Chapter 2
Probability and Random Variables

Philosophy is a game with objectives and no rules.
Mathematics is a game with rules and no objectives.

—Anonymous

Most signals we deal with in practice are random (unpredictable or erratic) and not deterministic. Random signals are encountered in one form or another in every practical communication system. They occur in communication both as information-conveying signal and as unwanted noise signal.

A **random quantity** is one having values which are regulated in some probabilistic way.

Thus, our work with random quantities must begin with the theory of probability, which is the mathematical discipline that deals with the statistical characterization of random signals and random processes. Although the reader is expected to have had at least one course on probability theory and random variables, this chapter provides a cursory review of the basic concepts needed throughout this book. The concepts include probabilities, random variables, statistical averages or mean values, and probability models. A reader already versed in these concepts may skip this chapter.

2.1 Probability Fundamentals

A fundamental concept in the probability theory is the idea of an *experiment*. An experiment (or trial) is the performance of an operation that leads to results called *outcomes*. In other words, an outcome is a result of performing the experiment once. An *event* is one or more outcomes of an experiment. The relationship between outcomes and events is shown in the Venn diagram of Fig. 2.1.

M.N.O. Sadiku and S.M. Musa, *Performance Analysis of Computer Networks*,
DOI 10.1007/978-3-319-01646-7_2, © Springer International Publishing Switzerland 2013

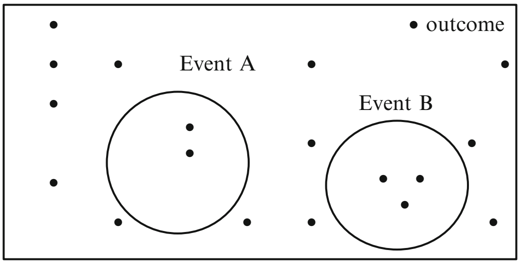


Fig. 2.1 Sample space illustrating the relationship between outcomes (*points*) and events (*circles*)

Thus,

An **experiment** consists of making a measurement or observation.
An **outcome** is a possible result of the experiment.
An **event** is a collection of outcomes.

An experiment is said to be *random* if its outcome cannot be predicted. Thus a random experiment is one that can be repeated a number of times but yields unpredictable outcome at each trial. Examples of random experiments are tossing a coin, rolling a die, observing the number of cars arriving at a toll booth, and keeping track of the number of telephone calls at your home. If we consider the experiment of rolling a die and regard event A as the appearance of the number 4. That event may or may not occur for every experiment.

2.1.1 Simple Probability

We now define the probability of an event. The probability of event A is the number of ways event A can occur divided by the total number of possible outcomes. Suppose we perform n trials of an experiment and we observe that outcomes satisfying event A occur n_A times. We define the probability P(A) of event A occurring as

$$\boxed{P(A) = \lim_{n \to \infty} \frac{n_A}{n}} \tag{2.1}$$

This is known as the *relative frequency* of event A. Two key points should be noted from Eq. (2.1). First, we note that the probability P of an event is always a positive number and that

$$0 \leq P \leq 1 \tag{2.2}$$

where P = 0 when an event is not possible (never occurs) and P = 1 when the event is sure (always occurs). Second, observe that for the probability to have meaning, the number of trials n must be large.

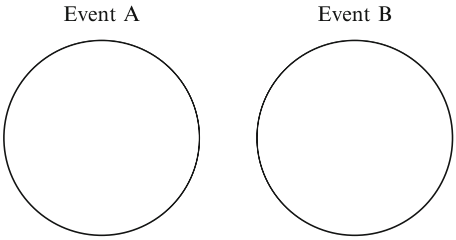


Fig. 2.2 Mutually exclusive or disjoint events

If events A and B are disjoint or mutually exclusive, it follows that the two events cannot occur simultaneously or that the two events have no outcomes in common, as shown in Fig. 2.2.

In this case, the probability that either event A or B occurs is equal to the sum of their probabilities, i.e.

$$\mathrm{P(A\ or\ B) = P(A) + P(B)} \tag{2.3}$$

To prove this, suppose in an experiments with n trials, event A occurs n_A times, while event B occurs n_B times. Then event A or event B occurs $n_A + n_B$ times and

$$P(A \text{ or B}) = \frac{n_A + n_B}{n} = \frac{n_A}{n} + \frac{n_B}{n} = P(A) + P(B) \tag{2.4}$$

This result can be extended to the case when all possible events in an experiment are A, B, C, . . ., Z. If the experiment is performed n times and event A occurs n_A times, event B occurs n_B times, etc. Since some event must occur at each trial,

$$n_A + n_B + n_C + \cdots + n_Z = n$$

Dividing by n and assuming n is very large, we obtain

$$P(A) + P(B) + P(C) + \cdots + P(Z) = 1 \tag{2.5}$$

which indicates that the probabilities of mutually exclusive events must add up to unity. A special case of this is when two events are complimentary, i.e. if event A occurs, B must not occur and vice versa. In this case,

$$P(A) + P(B) = 1 \tag{2.6}$$

or

$$\mathrm{P(A) = 1 - P(B)} \tag{2.7}$$

For example, in tossing a coin, the event of a head appearing is complementary to that of tail appearing. Since the probability of either event is ½, their probabilities add up to 1.

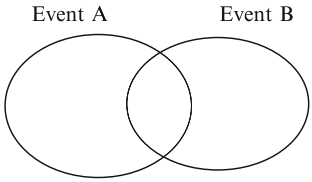


Fig. 2.3 Non-mutually exclusive events

2.1.2 Joint Probability

Next, we consider when events A and B are not mutually exclusive. Two events are non-mutually exclusive if they have one or more outcomes in common, as illustrated in Fig. 2.3.

The probability of the union event A or B (or A + B) is

$$\mathrm{P(A + B) = P(A) + P(B) - P(AB)} \tag{2.8}$$

where P(AB) is called the *joint probability* of events A and B, i.e. the probability of the intersection or joint event AB.

2.1.3 Conditional Probability

Sometimes we are confronted with a situation in which the outcome of one event depends on another event. The dependence of event B on event A is measured by the *conditional probability* P(B|A) given by

$$P(B|A) = \frac{P(AB)}{P(A)} \tag{2.9}$$

where P(AB) is the joint probability of events A and B. The notation B|A stands "B given A." In case events A and B are mutually exclusive, the joint probability P(AB) = 0 so that the conditional probability P(B|A) = 0. Similarly, the conditional probability of A given B is

$$\boxed{P(A|B) = \frac{P(AB)}{P(B)}} \tag{2.10}$$

From Eqs. (2.9) and (2.10), we obtain

$$\mathrm{P(AB) = P(B|A)P(A) = P(A|B)P(B)} \tag{2.11}$$

Eliminating P(AB) gives

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)} \tag{2.12}$$

which is a form of *Bayes' theorem*.

2.1.4 Statistical Independence

Lastly, suppose events A and B do not depend on each other. In this case, events A and B are said to be *statistically independent*. Since B has no influence of A or vice versa,

$$\mathrm{P(A|B) = P(A), \qquad P(B|A) = P(B)} \tag{2.13}$$

From Eqs. (2.11) and (2.13), we obtain

$$\mathrm{P(AB) = P(A)P(B)} \tag{2.14}$$

indicating that the joint probability of statistically independent events is the product of the individual event probabilities. This can be extended to three or more statistically independent events

$$\mathrm{P(ABC\ldots) = P(A)P(B)P(C)\ldots} \tag{2.15}$$

Example 2.1 Three coins are tossed simultaneously. Find: (a) the probability of getting exactly two heads, (b) the probability of getting at least one tail.

Solution

If we denote HTH as a head on the first coin, a tail on the second coin, and a head on the third coin, the $2^3 = 8$ possible outcomes of tossing three coins simultaneously are the following:

$$\mathrm{HHH, HTH, HHT, HTT, THH, TTH, THT, TTT}$$

The problem can be solved in several ways

Method 1: (Intuitive approach)

(a) Let event A correspond to having exactly two heads, then

$$\mathrm{Event\ A = \{HHT, HTH, THH\}}$$

Since we have eight outcomes in total and three of them are in event A, then

$$\mathrm{P(A)} = 3/8 = 0.375$$

Table 2.1 For Example 2.2; number of capacitors with given values and voltage ratings

	Voltage rating			
Capacitance	10 V	50 V	100 V	Total
4 pF	9	11	13	33
12 pF	12	16	8	36
20 pF	10	14	7	31
Total	31	41	28	100

(b) Let B denote having at least one tail,

$$\text{Event B} = \{\text{HTH}, \text{HHT}, \text{HTT}, \text{THH}, \text{TTH}, \text{THT}, \text{TTT}\}$$

Hence,

$$P(B) = 7/8 = 0.875$$

Method 2: (Analytic approach) Since the outcome of each separate coin is statistically independent, with head and tail equally likely,

$$P(H) = P(T) = ½$$

(a) Event consists of mutually exclusive outcomes. Hence,

$$P(A) = P(\text{HHT}, \text{HTH}, \text{THH}) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$$
$$= \frac{3}{8} = 0.375$$

(b) Similarly,

$$P(B) = (\text{HTH}, \text{HHT}, \text{HTT}, \text{THH}, \text{TTH}, \text{THT}, \text{TTT})$$
$$= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + \text{in seven places} = \frac{7}{8} = 0.875$$

Example 2.2 In a lab, there are 100 capacitors of three values and three voltage ratings as shown in Table 2.1. Let event A be drawing 12 pF capacitor and event B be drawing a 50 V capacitor. Determine: (a) P(A) and P(B), (b) P(AB), (c) P(A|B), (d) P(B|A).

Solution

(a) From Table 2.1,

$$P(A) = P(12\ \text{pF}) = 36/100 = 0.36$$

and

$$P(B) = P(50\ \text{V}) = 41/100 = 0.41$$

(b) From the table,

$$\mathrm{P(AB) = P(12\ pF, 50\ V) = 16/100 = 0.16}$$

(c) From the table

$$\mathrm{P(A|B) = P(12\ pF|50\ V) = 16/41 = 0.3902}$$

Check: From Eq. (2.10),

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{16/100}{41/100} = 0.3902$$

(d) From the table,

$$\mathrm{P(B|A) = P(50\ V|12\ pF) = 16/36 = 0.4444}$$

Check: From Eq. (2.9),

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{16/100}{36/100} = 0.4444$$

2.2 Random Variables

Random variables are used in probability theory for at least two reasons [1, 2]. First, the way we have defined probabilities earlier in terms of events is awkward. We cannot use that approach in describing sets of objects such as cars, apples, and houses. It is preferable to have numerical values for all outcomes. Second, mathematicians and communication engineers in particular deal with random processes that generate numerical outcomes. Such processes are handled using random variables.

The term “random variable” is a misnomer; a random variable is neither random nor a variable. Rather, it is a function or rule that produces numbers from the outcome of a random experiment. In other words, for every possible outcome of an experiment, a real number is assigned to the outcome. This outcome becomes the value of the random variable. We usually represent a random variable by an uppercase letters such as X, Y, and Z, while the value of a random variable (which is fixed) is represented by a lowercase letter such as x, y, and z. Thus, X is a function that maps elements of the sample space S to the real line $-\infty \leq x \leq \infty$, as illustrated in Fig. 2.4.

A **random variable** X is a single-valued real function that assigns a real value X(x) to every point x in the sample space.

Random variable X may be either discrete or continuous. X is said to be discrete random variable if it can take only discrete values. It is said to be continuous if it

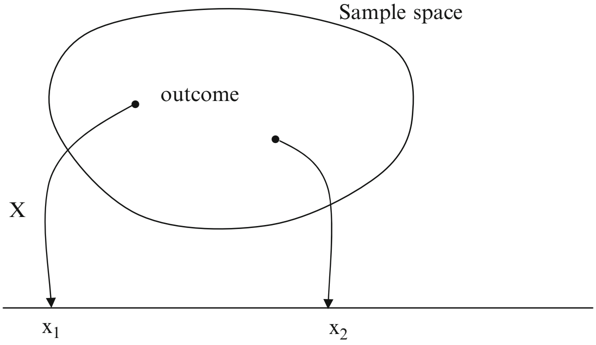


Fig. 2.4 Random variable X maps elements of the sample space to the real line

takes continuous values. An example of a discrete random variable is the outcome of rolling a die. An example of continuous random variable is one that is Gaussian distributed, to be discussed later.

2.2.1 *Cumulative Distribution Function*

Whether X is discrete or continuous, we need a probabilistic description of it in order to work with it. All random variables (discrete and continuous) have a cumulative distribution function (CDF).

> The **cumulative distribution function** (CDF) is a function given by the probability that the random variable X is less than or equal to x, for every value x.

Let us denote the probability of the event $X \leq x$, where x is given, as $P(X \leq x)$. The *cumulative distribution function* (CDF) of X is given by

$$F_X(x) = P(X \leq x), \quad -\infty \leq x \leq \infty \tag{2.16}$$

for a continuous random variable X. Note that $F_X(x)$ does not depend on the random variable X, but on the assigned value of X. $F_X(x)$ has the following five properties:

1. $$F_X(-\infty) = 0 \tag{2.17a}$$
2. $$F_X(\infty) = 1 \tag{2.17b}$$
3. $$0 \leq F_X(x) \leq 1 \tag{2.17c}$$
4. $$F_X(x_1) \leq F_X(x_2), \quad \text{if } x_1 < x_2 \tag{2.17d}$$
5. $$P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1) \tag{2.17e}$$

The first and second properties show that the $F_X(-\infty)$ includes no possible events and $F_X(\infty)$ includes all possible events. The third property follows from the fact that $F_X(x)$ is a probability. The fourth property indicates that $F_X(x)$ is a nondecreasing function. And the last property is easy to prove since

$$P(X \leq x_2) = P\big(X \leq x_1\big) + P\big(x_1 < X \leq x_2\big)$$

or

$$P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F_X(x_2) - F_X(x_1) \tag{2.18}$$

If X is discrete, then

$$F_X(x) = \sum_{i=0}^{N} P(x_i) \tag{2.19}$$

where $P(x_i) = P(X = x_i)$ is the probability of obtaining event x_i, and N is the largest integer such that $x_N \leq x$ and $N \leq M$, and M is the total number of points in the discrete distribution. It is assumed that $x_1 < x_2 < x_3 < \cdots < x_M$.

2.2.2 Probability Density Function

It is sometimes convenient to use the derivative of $F_X(x)$, which is given by

$$f_X(x) = \frac{dF_x(x)}{dx} \tag{2.20a}$$

or

$$F_X(x) = \int_{-\infty}^{x} f_X(x)dx \tag{2.20b}$$

where $f_X(x)$ is known as the *probability density function* (PDF). Note that $f_X(x)$ has the following properties:

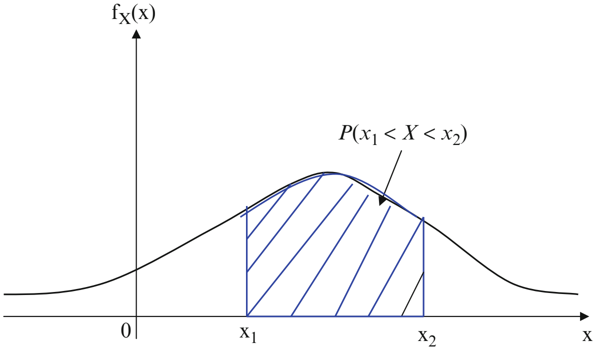
1. $$f_X(x) \geq 0 \tag{2.21a}$$

2. $$\int_{-\infty}^{\infty} f_X(x)dx = 1 \tag{2.21b}$$

3. $$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f_X(x)dx \tag{2.21c}$$

Properties 1 and 2 follows from the fact that $F_X(-\infty) = 0$ and $F_X(\infty) = 1$

Fig. 2.5 A typical PDF



respectively. As mentioned earlier, since $F_X(x)$ must be nondecreasing, its derivative $f_X(x)$ must always be nonnegative, as stated by Property 1. Property 3 is easy to prove. From Eq. (2.18),

$$\begin{aligned} P(x_1 < X \leq x_2) &= F_X(x_2) - F_X(x_1) \\ &= \int_{-\infty}^{x_2} f_X(x)dx - \int_{-\infty}^{x_1} f_X(x)dx = \int_{x_1}^{x_2} f_X(x)dx \end{aligned} \tag{2.22}$$

which is typically illustrated in Fig. 2.5 for a continuous random variable.

For discrete X,

$$f_X(x) = \sum_{i=1}^{M} P(x_i)\delta(x - x_i) \tag{2.23}$$

where M is the total number of discrete events, $P(x_i) = P(x = x_i)$, and $\delta(x)$ is the impulse function. Thus,

> The **probability density function** (PDF) of a continuous (or discrete) random variable is a function which can be integrated (or summed) to obtain the probability that the random variable takes a value in a given interval.

2.2.3 Joint Distribution

We have focused on cases when a single random variable is involved. Sometimes several random variables are required to describe the outcome of an experiment. Here we consider situations involving two random variables X and Y; this may be extended to any number of random variables. The *joint cumulative distribution function* (joint cdf) of X and Y is the function

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) \tag{2.24}$$

where $-\infty < x < \infty$, $-\infty < y < \infty$. If $F_{XY}(x,y)$ is continuous, the *joint probability density function* (joint PDF) of X and Y is given by

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} \tag{2.25}$$

where $f_{XY}(x,y) \geq 0$. Just as we did for a single variable, the probability of event $x_1 < X \leq x_2$ and $y_1 < Y \leq y_2$ is

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F_{XY}(x,y) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{XY}(x,y)dxdy \tag{2.26}$$

From this, we obtain the case where the entire sample space is included as

$$F_{XY}(\infty,\infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1 \tag{2.27}$$

since the total probability must be unity.

Given the joint CDF of X and Y, we can obtain the individual CDFs of the random variables X and Y. For X,

$$F_X(x) = P(X \leq x, -\infty < Y < \infty) = F_{XY}(x,\infty) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(x,y)dxdy \tag{2.28}$$

and for Y,

$$\begin{aligned} F_Y(y) &= P(-\infty < x < \infty, y \leq Y) = F_{XY}(\infty,y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{y} f_{XY}(x,y)dxdy \end{aligned} \tag{2.29}$$

$F_X(x)$ and $F_Y(y)$ are known as the *marginal cumulative distribution functions* (marginal CDFs).

Similarly, the individual PDFs of the random variables X and Y can be obtained from their joint PDF. For X,

$$f_X(x) = \frac{dF_X(x)}{dx} = \int_{-\infty}^{\infty} f_{XY}(x,y)dy \tag{2.30}$$

and for Y,

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \int_{-\infty}^{\infty} f_{XY}(x,y)dx \tag{2.31}$$

$f_X(x)$ and $f_Y(y)$ are known as the *marginal probability density functions* (marginal PDFs).

As mentioned earlier, two random variables are independent if the values taken by one do not affect the other. As a result,

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y) \tag{2.32}$$

or

$$F_{XY}(x, y) = F_X(x)F_Y(y) \tag{2.33}$$

This condition is equivalent to

$$f_{XY}(x, y) = f_X(x)f_Y(y) \tag{2.34}$$

Thus, two random variables are independent when their joint distribution (or density) is the product of their individual marginal distributions (or densities).

Finally, we may extend the concept of conditional probabilities to the case of continuous random variables. The conditional probability density function (conditional PDF) of X given the event $Y = y$ is

$$f_X(x|Y = y) = \frac{f_{XY}(x, y)}{f_Y(y)} \tag{2.35}$$

where $f_Y(y)$ is the marginal PDF of Y. Note that $f_X(x|Y = y)$ is a function of x with y fixed. Similarly, the conditional PFD of Y given $X = x$ is

$$f_Y(y|X = x) = \frac{f_{XY}(x, y)}{f_X(x)} \tag{2.36}$$

where $f_X(x)$ is the marginal PDF of X. By combining Eqs. (2.34) and (2.36), we get

$$f_Y(y|X = x) = \frac{f_X(x|Y = y)f_Y(y)}{f_X(x)} \tag{2.37}$$

which is Bayes' theorem for continuous random variables. If X and Y are independent, combining Eqs. (2.34)–(2.36) gives

$$f_X(x|Y = y) = f_X(x) \tag{2.38a}$$

$$f_Y(y|X = x) = f_Y(y) \tag{2.38b}$$

indicating that one random variable has no effect on the other.

Example 2.3 An analog-to-digital converter is an eight-level quantizer with the output of 0, 1, 2, 3, 4, 5, 6, 7. Each level has the probability given by

$$P(X = x) = 1/8, \quad x = 0, 1, 2, \ldots 7$$

(a) Sketch $F_X(x)$ and $f_X(x)$. (b) Find $P(X \le 1)$, $P(X > 3)$, (c) Determine $P(2 \le X \le 5)$.

Solution

(a) The random variable is discrete. Since the values of x are limited to $0 \le x \le 7$,

$$F_X(-1) = P(X < -1) = 0$$

$$F_X(0) = P(X \le 0) = 1/8$$

$$F_X(1) = P(X \le 1) = P(X = 0) + P(X = 1) = 2/8$$

$$F_X(2) = P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2) = 3/8$$

Thus, in general

$$F_X(i) = \begin{cases} (i+1)/8, & 2 \le i \le 7 \\ 1, & i > 7 \end{cases} \tag{2.3.1}$$

The distribution function is sketched in Fig. 2.6a. Its derivative produces the PDF, which is given by

$$f_X(x) = \sum_{i=0}^{7} \delta(x - i)/8 \tag{2.3.2}$$

and sketched in Fig. 2.6b.

(b) We already found $P(X \le 1)$ as

$$P(X \le 1) = P(X = 0) + P(X = 1) = 1/4$$

$$P(X > 3) = 1 - P(X \le 3) = 1 - F_X(3)$$

But

$$F_X(3) = P(X \le 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 4/8$$

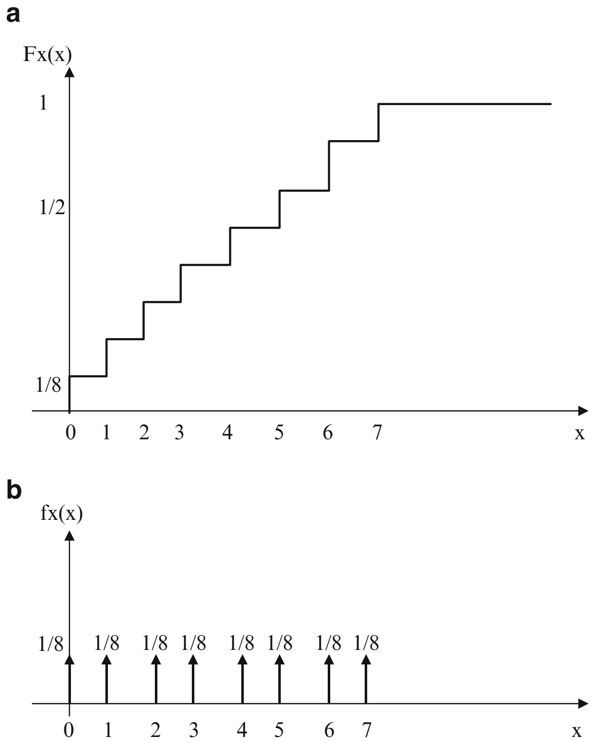
We can also obtain this from Eq. (2.3.1). Hence,

$$P(X > 3) = 1 - 4/8 = \frac{1}{2}.$$

(c) For $P(2 \le X \le 5)$, using Eq. (2.3.1)

$$P(2 \le X \le 5) = F_X(5) - F_X(2) = 5/8 - 2/8 = 3/8.$$

Fig. 2.6 For Example 2.3: (**a**) distribution function of X, (**b**) probability density function of X



Example 2.4 The CDF of a random variable is given by

$$F_X(x) = \begin{cases} 0, & x < 1 \\ \dfrac{x-1}{8}, & 1 \le x < 9 \\ 1, & x \ge 9 \end{cases}$$

(a) Sketch $F_X(x)$ and $f_X(x)$. (b) Find $P(X \le 4)$ and $P(2 < X \le 7)$.

Solution

(a) In this case, X is a continuous random variable. $F_X(x)$ is sketched in Fig. 2.7a. We obtain the PDF of X by taking the derivative of $F_X(x)$, i.e.

$$f_X(x) = \begin{cases} 0, & x < 1 \\ \dfrac{1}{8}, & 1 \le x < 9 \\ 0, & x \ge 9 \end{cases}$$

which is sketched in Fig. 2.7b. Notice that $f_X(x)$ satisfies the requirement of a probability because the area under the curve in Fig. 2.7b is unity. A random number having a PDF such as shown in Fig. 2.7b is said to be *uniformly distributed* because $f_X(x)$ is constant within 1 and 9.

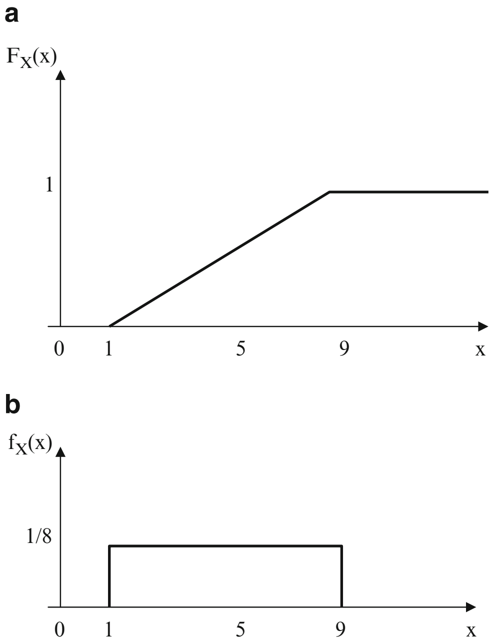


Fig. 2.7 For Example 2.4: (**a**) CDF, (**b**) PDF

(b) $P(X \leq 4) = F_X(4) = 3/8$

$$P(2 < x \leq 7) = F_X(7) - F_X(2) = 6/8 - 1/8 = 5/8$$

Example 2.5 Given that two random variables have the joint PDF

$$f_{XY}(x,y) = \begin{cases} ke^{-(x+2y)}, & 0 \leq x \leq \infty, 0 \leq y \leq \infty \\ 0, & \text{otherwise} \end{cases}$$

(a) Evaluate k such that the PDF is a valid one. (b) Determine $F_{XY}(x,y)$. (c) Are X and Y independent random variables? (d) Find the probabilities that $X \leq 1$ and $Y \leq 2$. (e) Find the probability that $X \leq 2$ and $Y > 1$.

Solution

(a) In order for the given PDF to be valid, Eq. (2.27) must be satisfied, i.e.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1$$

so that

$$1 = \int_0^{\infty}\int_0^{\infty} ke^{-(x+2y)}dxdy = k\int_0^{\infty} e^{-x}dx\int_0^{\infty} e^{-2y}dy = k(1)\left(\frac{1}{2}\right)$$

Hence, k = 2.

(b) $$F_{XY}(x,y) = \int_0^x\int_0^y 2e^{-(x+2y)}dxdy = 2\int_0^x e^{-x}dx\int_0^y e^{-2y}dy = (e^{-x}-1)\left(e^{-2y}-1\right)$$
$$= F_X(x)F_Y(y)$$

(c) Since the joint CDF factors into individual CDFs, we conclude that the random variables are independent.

(d) $$P(X \le 1, Y \le 2) = \int_{x=0}^{1}\int_{y=0}^{2} f_{XY}(x,y)dxdy$$
$$= 2\int_0^1 e^{-x}dx\int_0^2 e^{-2y}dy = \left(1-e^{-1}\right)\left(1-e^{-4}\right) = 0.6205$$

(e) $$P(X \le 2, Y > 1) = \int_{x=0}^{2}\int_{y=1}^{\infty} f_{XY}(x,y)dxdy$$
$$= 2\int_0^2 e^{-x}dx\int_1^{\infty} e - 2ydy = \left(e^{-2}-1\right)\left(e^{-2}\right) = 0.117$$

2.3 Operations on Random Variables

There are several operations that can be performed on random variables. These include the expected value, moments, variance, covariance, correlation, and transformation of the random variables. The operations are very important in our study of computer communications systems. We will consider some of them in this section, while others will be covered in later sections. We begin with the mean or average values of a random variable.

2.3.1 Expectations and Moments

Let X be a discrete random variable which takes on M values x_1, x_2, $x_{.3}$, $\cdots$, x_M that respectively occur n_1, n_2, $n_{.3}$, $\cdots$, n_M in n trials, where n is very large. The statistical average (mean or expectation) of X is given by

$$\overline{X} = \frac{n_1x_1 + n_2x_2 + n_3x_3 + \cdots + n_Mx_M}{n} = \sum_{i=1}^{M} x_i \frac{n_i}{n} \tag{2.39}$$

But by the relative-frequency definition of probability in Eq. (2.1), $n_i/n = P(x_i)$. Hence, the mean or expected value of the discrete random variable X is

$$\boxed{\overline{X} = E[X] = \sum_{i=0}^{\infty} x_i P(x_i)} \tag{2.40}$$

where E stands for the expectation operator.

If X is a continuous random variable, we apply a similar argument. Rather than doing that, we can replace the summation in Eq. (2.40) with integration and obtain

$$\boxed{\overline{X} = E[X] = \int_{-\infty}^{\infty} x f_X(x) dx} \tag{2.41}$$

where $f_X(x)$ is the PDF of X.

In addition to the expected value of X, we are also interested in the expected value of functions of X. In general, the expected value of a function g(X) of the random variable X is given by

$$\overline{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \tag{2.42}$$

for continuous random variable X. If X is discrete, we replace the integration with summation and obtain

$$\overline{g(X)} = E[g(X)] = \sum_{i=1}^{M} g(x_i) P(x_i) \tag{2.43}$$

Consider the special case when $g(x) = X^n$. Equation (2.42) becomes

$$\overline{X^n} = E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx \tag{2.44}$$

$E(X^n)$ is known as the *nth moment* of the random variable X. When n = 1, we have the first moment $\overline{X}$ as in Eq. (2.42). When n = 2, we have the second moment $\overline{X^2}$ and so on.

2.3.2 Variance

The moments defined in Eq. (2.44) may be regarded as moments about the origin, We may also define central moments, which are moments about the mean value $m_X = E(X)$ of X. If X is a continuous random variable,

$$E[(X - m_X)^n\} = \int_{-\infty}^{\infty} (x - m_X)^n f_X(x)dx \tag{2.45}$$

It is evident that the central moment is zero when n = 1. When n = 2, the second central moment is known as the *variance* ${\sigma_X}^2$ of X, i.e.

$$\mathrm{Var}(X) = \sigma_X^2 = E\left[(X - m_X)^2\right\} = \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x)dx \tag{2.46}$$

If X is discrete,

$$\mathrm{Var}(X) = \sigma_X^2 = E\left[(X - m_x)^2\right] = \sum_{i=0}^{\infty} (x_i - m_X)^2 P(x_i) \tag{2.47}$$

The square root of the variance (i.e. σ_X) is called the *standard deviation* of X. By expansion,

$$\begin{aligned}\sigma_X^2 &= E\left[(X - m_X)^2\right] = E\left[X^2 - 2m_X X + m_X^2\right] = E\left[X^2\right] - 2m_X E[X] + m_X^2 \\ &= E\left[X^2\right] - m_X^2\end{aligned} \tag{2.48}$$

or

$$\boxed{\sigma_X^2 = E\left[X^2\right] - m_X^2} \tag{2.49}$$

Note that from Eq. (2.48) that if the mean $m_X = 0$, the variance is equal to the second moment $E[X^2]$.

2.3.3 Multivariate Expectations

We can extend what we have discussed so far for one random variable to two or more random variables. If g(X,Y) is a function of random variables X and Y, its expected value is

$$\overline{g(X,Y)} = E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y)dx\, dy \tag{2.50}$$

Consider a special case in which g(X,Y) = X + Y, where X and Y need not be independent, then

$$\overline{X+Y} = \overline{X} + \overline{Y} = m_X + m_Y \tag{2.51}$$

indicating the mean of the sum of two random variables is equal to the sum of their individual means. This may be extended to any number of random variables.

Next, consider the case in which g(X,Y) = XY, then

$$\overline{XY} = E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{XY}(x)dx\,dy \tag{2.52}$$

If X and Y are independent,

$$\overline{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dx\,dy = \int_{-\infty}^{\infty} xf_X(x)dx \int_{-\infty}^{\infty} yf_Y(y)dy = m_X m_Y \tag{2.53}$$

implying that the mean of the product of two independent random variables is equal to the product of their individual means.

2.3.4 Covariance and Correlation

If we let $g(X,Y) = X^nY^k$, the generalized moments are defined as

$$E\left[X^nY^k\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^ny^kf_{XY}(x)dx\,dy \tag{2.54}$$

We notice that Eq. (2.50) is a special case of Eq. (2.54). The joint moments in Eqs. (2.52) and (2.54) are about the origin. The generalized central moments are defined by

$$E\left[(X-m_X)^n(Y-m_Y)^k\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-m_X)^n(y-m_Y)^kf_{XY}(x)dx\,dy \tag{2.55}$$

The sum of n and k is the order of the moment. Of particular importance is the second central moment (when n = k = 1) and it is called *covariance* of X and Y, i.e.

$$\mathrm{Cov(X, Y)} = E[(X-m_X)(Y-m_Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-m_X)(y-m_Y)f_{XY}(x)dx\,dy$$

or

Table 2.2 For Example 2.6

No. of failures	0	1	2	3	4	5
Probability	0.2	0.33	0.25	0.15	0.05	0.02

$$\boxed{\mathrm{Cov}(X,Y) = E(XY) - m_X m_Y} \tag{2.56}$$

Their *correlation coefficient* ρ_{XY} is given by

$$\boxed{\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}} \tag{2.57}$$

where $-1 \le \rho_{XY} \le 1$. Both covariance and correlation coefficient serve as measures of the interdependence of X and Y. $\rho_{XY} = 1$ when Y = X and $\rho_{XY} = -1$ when Y = −X. Two random variables X and Y are said to be *uncorrelated* if

$$\mathrm{Cov(X, Y)} = 0 \rightarrow \mathrm{E[XY] = E[X]E[Y]} \tag{2.58}$$

and they are *orthogonal* if

$$\mathrm{E[XY]} = 0 \tag{2.59}$$

If X and Y are independent, we can readily show that Cov(X,Y) = 0 = ρ_{XY}. This indicates that when two random variables are independent, they are also uncorrelated.

Example 2.6 A complex communication system is checked on regular basis. The number of failures of the system in a month of operation has the probability distribution given in Table 2.2. (a) Find the average number and variance of failures in a month. (b) If X denotes the number of failures, determine mean and variance of Y = X + 1.

Solution

(a) Using Eq. (2.40)

$$\begin{aligned}\overline{X} &= m_X = \sum_{i=1}^{M} x_i P(x_i) \\ &= 0(0.2) + 1(0.33) + 2(0.25) + 3(0.15) + 4(0.05) + 5(0.02) \\ &= 1.58\end{aligned}$$

To get the variance, we need the second moment.

$$\begin{aligned}\overline{X^2} &= E(X^2) = \sum_{i=1}^{M} x_i^2 P(x_i) \\ &= 0^2(0.2) + 1^2(0.33) + 2^2(0.25) + 3^2(0.15) + 4^2(0.05) + 5^2(0.02) \\ &= 3.98\end{aligned}$$

$$\mathrm{Var}(X) = \sigma_X^2 = E\left[X^2\right] - m_X^2 = 3.98 - 1.58^2 = 1.4836$$

(b) If Y = X + 1, then

$$\overline{Y} = m_Y = \sum_{i=1}^{M}(x_i + 1)P(x_i)$$
$$= 1(0.2) + 2(0.33) + 3(0.25) + 4(0.15) + 5(0.05) + 6(0.02)$$
$$= 2.58$$

Similarly,

$$\overline{Y^2} = E(Y^2) = \sum_{i=1}^{M}(x_i + 1)^2 P(x_i)$$
$$= 1^2(0.2) + 2^2(0.33) + 3^2(0.25) + 4^2(0.15) + 5^2(0.05) + 6^2(0.02)$$
$$= 8.14$$

$$\text{Var}(Y) = \sigma_y^2 = E[Y^2] - m_Y^2 = 8.14 - 2.58^2 = 1.4836$$

which is the same as Var(X). This should be expected because adding a constant value of 1 to X does not change its randomness.

Example 2.7 Given a continuous random variable X with PDF

$$f_X(x) = 2e^{-2x}u(x)$$

(a) Determine E(X) and E(X^2). (b) Assuming that Y = 3X + 1, calculate E(Y) and Var(Y).

Solution

(a) Using Eq. (2.41),

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^{\infty} x(2e^{-2x}) dx$$
$$= 2\left[\frac{e^{-2x}}{4}(-2x - 1)\right]_0^{\infty} = \frac{1}{2}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^{\infty} x^2 (2e^{-2x}) dx$$
$$= 2\left[\frac{e^{-2x}}{-8}(4x^2 + 4x + 2)\right]_0^{\infty} = \frac{1}{2}$$

$$\text{Var(X)} = E(X^2) - [E(X)]^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

(b) Rather than carrying out a similar complex integration, we can use common sense or intuitive argument to obtain E(Y) and E(Y^2). Since Y is linearly dependent on X and the mean value of 1 is 1,

$$\mathrm{E(Y) = E(3X + 1) = 3E(X) + E(1) = 3/2 + 1 = 5/2.}$$

Since the 1 in Y = 3X + 1 is constant, it does not affect the Var(Y). And because a square factor is involved in the calculation of variance,

$$\mathrm{Var(Y) = 3^2 Var(X) = 9/4.}$$

We would have got the same thing if we have carried the integration in Eq. (2.45). To be sure this is the case,

$$E\left(Y^2\right) = \int_{-\infty}^{\infty} (3x+1)^2 f_X(x)dx = \int_{-\infty}^{\infty} \left(9x^2 + 6x + 1\right) f_X(x)dx$$
$$= 9E\left(X^2\right) + 6E\left(X\right) + E\left(1\right) = \frac{9}{2} + \frac{6}{2} + 1 = \frac{17}{2}$$

$$\mathrm{Var}(Y) = E\left(Y^2\right) - E^2(Y) = \frac{17}{2} - \frac{25}{4} = \frac{9}{4}$$

confirming our intuitive approach.

Example 2.8 X and Y are two random variables with joint PDF given by

$$f_{XY}(x,y) = \begin{cases} x+y, & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

(a) Find E(X + Y) and E(XY). (b) Compute Cov(X,Y) and ρ_{XY}. (c) Determine whether X and Y are uncorrelated and/or orthogonal.

Solution

(a)

$$\overline{X+Y} = E[X+Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y) f_{XY}(x) dxdy = \int_0^1\int_0^1 (x+y)(x+y)dxdy$$
$$= \int_0^1\int_0^1 \left(x^2 + 2xy + y^2\right)dxdy = \int_0^1 \left[\frac{x^3}{3} + x^2 y + xy^2\right]_{x=0}^{x=1} dy = \int_0^1 \left(\frac{1}{3} + y + y^2\right) dy$$
$$= \left[\frac{1}{3}y + \frac{y^2}{2} + \frac{y^3}{3}\right]_0^1 = \frac{7}{6}$$

An indirect way of obtaining this result is using Eq. (2.51) but that will require that we first find the marginal PDFs $f_X(x)$ and $f_Y(y)$.
Similarly,

$$\overline{XY} = E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{XY}(x)dxdy = \int_0^1\int_0^1 xy(x+y)dxdy$$
$$= \int_0^1\int_0^1 \left(x^2y + xy^2\right)dxdy = \int_0^1 \left[\frac{x^3}{3}y + \frac{x^2}{2}y^2\right]_{x=0}^{x=1} dy = \int_0^1 \left(\frac{1}{3}y + \frac{1}{2}y^2\right)dy$$
$$= \left[\frac{y^2}{6} + \frac{y^3}{6}\right]_0^1 = \frac{1}{3}$$

(b) To find Cov(X,Y), we need the marginal PDFs.

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \begin{cases} \int_0^1 (x+y)dy = \left[xy + \frac{y^2}{2}\right]_0^1 = x + \frac{1}{2} \\ 0, \quad \text{otherwise} \end{cases}$$

$$m_X = \int_0^1 xf_X(x)dx = \int_0^1 x\left(x + \frac{1}{2}\right)dx = \left[\frac{x^3}{3} + \frac{x^2}{4}\right]_0^1 = \frac{7}{12}$$

Due to the symmetry of the joint PDF, $m_Y = 7/12$.

$$E\left[X^2\right] = \int_0^1 x^2\left(x + \frac{1}{2}\right)dx = \left[\frac{x^4}{4} + \frac{x^6}{6}\right]_0^1 = \frac{5}{12}$$

$$\sigma_X^2 = E\left[X^2\right] - m_X^2 = \frac{5}{12} - \frac{49}{144} = \frac{11}{144}$$

$$\text{Cov}(X,Y) = E(XY) - m_X m_Y = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}$$

Similarly, $\sigma_Y^2 = \frac{11}{144}$. Thus,

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\frac{-1}{144}}{\frac{11}{144}} = -\frac{1}{11}$$

(c) Since $E[XY] = \frac{1}{3} \neq m_X m_Y$, X and Y are correlated. Also, since $E[XY] \neq 0$, they are not orthogonal.

2.4 Discrete Probability Models

Based on experience and usage, several probability distributions have been developed by engineers and scientists as models of physical phenomena. These distributions often arise in communication problems and deserve special attention. It is needless to say that each of these distributions satisfies the axioms of probability covered in Sect. 2.1. In this section, we discuss four discrete probability distributions; continuous probability distributions will be covered in the next section. In fact, some of these distributions have already been considered earlier in the chapter. In this and the next section, we will briefly consider their CDF, PDF, and their parameters such as mean and variance [3–5].

2.4.1 Bernoulli Distribution

A Bernoulli trial is an experiment that has two possible outcomes. Examples are tossing a coin with the two outcomes (heads and tails) and the output of half-wave rectifier which is 0 or 1. Let us denote the outcome of ith trial as 0 (failure) or 1 (success) and let X be a Bernoulli random variable with $P(X = 1) = p$ and $P(X = 0) = 1 - p$. Then the probability mass function (PMF) of X is given by

$$P(x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.60}$$

which is illustrated in Fig. 2.8.

The parameters of the Bernoulli distribution are easily obtained as

$$E[X] = p \tag{2.61a}$$

$$E\left[X^2\right] = p \tag{2.61b}$$

$$Var(X) = p(1 - p) \tag{2.61c}$$

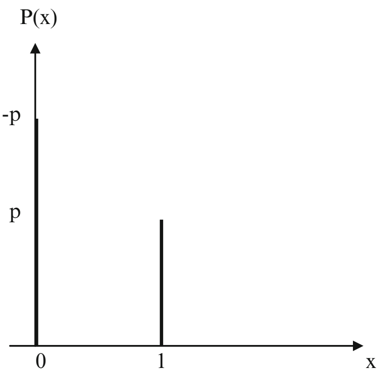


Fig. 2.8 Probability mass function of the Bernoulli distribution

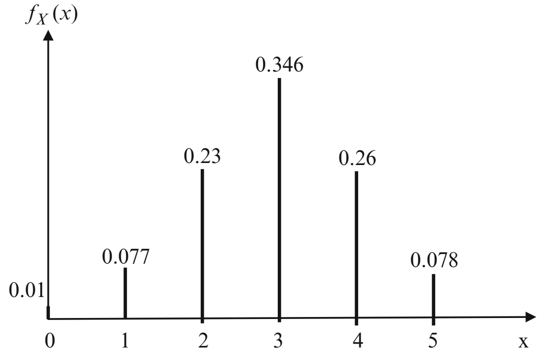


Fig. 2.9 PDF for binomial distribution with n = 5 and p = 0.6

2.4.2 Binomial Distribution

This is an extension of Bernoulli distribution. A random variable follows a Binomial distribution when: (1) n Bernoulli trials are involved, (2) the n trials are independent of each other, and (3) the probabilities of the outcome remain constant as p for success and q = 1 − p for failure. The random variable X for Binomial distribution represents the number of successes in n Bernoulli trials.

In order to find the probability of k successes in n trials, we first define different ways of combining k out of n things, which is

$$^nC_k = \binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{2.62}$$

Note that $\binom{n}{k} = \binom{n}{n-k}$. Hence, the probability of having k successes in n trials is

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{2.63}$$

since there are k successes each with probability p and n − k failures each with probability q = 1 − p and all the trials are independent of each other. If we let x = k, where k = 0, 1, 2, . . ., n, the PDF of the Binomial random variable X is

$$f_X(x) = \sum_{k=0}^{n} P(k)\delta(x-k) \tag{2.64}$$

which is illustrated in Fig. 2.9 for n = 5 and p = 0.6.

From $f_X(x)$, we can obtain the mean and variance for X as

$$\mathrm{E(X) = np} \tag{2.65a}$$

$$\mathrm{Var(X) = npq = np(1-p)} \tag{2.65b}$$

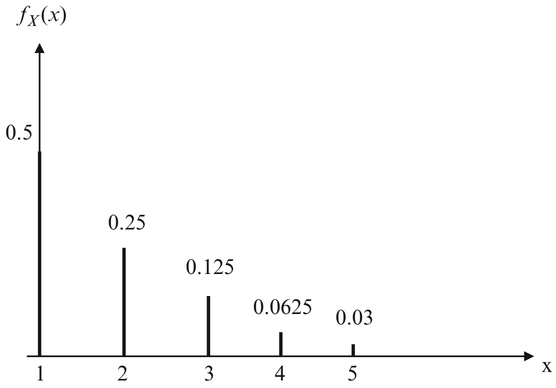


Fig. 2.10 PDF of a geometric distribution with p = 0.5 and n = 5

2.4.3 Geometric Distribution

The geometric distribution is related to Bernoulli trials. A geometric random variable represents the number of Bernoulli trials required to achieve the first success. Thus, a random variable X has a geometric distribution if it takes the values of 1, 2, 3, . . . with probability

$$P(k) = pq^{k-1}, \quad k = 1, 2, 3, \ldots \tag{2.66}$$

where p = probability of success ($0 < p < 1$) and $q = 1 - p$ = probability of failure. This forms a geometric sequence so that

$$\sum_{k=1}^{\infty} pq^{k-1} = \frac{p}{1-q} = 1 \tag{2.67}$$

Figure 2.10 shows the PDF of the geometric random variable for p = 0.5 and x = k = 1, 2, . . . 5.

The mean and variance of the geometric distribution are

$$E(X) = \frac{1}{p} \tag{2.68a}$$

$$\mathrm{Var(X)} = \frac{\mathrm{q}}{\mathrm{p}^2} \tag{2.68b}$$

The geometric distribution is somehow related to binomial distribution. They are both based on independent Bernoulli trials with equal probability of success p. However, a geometric random variable is the number of trials required to achieve the first success, whereas a binomial random variable is the number of successes in n trials.

2.4.4 Poisson Distribution

The Poisson distribution is perhaps the most important discrete probability distribution in engineering. It can be obtained as a special case of Binomial distribution when n is very large and p is very small. Poisson distribution is commonly used in engineering to model problems such as queueing (birth-and-death process or waiting on line), radioactive experiments, the telephone calls received at an office, the emission of electrons from a cathode, and natural hazards (earthquakes, hurricanes, or tornados). A random variable X has a Poisson distribution with parameter λ if it takes the values 0, 1, 2, . . . with

$$P(k) = \frac{\lambda^k}{k!}e^{-\lambda}, \qquad k = 0, 1, 2, \cdots \tag{2.69}$$

The corresponding PDF is

$$f_X(x) = \sum_{k=0}^{\infty} P(k)\delta(x-k) \tag{2.70}$$

which is shown in Fig. 2.11 for λ = 2.

The mean and variance of X are

$$E[X] = \lambda \tag{2.71a}$$

$$\mathrm{Var(X)} = \lambda \tag{2.71b}$$

Note from Eq. (2.71a) that the parameter λ represents the average rate of occurrence of X. A summary of the properties of the four discrete probability distributions is provided in Table 2.3.

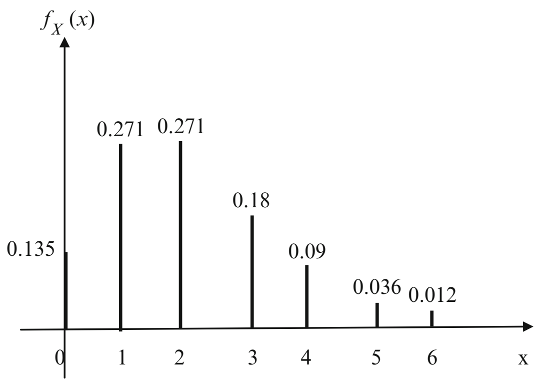


Fig. 2.11 PDF for Poisson distribution with λ = 2

Table 2.3 Properties of discrete probability distributions

Name	P(k)	PDF	Mean	Variance
Bernoulli	$P(x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \\ 0, & \text{otherwise} \end{cases}$	$f_X(x) = \sum_{k=0}^{1} P(k)\delta(x-k)$	p	p(1 − p)
Binomial	$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$	$f_X(x) = \sum_{k=0}^{n} P(k)\delta(x-k)$	np	np(1 − p)
Geometric	$P(k) = pq^{k-1}$	$f_X(x) = \sum_{k=0}^{n} P(k)\delta(x-k)$	1/p	q/p^2
Poisson	$P(k) = \frac{\lambda^k}{k!} e^{-\lambda}$	$f_X(x) = \sum_{k=0}^{\infty} P(k)\delta(x-k)$	λ	λ

Example 2.9 Verify Eq. (2.71).

Solution

First, we notice that

$$\sum_{k=0}^{\infty} P(k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda}\left(e^{\lambda}\right) = 1$$

We obtain the mean value of X as

$$E[X] = \sum_{k=0}^{\infty} kP(k) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = 0 + \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \lambda e^{-\lambda}$$

If we let n = k − 1, we get

$$E[X] = \lambda e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = \lambda e^{-\lambda}\left(e^{\lambda}\right) = \lambda$$

The second moment is handled the same way.

$$E\left[X^2\right] = \sum_{k=0}^{\infty} k^2 P(k) = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = 0 + \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!} \lambda e^{-\lambda}$$

Since, k = k − 1 + 1

$$E\left[X^2\right] = \sum_{k=1}^{\infty} (k-1+1) \frac{\lambda^{k-1}}{(k-1)!} \lambda e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda^2 + \lambda$$

Hence

$$\text{Var}(X) = E\left[X^2\right] - E^2[X] = \lambda^2 + \lambda - \lambda^2 = \lambda$$

as expected.

2.5 Continuous Probability Models

In this section, we consider five continuous probability distributions: uniform, exponential, Erlang, hyperexponential, and Gaussian distributions [3–5].

2.5.1 Uniform Distribution

This distribution, also known as *rectangular distribution*, is very important for performing pseudo random number generation used in simulation. It is also useful for describing quantizing noise that is generated in pulse-code modulation. It is a distribution in which the density is constant. It models random events in which every value between a minimum and maximum value is equally likely. A random variable X has a uniform distribution if its PDF is given by

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases} \tag{2.72}$$

which is shown in Fig. 2.12.

The mean and variance are given by

$$E(X) = \frac{b+a}{2} \tag{2.73a}$$

$$\mathrm{Var}(X) = \frac{(b-a)^2}{12} \tag{2.73b}$$

A special uniform distribution for which a = 0, b = 1, called the standard uniform distribution, is very useful in generating random samples from any probability distribution function. Also, if Y = Asin X, where X is a uniformly distributed random variable, the distribution of Y is said to be *sinusoidal distribution*.

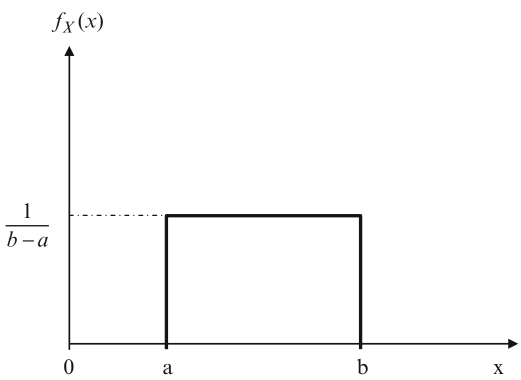


Fig. 2.12 PDF for a uniform random variable

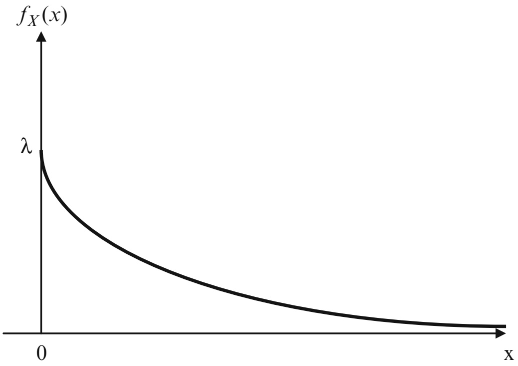


Fig. 2.13 PDF for an exponential random variable

2.5.2 *Exponential Distribution*

This distribution, also known as *negative exponential distribution*, is important because of its relationship to the Poisson distribution. It is frequently used in simulation of queueing systems to describe the interarrival or interdeparture times of customers at a server. Its frequent use is due to the lack of conditioning of remaining time on past time expended. This peculiar characteristic is known variably as Markov, *forgetfulness* or *lack of memory* property. For a given Poisson process, the time interval X between occurrence of events has an exponential distribution with the following PDF

$$f_X(x) = \lambda e^{-\lambda x} u(x) \tag{2.74}$$

which is portrayed in Fig. 2.13.

The mean and the variance of X are

$$E(X) = \frac{1}{\lambda} \tag{2.75a}$$

$$\mathrm{Var}(X) = \frac{1}{\lambda^2} \tag{2.75b}$$

2.5.3 *Erlang Distribution*

This is an extension of the exponential distribution. It is commonly used in queueing theory to model an activity that occurs in phases, with each phase being exponentially distributed. Let $X_1, X_2, \cdots, X_n$ be independent, identically distributed random variables having exponential distribution with mean 1/λ. Then their sum $X = X_1 + X_2 + \cdots X_n$ has n-stage Erlang distribution. The PDF of X is

$$f_X(x) = \frac{\lambda^k x^{k-1}}{(n-1)!} e^{-\lambda x} \tag{2.76}$$

with mean

$$E(X) = \frac{n}{\lambda} \tag{2.77a}$$

and variance

$$\mathrm{Var}(X) = \frac{n}{\lambda^2} \tag{2.77b}$$

2.5.4 Hyperexponential Distribution

This is another extension of the exponential distribution. Suppose X_1 and X_2 are two exponentially distributed random variables with means $1/\lambda_1$ and $1/\lambda_2$ respectively. If the random variable X assumes the value X_1 with probability p, and the value of X_2 with probability $q = 1 - p$, then the PFD of X is

$$f_X(x) = p\lambda_1 e^{-\lambda_1 x} + q\lambda_2 e^{-\lambda_2 x} \tag{2.78}$$

This is known as a two-stage hyperexponential distribution. Its mean and variance are given by

$$E(X) = \frac{p}{\lambda_1} + \frac{q}{\lambda_2} \tag{2.79}$$

$$\mathrm{Var}(X) = \frac{p(2-p)}{\lambda_1^2} + \frac{1-p^2}{\lambda_2^2} - \frac{2p(1-p)}{\lambda_1 \lambda_2} \tag{2.80}$$

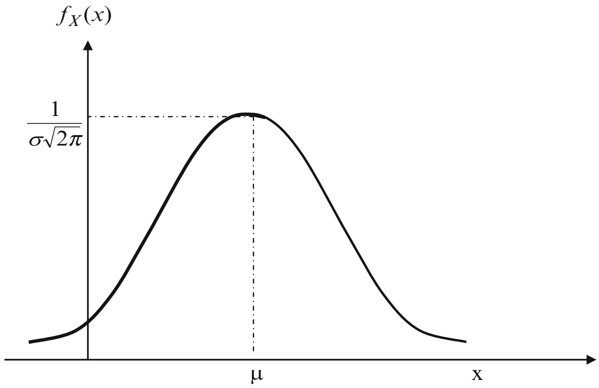
2.5.5 Gaussian Distribution

This distribution, also known as *normal* distribution, is the most important probability distribution in engineering. It is used to describe phenomena with symmetric variations above and below the mean μ. A random variable X with Gaussian distribution has its PDF of the form

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \qquad -\infty < x < \infty \tag{2.81}$$

where the mean

Fig. 2.14 PDF for an Gaussian random variable



$$E(X) = \mu \tag{2.82a}$$

and the variance

$$\text{Var}(X) = \sigma^2 \tag{2.82b}$$

are themselves incorporated in the PDF. Figure 2.14 shows the Gaussian PDF.

It is a common practice to use the notation $X \approx N(\mu,\sigma^2)$ to denote a normal random variable X with mean μ and variance σ^2. When $\mu = 0$ and $\sigma = 1$, we have X = N(0,1), and the *normalized* or *standard normal* distribution function with

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{2.83}$$

which is widely tabulated.

It is important that we note the following points about the normal distribution which make the distribution the most prominent in probability and statistics and also in communication.

1. The binomial probability function with parameters n and p is approximated by a Gaussian PDF with $\mu = np$ and $\sigma^2 = np(1 - p)$ for large n and finite p.
2. The Poisson probability function with parameter λ can be approximated by a normal distribution with $\mu = \sigma^2 = \lambda$ for large λ.
3. The normal distribution is useful in characterizing the uncertainty associated with the estimated values. In other words, it is used in performing statistical analysis on simulation output.
4. The justification for the use of normal distribution comes from the *central limit theorem*.

The **central limit theorem** states that the distribution of the sum of n independent random variables from any distribution approaches a normal distribution as n becomes large.

(We will elaborate on the theorem a little later.) Thus the normal distribution is used to model the cumulative effect of many small disturbances each of which contributes to the stochastic variable X. It has the advantage of being mathematically tractable. Consequently, many statistical analysis such as those of regression and variance have been derived assuming a normal density function. In several communication applications, we assume that noise is Gaussian distributed in view of the central limit theorem because noise is due to the sum of several random parameters. A summary of the properties of the five continuous probability distributions is provided in Table 2.4.

Example 2.10 Let X be a Gaussian random variable. (a) Find E[X], E[X^2], and Var(X). (b) Calculate P(a < X < b).

Solution

(a) By definition,

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx \tag{2.10.1}$$

Let y = (x − μ)/σ so that

$$E[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma y + \mu) e^{-y^2/2} dy = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-y^2/2} dy + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy$$
$$= 0 + \mu \tag{2.10.2}$$

Notice the first integral on the right-hand side is zero since the integrand is an odd function and the second integral gives μ since it represents the PDF of a Gaussian random variable N(0,1). Hence,

$$E[X] = \mu \tag{2.10.3}$$

Similarly,

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx$$

Again, we let y = (x − μ)/σ so that

$$E[X^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma y + \mu)^2 e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^2 y^2 e^{-y^2/2} dy + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 2\sigma\mu y e^{-y^2/2} dy$$
$$+ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mu^2 e^{-y^2/2} dy \tag{2.10.4}$$

We can evaluate the first integral on the right-hand side by parts. The second integral is zero because the integrand is an odd function of y. The third integral yields μ^2 since it represents the PDF of a Gaussian random variable N(0,1). Thus,

Table 2.4 Properties of continuous probability distributions

Name	PDF	CDF	Mean	Variance
Uniform	$f_X(x) = \frac{1}{b-a}$	$F_X(x) = \frac{x-a}{b-a}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$
Exponential	$f_X(x) = \lambda e^{-\lambda x} u(x)$	$F_X(x) = 1 - e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Erlang	$f_X(x) = \frac{\lambda^k x^{k-1}}{(n-1)!} e^{-\lambda x}$	$F_X(x) = 1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
Hyperexponential	$f_X(x) = p\lambda_1 e^{-\lambda_1 x} + q\lambda_2 e^{-\lambda_2 x}$	$\frac{p}{\lambda_1} + \frac{q}{\lambda_2}$	$F_X(x) = p\left(1 - e^{-\lambda_1 t}\right) + q\left(1 - e^{-\lambda_2 t}\right)$	$\frac{p(2-p)}{\lambda_1^2} + \frac{1-p^2}{\lambda_2^2} - \frac{2p(1-p)}{\lambda_1 \lambda_2}$
Gaussian	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$	$F_X(x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$	μ	σ^2

Where erf(.) is the error function to be discussed in Example 2.10

$$E\left[X^2\right] = \frac{\sigma^2}{\sqrt{2\pi}}\left[ye^{-y^2/2}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-y^2/2}dy \right] + 2\sigma\mu(0) + \mu^2 = \sigma^2 + \mu^2 \quad (2.10.5)$$

and

$$\text{Var}(X) = E\left[X^2\right] - E^2[X] = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

We have established that for any real and finite number a and b, the following three integrals hold.

$$\int_{-\infty}^{\infty} \frac{1}{b\sqrt{2\pi}} \exp\left[-\frac{(x-a)^2}{2b^2}\right] dx = 1 \quad (2.10.6a)$$

$$\int_{-\infty}^{\infty} \frac{x}{b\sqrt{2\pi}} \exp\left[-\frac{(x-a)^2}{2b^2}\right] dx = a \quad (2.10.6b)$$

$$\int_{-\infty}^{\infty} \frac{x^2}{b\sqrt{2\pi}} \exp\left[-\frac{(x-a)^2}{2b^2}\right] dx = a^2 + b^2 \quad (2.10.6c)$$

(b) To determine the Gaussian probability, we need the CDF of the Gaussian random variable X.

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

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx - \int_{x}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx$$

The value of the first integral is 1 since we are integrating the Gaussian PDF over its entire domain. For the second integral, we substitute

$$z = \frac{(x-\mu)}{\sigma\sqrt{2}}, \qquad dz = \frac{dx}{\sigma\sqrt{2}}$$

and obtain

$$F_X(x) = 1 - \int_{x}^{\infty} \frac{1}{\sqrt{\pi}} e^{-z^2} dz \quad (2.10.7)$$

We define *error function* as

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{2.10.8}$$

and the complimentary error function as

$$erfc(x) = 1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-z^2} dz \tag{2.10.9}$$

Hence, from Eqs. (2.10.7)–(2.10.9),

$$F_X(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x - \mu}{\sigma\sqrt{2}}\right)\right] \tag{2.10.10}$$

and

$$P(a < x < b) = F_X(b) - F_X(a) = \frac{1}{2}\text{erf}\left(\frac{\text{b} - \mu}{\sigma\sqrt{2}}\right) - \frac{1}{2}\text{erf}\left(\frac{\text{a} - \mu}{\sigma\sqrt{2}}\right) \tag{2.10.11}$$

Note that the definition of erf(x) varies from one book to another. Based on its definition in Eq. (2.10.8), some tabulated values are presented in Table 2.5. For example, given a Gaussian distribution with mean 0 and variance 2, we use the table to obtain

$$P(1 < x < 2) = \frac{1}{2}\text{erf}(1) - \frac{1}{2}\text{erf}(0.5) = 0.1611$$

2.6 Transformation of a Random Variable

It is sometimes required in system analysis that we obtain the PDF $f_Y(y)$ of the output random variable Y given that the PDF $f_X(x)$ for the input random variable X is known and the input-output transformation function

$$\text{Y} = \text{g}(\text{X}) \tag{2.84}$$

is provided. If we assume that g(X) is continuous or piecewise continuous, then Y will be a random variable. Our goal is to get $f_Y(y)$. We begin with the distribution of Y.

$$F_Y(y) = P[Y \le y] = P[g(X) \le y] = P\left[X \le g^{-1}(y)\right] = F_X\left(g^{-1}(y)\right)$$

Hence

Table 2.5 Error function

x	erf(x)	x	erf(x)
0.00	0.00000	1.10	0.88021
0.05	0.05637	1.15	0.89612
0.10	0.11246	1.20	0.91031
0.15	0.16800	1.25	0.92290
0.20	0.22270	1.30	0.93401
0.25	0.27633	1.35	0.94376
0.30	0.32863	1.40	0.95229
0.35	0.37938	1.45	0.95970
0.40	0.42839	1.50	0.96611
0.45	0.47548	1.55	0.97162
0.50	0.52050	1.60	0.97635
0.55	0.56332	1.65	0.98038
0.60	0.60386	1.70	0.98379
0.65	0.64203	1.75	0.98667
0.70	0.67780	1.80	0.98909
0.75	0.71116	1.85	0.99111
0.80	0.74210	1.90	0.99279
0.85	0.77067	1.95	0.99418
0.90	0.79691	2.00	0.99532
0.95	0.82089	2.50	0.99959
1.00	0.84270	3.00	0.99998
1.05	0.86244	3.30	1.0

$$f_Y(y) = \frac{d}{dy} F_X\left(g^{-1}(y)\right) = \frac{d}{dx} F_X\left(g^{-1}(y)\right) \frac{dx}{dy}$$

or

$$\boxed{f_Y(y) = \frac{f_X(x)}{\left|\frac{dy}{dx}\right|}} \tag{2.85}$$

where $x = g^{-1}(y)$. In case $Y = g(X)$ has a finite number of roots $X_1, X_2, \ldots, X_n$ such that

$$Y = g(X_1) = g(X_2) = \cdots = g(X_n)$$

then the PDF of y becomes

$$f_X(y) = \frac{f_X(x_1)}{\left|\frac{dy}{dx_1}\right|} + \frac{f_X(x_2)}{\left|\frac{dy}{dx_2}\right|} + \cdots + \frac{f_X(x_n)}{\left|\frac{dy}{dx_n}\right|} \tag{2.86}$$

Once the PDF of Y is determined, we can find its mean and variance using the regular approach.

Example 2.11 Suppose that X is a Gaussian random variable with mean 3 and variance 4 and Y = 3X − 1. Find the PDF of Y and its mean and variance.

Solution

With $\mu = 3$ and $\sigma^2 = 4$, the PDF of X is obtained using Eq. (2.81) as

$$f_X(x) = \frac{1}{2\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-3}{2}\right)^2\right]$$

Since Y = g(X) = 3X − 1, X = (Y + 1)/3 and

$$\frac{dy}{dx} = 3$$

Hence,

$$f_Y(y) = \frac{f_X(x)}{3} = \frac{1}{3} f_X\left(\frac{y+1}{3}\right) = \frac{1}{6\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\frac{y+1}{3}-3}{2}\right)^2\right]$$

or

$$f_Y(y) = \frac{1}{6\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y-8}{6}\right)^2\right]$$

Comparing this with Eq. (2.81) indicates that Y has a Gaussian distribution with mean 8 and variance $6^2 = 36$. We can easily check this.

$$\mathrm{E[Y]} = \mathrm{E[3X - 1]} = 3\mathrm{E[X]} - 1 = 3 \times 3 - 1 = 8$$

$$\mathrm{Var(Y)} = 3^2\,\mathrm{Var(X)} = 9 \times 4 = 36.$$

2.7 Generating Functions

It is sometimes more convenient to work with generating functions. A probability generating function, often called the *z-transform*, is a tool for manipulating infinite series. Generating functions are important for at least two reasons. First, they may have a closed form. Second, they may be used to generate probability distribution and the moments of the distributions.

If $p_0, p_1, p_2, \cdots$ form a probability distribution, the probability generating function is

$$G(z) = E\left[z^i\right] = \sum_{i=0}^{\infty} z^i p_i \tag{2.87}$$

Notice that G(1) = 1 since the probabilities must sum up to 1. The generating function G(z) contains all the information that the individual probabilities have. We can find the individual probabilities from G(z) by repeated differentiation as

$$p_n = \frac{1}{n!} \frac{d^n G(z)}{dz^n}\bigg|_{z=0} \tag{2.88}$$

The moments of the random variable can be obtained from G(z). For example, for the first moment,

$$E[X] = \sum_{i=0}^{\infty} ip_i = \sum_{i=0}^{\infty} ip_i z^{i-1}\bigg|_{z=1} = \frac{d}{dz}\sum_{i=0}^{\infty} p_i z^i\bigg|_{z=1} = G'(1) \tag{2.89}$$

For the second moment,

$$\begin{aligned} E\left[X^2\right] &= \sum_{i=0}^{\infty} i^2 p_i = \sum_{i=0}^{\infty} i(i-1)p_i + \sum_{i=0}^{\infty} ip_i \\ &= \sum_{i=0}^{\infty} i(i-1)p_i z^{i-2}\bigg|_{z=1} + \sum_{i=0}^{\infty} ip_i z^{i-1}\bigg|_{z=1} \\ &= G''(1) + G'(1) \end{aligned} \tag{2.90}$$

Example 2.12 Find the generating function for geometric distribution.

Solution

For geometric distribution, q = 1 − p and $p_i = pq^{i-1}$. Hence,

$$G(z) = \sum_{i=1}^{\infty} pq^{i-1} z^i = pz \sum_{i=1}^{\infty} (qz)^{i-1} = \frac{pz}{1-qz}$$

For $n \geq 1$,

$$\frac{d^n G(z)}{dz^n} = \frac{n! pq^{n-1}}{(1-qz)^{n+1}}$$

Thus,

$$E[X] = G'(1) = \frac{p}{(1-q)^2} = \frac{1}{p}$$

and

$$E\left[X^2\right] = G'(1) + G''(1) = \frac{1}{p} + \frac{2q}{p^2} = \frac{1+q}{p^2}$$

so that variance is

$$\mathrm{Var(X)} = \mathrm{E}\left[\mathrm{X}^2\right] - E^2[X] = \frac{q}{p^2}$$

2.8 Central Limit Theorem

This is a fundamental result in probability theory. The theorem explains why many random variables encountered in nature have distributions close to the Gaussian distribution. To derive the theorem, consider the binomial function

$$B(M) = \frac{N!}{M!(N-M)!} p^M q^{N-M} \tag{2.91}$$

which is the probability of M successes in N independent trials. If M and N − M are large, we may use Stirling's formula

$$n! \cong n^n e^{-n} \sqrt{2n\pi} \tag{2.92}$$

Hence,

$$B(M) = f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{2.93}$$

which is a normal distribution, $\mu = Np$ and $\sigma = \sqrt{Npq}$. Thus, as $N \to \infty$, the sum of a large number of random variables tends to be normally distributed. This is known as the *central limit theorem.*

The **central limit theorem** states that the PDF of the sum of a large number of individual random variables approaches a Gaussian (normal) distribution regardless of whether or not the distribution of the individual variables are normal.

Although the derivation above is based on binomial distribution, the central limit theorem is true for all distributions. A simple consequence of the theorem is that any random variable which is the sum of n independent identical random variables approximates a normal random variable as n becomes large.

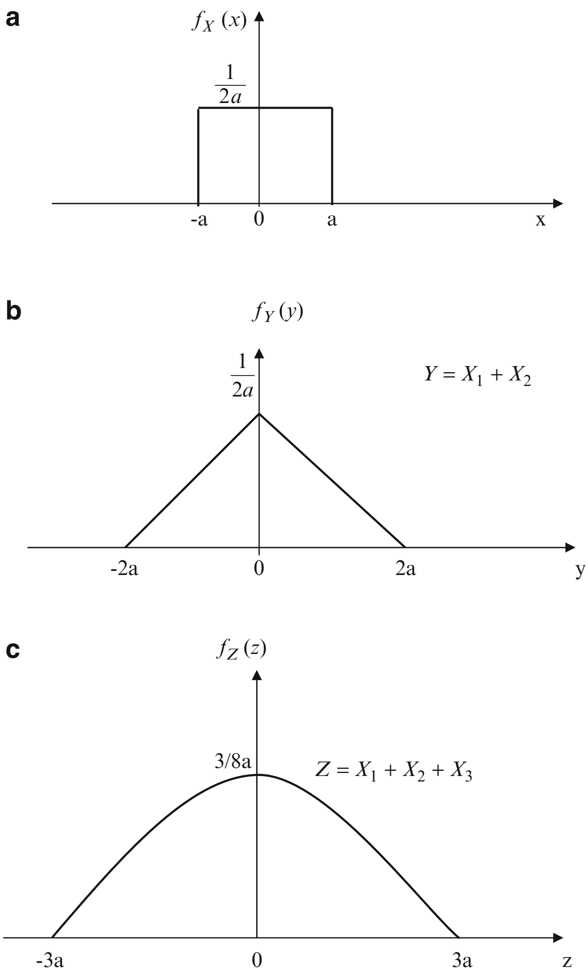


Fig. 2.15 (**a**) PDF of uniform random variable X, (**b**) PDF of $Y = X_1 + X_2$, (**c**) PDF of $Z = X_1 + X_2 + X_3$

Example 2.13 This example illustrates the central limit theorem. If $X_1, X_2, X_3, \ldots X_n$ are n dependent random variables and $c_1, c_2, c_3, \ldots, c_n$ are constants, then

$$X = c_1X_1 + c_2X_2 + c_3X_3 + \ldots + c_nX_n$$

is a Gaussian random variable as n becomes large.

Solution

To make things simple, let use assume that $X_1, X_2, X_3, \ldots X_n$ are identical uniform variable with one of them as shown in Fig. 2.15a. For the sum $Y = X_1 + X_2$, the PDF of y is a convolution of the PDF in Fig. 2.15a with itself, i.e.

$$f_Y(y) = \int_{-\infty}^{\infty} f_X(x) f_X(y - x) dx$$

By performing the convolution, we obtain the joint PDF in Fig. 2.15b. In the same way, for the sum $Z = X_1 + X_2 + X_3$, the PDF of Z is the convolution of the PDF in Fig. 2.15a with that in Fig. 2.15b, i.e.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(\lambda) f_Y(\lambda - z) d\lambda$$

which results in Fig. 2.15c. With only three terms, the PDF of the sum is already approaching Gaussian PDF. According to the central limit theorem, as more terms are added, the PDF becomes Gaussian.

2.9 Computation Using MATLAB

MATLAB is a useful tool for handling and demonstrating some of the concepts covered in this chapter. For example, the MATLAB commands **mean**, **std**, **cov**, and **corrcoef** can be used to find the average/mean value, standard deviation, covariance, and correlation coefficient respectively. We will illustrate with examples how MATLAB can be used.

2.9.1 Performing a Random Experiment

Suppose we want to carry out the random experiment of tossing a die, we can use the MATLAB command **unidrnd** to generate as many trials as possible, with each trial yield randomly 1, 2, . . .6.

We use this command to generate a 12 × 12 matrix with numbers that are uniformly distributed between 1 and 6 as follows.

```
>> x = unidrnd(6,12,12)
x =
   5   3   5   4   6   3   5   6   4   3   1   5
   3   4   1   2   5   3   4   5   1   3   4   1
   1   2   6   6   5   6   5   3   2   2   5   3
   4   4   6   3   5   4   1   4   4   3   2   4
   3   1   4   5   5   2   3   6   2   2   2   5
   1   4   2   1   2   3   3   4   4   3   5   4
   4   5   6   3   2   4   1   2   2   4   5   3
```

```
    5   6   4   3   4   1   5   3   4   5   6   3
    5   5   5   3   4   5   6   2   4   2   3   1
    6   5   3   4   1   6   3   3   3   1   6   3
    6   3   6   4   1   4   6   3   4   3   3   3
    1   4   1   1   2   1   1   3   6   3   5   2
>> x1 = mean(x)
x1 =
  Columns 1 through 10
   3.6667  3.8333  4.0833  3.2500  3.5000  3.5000  3.5833
3.6667  3.3333  2.8333
  Columns 11 through 12
   3.9167  3.0833
>> x2 = mean(x1)
x2 =
   3.5208
>> y1 = std(x)
y1 =
  Columns 1 through 10
   1.8749  1.4035  1.9287  1.4848  1.7838  1.6787  1.8809
1.3707  1.3707  1.0299
  Columns 11 through 12
   1.6765  1.3114
>> y2 = std(y1)
y2 =
   0.2796
```

From 144 outcomes above, we tabulate the results as shown in Table 2.6. We expect $P(x_i) = 1/6 = 0.1667$ for all $i = 1, 2, \ldots 6$ but it is not quite so because the number of trials is not large enough. We have chosen 144 to make the result manageable. If higher number of trials is selected, the results would be more accurate. We also find the mean value to be 3.5208 instead of 3.5 and the standard deviation to be 0.2796.

2.9.2 Plotting PDF

MATLAB can also be used in plotting the cumulative distribution functions (CDF) or probability density function (PDF) of a random variable. The MATLAB commands for the CDF and PDF for various types of random variables we considered in Sects. 2.4 and 2.5 are provided in Table 2.7. One may use the **help** command to get assistance on how to use any of these commands.

For example, we will use MATLAB code to plot PDF or P(x) for Binomial distribution for cases (1) $p = 0.6$, $n = 20$, (2) $p = 0.6$, $n = 100$ by using the command **binopdf**. The MATLAB commands are:

Table 2.6 Outcomes of the experiment of tossing a die

Number (i)	1	2	3	4	5	6
No. of occurrence	20	18	34	29	25	18
$P(x_i)$	0.1389	0.1250	0.2361	0.2014	0.1736	0.1250

Table 2.7 MATLAB commands for commom CDFs and PDFs

Name	CDF	PDF
Binomial	**binocdf**	**binopdf**
Poisson	**poisscdf**	**poisspdf**
Geometric	**geocdf**	**geopdf**
Uniform (discrete)	**unidcdf**	**unidpdf**
Uniform (continuous)	**unifcdf**	**unifpdf**
Exponential	**expcdf**	**exppdf**
Gaussian (Normal)	**normcdf**	**normpdf**
Rayleigh	**raylcdf**	**raylpdf**

```
>> n = 20; % later change n to 100
>> p = 0.6;
>> x = 1:n;
>> y = binopdf(x,n,p);
>> stem(x,y); %plots the discrete distribution
```

The two cases are shown in Fig. 2.16. Notice that as n increases, the distribution approaches Gaussian distribution, as expected.

MATLAB can also be used to plot the CDF or PDF when there is no MATLAB command. For example, suppose we are given a joint PDF for random variables X and Y as

$$f_{XY}(x,y) = \frac{1}{2\pi}\exp\left[-\left(x^2+y^2\right)/2\right], \quad -\infty < x < \infty, \quad -\infty < y < \infty \tag{2.94}$$

Since the computer cannot possibly cover the entire domain of the PDF, we may restrict x and y to [−4,4]. The following MATLAB code can be used to plot the PDF in Eq. (2.94) as shown in Fig. 2.17.

```
[x,y] = meshgrid(-4:0.2:4,-4:0.2:4); % defines grid
f = exp(-(x.^2 + y.^2)/2)/(2*pi); % pdf to be plotted
surf(x,y,f) % creates 3-D plot
xlabel('x'); ylabel('y'); zlabel('pdf');
```

2.9.3 Gaussian Function

As mentioned earlier, the Gaussian distribution is the most important PDF in communications. We can use MATLAB commands **normpdf** and **normcdf** to

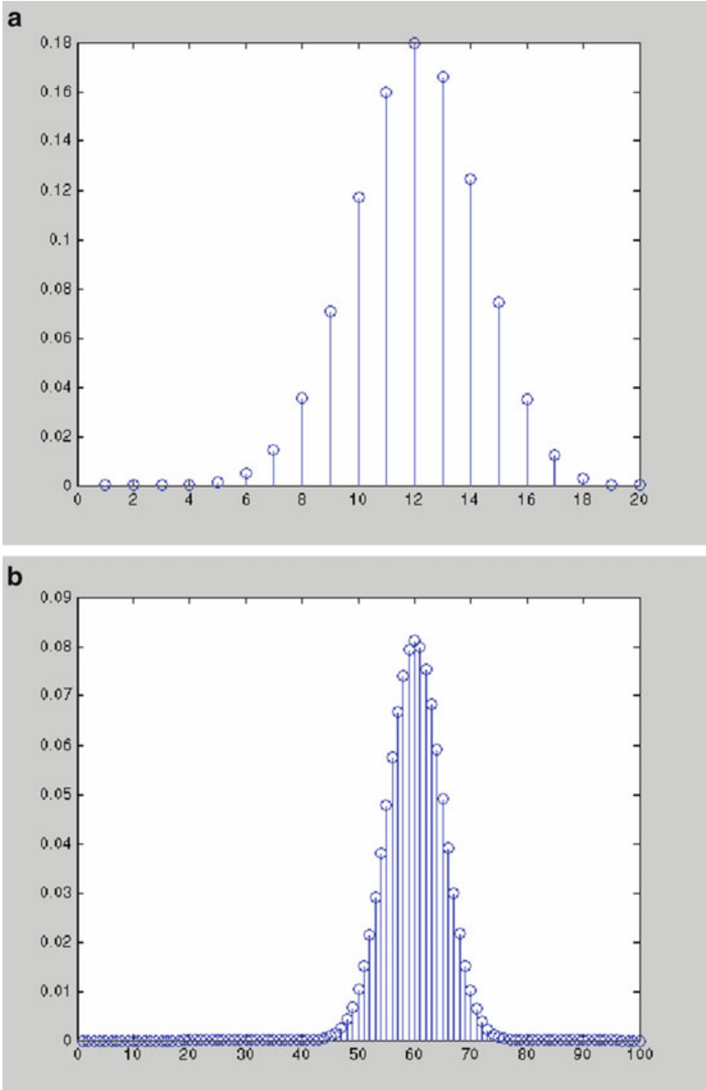


Fig. 2.16 Plot PDF for Binomial distribution for cases (**a**) p = 0.6, n = 20, (**b**) p = 0.6, n = 100

plot the PDF and CDF of the Gaussian distribution. In Sect. 2.5, we defined CDF of the Gaussian random variable X as

$$F_X(x) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right) \tag{2.95}$$

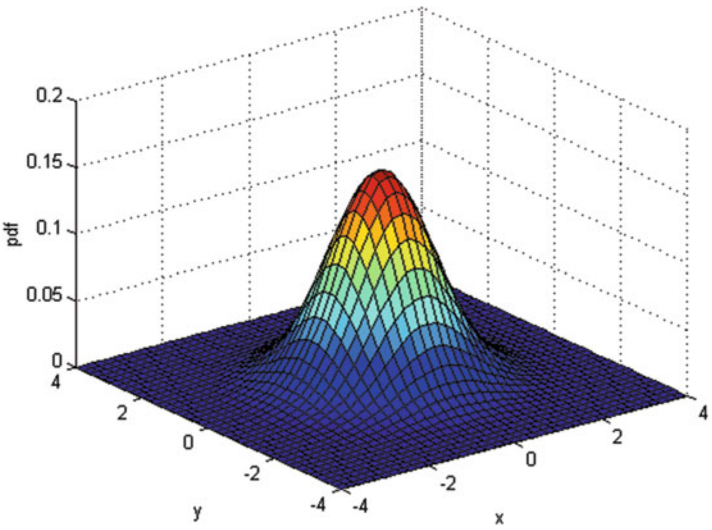


Fig. 2.17 The plot of the joint PDF in Eq. (2.10.5)

where erf(.) is the error function defined as

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{2.96}$$

The MATLAB command **erf** for the error function evaluates the integral in Eq. (2.96). Hence

$$P(a < x < b) = F_X(b) - F_X(a) = \frac{1}{2}\mathrm{erf}\left(\frac{b-\mu}{\sigma\sqrt{2}}\right) - \frac{1}{2}\mathrm{erf}\left(\frac{a-\mu}{\sigma\sqrt{2}}\right)$$

For example, given a Gaussian distribution with mean 0 and variance 2

$$P(1 < x < 2) = \frac{1}{2}\mathrm{erf}(1) - \frac{1}{2}\mathrm{erf}(0.5)$$

Rather than using Table 2.5 to figure this out, we can use MATLAB.

```
> > P = 0.5*(erf(1) - erf(0.5))
P =
0.1611
```

i.e. P(1 < x < 2) = 0.1611, in agreement with what we got in Example 2.10. MATLAB becomes indispensable when the value of erf(x) is not tabulated.

2.10 Summary

1. The probability of an event is the measure of how likely the event will occur as a result of a random experiment. A random experiment is one in which all the outcomes solely depend on *chance*, i.e., each outcome is equally likely to happen.
2. The relative-frequency definition of the probability of an event A assumes that if an experiment is repeated for a large number of times n and event A occurs n_A times,

$$P(A) = \frac{n_A}{n}$$

3. A random variable is a real-value function defined over a sample space. A discrete random variable is one which may take on only a countable number of distinct values such as 0, 1, 2, 3, . . .

   A continuous random variable is one which takes an infinite number of possible values.
4. The cumulative distribution function (CDF) $F_X(x)$ of a random variable X is defined as the probability $P(X \leq x)$ and $F_X(x)$ lies between 0 and 1.
5. The probability density function (PDF) $f_X(x)$ of a random variable X is the derivative of the CDF $F_X(x)$, i.e.

$$f_X(x) = \frac{dF_X(x)}{dx} \quad \longleftrightarrow \quad F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$

   Note that $f_X(x)dx$ is the probability of a random variable X lying within dx of x.
6. The joint CDF $F_{XY}(x,y)$ of two random variables X and Y is the probability $P(X \leq x, Y \leq y)$, while the joint PDF $f_{XY}(x,y)$ is the second partial derivative of the joint CDF with respect to x and y. The PDF of X alone (the marginal PDF) is obtained by integrating the joint PDF $f_{XY}(x,y)$ over all y. The joint CDF or PDF of two independent random variables are factors.
7. The mean value of a random variable X is

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)dx \text{ if X is continuous}$$

   or

$$E(X) = \sum_{i=1}^{M} x_i P(x_i) \text{ if X is discrete}$$

8. The variance of random variable X is

$$\mathrm{Var(x)} = \sigma_X^2 = E\left[X^2\right] - E^2(X)$$

where σ_X is the standard deviation of the random variable; σ_X is a measure of the width of its PDF.

9. Table 2.3 summarizes the P(k), PDF, mean, and variance of common discrete probability distributions: Bernoulli, binomial, geometric, and Poisson.
10. Table 2.4 summarizes the CDF, PDF, mean, and variance of common continuous probability distributions: uniform, exponential, Erlang, hyperexponential, and Gaussian.
11. The central limit theorem is the usual justification for using the Gaussian distribution for modeling. It states that the sum of independent samples from any distribution approaches the Gaussian distribution as the sample size becomes large.
12. MATLAB can be used to plot or generate CDF and PDF, perform random experiments, and determine mean and standard deviation of a given random variable.

For more information on the material covered in this chapter, see [6, 7].

Problems

2.1 An experiment consists of throwing two dice simultaneously. (a) Calculate the probability of having a 2 and a 5 appearing together. (b) What is the probability of the sum being 8.

2.2 A circle is split into ten equal sectors which are numbered 1–10. When the circle is rotated about its center, a pointer indicates where it stops (like a wheel of fortune). Determine the probability: (a) of stopping at number 8, (b) of stopping at an odd number, (c) of stopping at numbers 1, 4, or 6, (d) of stopping at a number greater than 4.

2.3 A jar initially contains four white marbles, three green marbles, and two red marbles. Two marbles are drawn randomly one after the other without replacement. (a) Find the probability that the two marbles are red. (b) Calculate the probability that the two marbles have marching colors.

2.4 The telephone numbers are selected randomly from a telephone directory and the first digit (k) is observed. The result of the observation for 100 telephone numbers is shown below.

k	0	1	2	3	4	5	6	7	8	9
N_k	0	2	18	11	20	13	19	15	1	1

What is the probability that a phone number: (a) starts with 6? (b) begins with an odd number?

2.5 A class has 50 students. Suppose 20 of them are Chinese and 4 of the Chinese students are female. Let event A denote "student is Chinese" and event B denote "student is female." Find: (a) P(A), (b) P(AB), (c) P(B|A).

2.6 In a particular city, voters registration follows the tabulated statistics below. What is the probability that a person selected at random will be a male given that the person is also a Republican?

	Male (%)	Female (%)
Democrat	26	28
Republican	20	13
Independent	12	12

2.7 For three events A, B, and C, show that

$$P(A+B+C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

2.8 A continuous random variable X has the following PDF

$$f_X(x) = \begin{cases} kx, & 1 < x < 4 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find the value of constant k.
(b) Obtain $F_X(x)$.
(c) Evaluate $P(X \leq 2.5)$.

2.9 A random variable has a PDF given by

$$f_X(x) = \begin{cases} \dfrac{1}{2\sqrt{x}}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Find the corresponding $F_X(x)$ and $P(0.5 < x < 0.75)$.

2.10 A Cauchy random variable X has PDF

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty$$

Find the corresponding CDF.

2.11 A joint PDF is given by

$$f_{XY}(x,y) = ke^{-(2x+3y)/6}u(x)u(y)$$

(a) Determine the value of the constant k such that the PDF is valid.
(b) Obtain the corresponding CDF $F_{XY}(x,y)$.
(c) Calculate the marginal PDFs $f_X(x)$ and $f_Y(y)$.
(d) Find $P(X \leq 3, Y > 2)$ and $P(0 < X < 1, 1 < Y < 3)$.

2.12 X and Y are random variables which assume values 0 and 1 according to the probabilities in the table below. Find Cov(X,Y).

		X		
		0	1	Total
Y	0	0.3	0.4	0.7
	1	0.1	0.2	0.3
	Total	0.4	0.6	1.0

2.13 The random variables X and Y have joint PDF as

$$f_{XY}(x,y) = \begin{cases} \frac{1}{4}, & 0 < x < 2, \quad 0 < y < 2 \\ 0, & \text{otherwise} \end{cases}$$

Find: (a) E[X + Y], (b) E[XY].

2.14 Given that a is a constant, show that

(a) $\mathrm{Var}(aX) = a^2\mathrm{Vax}(\mathrm{X})$
(b) $\mathrm{Var}(\mathrm{X} + \mathrm{a}) = \mathrm{Var}(\mathrm{X})$

2.15 If X and Y are two independent random variables with mean μ_X and μ_{Y} and variances σ_X^2 and σ_{y}^2 respectively, show that

$$\mathrm{Var}[XY] = \sigma_X^2\sigma_y^2 + \sigma_X^2\mu_y^2 + \mu_X^2\sigma_y^2$$

2.16 Let $f(x) = \begin{cases} e^{-\alpha x}(\beta x + \gamma), & x > 0 \\ 0, & \text{otherwise} \end{cases}$

Find the conditions for α, β, and γ so that f(x) is a probability density function.

2.17 Given the joint PDF of random variables X and Y as

$$f_{XY}(x,y) = \begin{cases} \frac{1}{2}(x + 3y), & 0 < x < 1, \quad 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find E[X + Y] and E[XY].
(b) Calculate Cov(X,Y) and ρ_{XY}.
(c) Are X and Y uncorrelated? Are they orthogonal?

2.18 The joint PDF of two random variables X and Y is

$$f_{XY}(x,y) = ye^{-y(x+1)}u(x)u(y)$$

(a) Find the marginal PDFs $f_X(x)$ and $f_Y(y)$.
(b) Are X and Y independent?
(c) Calculate the mean and variance of X.
(d) Determine $P(X < Y)$.

2.19 Given the joint PDF

$$f_{XY}(x,y) = \begin{cases} k(x+xy), & 0<x<2, \quad 0<y<2 \\ 0, & \text{otherwise} \end{cases}$$

(a) Evaluate k.
(b) Determine $P(X < 1, y > 1)$.
(c) Find $F_{XY}(0.5,1.5)$.
(d) Obtain $F_Y(y|X = x)$.
(e) Calculate Cov(X,Y).

2.20 The *skew* is defined as the third moment taken about the mean, i.e.

$$\text{skew}(X) = E\left[(X - m_x)^3\right] = \int_{-\infty}^{\infty} (x - m_x)^3 f_X(x)dx$$

Given that a random variable X has a PDF

$$f_X(x) = \begin{cases} \frac{1}{6}(8-x), & 4 < x < 10 \\ 0, & \text{otherwise} \end{cases}$$

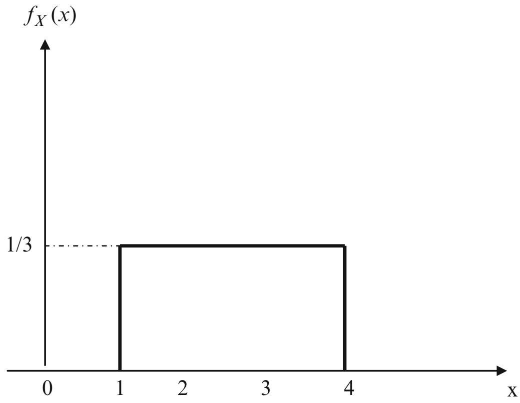
find skew(X).

2.21 Refer to the previous problem for the definition of skewness. Calculate skew(X), where X is a random variable with the following distributions:

(a) Binomial with parameters n and p
(b) Poisson with parameter λ.
(c) Uniform on the interval (a,b).
(d) Exponential with parameter α.

2.22 There are four resistors in a circuit and the circuit will fail if two or more resistors are defective. If the probability of a resistor being defective is 0.005, calculate the probability that the circuit does not fail.

Fig. 2.18 For Prob. 2.27



2.23 Let X be a binomial random variable with p = 0.5 and n = 20. Find $P(4 \le X \le 7)$.
Hint: $P(4 \le X \le 7) = P(X = 4) + P(4 < X \le 7)$.

2.24 The occurrence of earthquakes can be modeled by a Poisson process. If the annual rate of occurrence of earthquakes in a particular area is 0.02, calculate the probability of having exactly one earthquake in 2 years.

2.25 The number of cars arriving at a toll booth during any time interval T (in minutes) follows Poisson distribution with parameter T/2. Calculate the probability that it takes more than 2 min for the first car to arrive at the booth.

2.26 A uniform random variable X has E[X] = 1 and Var(X) = 1/2. Find its PDF and determine $P(X > 1)$.

2.27 Two independent random variables are uniformly distributed, each having the PDF shown in Fig. 2.18. (a) Calculate the mean and variance of each. (b) Determine the PDF of the sum of the two random variables.

2.28 A continuous random variable X may take any value with equal probability within the interval range 0 to α. Find E[X], E[X^2], and Var(X).

2.29 A random variable X with mean 3 follows an exponential distribution. (a) Calculate P(X < 1) and P(X > 1.5). (b) Determine λ such that P(X < λ) = 0.2.

2.30 A zero-mean Gaussian random variable has a variance of 9. Find *a* such that P(|X| > a) < 0.01.

2.31 A random variable T represents the lifetime of an electronic component. Its PDF is given by

$$f_T(t) = \frac{t}{\alpha^2}\exp\left[-\frac{t^2}{\alpha^2}\right]u(t)$$

where $\alpha = 10^3$. Find E[T] and Var(T).

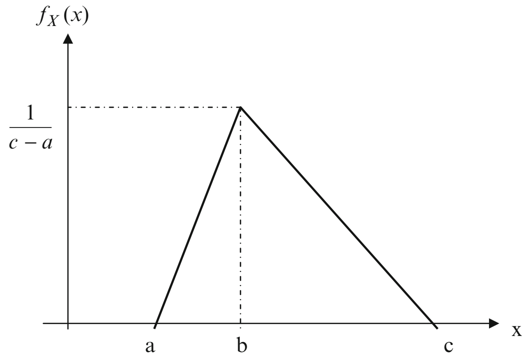


Fig. 2.19 For Prob. 2.33

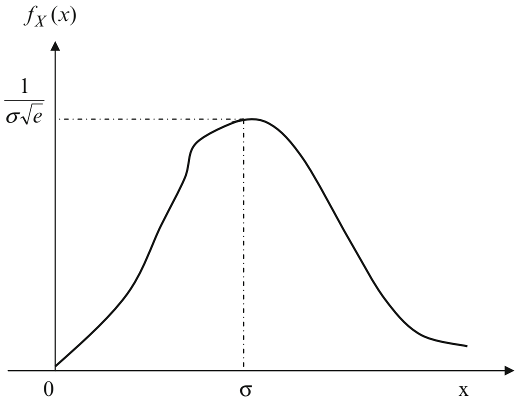


Fig. 2.20 PDF of a Rayleigh random variable for Prob. 2.37

2.32 A measurement of a noise voltage produces a Gaussian random signal with zero mean and variance 2×10^{-11} V^2. Find the probability that a sample measurement exceeds 4 μV.

2.33 A random variable has triangular PDF as shown Fig. 2.19. Find E[X] and Var(X).

2.34 A transformation between X and Y is defined by $Y = e^{-3X}$. Obtain the PDF of Y if:
(a) X is uniformly distributed between −1 and 1, (b) $f_X(x) = e^{-x}u(x)$.

2.35 If $f_X(x) = \alpha e^{-\alpha x}$, $\quad 0 < x < \infty$ and Y = 1/X, find $f_Y(y)$.

2.36 Let X be a Gaussian random variable with mean μ and variance σ^2. (a) Find the PDF of $Y = e^X$. (b) Determine the PDF of $Y = X^2$.

2.37 If X and Y are two independent Gaussian random variables each with zero mean and the same variance σ, show that random variable $R = \sqrt{X^2 + Y^2}$ has a Rayleigh distribution as shown Fig. 2.20. Hint: The joint PDF is $f_{XY}(x,y) = f_X(x) f_Y(y)$ and $f_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} u(r)$.

2.38 Obtain the generating function for Poisson distribution.

2.39 A queueing system has the following probability of being in state n (n = number of customers in the system)

$$p_n = (1 - \rho)\rho^n, \quad n = 0, 1, 2, \cdots$$

(a) Find the generating function G(z). (b) Use G(z) to find the mean number of customers in the system.

2.40 Use MATLAB to plot the joint PDF of random variables X and Y given by

$$f_{XY}(x, y) = xye^{-\left(x^2+y^2\right)}, \quad 0 < x < \infty, 0 < y < \infty$$

Limit x and y to (0,4).

2.41 Use MATLAB to plot the binomial probabilities

$$P(k) = \binom{k}{n} 2^{-k}$$

as a function of n for: (a) k = 5, (b) k = 10.

2.42 Error in data transmission occurs due to white Gaussian noise. The probability of an error is given by

$$P = \frac{1}{2}[1 - \mathrm{erf}(x)]$$

where x is a measure of the signal-to-noise ratio. Use MATLAB to plot P over 0 < x <1.

2.43 Plot the PDF of Gaussian distribution with mean 2 and variance 4 using MATLAB.

2.44 Using the MATLAB command **rand**, one can generate random numbers uniformly distributed on the interval (0,1). Generate 10,000 such numbers and compute the mean and variance. Compare your result with that obtained using E[X] = (a + b)/2 and Var(X) = $(b - a)^2/12$.

References

1. G. R. Grimmett and D.R. Stirzaker, *Probability and Random Processes*. Oxford: Oxford University Press, 2001, pp. 26–45.
2. X. R. Li, Probability, *Random Signals, and Statistics*. Boca Raton, FL: CRC Press, 1999, pp. 65–143.
3. R. Jain, *The Art of Computer Systems Performance Analysis*. New York: John Wiley & Sons, 1991, pp. 483–501.
4. R. Nelson, *Probability, Stochastic Processes, and Queueing Theory*. New York: Springer-Verlag, 1995, pp. 101–165.

5. P. G. Harrison and N. M. Patel, *Performance Modelling of Communication Networks and Computer Architecture*. Wokingham, UK: Addison-Wesley, 1992, pp. 19–48.
6. R. Goodman, *Introduction to Stochastic Models*. Mineola, NY: Dover Publications, 2^{nd} ed., 2006.
7. O. C. Ibe, *Markov Processes for Stochastic Modeling*. Burlington, MA: Elsevier Academic Press, 2009.