Chapter 3

Continuous Random Variables

3.1 INTRODUCTION

So far, we have considered discrete random variables and their distributions. In applications, such random variables denote the number of objects of a certain type, such as the number of job arrivals to a file server in one minute or the number of calls into a telephone exchange in one minute.

Many situations, both applied and theoretical, require the use of random variables that are "continuous" rather than discrete. As described in the last chapter, a random variable is a real-valued function on the sample space S. When the sample space S is nondenumerable (as mentioned in Section 1.7), not every subset of the sample space is an event that can be assigned a probability. As before, let $\mathcal{F}$ denote the class of measurable subsets of S. Now if X is to be a random variable, it is natural to require that $P(X \leq x)$ be well defined for every real number x. In other words, if X is to be a random variable defined on a probability space $(S, \mathcal{F}, P)$, we require that $\{s|X(s) \leq x\}$ be an event (i.e., a member of $\mathcal{F}$). We are, therefore, led to the following extension of our earlier definition.

Definition (Random Variable). A random variable X on a probability space $(S, \mathcal{F}, P)$ is a function $X : S \rightarrow \Re$ that assigns a real number $X(s)$ to each sample point $s \in S$, such that for every real number x, the set of sample points $\{s|X(s) \leq x\}$ is an event, that is, a member of $\mathcal{F}$.

Definition (Distribution Function). The (cumulative) distribution function or CDF F_X of a random variable X is defined to be the function

$$F_X(x) = P(X \leq x), \qquad -\infty < x < \infty.$$

Probability and Statistics with Reliability, Queuing, and Computer Science Applications, Second Edition. Kishor S. Trivedi.

Companion Website: www.wiley.com/go/trivedi/probabilityandstatistics2e

The subscript X is used here to indicate the random variable under consideration. When there is no ambiguity, the subscript will be dropped and $F_X(x)$ will be denoted by $F(x)$.

As we saw in Chapter 2, the distribution function of a discrete random variable grows only by jumps. By contrast, the distribution function of a continuous random variable has no jumps but grows continuously. Thus, a **continuous random variable** X is characterized by a distribution function $F_X(x)$ that is a continuous function of x for all $-\infty < x < \infty$. Most continuous random variables that we encounter will have an absolutely continuous distribution function, $F(x)$, that is, one for which the derivative, $dF(x)/dx$, exists everywhere (except perhaps at a finite number of points). Such a random variable is called **absolutely continuous**. Thus, for instance, the continuous uniform distribution, given by

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \le x < 1, \\ 1, & x \ge 1, \end{cases}$$

possesses a derivative at all points except at $x = 0$ and $x = 1$. Therefore, it is an absolutely continuous distribution. All continuous random variables that we will study are absolutely continuous and hence the adjective will be dropped.

Definition (Probability Density Function). For a continuous random variable, X, $f(x) = dF(x)/dx$ is called the **probability density function** (pdf or density function) of X.

The pdf enables us to obtain the CDF by integrating the pdf:

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)\ dt, \qquad -\infty < x < \infty.$$

The analogy with (2.2) is clear, with the sum being replaced by an integral. We can also obtain other probabilities of interest such as

$$\begin{aligned} P(X \in (a, b]) &= P(a < X \le b) \\ &= P(X \le b) - P(X \le a) \\ &= \int_{-\infty}^{b} f_X(t)\ dt - \int_{-\infty}^{a} f_X(t)\ dt \\ &= \int_{a}^{b} f_X(t)\ dt. \end{aligned}$$

The pdf, $f(x)$, satisfies the following properties:

(f1) $f(x) \ge 0$ for all x.

(f2) $\displaystyle\int_{-\infty}^{\infty} f(x)\ dx = 1.$

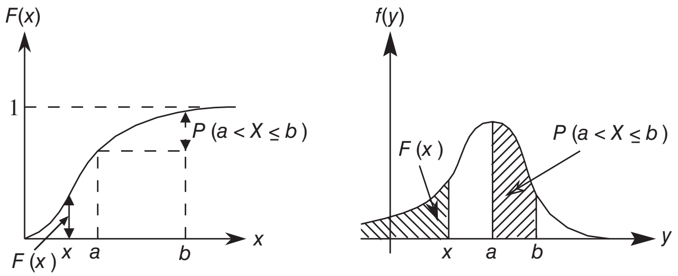

Figure 3.1. Relation between CDF and pdf

It should be noted that, unlike the pmf, the values of the pdf are not probabilities, and thus it is perfectly acceptable if $f(x) > 1$ at a point x.

As is the case for the CDF of a discrete random variable, the CDF of a continuous random variable, $F(x)$, satisfies the following properties:

(F1) $0 \le F(x) \le 1, \quad -\infty < x < \infty.$

(F2) $F(x)$ is a monotone increasing function of x.

(F3) $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to +\infty} F(x) = 1.$

Unlike the CDF of a discrete random variable, the CDF of a continuous random variable does not have any jumps. Therefore, the probability associated with the event $[X = c] = \{s | X(s) = c\}$ is zero:

(F4′) $P(X = c) = P(c \le X \le c) = \int_c^c f_X(y)\, dy = 0.$

This does not imply that the set $\{s | X(s) = c\}$ is empty, but that the probability assigned to this set is zero. As a consequence of the fact that $P(X = c) = 0$, we have

$$\begin{aligned} P(a \le X \le b) &= P(a < X \le b) = P(a \le X < b) \\ &= P(a < X < b) \\ &= \int_a^b f_X(x)\ dx \\ &= F_X(b) - F_X(a). \end{aligned} \tag{3.1}$$

The relation between the functions f and F is illustrated in Figure 3.1. Probabilities are represented by areas under the pdf curve. The total area under the curve is unity.

Example 3.1

The time (measured in years), X, required to complete a software project has a pdf of the form:

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

Since f_X satisfies property (f1), we know $k \geq 0$. In order for f_X to be a pdf, it must also satisfy property (f2); hence

$$\int_0^1 kx(1-x)\ dx = k\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Bigg|_0^1 = 1.$$

Therefore

$$k = 6.$$

Now the probability that the project will be completed in less than four months is given by

$$P(X < \tfrac{4}{12}) = F_X(\tfrac{1}{3}) = \int_0^{1/3} f_X(x)\ dx = \tfrac{7}{27}$$

or about 26 percent chance.

♯

Most random variables we consider will either be discrete (as in Chapter 2) or continuous, but *mixed* random variables do occur sometimes. For example, there may be a nonzero probability, p_0, of initial failure of a component at time 0 due to manufacturing defects. In this case, the time to failure, X, of the component is neither discrete nor a continuous random variable. The CDF of such a modified exponential random variable X with a mass at origin (shown in Figure 3.2) is then

$$F_X(x) = \begin{cases} 0, & x < 0, \\ p_0, & x = 0, \\ p_0 + (1-p_0)(1-e^{-\lambda x}), & x > 0. \end{cases} \tag{3.2}$$

The CDF of a mixed random variable satisfies properties (F1)–(F3) but it does not satisfy property (F4) of Chapter 2 or the property (F4′) above.

The distribution function of a mixed random variable can be written as a linear combination of two distribution functions, denoted by $F^{(d)}(.)$ and $F^{(c)}(.)$, which are discrete and continuous, respectively, so that for every real number x

$$F_X(x) = \alpha_d\ F^{(d)}(x) + \alpha_c\ F^{(c)}(x)$$

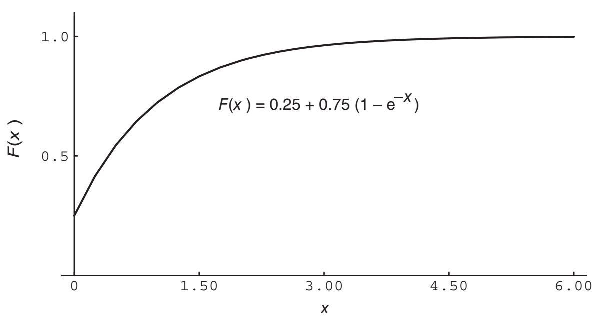

Figure 3.2. CDF of a mixed random variable

where $0 \le \alpha_d,\ \alpha_c \le 1$ and $\alpha_d + \alpha_c = 1$. Thus the mixed distribution (3.2) can be represented in this way if we let $F^{(d)}(x)$ as the unit step function, $F^{(c)}(x) = 1 - e^{-\lambda x}$, $\alpha_d = p_0$, and $\alpha_c = 1 - p_0$. (A unified treatment of discrete, continuous, and mixed random variables can also be given through the use of Riemann–Stieltjes integrals [BREI 1968, RUDI 1964].)

Problems

1. Find the value of the constant k so that

$$f(x) = \begin{cases} kx^2(1 - x^3), & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

   is a proper density function of a continuous random variable.

2. Let X be a continuous random variable denoting the time to failure of a component. Suppose the distribution function of X is $F(x)$. Use this distribution function to express the probability of the following events:

   (a) $9 < X < 90$.
   (b) $X < 90$.
   (c) $X > 90$, given that $X > 9$.

3. Consider a random variable X defined by the CDF:

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{2}\sqrt{x} + \frac{1}{2}(1 - e^{-\sqrt{x}}), & 0 \le x \le 1, \\ \frac{1}{2} + \frac{1}{2}(1 - e^{-\sqrt{x}}), & x > 1. \end{cases}$$

   Show that this function satisfies properties (F1)–(F3) and (F4′). Note that $F_X(x)$ is a continuous function but it does not have a derivative at $x = 1$. (That is, the pdf of X has a discontinuity at $x = 1$.) Plot the CDF and the pdf of X.

4. See Hamming [HAMM 1973]. Consider a normalized floating-point number in base (or radix) β so that the mantissa, X, satisfies the condition $1/\beta \le X < 1$. Experience shows that X has the following **reciprocal density**:

$$f_X(x) = \frac{k}{x}, \quad k > 0.$$

   Determine

   (a) The value of k.
   (b) The distribution function of X.
   (c) The probability that the leading digit of X is i for $1 \le i < \beta$.

3.2 THE EXPONENTIAL DISTRIBUTION

This distribution, sometimes called the **negative exponential distribution**, occurs in applications such as reliability theory and queuing theory. Reasons for its use include its memoryless property (and resulting analytical tractability) and its relation to the (discrete) Poisson distribution. Thus the following random variables will often be modeled as exponential:

1. Time between two successive job arrivals to a file server (often called **interarrival time**).
2. Service time at a server in a queuing network; the server could be a resource such as a CPU, an I/O device, or a communication channel.
3. Time to failure (lifetime) of a component.
4. Time required to repair a component that has malfunctioned.

Note that the assertion "Above distributions are exponential" is not a given fact but an assumption. Experimental verification of this assumption must be sought before relying on the results of the analysis (see Chapter 10 for further elaboration on this topic).

The **exponential distribution function**, shown in Figure 3.3, is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{if } 0 \le x < \infty, \\ 0, & \text{otherwise.} \end{cases} \tag{3.3}$$

If a random variable X possesses CDF given by equation (3.3), we use the notation $X \sim EXP(\lambda)$, for brevity. The pdf of X has the shape shown in Figure 3.4 and is given by

$$f(x) = \begin{cases} \lambda\, e^{-\lambda x}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

While specifying a pdf, usually we state only the nonzero part, and it is understood that the pdf is zero over any unspecified region. Since $\lim_{x\to\infty} F(x) = 1$, it follows that the total area under the exponential pdf is unity. Also

$$\begin{aligned} P(X \ge t) &= \int_t^\infty f(x)\, dx \\ &= e^{-\lambda t} \end{aligned} \tag{3.5}$$

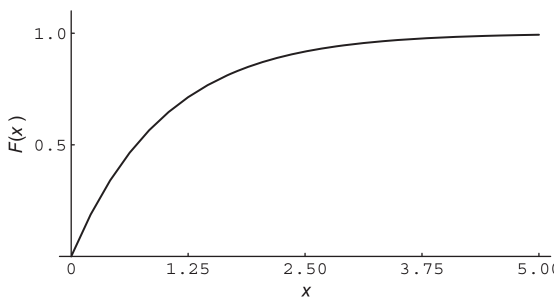

Figure 3.3. The CDF of an exponentially distributed random variable ($\lambda = 1$)

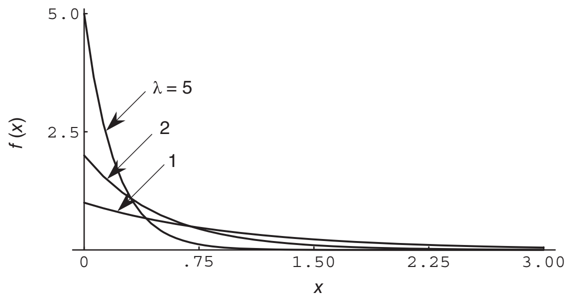

Figure 3.4. Exponential pdf

and

$$\begin{aligned} P(a \le X \le b) &= F(b) - F(a) \\ &= e^{-\lambda a} - e^{-\lambda b}. \end{aligned}$$

Now let us investigate the **memoryless property** of the exponential distribution. Suppose we know that X exceeds some given value t; that is, $X > t$. For example, let X be the lifetime of a component, and suppose we have observed that this component has already been operating for t hours. We may then be interested in the distribution of $Y = X - t$, the remaining (residual) lifetime. Let the conditional probability of $Y \le y$, given that $X > t$, be denoted by $G_Y(y|t)$. Thus, for $y \ge 0$, we have

$$\begin{aligned} G_Y(y|t) &= P(Y \le y|X > t) \\ &= P(X - t \le y|X > t) \\ &= P(X \le y + t|X > t) \\ &= \frac{P(X \le y + t \text{ and } X > t)}{P(X > t)} \\ &\quad \text{(by the definition of conditional probability)} \\ &= \frac{P(t < X \le y + t)}{P(X > t)}. \end{aligned}$$

Thus (see Figure 3.5)

$$\begin{aligned} G_Y(y|t) &= \frac{\int_t^{y+t} f(x)\ dx}{\int_t^{\infty} f(x)\ dx} \\ &= \frac{\int_t^{y+t} \lambda e^{-\lambda x}\ dx}{\int_t^{\infty} \lambda e^{-\lambda x}\ dx} \\ &= \frac{e^{-\lambda t}\ (1 - e^{-\lambda y})}{e^{-\lambda t}} \\ &= 1 - e^{-\lambda y}. \end{aligned}$$

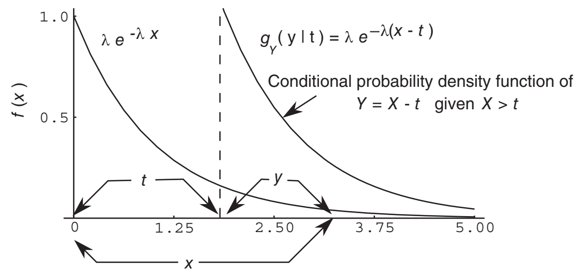

Figure 3.5. Memoryless property of the exponential distribution ($\lambda = 1$)

Thus, $G_Y(y|t)$ is independent of t and is identical to the original exponential distribution of X. The distribution of the remaining life does not depend on how long the component has been operating. The component does not "age" (it is as good as new or it "forgets" how long it has been operating), and its eventual breakdown is the result of some suddenly appearing failure, not of gradual deterioration.

If the interarrival times are exponentially distributed, then the memoryless property implies that the time we must wait for a new arrival is statistically independent of how long we have already spent waiting for it.

If X is a nonnegative continuous random variable with the memoryless property, then we can show that the distribution of X must be exponential:

$$\frac{P(t < X \le y + t)}{P(X > t)} = P(X \le y) = P(0 < X \le y),$$

or

$$F_X(y + t) - F_X(t) = [1 - F_X(t)][F_X(y) - F_X(0)].$$

Since $F_X(0) = 0$, we rearrange this equation to get

$$\frac{F_X(y + t) - F_X(y)}{t} = \frac{F_X(t)[1 - F_X(y)]}{t}.$$

Taking the limit as t approaches zero, we get

$$F'_X(y) = F'_X(0)[1 - F_X(y)],$$

where F'_X denotes the derivative of F_X. Let $R_X(y) = 1 - F_X(y)$; then the preceding equation reduces to

$$R'_X(y) = R'_X(0)R_X(y).$$

The solution to this differential equation is given by

$$R_X(y) = Ke^{R'_X(0)y},$$

where K is a constant of integration and $-R'_X(0) = F'_X(0) = f_X(0)$, the pdf evaluated at 0. Noting that $R_X(0) = 1$, and denoting $f_X(0)$ by the constant λ, we get

$$R_X(y) = e^{-\lambda y}$$

and hence

$$F_X(y) = 1 - e^{-\lambda y}, \quad y > 0.$$

Therefore X must have the exponential distribution.

The exponential distribution can be obtained from the Poisson distribution by considering the interarrival times rather than the number of arrivals.

Example 3.2

Let the discrete random variable N_t denote the number of jobs arriving to a file server in the interval $(0, t]$. Let X be the time of the next arrival. Further assume that N_t is Poisson distributed with parameter λt, so that λ is the arrival rate. Then

$$\begin{aligned} P(X > t) &= P(N_t = 0) \\ &= \frac{e^{-\lambda t}(\lambda t)^0}{0!} \\ &= e^{-\lambda t} \end{aligned}$$

and

$$F_X(t) = 1 - e^{-\lambda t}.$$

Therefore, the time to the next arrival is exponentially distributed. More generally, it can be shown that the interarrival times of Poisson events are exponentially distributed [BHAT 1984, p. 197].

♯

Example 3.3

Consider a Web server with an average rate of requests $\lambda = 0.1$ jobs per second. Assuming that the number of arrivals per unit time is Poisson distributed, the interarrival time, X, is exponentially distributed with parameter λ. The probability that an interval of 10 seconds elapses without requests is then given by

$$\begin{aligned} P(X \geq 10) &= \int_{10}^{\infty} 0.1e^{-0.1t}\ dt = \lim_{t\to\infty}[-e^{-0.1t}] - (-e^{-1}) \\ &= e^{-1} = 0.368. \end{aligned}$$

♯

Problems

1. Jobs arriving to a compute server have been found to require CPU time that can be modeled by an exponential distribution with parameter $1/140\ \text{ms}^{-1}$. The CPU scheduling discipline is quantum-oriented so that a job not completing within a quantum of 100 ms will be routed back to the tail of the queue of waiting jobs. Find the probability that an arriving job is forced to wait for a second quantum. Of the 800 jobs coming in during a day, how many are expected to finish within the first quantum?