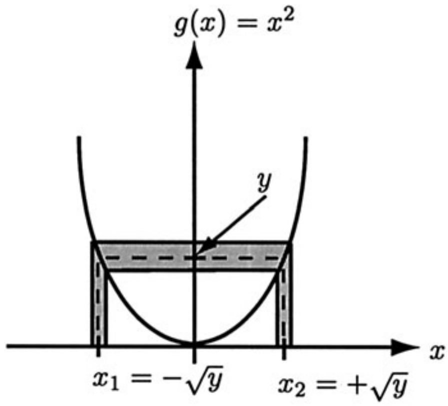


Figure 10.23: Solutions for x in $y = g(x) = x^2$.

(since both the x-intervals map into the same y-interval and the x-intervals are disjoint) to yield

$$p_Y(y) = p_X(g_1^{-1}(y)) \left| \frac{dg_1^{-1}(y)}{dy} \right| + p_X(g_2^{-1}(y)) \left| \frac{dg_2^{-1}(y)}{dy} \right|. \tag{10.32}$$

Example 10.7 – PDF of $Y = X^2$ for $X \sim \mathcal{N}(0, 1)$

Since $-\infty < X < \infty$, we must have $Y \geq 0$. Next because $g_1^{-1}(y) = -\sqrt{y}$ and $g_2^{-1}(y) = \sqrt{y}$ we have from (10.32)

$$p_Y(y) = \begin{cases} p_X(-\sqrt{y}) \left| -\frac{1}{2\sqrt{y}} \right| + p_X(\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right| & y \geq 0 \\ 0 & y < 0 \end{cases}$$

which reduces to

$$p_Y(y) = \begin{cases} \left[\frac{1}{\sqrt{2\pi}}\exp(-y/2)\right]\frac{1}{2\sqrt{y}} + \left[\frac{1}{\sqrt{2\pi}}\exp(-y/2)\right]\frac{1}{2\sqrt{y}} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

$$= \begin{cases} \frac{1}{\sqrt{2\pi y}}\exp(-y/2) & y \geq 0 \\ 0 & y < 0. \end{cases}$$

This is shown in Figure 10.24 and should be compared to Figure 2.10 in which this PDF was estimated (see also Problem 10.59). Note that the PDF is undefined at

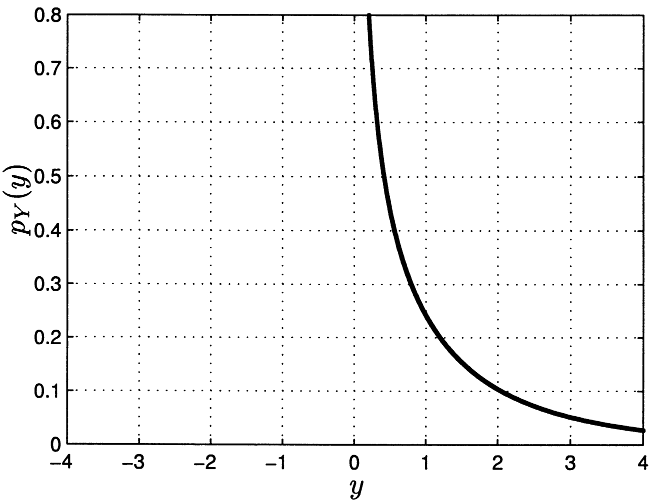


Figure 10.24: PDF for $Y = X^2$ for $X \sim \mathcal{N}(0,1)$.

$y = 0$ since $p_Y(0) \to \infty$, although the total area under the PDF is finite and of course is equal to 1. Also, $Y \sim \chi_1^2$ as can be seen by referring to (10.12) with $N = 1$.

◇

In general, if $y = g(x)$ has solutions $x_i = g_i^{-1}(y)$ for $i = 1, 2, \ldots, M$, then

$$p_Y(y) = \sum_{i=1}^{M} p_X(g_i^{-1}(y)) \left| \frac{dg_i^{-1}(y)}{dy} \right|. \tag{10.33}$$

An alternative means of finding the PDF of a transformed random variable is to first find the CDF and then differentiate it (see (10.26)). We illustrate this approach by redoing the previous example.

Example 10.8 – CDF approach to determine PDF of $Y = X^2$ for $X \sim \mathcal{N}(0,1)$

First we determine the CDF of Y in terms of the CDF for X as

$$\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= P[X^2 \leq y] \\ &= P[-\sqrt{y} \leq X \leq \sqrt{y}] \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}). \qquad \text{(from (10.25))} \end{aligned}$$

Then, differentiating we have

$$\begin{aligned} p_Y(y) &= \frac{dF_Y(y)}{dy} \\ &= \frac{d}{dy}\left[F_X(\sqrt{y}) - F_X(-\sqrt{y})\right] \\ &= p_X(\sqrt{y})\frac{d\sqrt{y}}{dy} - p_X(-\sqrt{y})\frac{d(-\sqrt{y})}{dy} \quad \text{(from (10.25) and chain rule of calculus)} \\ &= p_X(\sqrt{y})\frac{1}{2\sqrt{y}} + p_X(-\sqrt{y})\frac{1}{2\sqrt{y}} \\ &= \begin{cases} p_X(\sqrt{y})\frac{1}{\sqrt{y}} & y \geq 0 \\ 0 & y < 0 \end{cases} \quad \text{(since } p_X(-x) = p_X(x) \text{ for } X \sim \mathcal{N}(0,1)) \\ &= \begin{cases} \frac{1}{\sqrt{2\pi y}}\exp(-y/2) & y \geq 0 \\ 0 & y < 0. \end{cases} \end{aligned}$$

◊

10.8 Mixed Random Variables

We have so far described two types of random variables, the discrete random variable and the continuous random variable. The sample space for a discrete random variable consists of a countable (either finite or infinite) set of points while that for a continuous random variable has an infinite and uncountable set of points. The points in $\mathcal{S}_X$ for a discrete random variable have a nonzero probability while those for a continuous random variable have a zero probability. In some physical situations, however, we wish to assign a nonzero probability to some points but not others. As an example, consider an experiment in which a fair coin is tossed. If it comes up heads, we generate the outcome of a continuous random variable $X \sim \mathcal{N}(0,1)$ and if it comes up tails we set $X = 0$. Then, the possible outcomes are $-\infty < x < \infty$ and the probability of any point except $x = 0$ has a zero probability of occurring. However, the point $x = 0$ occurs with a probability of $1/2$ since the probability of a tail is $1/2$. A typical sequence of outcomes is shown in Figure 10.25. One could

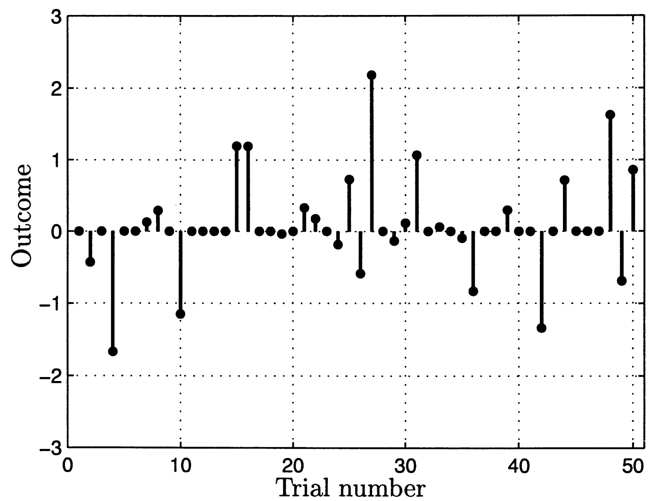


Figure 10.25: Sequence of outcomes for mixed random variable – $X = 0$ with nonzero probability.

define a random variable as

$$\begin{aligned} X &\sim \mathcal{N}(0,1) && \text{if heads} \\ X &= 0 && \text{if tails} \end{aligned}$$

which is neither a discrete nor a continuous random variable. To find its CDF we use the law of total probability to yield

$$\begin{aligned} F_X(x) &= P[X \leq x] \\ &= P[X \leq x|\text{heads}]P[\text{heads}] + P[X \leq x|\text{tails}]P[\text{tails}] \\ &= \begin{cases} \Phi(x)\frac{1}{2} + 0(\frac{1}{2}) & x < 0 \\ \Phi(x)\frac{1}{2} + 1(\frac{1}{2}) & x \geq 0 \end{cases} \end{aligned}$$

which can be written more succinctly using the unit step function. The unit step function is defined as $u(x) = 1$ for $x \geq 0$ and $u(x) = 0$ for $x < 0$. With this definition the CDF becomes

$$F_X(x) = \frac{1}{2}\Phi(x) + \frac{1}{2}u(x) \qquad -\infty < x < \infty.$$

The CDF is shown in Figure 10.26. Note the jump at $x = 0$, indicative of the contribution of the discrete part of the random variable. The CDF is continuous for all $x \neq 0$ but has a jump at $x = 0$ of $1/2$. It corresponds to neither a discrete random variable, whose CDF consists *only* of jumps, nor a continuous random variable, whose CDF is continuous everywhere. Hence, it is called a *mixed random variable.*

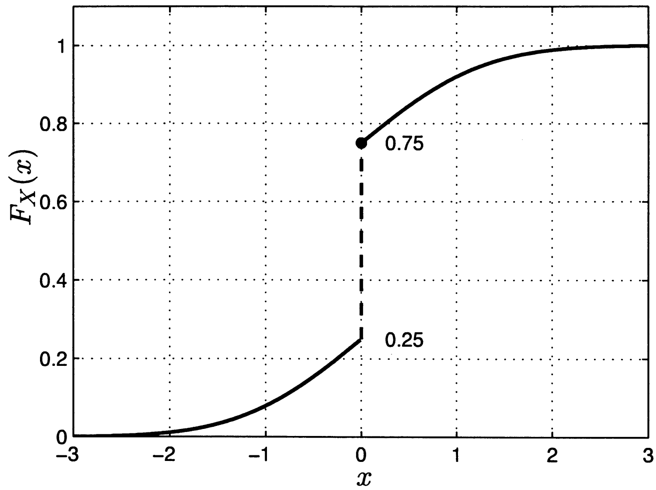


Figure 10.26: CDF for mixed random variable.

Its CDF is in general continuous except for a countable number of jumps (either finite or infinite). As usual it is right-continuous at the jump.

Strictly speaking, a mixed random variable does not have a PMF or a PDF. However, by the use of the Dirac delta function (also called an *impulse*), we can define a PDF which may then be used to find the probability of an interval via integration by using (10.4). To first find the PDF we attempt to differentiate the CDF

$$p_X(x) = \frac{d}{dx}\left[\frac{1}{2}\Phi(x) + \frac{1}{2}u(x)\right].$$

The difficulty encountered is that $u(x)$ is discontinuous at $x = 0$ and thus formally its derivative does not exist there. We can, however, *define* a derivative for the purposes of probability calculations as well as for conceptualization. To do so requires the introduction of the Dirac delta function $\delta(x)$ which is defined as (see also Appendix D)

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

The function $\delta(x)$ is usually thought of as a very narrow pulse with a very large amplitude which is centered at $x = 0$. It has the property that $\delta(t) = 0$ for all $t \neq 0$ but

$$\int_{-\epsilon}^{\epsilon} \delta(t)dt = 1$$

for ϵ a small positive number. Hence, the area under the narrow pulse is one. Using this definition we can now differentiate the CDF to find that

$$p_X(x) = \frac{1}{2}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) + \frac{1}{2}\delta(x) \tag{10.34}$$

which is shown in Figure 10.27. This may be thought of as a *generalized PDF*. Note

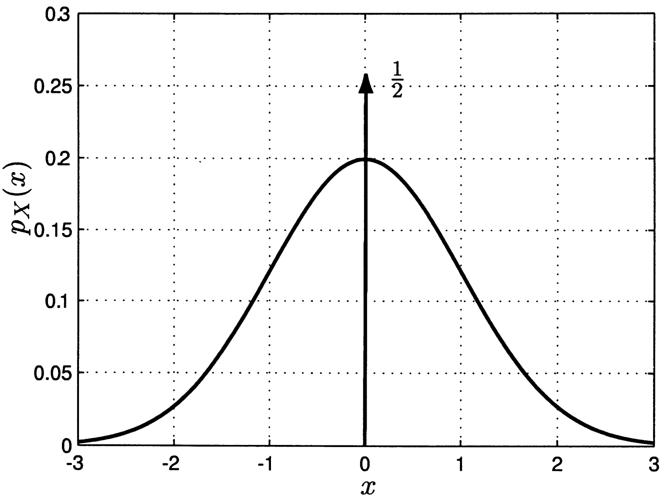


Figure 10.27: PDF for mixed random variable.

that it is the *strength*, which is defined as the area under the approximating narrow pulse, that is equal to 1/2. The amplitude is theoretically infinite. The CDF can be recovered using (10.16) and the result that

$$u(x) = \int_{-\infty}^{x^+} \delta(t)dt$$

where x^+ means that the integration interval is $(-\infty, x+\epsilon]$ for ϵ a small positive number. Thus, the impulse should be included in the integration interval if $x = 0$ so that $u(0) = 1$ according to the definition of the unit step function.

When do we include the impulse in the integration interval?

For a mixed random variable the presence of impulses in the PDF requires a modification to (10.4). This is because an endpoint of the interval can have a nonzero probability. As a result, the probabilities $P[0 < X < 1]$ and $P[0 \leq X < 1]$ will be different if there is an impulse at $x = 0$. Specifically, consider the computation of $P[0 \leq X < 1]$ and note that the probability of $X = 0$ should be included. Therefore, if there is an impulse at $x = 0$, the area under the PDF should include the contribution of the impulse. Thus, the integration interval should be chosen as $[0^-, 1]$ so that

$$P[0 \leq X < 1] = \int_{0-}^{1} p_X(x)dx.$$

The more general modifications to (10.4) are

$$\begin{aligned}
P[a \leq X \leq b] &= \int_{a^-}^{b^+} p_X(x)dx \\
P[a < X \leq b] &= \int_{a^+}^{b^+} p_X(x)dx \\
P[a \leq X < b] &= \int_{a^-}^{b^-} p_X(x)dx \\
P[a < X < b] &= \int_{a^+}^{b^-} p_X(x)dx
\end{aligned}$$

where x^- is a number slightly less than x and x^+ is a number slightly greater than x. Of course, if the PDF does not have any impulses at $x = a$ or $x = b$, then all the integrals above will be the same and, therefore there is no need to choose between them. See also Problem 10.51.

Continuing with our example, let's say we wish to determine $P[-2 \leq X \leq 2]$. Then, using (10.4) since the impulse does not occur at one of the interval endpoints, and our generalized PDF of (10.34) yields

$$\begin{aligned}
P[-2 \leq X \leq 2] &= \int_{-2}^{2} p_X(x)dx \\
&= \int_{-2}^{2} \left[\frac{1}{2}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) + \frac{1}{2}\delta(x)\right] dx \\
&= \frac{1}{2}\int_{-2}^{2} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) dx + \frac{1}{2}\int_{-2}^{2} \delta(x)dx \\
&= \frac{1}{2}[Q(-2) - Q(2)] + \frac{1}{2} \\
&= \frac{1}{2}[1 - 2Q(2)] + \frac{1}{2} = 1 - Q(2).
\end{aligned}$$

Alternatively, we could have obtained this result using $P[-2 \leq X \leq 2] = F_X(2) - F_X(-2)$ with $F_X(x) = (1/2)(1 - Q(x)) + (1/2)u(x)$.

Mixed random variables often arise as a result of a transformation of a continuous random variable. A final example follows.

Example 10.9 – PDF for amplitude-limited Rayleigh random variable

Consider a Rayleigh random variable whose PDF is given by (10.14) that is input to a device that limits its output. One might envision a physical quantity such as temperature and the device being a thermometer which can only read temperatures up to a maximum value. All temperatures above this maximum value are read as the maximum. Then the effect of the device can be represented by the transformation

$$y = g(x) = \begin{cases} x & 0 \leq x < x_{\text{max}} \\ x_{\text{max}} & x \geq x_{\text{max}} \end{cases}$$

which is shown in Figure 10.28. The PDF of Y is zero for $y < 0$ since X can only

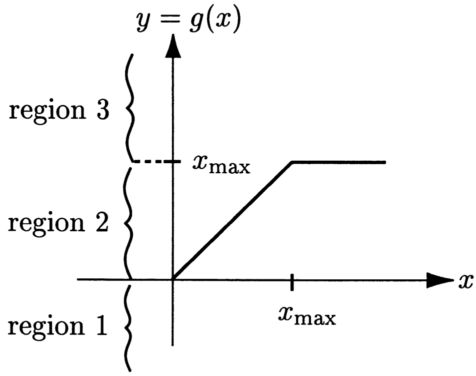


Figure 10.28: Amplitude limiter.

take on nonnegative values. For $0 \leq y < x_{\text{max}}$ it is seen from Figure 10.28 that $g^{-1}(y) = y$. Finally, for $y \geq x_{\text{max}}$ we have from Figure 10.28 the infinite number of solutions $x \in [x_{\text{max}}, \infty)$. Thus, we have for region 1 or for $y < 0$ that $p_Y(y) = 0$. For region 2 or for $0 \leq y < x_{\text{max}}$ where $g^{-1}(y) = y$, we have from (10.30)

$$\begin{aligned} p_Y(y) &= p_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| \\ &= p_X(y). \end{aligned}$$

For region 3 which is $y \geq x_{\text{max}}$, we note that Y cannot exceed x_{max} and so $y = x_{\text{max}}$ is the only possible value for y in region 3. The probability of $Y = x_{\text{max}}$ is equal to the probability that $X \geq x_{\text{max}}$. In particular, it is

$$P[Y = x_{\text{max}}] = \int_{x_{\text{max}}}^{\infty} p_X(x)dx \tag{10.35}$$

since from Figure 10.28 the x-interval $[x_{\text{max}}, \infty)$ is mapped into the y-point given by $y = x_{\text{max}}$. Since the probability of Y at the point $y = x_{\text{max}}$ is nonzero, we represent its contribution to the PDF by using an impulse as

$$p_Y(y) = \left[\int_{x_{\text{max}}}^{\infty} p_X(x)dx\right] \delta(y - x_{\text{max}}) \qquad y = x_{\text{max}}.$$

In summary, the PDF of the transformed random variable is

$$p_Y(y) = \begin{cases} 0 & y < 0 \\ p_X(y) & 0 \le y < x_{\max} \\ \left[\int_{x_{\max}}^{\infty} p_X(x)dx\right] \delta(y - x_{\max}) & y = x_{\max} \\ 0 & y > x_{\max} \, . \end{cases}$$

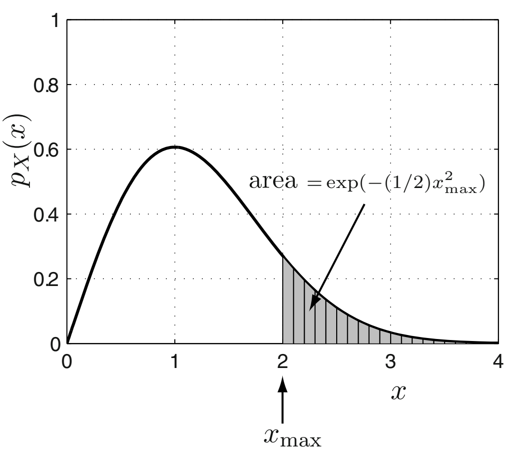
It is seen to be the PDF of a mixed random variable in that it contains an impulse. Finally, for $x \ge 0$ the Rayleigh PDF is for $\sigma^2 = 1$

$$p_X(x) = x \exp\left(-\frac{1}{2}x^2\right)$$

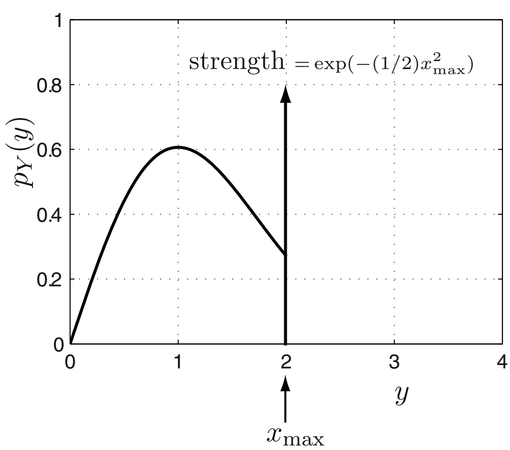
so that the PDF of Y becomes

$$\begin{aligned} p_Y(y) &= \begin{cases} 0 & y < 0 \\ y \exp\left(-\frac{1}{2}y^2\right) & 0 \le y < x_{\max} \\ \left[\int_{x_{\max}}^{\infty} x \exp\left(-\frac{1}{2}x^2\right) dx\right] \delta(y - x_{\max}) & y = x_{\max} \\ 0 & y > x_{\max} \, . \end{cases} \\ &= \begin{cases} 0 & y < 0 \\ y \exp\left(-\frac{1}{2}y^2\right) & 0 \le y < x_{\max} \\ \exp\left(-\frac{1}{2}x_{\max}^2\right) \delta(y - x_{\max}) & y = x_{\max} \\ 0 & y > x_{\max}. \end{cases} \end{aligned}$$

This is plotted in Figure 10.29b.



(a) PDF of X – continuous random variable



(b) PDF of $Y = g(X)$ – mixed random variable

Figure 10.29: PDFs before and after transformation of Figure 10.28.

$\diamondsuit$

In general, if a random variable X can take on a continuum of values as well as discrete values $\{x_1, x_2 \ldots\}$ with corresponding nonzero probabilities $\{p_1, p_2, \ldots\}$, then the PDF of the mixed random variable X can be written in the succinct form

$$p_X(x) = p_c(x) + \sum_{i=1}^{\infty} p_i \delta(x - x_i) \tag{10.36}$$

where $p_c(x)$ represents the contribution to the PDF of the continuous part (its integral must be < 1) and must satisfy $p_c(x) \geq 0$. To be a valid PDF we require that

$$\int_{-\infty}^{\infty} p_c(x)dx + \sum_{i=1}^{\infty} p_i = 1.$$

For solely discrete random variables we can use the generalized PDF

$$p_X(x) = \sum_{i=1}^{\infty} p_i \delta(x - x_i)$$

or equivalently the PMF

$$p_X[x_i] = p_i \qquad i = 1, 2, \ldots$$

to perform probability calculations.

10.9 Computer Simulation

In simulating the outcome of a discrete random variable X we saw in Figure 5.14 that first an outcome of a $U \sim \mathcal{U}(0,1)$ random variable is generated and then mapped into a value of X. The mapping needed was the inverse of the CDF. This result is also valid for a continuous random variable so that $X = F_X^{-1}(U)$ is a random variable with CDF $F_X(x)$. Stated another way, we have that $U = F_X(X)$ or if a random variable is transformed according to its CDF, the transformed random variable $U \sim \mathcal{U}(0,1)$. This latter transformation is termed *the probability integral transformation.* The transformation $X = F_X^{-1}(U)$ is called the *inverse probability integral transformation.* Before proving these results we give an example.

Example 10.10 – Probability integral transformation of exponential random variable

Since the exponential PDF is given for $\lambda = 1$ by

$$p_X(x) = \begin{cases} \exp(-x) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

the CDF is from (10.16)

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ 1 - \exp(-x) & x > 0 \,. \end{cases}$$

The probability integral transformation asserts that $Y = g(X) = F_X(X)$ has a $\mathcal{U}(0,1)$ PDF. Considering the transformation $g(x) = 1 - \exp(-x)$ for $x > 0$ and zero otherwise, we have that $y = 1 - \exp(-x)$ and, therefore the unique solution for x is $x = -\ln(1-y)$ for $0 < y < 1$ and zero otherwise. Hence,

$$g^{-1}(y) = \begin{cases} -\ln(1-y) & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

and using (10.30), we have for $0 < y < 1$

$$\begin{aligned} p_Y(y) &= p_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| \\ &= \exp\left[-\left(-\ln(1-y)\right)\right] \left| \frac{1}{1-y} \right| \\ &= 1. \end{aligned}$$

Finally, then

$$p_Y(y) = \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

which is the PDF of a $\mathcal{U}(0,1)$ random variable.

◇

To summarize our results we have the following theorem.

Theorem 10.9.1 (Inverse Probability Integral Transformation) *If a continuous random variable X is given as $X = F_X^{-1}(U)$, where $U \sim \mathcal{U}(0,1)$, then X has the PDF $p_X(x) = dF_X(x)/dx$.*

Proof:
Let $V = F_X^{-1}(U)$ and consider the CDF of V.

$$\begin{aligned} F_V(v) &= P[V \le v] = P[F_X^{-1}(U) \le v] \\ &= P[U \le F_X(v)] \quad (F_X \text{ is monotonically increasing – see Problem 10.58}) \\ &= \int_0^{F_X(v)} p_U(u)du \\ &= \int_0^{F_X(v)} 1\, du \\ &= F_X(v). \end{aligned}$$

Hence, the CDFs of V and X are equal and therefore the PDF of $V = F_X^{-1}(U)$ is $p_X(x)$.

△

Another example follows.

Example 10.11 – Computer generation of outcome of Laplacian random variable

The Laplacian random variable has a PDF

$$p_X(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left[-\sqrt{\frac{2}{\sigma^2}}|x|\right] \qquad -\infty < x < \infty$$

and therefore its CDF is found as

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\sigma^2}} \exp\left[-\sqrt{\frac{2}{\sigma^2}}|t|\right] dt.$$

For $x < 0$ we have

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\sigma^2}} \exp\left[\sqrt{\frac{2}{\sigma^2}}t\right] dt \\ &= \frac{1}{2} \exp\left[\sqrt{\frac{2}{\sigma^2}}t\right]\Bigg|_{-\infty}^{x} \\ &= \frac{1}{2} \exp\left[\sqrt{\frac{2}{\sigma^2}}x\right] \end{aligned}$$

and for $x \geq 0$ we have

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{0} \frac{1}{\sqrt{2\sigma^2}} \exp\left[\sqrt{\frac{2}{\sigma^2}}t\right] dt + \int_{0}^{x} \frac{1}{\sqrt{2\sigma^2}} \exp\left[-\sqrt{\frac{2}{\sigma^2}}t\right] dt \\ &= \frac{1}{2} - \frac{1}{2} \exp\left[-\sqrt{\frac{2}{\sigma^2}}t\right]\Bigg|_{0}^{x} \qquad \text{(first integral is 1/2 since } p_X(-x) = p_X(x)) \\ &= 1 - \frac{1}{2} \exp\left[-\sqrt{\frac{2}{\sigma^2}}x\right]. \end{aligned}$$

By letting $y = F_X(x)$, we have

$$y = \begin{cases} \frac{1}{2} \exp\left[\sqrt{\frac{2}{\sigma^2}}x\right] & x < 0 \\ 1 - \frac{1}{2} \exp\left[-\sqrt{\frac{2}{\sigma^2}}x\right] & x \geq 0\,. \end{cases}$$

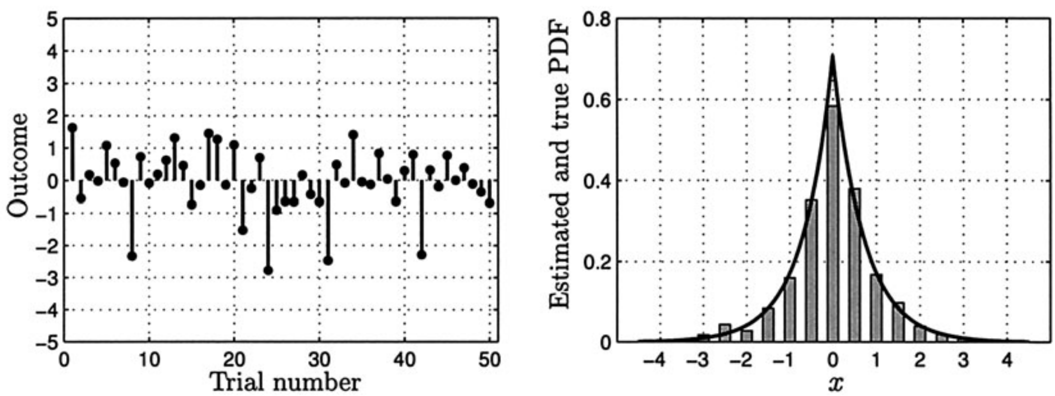
We note that if $x < 0$, then $0 < y < 1/2$, and if $x \geq 0$, then $1/2 \leq y < 1$. Thus, solving for x to produce $F_X^{-1}(y)$ yields

$$x = \begin{cases} \sqrt{\sigma^2/2}\ln(2y) & 0 < y < 1/2 \\ \sqrt{\sigma^2/2}\ln\left(\frac{1}{2(1-y)}\right) & 1/2 \leq y < 1\,. \end{cases}$$

Finally to generate the outcome of a Laplacian random variable we can use

$$x = \begin{cases} \sqrt{\sigma^2/2}\ln(2u) & 0 < u < 1/2 \\ \sqrt{\sigma^2/2}\ln\left(\frac{1}{2(1-u)}\right) & 1/2 \le u < 1 \end{cases} \tag{10.37}$$

where u is a realization of a $\mathcal{U}(0,1)$ random variable. An example of the outcomes of a Laplacian random variable with $\sigma^2 = 1$ is shown in Figure 10.30a. In Figure 10.30b the true PDF (the solid curve) along with the estimated PDF (the bar plot) is shown based on $M = 1000$ outcomes. The estimate of the PDF was accomplished by



(a) First 50 outcomes

(b) True PDF and estimated PDF based on 1000 outcomes

Figure 10.30: Computer generation of Laplacian random variable outcomes using inverse probability integral transformation.

the procedure described in Example 2.1 (see Figure 2.7 for the code for a Gaussian PDF). We can now justify that procedure. Since from Section 10.3 we have

$$p_X(x_0) \approx \frac{P[x_0 - \Delta x/2 \le X \le x_0 + \Delta x/2]}{\Delta x}$$

and

$$P[x_0 - \Delta x/2 \le X \le x_0 + \Delta x/2] \approx \frac{\text{Number of outcomes in } [x_0 - \Delta x/2, x_0 + \Delta x/2]}{M}$$

we use as our PDF estimator

$$\hat{p}_X(x_0) = \frac{\text{Number of outcomes in } [x_0 - \Delta x/2, x_0 + \Delta x/2]}{M\Delta x}. \tag{10.38}$$

In Figure 10.30b we have chosen the *bins* or intervals to be $[-4.25, -3.75]$, $[-3.75, -3.25]$, $\ldots$, $[3.75, 4.25]$ so that $\Delta x = 0.5$. We have therefore estimated $p_X(-4), p_X(-3.5), \ldots,$

$p_X(4)$. To estimate the PDF at more points we would have to decrease the *binwidth* or Δx. However, in doing so we cannot make it too small. This is because as the binwidth decreases, the probability of an outcome falling within the bin also decreases. As a result, fewer of the outcomes will occur within each bin, resulting in a poor estimate. The only way to remedy this situation is to increase the number of trials M. What do you suppose would happen if we wanted to estimate $p_X(5)$? The MATLAB code for producing the PDF estimate is given below.

```
% Assume outcomes are in x, which is M x 1 vector
M=1000;
bincenters=[-4:0.5:4]'; % set binwidth = 0.5
bins=length(bincenters);
h=zeros(bins,1);
for i=1:length(x)  % count outcomes in each bin
  for k=1:bins
    if x(i)>bincenters(k)-0.5/2. ...
      & x(i)<bincenters(k)+0.5/2
      h(k,1)=h(k,1)+1;
    end
  end
end
pxest=h/(M*0.5); % see (10.38)
```

The CDF can be estimated by using

$$\hat{F}_X(x) = \frac{\text{Number of outcomes} \le x}{M} \tag{10.39}$$

and is the same for either a discrete or a continuous random variable. See also Problems 10.60–62.

◇

10.10 Real-World Example - Setting Clipping Levels for Speech Signals

In order to communicate speech over a transmission channel it is important to make sure that the equipment does not "clip" the speech signal. Commercial broadcast stations commonly use VU meters to monitor the power of the speech. If the power becomes too large, then the amplifier gains are manually decreased. Clipped speech sounds distorted and is objectionable. In other situations, the amplifier gains must be set automatically, as for example, in telephone speech transmission. This is necessary so that the speech, if transmitted in an analog form, is not distorted at the receiver, and if transmitted in a digital form is not clipped by an analog-to-digital convertor. To determine the highest amplitude of the speech signal that can be expected to occur a common model is to use a Laplacian PDF for the amplitudes [Rabiner and Schafer 1978]. Hence, most of the amplitudes are near zero but larger level ones are possible according to

$$p_X(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left[-\sqrt{\frac{2}{\sigma^2}}|x|\right] \qquad -\infty < x < \infty.$$

As seen in Figure 10.10, the width of the PDF increases as σ^2 increases. In effect, σ^2 measures the width of the PDF and is actually its variance (to be shown in Problem 11.34). The parameter σ^2 is also a measure of the speech power. In order to avoid excessive clipping we must be sure that an amplifier can accommodate a high level, even if it occurs rather infrequently. A design requirement might then be to transmit a speech signal without clipping 99% of the time. A model for a clipper is shown in Figure 10.31. As long as the input signal, i.e., x, remains in the interval

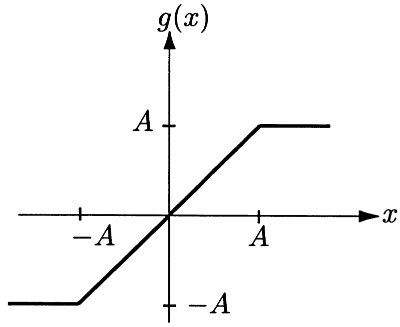


Figure 10.31: Clipper input-output characteristics.

$-A \leq x \leq A$, the output will be the same as the input and no clipping takes place. However, if $x > A$, the output will be limited to A and similarly if $x < -A$. Clipping will then occur whenever $|x| > A$. To satisfy the design requirement that clipping should not occur for 99% of the time, we should choose A (which is a characteristic of the amplifier or analog-to-digital convertor) so that $P_{\text{clip}} \leq 0.01$. But

$$P_{\text{clip}} = P[X > A \text{ or } X < -A]$$

and since the Laplacian PDF is symmetric about $x = 0$ this is just

$$\begin{aligned} P_{\text{clip}} &= 2P[X > A] = 2\int_A^\infty \frac{1}{\sqrt{2\sigma^2}} \exp\left[-\sqrt{\frac{2}{\sigma^2}}x\right] dx \\ &= 2\left[-\frac{1}{2}\exp\left[-\sqrt{\frac{2}{\sigma^2}}x\right]\Bigg|_A^\infty\right] \\ &= \exp\left[-\sqrt{\frac{2}{\sigma^2}}A\right]. \end{aligned} \qquad (10.40)$$

Hence, if this probability is to be no more than 0.01, we must have

$$\exp\left[-\sqrt{\frac{2}{\sigma^2}}A\right] \leq 0.01$$

or solving for A produces the requirement that

$$A \geq \sqrt{\frac{\sigma^2}{2}} \ln\left(\frac{1}{0.01}\right). \tag{10.41}$$

It is seen that as the speech power σ^2 increases, so must the clipping level A. If the clipping level is fixed, then speech with higher powers will be clipped more often. As an example, consider a speech signal with $\sigma^2 = 1$. The Laplacian model outcomes are shown in Figure 10.32 along with a clipping level of $A = 1$. According to (10.40)

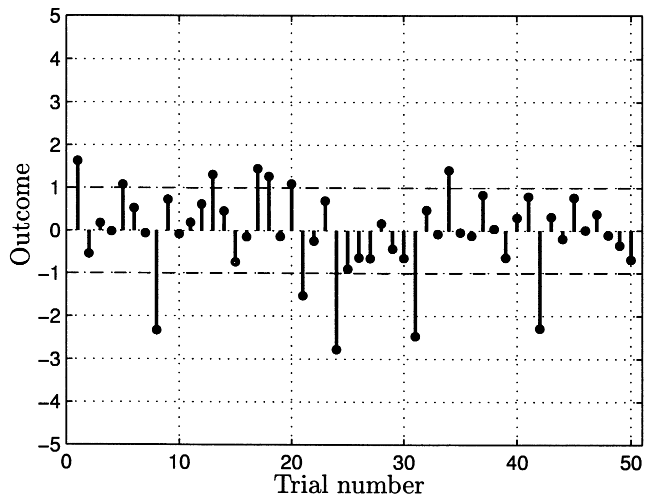


Figure 10.32: Outcomes of Laplacian random variable with $\sigma^2 = 1$ – model for speech amplitudes.

the probability of clipping is $\exp(-\sqrt{2}) = 0.2431$. Since there are 50 outcomes in Figure 10.32 we would expect about $50 \cdot 0.2431 \approx 12$ instances of clipping. From the figure we see that there are exactly 12. To meet the specification we should have that

$$A \geq \sqrt{1/2} \ln\left(\frac{1}{0.01}\right) = 3.25.$$

As seen from Figure 10.32 there are no instances of clipping for $A = 3.25$. In order to set the appropriate clipping level A, we need to know σ^2. In practice, this too must be estimated since different speakers have different volumes and even the same speaker will exhibit a different volume over time!

References

Abramowitz, M., I.A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965.

Capinski, M., P.E. Kopp, *Measure, Integral, and Probability*, Springer-Verlag, New York, 2004.

Johnson, N.L., S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions, Vols. 1,2*, John Wiley & Sons, New York, 1994.

Members of Technical Staff, *Transmission Systems for Communications*, Western Electric Co., Inc., Winston-Salem, NC, 1970.

Rabiner, L.R., R.W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, NJ, 1978.

Widder, D.A., *Advanced Calculus*, Dover, New York, 1989.

Problems

10.1 (w) Are the following random variables continuous or discrete?

- **a.** Temperature in degrees Fahrenheit
- **b.** Temperature rounded off to nearest 1°
- **c.** Temperature rounded off to nearest $1/2^\circ$
- **d.** Temperature rounded off to nearest $1/4^\circ$

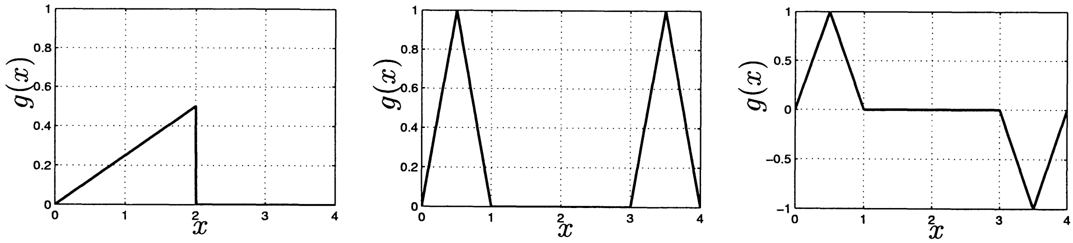
10.2 (⌣̈) (w) The temperature in degrees Fahrenheit is modeled as a uniform random variable with $T \sim \mathcal{U}(20, 60)$. If T is rounded off to the nearest $1/2^\circ$ to form $\hat{T}$, what is $P[\hat{T} = 30^\circ]$? What can you say about the use of a PDF versus a PMF to describe the probabilistic outcome of a physical experiment?

10.3 (w) A wedge of cheese as shown in Figure 10.5 is sliced from $x = a$ to $x = b$. If $a = 0$ and $b = 0.2$, what is the mass of cheese in the wedge? How about if $a = 1.8$ and $b = 2$?

10.4 (⌣̈) (w) Which of the functions shown in Figure 10.33 are valid PDFs? If a function is not a PDF, why not?

10.5 (f) Determine the value of c to make the following function a valid PDF

$$g(x) = \begin{cases} c(1 - |x/5|) & |x| < 5 \\ 0 & \text{otherwise.} \end{cases}$$



(a) (b) (c)

Figure 10.33: Possible PDFs for Problem 10.4.

10.6 (☺) (w) A Gaussian mixture PDF is defined as

$$p_X(x) = \alpha_1 \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{1}{2\sigma_1^2}x^2\right) + \alpha_2 \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{1}{2\sigma_2^2}x^2\right)$$

for $\sigma_1^2 \neq \sigma_2^2$. What are the possible values for α_1 and α_2 so that this is a valid PDF?

10.7 (w) Find the area under the curves given by the following functions:

$$g_1(x) = \begin{cases} x & 0 \le x < 1 \\ 1+x & 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$$g_2(x) = \begin{cases} x & 0 \le x \le 1 \\ 1+x & 1 < x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

and explain your results.

10.8 (w) A memory chip has a projected lifetime X in days that is modeled as $X \sim \exp(0.001)$. What is the probability that it will fail within one year?

10.9 (t) In this problem we prove that the Gaussian PDF integrates to one. First we let

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx$$

and write I^2 as the iterated integral

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) dy dx.$$

Next, convert (x, y) into polar coordinates and evaluate the expression to prove that $I^2 = 1$. Finally, you can conclude that $I = 1$ (why?).

10.10 (f,c) If $X \sim \mathcal{N}(\mu, \sigma^2)$, find $P[X > \mu + a\sigma]$ for $a = 1, 2, 3$, where $\sigma = \sqrt{\sigma^2}$.

10.11 (t) The *median* of a PDF is defined as the point $x = \text{med}$ for which $P[X \leq \text{med}] = 1/2$. Prove that if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\text{med} = \mu$.

10.12 (☺) (w) A constant or DC current source that outputs 1 amp is connected to a resistor of nominal resistance of 1 ohm. If the resistance value can vary according to $R \sim \mathcal{N}(1, 0.1)$, what is the probability that the voltage across the resistor will be between 0.99 and 1.01 volts?

10.13 (w) An analog-to-digital convertor can convert voltages in the range $[-3, 3]$ volts to a digital number. Outside this range, it will "clip" a positive voltage at the highest positive level, i.e., $+3$, or a negative voltage at the most negative level, i.e., -3. If the input to the convertor is modeled as $X \sim \mathcal{N}(\mu, 1)$, how should μ be chosen to minimize the probability of clipping?

10.14 (☺) (f) Find $P[X > 3]$ for the two PDFs given by the Gaussian PDF with $\mu = 0, \sigma^2 = 1$ and the Laplacian PDF with $\sigma^2 = 1$. Which probability is larger and why? Plot both PDFs.

10.15 (f) Verify that the Cauchy PDF given in (10.9) integrates to one.

10.16 (t) Prove that $\Gamma(z+1) = z\Gamma(z)$ by using integration by parts (see Appendix B and Problem 11.7).

10.17 (☺) (f) The arrival time in minutes of the Nth person at a ticket counter has a PDF that is Erlang with $\lambda = 0.1$. What is the probability that the first person will arrive within the first 5 minutes of the opening of the ticket counter? What is the probability that the first two persons will arrive within the first 5 minutes of opening?

10.18 (f) A person cuts off a wedge of cheese as shown in Figure 10.5 starting at $x = 0$ and ending at some value $x = x_0$. Determine the mass of the wedge as a function of the value x_0. Can you relate this to the CDF?

10.19 (☺) (f) Determine the CDF for the Cauchy PDF.

10.20 (f) If $X \sim \mathcal{N}(0, 1)$ find the probability that $|X| \leq a$, where $a = 1, 2, 3$. Also, plot the PDF and shade in the corresponding areas under the PDF.

10.21 (f,c) If $X \sim \mathcal{N}(0, 1)$, determine the number of outcomes out of 1000 that you would expect to occur within the interval $[1, 2]$. Next conduct a computer simulation to carry out this experiment. How many outcomes actually occur within this interval?

10.22 (☺) (w) If $X \sim \mathcal{N}(\mu, \sigma^2)$, find the CDF of X in terms of $\Phi(x)$.

10.23 (t) If a PDF is symmetric about $x = 0$ (also called an *even function*), prove that $F_X(-x) = 1 - F_X(x)$. Does this property hold for a Gaussian PDF with $\mu = 0$? Hint: See Figure 10.16.

10.24 (t) Prove that if $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$P[X > a] = Q\left(\frac{a-\mu}{\sigma}\right)$$

where $\sigma = \sqrt{\sigma^2}$.

10.25 (t) Prove the properties of the Q function given by (10.19)–(10.22).

10.26 (f) Plot the function $Q(A/2)$ versus A for $0 \leq A \leq 5$ to verify the true probability of error as shown in Figure 2.15.

10.27 (c) If $X \sim \mathcal{N}(0, 1)$, evaluate $P[X > 4]$ and then verify your results using a computer simulation. How easy do you think it would be to determine $P[X > 7]$ using a computer simulation? (See Section 11.10 for an alternative approach.)

10.28 (⌣̇̇) (w) A survey is taken of the incomes of a large number of people in a city. It is determined that the income in dollars is distributed as $X \sim \mathcal{N}(50000, 10^8)$. What percentage of the people have incomes above \$70,000?

10.29 (w) In Chapter 1 an example was given of the length of time in minutes an office worker spends on the telephone in a given 10-minute period. The length of time T was given as $\mathcal{N}(7, 1)$ as shown in Figure 1.5. Determine the probability that a caller is on the telephone more than 8 minutes by finding $P[T > 8]$.

10.30 (⌣̇̇) (w) A population of high school students in the eastern United States score X points on their SATs, where $X \sim \mathcal{N}(500, 4900)$. A similar population in the western United States score X points, where $X \sim \mathcal{N}(525, 3600)$. Which group is more likely to have scores above 700?

10.31 (f) Verify the numerical results given in (1.3).

10.32 (f) In Example 2.2 we asserted that $P[X > 2]$ for a standard normal random variable is 0.0228. Verify this result.

10.33 (⌣̇̇) (w) Is the following function a valid CDF?

$$F_X(x) = \frac{1}{1 + \exp(-x)} \qquad -\infty < x < \infty.$$

10.34 (f) If $F_X(x) = (2/\pi)\arctan(x)$ for $0 \leq x < \infty$, determine $P[0 \leq X \leq 1]$.

10.35 (t) Prove that (10.25) is true.

10.36 (⌣) (w) Professor Staff always scales his test scores. He adds a number of points c to each score so that 50% of the class get a grade of C. A C is given if the score is between 70 and 80. If the scores have the distribution $\mathcal{N}(65, 38)$, what should c be? Hint: There are two possible solutions to this problem but the students will prefer only one of them.

10.37 (w) A Rhode Island weatherman says that he can accurately predict the temperature for the following day 95% of the time. He makes his prediction by saying that the temperature will be between T_1° Fahrenheit and T_2° Fahrenheit. If he knows that the actual temperature is a random variable with PDF $\mathcal{N}(50, 10)$, what should his prediction be for the next day?

10.38 (f) For the CDF given in Figure 10.14 find the PDF by differentiating. What happens at $x = 1$ and $x = 2$?

10.39 (f,c) If $Y = \exp(X)$, where $X \sim \mathcal{U}(0, 1)$, find the PDF of Y. Next generate realizations of X on a computer and transform them according to $\exp(X)$ to yield the realizations of Y. Plot the x's and y's in a similar manner to that shown in Figure 10.22 and discuss your results.

10.40 (⌣) (f) Find the PDF of $Y = X^4 + 1$ if $X \sim \exp(\lambda)$.

10.41 (w) Find the constants a and b so that $Y = aX + b$, where $X \sim \mathcal{U}(0, 1)$, yields $Y \sim \mathcal{U}(2, 6)$.

10.42 (f) If $Y = aX$, find the PDF of Y if the PDF of X is $p_X(x)$. Next, assume that $X \sim \exp(1)$ and find the PDFs of Y for $a > 1$ and $0 < a < 1$. Plot these PDFs and explain your results.

10.43 (⌣) (f) Find a general formula for the PDF of $Y = |X|$. Next, evaluate your formula if X is a standard normal random variable.

10.44 (f) If $X \sim \mathcal{N}(0, 1)$ is transformed according to $Y = \exp(X)$, determine $p_Y(y)$ by using the CDF approach. Compare your results to those given in Example 10.6. Hint: You will need Leibnitz's rule

$$\frac{d}{dy} \int_a^{g(y)} p(x)dx = p(g(y)) \frac{dg(y)}{dy}.$$

10.45 (w) A random voltage X is input to a full wave rectifier that produces at its output the absolute value of the voltage. If X is a standard normal random variable, what is the probability that the output of the rectifier will exceed 2?

10.46 (⌣) **(f,c)** If $Y = X^2$, where $X \sim \mathcal{U}(0,1)$, determine the PDF of Y. Next perform a computer simulation using the realizations of Y (obtained as $y_m = x_m^2$, where x_m is the mth realization of X) to estimate the PDF $p_Y(y)$. Do your theoretical results match the simulated results?

10.47 (w) If a discrete random variable X has a Ber(p) *PMF*, find the *PDF* of X using impulses. Next find the CDF of X by integrating the PDF.

10.48 (t) In this problem we point out that the use of impulses or Dirac delta functions serves mainly as a tool to allow sums to be written as integrals. For example, the sum

$$S = \sum_{i=1}^{N} a_i$$

can be written as the integral

$$S = \int_{-\infty}^{\infty} g(x)dx$$

if we define $g(x)$ as

$$g(x) = \sum_{i=1}^{N} a_i \delta(x-i).$$

Verify that this is true and show how it applies to computing probabilities of events of discrete random variables by using integration.

10.49 (f) Evaluate the expression

$$\int_1^{2^-} \left(\frac{1}{2}\delta(x-2) + \frac{3}{8}\delta(x-4) + \frac{1}{8}\delta(x-3/2)\right) dx.$$

Could the integrand represent a PDF? If it does, what does this integral represent?

10.50 (w) Plot the PDF and CDF if

$$p_X(x) = \frac{1}{2}\exp(-x)u(x) + \frac{1}{4}\delta(x+1) + \frac{1}{4}\delta(x-1).$$

10.51 (⌣) **(w)** For the PDF given in Problem 10.50 determine the following: $P[-2 \leq X \leq 2]$, $P[-1 \leq X \leq 1]$, $P[-1 < X \leq 1]$, $P[-1 < X < 1]$, $P[-1 \leq X < 1]$.

10.52 (f) Find and plot the PDF of the transformed random variable

$$Y = \begin{cases} 2X & 0 \leq X < 1 \\ 2 & X \geq 1 \end{cases}$$

where $X \sim \exp(1)$.

10.53 (f) Find the PDF representation of the PMF of a bin(3, 1/2) random variable. Plot the PMF and the PDF.

10.54 (⌣) (f) Determine the function g so that $X = g(U)$, where $U \sim \mathcal{U}(0,1)$, has a Rayleigh PDF with $\sigma^2 = 1$.

10.55 (f) Find a transformation so that $X = g(U)$, where $U \sim \mathcal{U}(0,1)$, has the PDF shown in Figure 10.34.

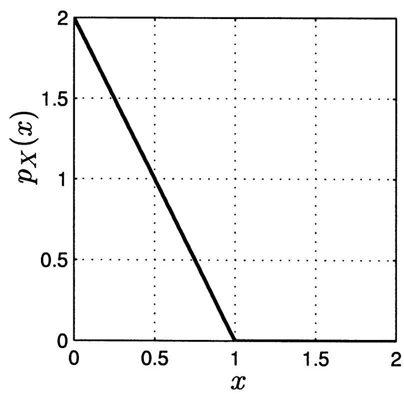


Figure 10.34: PDF for Problem 10.55

10.56 (c) Verify your results in Problem 10.55 by generating realizations of the random variable whose PDF is shown in Figure 10.34. Next estimate the PDF and compare it to the true PDF.

10.57 (t) A monotonically increasing function $g(x)$ is defined as one for which if $x_2 \geq x_1$, then $g(x_2) \geq g(x_1)$. A monotonically decreasing function is one for which if $x_2 \geq x_1$, then $g(x_2) \leq g(x_1)$. It can be shown that if $g(x)$ is differentiable, then a function is monotonically increasing (decreasing) if $dg(x)/dx \geq 0$ ($dg(x)/dx \leq 0$) for all x. Which of the following functions are monotonically increasing or decreasing: $\exp(x)$, $\ln(x)$, and $1/x$?

10.58 (t) Explain why the values of x for which the inequality $x \geq x_0$ is true do not change if we take the logarithm of both sides to yield $\ln(x) \geq \ln(x_0)$. Would the inequality still hold if we inverted both sides or equivalently applied the function $g(x) = 1/x$ to both sides? Hint: See Problem 10.57.

10.59 (w) Compare the true PDF given in Figure 10.24 with the estimated PDF shown in Figure 2.10. Are they the same and if not, why not?

10.60 (c) Generate on the computer realizations of the random variable $X \sim \mathcal{N}(1,4)$. Estimate the PDF and compare it to the true one.

10.61 (c) Determine the PDF of $Y = X^3$ if $X \sim \mathcal{U}(0,1)$. Next generate realizations of X on the computer, apply the transformation $g(x) = x^3$ to each realization to yield realizations of Y, and finally estimate the PDF of Y from these realizations. Does it agree with the true PDF?

10.62 (c) For the random variable Y described in Problem 10.61 determine the CDF. Then, generate realizations of Y, estimate the CDF, and compare it to the true one.

Appendix 10A

Derivation of PDF of a Transformed Continuous Random Variable

The proof uses the CDF approach as described in Section 10.7. It assumes that g is a one-to-one function. If $Y = g(X)$, where g is a one-to-one and monotonically *increasing* function, then there is a single solution for x in $y = g(x)$. Thus,

$$\begin{aligned} F_Y(y) &= P[g(X) \le y] \\ &= P[X \le g^{-1}(y)] \\ &= F_X(g^{-1}(y)]. \end{aligned}$$

But $p_Y(y) = dF_Y(y)/dy$ so that

$$\begin{aligned} p_Y(y) &= \frac{d}{dy} F_X(g^{-1}(y)) \\ &= \left. \frac{dF_X(x)}{dx} \right|_{x=g^{-1}(y)} \frac{dg^{-1}(y)}{dy} \qquad \text{(chain rule of calculus)} \\ &= p_X(g^{-1}(y)) \frac{dg^{-1}(y)}{dy}. \end{aligned}$$

If $g(x)$ is one-to-one and monotonically *decreasing*, then

$$\begin{aligned} F_Y(y) &= P[g(X) \le y] \\ &= P[X \ge g^{-1}(y)] \\ &= 1 - P[X \le g^{-1}(y)] \quad (\text{since } P[X = g^{-1}(y)] = 0) \\ &= 1 - F_X(g^{-1}(y)) \end{aligned}$$

and

$$\begin{aligned} p_Y(y) &= \frac{dF_Y(y)}{dy} = -\frac{d}{dy}F_X(g^{-1}(y)) \\ &= -p_X(g^{-1}(y))\frac{dg^{-1}(y)}{dy}. \end{aligned}$$

Note that if g is montonically decreasing, then g^{-1} is also montonically decreasing. Hence, $dg^{-1}(y)/dy$ will be negative. Thus, both cases can be subsumed by the formula

$$p_Y(y) = p_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|.$$

Appendix 10B

MATLAB Subprograms to Compute Q and Inverse Q Functions

```
% Q.m
%
% This program computes the right-tail probability
% (complementary cumulative distribution function) for
% a N(0,1) random variable.
%
% Input Parameters:
%
%   x - Real column vector of x values
%
% Output Parameters:
%
%   y - Real column vector of right-tail probabilities
%
% Verification Test Case:
%
% The input x=[0 1 2]'; should produce y=[0.5 0.1587 0.0228]'.
%
  function y=Q(x)
  y=0.5*erfc(x/sqrt(2)); % complementary error function
```

```
% Qinv.m
```

```
%
% This program computes the inverse Q function or the value
% which is exceeded by a N(0,1) random variable with a
% probability of x.
%
% Input Parameters:
%
%   x - Real column vector of right-tail probabilities
%       (in interval [0,1])
%
% Output Parameters:
%
%   y - Real column vector of values of random variable
%
% Verification Test Case:
%
% The input x=[0.5 0.1587 0.0228]'; should produce
% y=[0 0.9998 1.9991]'.
%
  function y=Qinv(x)
  y=sqrt(2)*erfinv(1-2*x);  % inverse error function
```