CHAPTER 5

Pairs of Random Variables

Many random experiments involve several random variables. In some experiments a number of different quantities are measured. For example, the voltage signals at several points in a circuit at some specific time may be of interest. Other experiments involve the repeated measurement of a certain quantity such as the repeated measurement ("sampling") of the amplitude of an audio or video signal that varies with time. In Chapter 4 we developed techniques for calculating the probabilities of events involving a single random variable *in isolation*. In this chapter, we extend the concepts already introduced to two random variables:

- We use the joint pmf, cdf, and pdf to calculate the probabilities of events that involve the *joint* behavior of two random variables;
- We use expected value to define joint moments that summarize the behavior of two random variables;
- We determine when two random variables are independent, and we quantify their degree of "correlation" when they are not independent;
- We obtain conditional probabilities involving a pair of random variables.

In a sense we have already covered all the fundamental concepts of probability and random variables, and we are "simply" elaborating on the case of two or more random variables. Nevertheless, there are significant analytical techniques that need to be learned, e.g., double summations of pmf's and double integration of pdf's, so we first discuss the case of two random variables in detail because we can draw on our geometric intuition. Chapter 6 considers the general case of vector random variables. Throughout these two chapters you should be mindful of the forest (fundamental concepts) and the trees (specific techniques)!

5.1 TWO RANDOM VARIABLES

The notion of a random variable as a mapping is easily generalized to the case where two quantities are of interest. Consider a random experiment with sample space S and event class $\mathcal{F}$. We are interested in a function that assigns a pair of real numbers

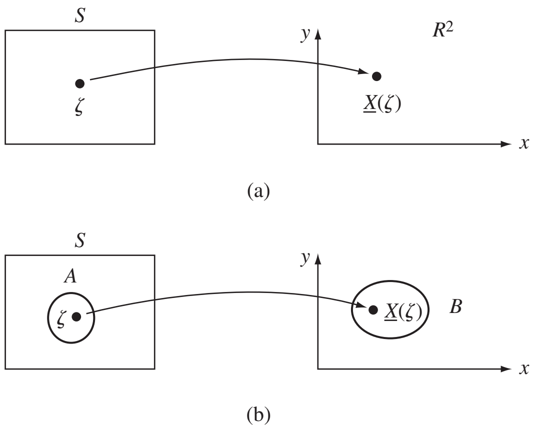


FIGURE 5.1
(a) A function assigns a pair of real numbers to each outcome in S. (b) Equivalent events for two random variables.

$\mathbf{X}(\zeta) = (X(\zeta), Y(\zeta))$ to each outcome ζ in S. Basically we are dealing with a vector function that maps S into R^2, the real plane, as shown in Fig. 5.1(a). We are ultimately interested in events involving the pair (X, Y).

Example 5.1

Let a random experiment consist of selecting a student's name from an urn. Let ζ denote the outcome of this experiment, and define the following two functions:

$$H(\zeta) = \text{height of student } \zeta \text{ in centimeters}$$
$$W(\zeta) = \text{weight of student } \zeta \text{ in kilograms}$$

$(H(\zeta), W(\zeta))$ assigns a pair of numbers to each ζ in S.

We are interested in events involving the pair (H, W). For example, the event $B = \{H \le 183, W \le 82\}$ represents students with height less that 183 cm (6 feet) and weight less than 82 kg (180 lb).

Example 5.2

A Web page provides the user with a choice either to watch a brief ad or to move directly to the requested page. Let ζ be the patterns of user arrivals in T seconds, e.g., number of arrivals, and listing of arrival times and types. Let $N_1(\zeta)$ be the number of times the Web page is directly requested and let $N_2(\zeta)$ be the number of times that the ad is chosen. $(N_1(\zeta), N_2(\zeta))$ assigns a pair of nonnegative integers to each ζ in S. Suppose that a type 1 request brings 0.001¢ in revenue and a type 2 request brings in 1¢. Find the event "revenue in T seconds is less than \$100."

The total revenue in T seconds is $0.001\ N_1 + 1N_2$, and so the event of interest is $B = \{0.001\,N_1 + 1\,N_2 < 10{,}000\}$.

Example 5.3

Let the outcome ζ in a random experiment be the length of a randomly selected message. Suppose that messages are broken into packets of maximum length M bytes. Let Q be the number of full packets in a message and let R be the number of bytes left over. $(Q(\zeta), R(\zeta))$ assigns a pair of numbers to each ζ in S. Q takes on values in the range $0, 1, 2, \ldots$, and R takes on values in the range $0, 1, \ldots, M - 1$. An event of interest may be $B = \{R < M/2\}$, "the last packet is less than half full."

Example 5.4

Let the outcome of a random experiment result in a pair $\underline{\zeta} = (\zeta_1, \zeta_2)$ that results from two independent spins of a wheel. Each spin of the wheel results in a number in the interval $(0, 2\pi]$. Define the pair of numbers (X, Y) in the plane as follows:

$$X(\underline{\zeta}) = \left(2 \ln \frac{2\pi}{\zeta_1}\right)^{1/2} \cos \zeta_2 \qquad Y(\underline{\zeta}) = \left(2 \ln \frac{2\pi}{\zeta_1}\right)^{1/2} \sin \zeta_2.$$

The vector function $(X(\underline{\zeta}), Y(\underline{\zeta}))$ assigns a pair of numbers in the plane to each $\underline{\zeta}$ in S. The square root term corresponds to a radius and to ζ_2 an angle.

We will see that (X, Y) models the noise voltages encountered in digital communication systems. An event of interest here may be $B = \{X^2 + Y^2 < r^2\}$, "total noise power is less than r^2."

The events involving a pair of random variables (X, Y) are specified by conditions that we are interested in and can be represented by regions in the plane. Figure 5.2 shows three examples of events:

$$A = \{X + Y \le 10\}$$
$$B = \{\min(X, Y) \le 5\}$$
$$\mathrm{C} = \{X^2 + Y^2 \le 100\}.$$

Event A divides the plane into two regions according to a straight line. Note that the event in Example 5.2 is of this type. Event C identifies a disk centered at the origin and

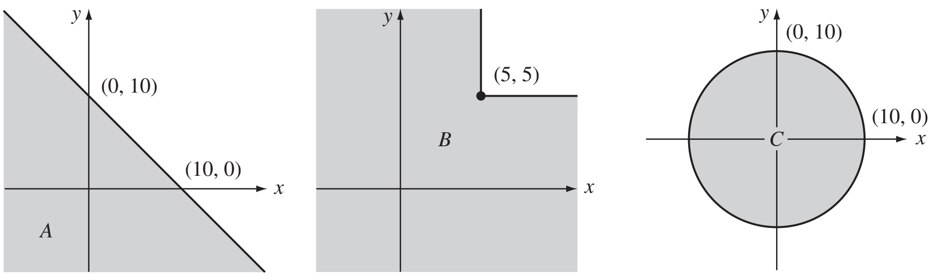


FIGURE 5.2
Examples of two-dimensional events.

it corresponds to the event in Example 5.4. Event B is found by noting that $\{\min(X, Y) \le 5\} = \{X \le 5\} \cup \{Y \le 5\}$, that is, the minimum of X and Y is less than or equal to 5 if either X and/or Y is less than or equal to 5.

To determine the probability that the pair $\mathbf{X} = (X, Y)$ is in some region B in the plane, we proceed as in Chapter 3 to find the equivalent event for B in the underlying sample space S:

$$A = \mathbf{X}^{-1}(B) = \{\zeta: (X(\zeta), Y(\zeta)) \text{ in } B\}. \tag{5.1a}$$

The relationship between $A = \mathbf{X}^{-1}(B)$ and B is shown in Fig. 5.1(b). If A is in $\mathcal{F}$, then it has a probability assigned to it, and we obtain:

$$P[X \text{ in } B] = P[A] = P[\{\zeta: (X(\zeta), Y(\zeta)) \text{ in } B\}]. \tag{5.1b}$$

The approach is identical to what we followed in the case of a single random variable. The only difference is that we are considering the *joint behavior of X and Y* that is induced by the underlying random experiment.

A scattergram can be used to deduce the joint behavior of two random variables. A scattergram plot simply places a dot at every observation pair (x, y) that results from performing the experiment that generates (X, Y). Figure 5.3 shows the scattergram for 200 observations of four different pairs of random variables. The pairs in Fig. 5.3(a) appear to be uniformly distributed in the unit square. The pairs in Fig. 5.3(b) are clearly confined to a disc of unit radius and appear to be more concentrated near the origin. The pairs in Fig. 5.3(c) are concentrated near the origin, and appear to have circular symmetry, but are not bounded to an enclosed region. The pairs in Fig. 5.3(d) again are concentrated near the origin and appear to have a clear linear relationship of some sort, that is, larger values of x tend to have linearly proportional increasing values of y. We later introduce various functions and moments to characterize the behavior of pairs of random variables illustrated in these examples.

The joint probability mass function, joint cumulative distribution function, and joint probability density function provide approaches to specifying the probability law that governs the behavior of the pair (X, Y). Our general approach is as follows. We first focus on events that correspond to rectangles in the plane:

$$B = \{X \text{ in } A_1\} \cap \{Y \text{ in } A_2\} \tag{5.2}$$

where A_k is a one-dimensional event (i.e., subset of the real line). We say that these events are of **product form**. The event B occurs when both $\{X \text{ in } A_1\}$ and $\{Y \text{ in } A_2\}$ occur jointly. Figure 5.4 shows some two-dimensional product-form events. We use Eq. (5.1b) to find the probability of product-form events:

$$P[B] = P[\{X \text{ in } A_1\} \cap \{Y \text{ in } A_2\}] \triangleq P[X \text{ in } A_1, Y \text{ in } A_n]. \tag{5.3}$$

By defining A appropriately we then obtain the joint pmf, joint cdf, and joint pdf of (X, Y).

5.2 PAIRS OF DISCRETE RANDOM VARIABLES

Let the vector random variable $\mathbf{X} = (X, Y)$ assume values from some countable set $S_{X,Y} = \{(x_j, y_k), j = 1, 2, \ldots, k = 1, 2, \ldots\}$. The **joint probability mass function** of $\mathbf{X}$ specifies the probabilities of the event $\{X = x\} \cap \{Y = y\}$:

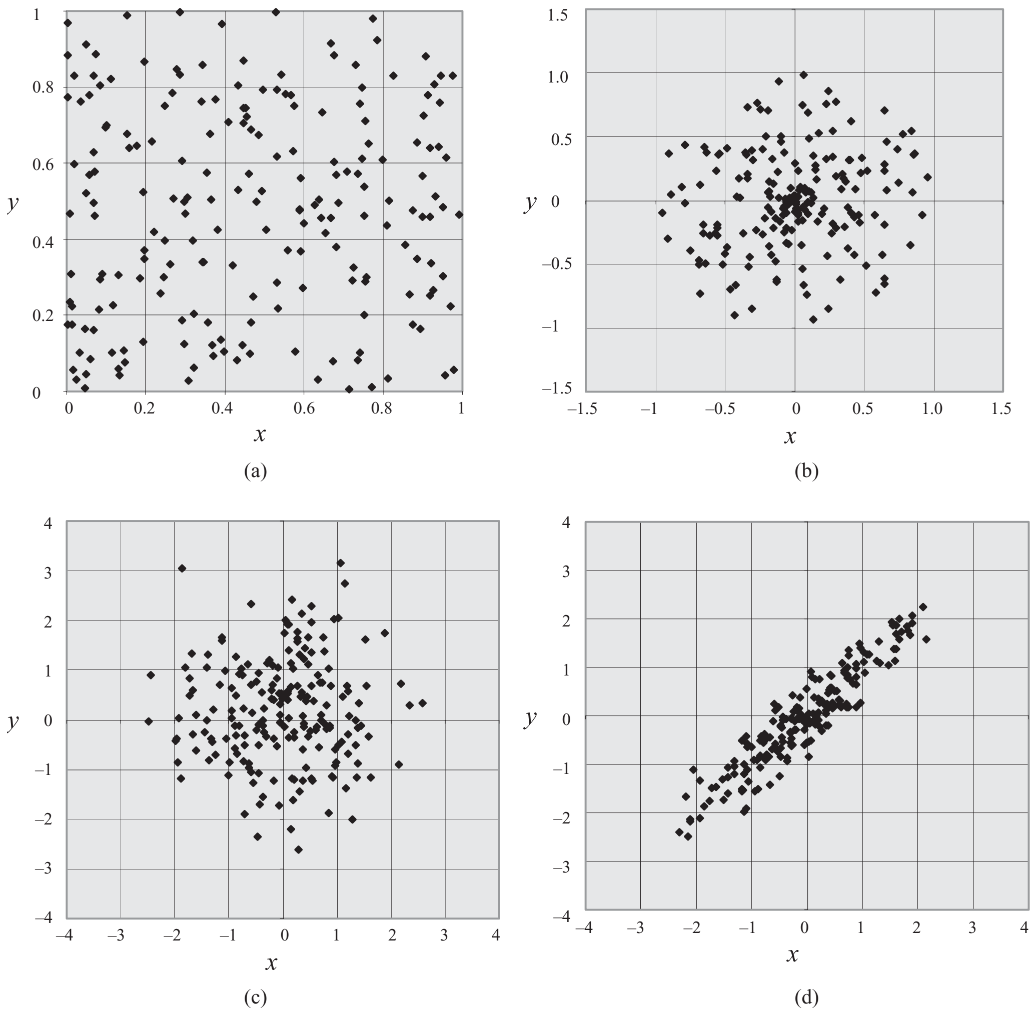


FIGURE 5.3
A scattergram for 200 observations of four different pairs of random variables.

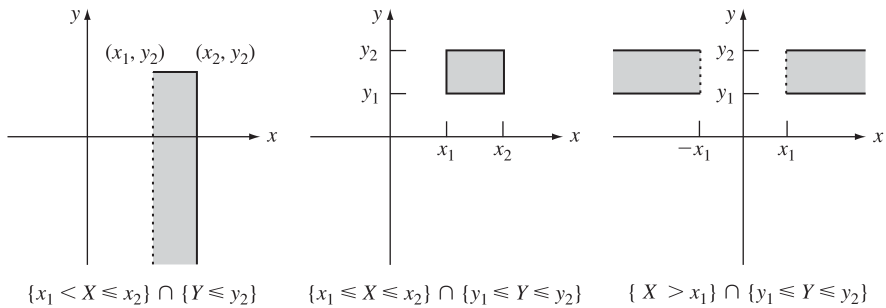


FIGURE 5.4
Some two-dimensional product-form events.

$$p_{X,Y}(x, y) = P[\{X = x\} \cap \{Y = y\}]$$
$$\triangleq P[X = x, Y = y] \qquad \text{for } (x, y) \in R^2. \tag{5.4a}$$

The values of the pmf on the set $S_{X,Y}$ provide the essential information:

$$p_{X,Y}(x_j, y_k) = P[\{X = x_j\} \cap \{Y = y_k\}]$$
$$\triangleq P[X = x_j, Y = y_k] \quad (x_j, y_k) \in S_{X,Y}. \tag{5.4b}$$

There are several ways of showing the pmf graphically: (1) For small sample spaces we can present the pmf in the form of a table as shown in Fig. 5.5(a). (2) We can present the pmf using arrows of height $p_{X,Y}(x_j, y_k)$ placed at the points $\{(x_j, y_k)\}$ in the plane, as shown in Fig. 5.5(b), but this can be difficult to draw. (3) We can place dots at the points $\{(x_j, y_k)\}$ and label these with the corresponding pmf value as shown in Fig. 5.5(c).

The probability of any event B is the sum of the pmf over the outcomes in B:

$$P[\mathbf{X} \text{ in } B] = \sum_{(x_j,y_k)} \sum_{\text{in } B} p_{X,Y}(x_j, y_k). \tag{5.5}$$

Frequently it is helpful to sketch the region that contains the points in B as shown, for example, in Fig. 5.6. When the event B is the entire sample space $S_{X,Y}$, we have:

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} p_{X,Y}(x_j, y_k) = 1. \tag{5.6}$$

Example 5.5

A packet switch has two input ports and two output ports. At a given time slot a packet arrives at each input port with probability 1/2, and is equally likely to be destined to output port 1 or 2. Let X and Y be the number of packets destined for output ports 1 and 2, respectively. Find the pmf of X and Y, and show the pmf graphically.

The outcome I_j for an input port j can take the following values: "n", no packet arrival (with probability 1/2); "a1", packet arrival destined for output port 1 (with probability 1/4); "a2", packet arrival destined for output port 2 (with probability 1/4). The underlying sample space S consists of the pair of input outcomes $\zeta = (I_1, I_2)$. The mapping for (X, Y) is shown in the table below:

ζ	(n, n)	(n, a1)	(n, a2)	(a1, n)	(a1, a1)	(a1, a2)	(a2, n)	(a2, a1)	(a2, a2)
X, Y	(0, 0)	(1, 0)	(0, 1)	(1, 0)	(2, 0)	(1, 1)	(0, 1)	(1, 1)	(0, 2)

The pmf of (X, Y) is then:

$$p_{X,Y}(0, 0) = P[\zeta = (\text{n}, \text{n})] = \frac{1}{2}\frac{1}{2} = \frac{1}{4},$$
$$p_{X,Y}(0, 1) = P[\zeta \in \{(\text{n}, \text{a2}), (\text{a2}, \text{n})\}] = 2*\frac{1}{8} = \frac{1}{4},$$

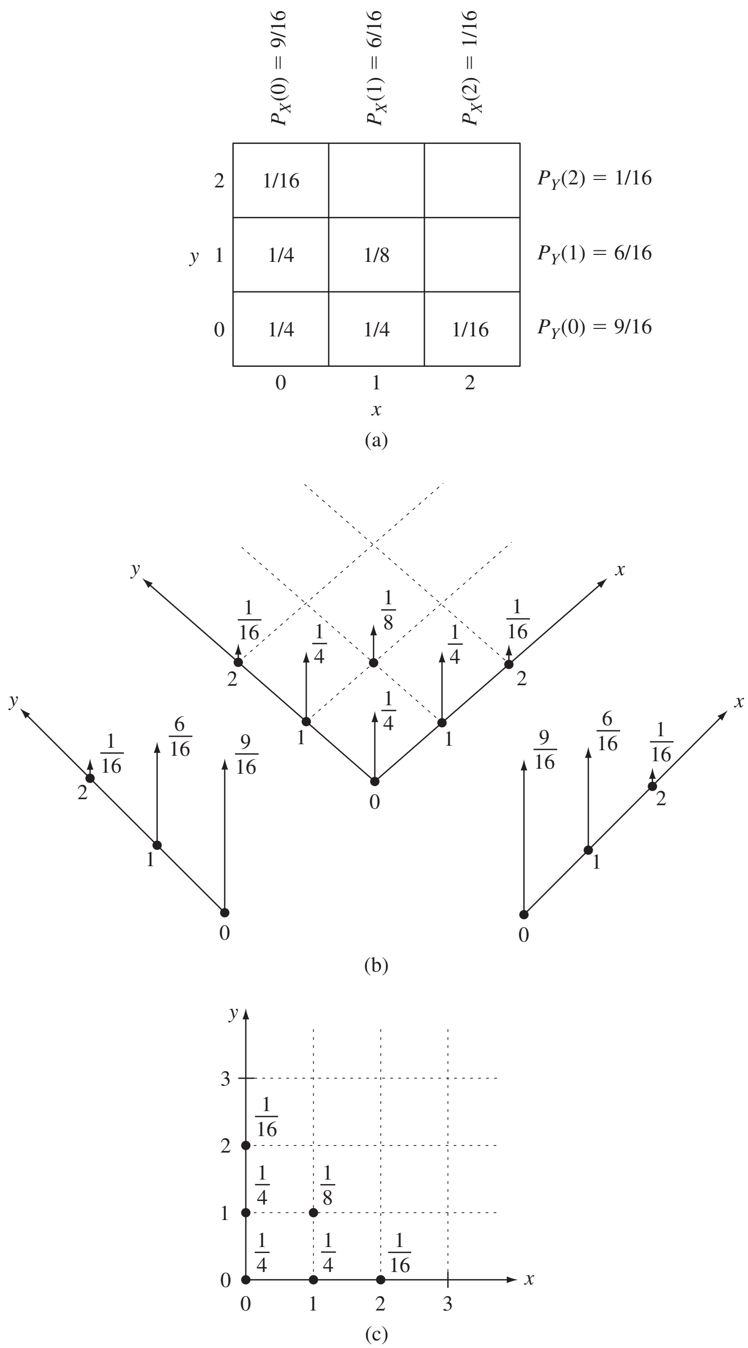


FIGURE 5.5
Graphical representations of pmf's: (a) in table format; (b) use of arrows to show height; (c) labeled dots corresponding to pmf value.

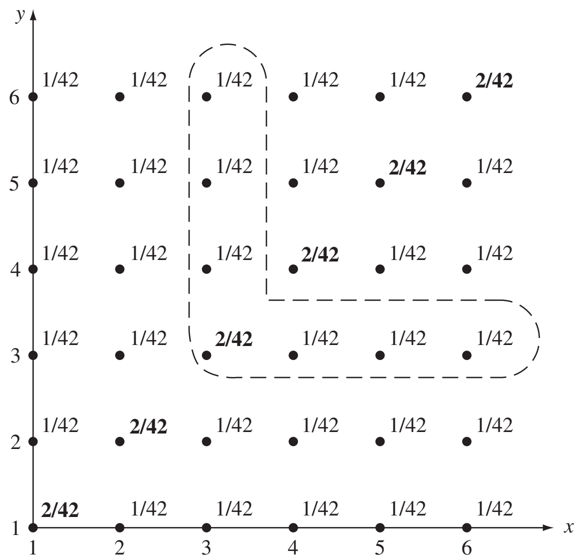


FIGURE 5.6
Showing the pmf via a sketch containing the points in B.

$$
\begin{aligned}
p_{X,Y}(1,0) &= P[\zeta \in \{(\text{n}, \text{a1}), (\text{a1}, \text{n})\}] = \frac{1}{4}, \\
p_{X,Y}(1,1) &= P[\zeta \in \{(\text{a1}, \text{a2}), (\text{a2}, \text{a1})\}] = \frac{1}{8}, \\
p_{X,Y}(0,2) &= P[\zeta = (\text{a2}, \text{a2})] = \frac{1}{16}, \\
p_{X,Y}(2,0) &= P[\zeta = (\text{a1}, \text{a1})] = \frac{1}{16}.
\end{aligned}
$$

Figure 5.5(a) shows the pmf in tabular form where the number of rows and columns accommodate the range of X and Y respectively. Each entry in the table gives the pmf value for the corresponding x and y. Figure 5.5(b) shows the pmf using arrows in the plane. An arrow of height $p_{X,Y}(j,k)$ is placed at each of the points in $S_{X,Y} = \{(0,0), (0,1), (1,0), (1,1), (0,2), (2,0)\}$. Figure 5.5(c) shows the pmf using labeled dots in the plane. A dot with label $p_{X,Y}(j,k)$ is placed at each of the points in $S_{X,Y}$.

Example 5.6

A random experiment consists of tossing two "loaded" dice and noting the pair of numbers (X, Y) facing up. The joint pmf $p_{X,Y}(j,k)$ for $j = 1, \ldots, 6$ and $k = 1, \ldots, 6$ is given by the two-dimensional table shown in Fig. 5.6. The (j,k) entry in the table contains the value $p_{X,Y}(j,k)$. Find the $P[\min(X, Y) = 3]$.

Figure 5.6 shows the region that corresponds to the set $\{\min(x, y) = 3\}$. The probability of this event is given by:

$$\begin{aligned} P[\min(X, Y) = 3] &= p_{X,Y}(6,3) + p_{X,Y}(5,3) + p_{X,Y}(4,3) \\ &\quad + p_{X,Y}(3,3) + p_{X,Y}(3,4) + p_{X,Y}(3,5) + p_{X,Y}(3,6) \\ &= 6\left(\frac{1}{42}\right) + \frac{2}{42} = \frac{8}{42}. \end{aligned}$$

5.2.1 Marginal Probability Mass Function

The joint pmf of $\mathbf{X}$ provides the information about the joint behavior of X and Y. We are also interested in the probabilities of events involving each of the random variables in isolation. These can be found in terms of the **marginal probability mass functions**:

$$\begin{aligned} p_X(x_j) &= P[X = x_j] \\ &= P[X = x_j, Y = \text{anything}] \\ &= P[\{X = x_j \text{ and } Y = y_1\} \cup \{X = x_j \text{ and } Y = y_2\} \cup \ldots] \\ &= \sum_{k=1}^{\infty} p_{X,Y}(x_j, y_k), \end{aligned} \tag{5.7a}$$

and similarly,

$$\begin{aligned} p_Y(y_k) &= P[Y = y_k] \\ &= \sum_{j=1}^{\infty} p_{X,Y}(x_j, y_k). \end{aligned} \tag{5.7b}$$

The marginal pmf's satisfy all the properties of one-dimensional pmf's, and they supply the information required to compute the probability of events involving the corresponding random variable.

The probability $p_{X,Y}(x_j, y_k)$ can be interpreted as the long-term relative frequency of the joint event $\{X = X_j\} \cap \{Y = Y_k\}$ in a sequence of repetitions of the random experiment. Equation (5.7a) corresponds to the fact that the relative frequency of the event $\{X = X_j\}$ is found by adding the relative frequencies of all outcome pairs in which X_j appears. In general, it is impossible to deduce the relative frequencies of pairs of values X and Y from the relative frequencies of X and Y in isolation. The same is true for pmf's: In general, knowledge of the marginal pmf's is insufficient to specify the joint pmf.

Example 5.7

Find the marginal pmf for the output ports (X, Y) in Example 5.2.

Figure 5.5(a) shows that the marginal pmf is found by adding entries along a row or column in the table. For example, by adding along the $x = 1$ column we have:

$$p_X(1) = P[X = 1] = p_{X,Y}(1,0) + p_{X,Y}(1,1) = \frac{1}{4} + \frac{1}{8} = \frac{3}{8}.$$

Similarly, by adding along the $y = 0$ row:

$$p_Y(0) = P[Y = 0] = p_{X,Y}(0,0) + p_{X,Y}(1,0) + p_{X,Y}(2,0) = \frac{1}{4} + \frac{1}{4} + \frac{1}{16} = \frac{9}{16}.$$

Figure 5.5(b) shows the marginal pmf using arrows on the real line.

Example 5.8

Find the marginal pmf's in the loaded dice experiment in Example 5.2.

The probability that $X = 1$ is found by summing over the first row:

$$P[X = 1] = \frac{2}{42} + \frac{1}{42} + \cdots + \frac{1}{42} = \frac{1}{6}.$$

Similarly, we find that $P[X = j] = 1/6$ for $j = 2, \ldots, 6$. The probability that $Y = k$ is found by summing over the kth column. We then find that $P[Y = k] = 1/6$ for $k = 1, 2, \ldots, 6$. Thus each die, in isolation, appears to be fair in the sense that each face is equiprobable. If we knew only these marginal pmf's we would have no idea that the dice are loaded.

Example 5.9

In Example 5.3, let the number of bytes N in a message have a geometric distribution with parameter $1 - p$ and range $S_N = \{0, 1, 2, \ldots\}$. Find the joint pmf and the marginal pmf's of Q and R.

If a message has N bytes, then the number of full packets is the quotient Q in the division of N by M, and the number of remaining bytes is the remainder R. The probability of the pair $\{(q, r)\}$ is given by

$$P[Q = q, R = r] = P[N = qM + r] = (1 - p)p^{qM+r}.$$

The marginal pmf of Q is

$$\begin{aligned} P[Q = q] &= P[N \text{ in}\{qM, qM + 1, \ldots, qM + (M - 1)\}] \\ &= \sum_{k=0}^{(M-1)} (1 - p)p^{qM+k} \\ &= (1 - p)p^{qM}\frac{1 - p^M}{1 - p} = (1 - p^M)(p^M)^q \qquad q = 0, 1, 2, \ldots \end{aligned}$$

The marginal pmf of Q is geometric with parameter p^M. The marginal pmf of R is:

$$\begin{aligned} P[R = r] &= P[N \text{ in}\{r, M + r, 2M + r, \ldots\}] \\ &= \sum_{q=0}^{\infty}(1 - p)p^{qM+r} = \frac{(1 - p)}{1 - p^M}p^r \quad r = 0, 1, \ldots, M - 1. \end{aligned}$$

R has a truncated geometric pmf. As an exercise, you should verify that all the above marginal pmf's add to 1.

5.3 THE JOINT CDF OF *X* AND *Y*

In Chapter 3 we saw that semi-infinite intervals of the form $(-\infty, x]$ are a basic building block from which other one-dimensional events can be built. By defining the cdf $F_X(x)$ as the probability of $(-\infty, x]$, we were then able to express the probabilities of other events in terms of the cdf. In this section we repeat the above development for two-dimensional random variables.

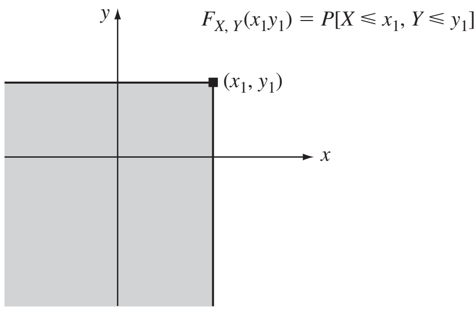


FIGURE 5.7
The joint cumulative distribution function is defined as the probability of the semi-infinite rectangle defined by the point (x_1, y_1).

A basic building block for events involving two-dimensional random variables is the semi-infinite rectangle defined by $\{(x, y): x \le x_1 \text{ and } y \le y_1\}$, as shown in Fig. 5.7. We also use the more compact notation $\{x \le x_1, y \le y_1\}$ to refer to this region. The **joint cumulative distribution function of *X* and *Y*** is defined as the probability of the event $\{X \le x_1\} \cap \{Y \le y_1\}$:

$$F_{X,Y}(x_1, y_1) = P[X \le x_1, Y \le y_1]. \tag{5.8}$$

In terms of relative frequency, $F_{X,Y}(x_1, y_1)$ represents the long-term proportion of time in which the outcome of the random experiment yields a point X that falls in the rectangular region shown in Fig. 5.7. In terms of probability "mass," $F_{X,Y}(x_1, y_1)$ represents the amount of mass contained in the rectangular region.

The joint cdf satisfies the following properties.

(i) The joint cdf is a nondecreasing function of x and y:

$$F_{X,Y}(x_1, y_1) \le F_{X,Y}(x_2, y_2) \qquad \text{if } x_1 \le x_2 \text{ and } y_1 \le y_2, \tag{5.9a}$$

(ii) $F_{X,Y}(x_1, -\infty) = 0, \qquad F_{X,Y}(-\infty, y_1) = 0, \qquad F_{X,Y}(\infty, \infty) = 1.$ (5.9b)

(iii) We obtain the **marginal cumulative distribution functions** by removing the constraint on one of the variables. The marginal cdf's are the probabilities of the regions shown in Fig. 5.8:

$$F_X(x_1) = F_{X,Y}(x_1, \infty) \quad \text{and} \quad F_Y(y_1) = F_{X,Y}(\infty, y_1). \tag{5.9c}$$

(iv) The joint cdf is continuous from the "north" and from the "east," that is,

$$\lim_{x \to a^+} F_{X,Y}(x, y) = F_{X,Y}(a, y) \quad \text{and} \quad \lim_{y \to b^+} F_{X,Y}(x, y) = F_{X,Y}(x, b). \tag{5.9d}$$

(v) The probability of the rectangle $\{x_1 < x \le x_2, y_1 < y \le y_2\}$ is given by:

$$P[x_1 < X \le x_2, y_1 < Y \le y_2] =$$
$$F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1). \tag{5.9e}$$

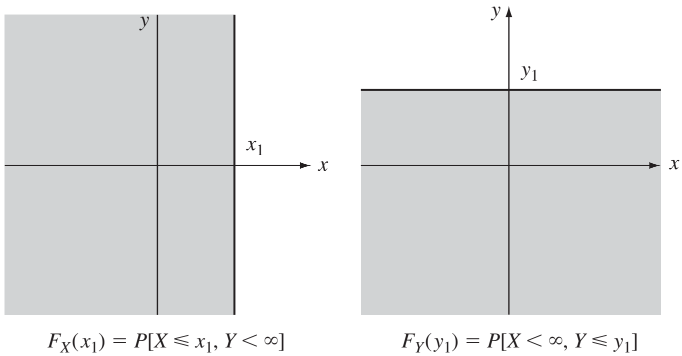


FIGURE 5.8
The marginal cdf's are the probabilities of these half-planes.

Property (i) follows by noting that the semi-infinite rectangle defined by (x_1, y_1) is contained in that defined by (x_2, y_2) and applying Corollary 7. Properties (ii) to (iv) are obtained by limiting arguments. For example, the sequence $\{x \le x_1 \text{ and } y \le -n\}$ is decreasing and approaches the empty set $\varnothing$, so

$$F_{X,Y}(x_1, -\infty) = \lim_{n\to\infty} F_{X,Y}(x_1, -n) = P[\varnothing] = 0.$$

For property (iii) we take the sequence $\{x \le x_1 \text{ and } y \le n\}$ which increases to $\{x \le x_1\}$, so

$$\lim_{n\to\infty} F_{X,Y}(x_1, n) = P[X \le x_1] = F_X(x_1).$$

For property (v) note in Fig. 5.9(a) that $B = \{x_1 < x \le x_2, y \le y_1\} = \{X \le x_2, Y \le y_1\} - \{X \le x_1, Y \le y_1\}$, so $P[B] = P[x_1 < X \le x_2, Y \le y_1] = F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_1)$. In Fig. 5.9(b), note that $F_{X,Y}(x_2, y_2) = P[A] + P[B] + F_{X,Y}(x_1, y_2)$. Property (v) follows by solving for $P[A]$ and substituting the expression for $P[B]$.

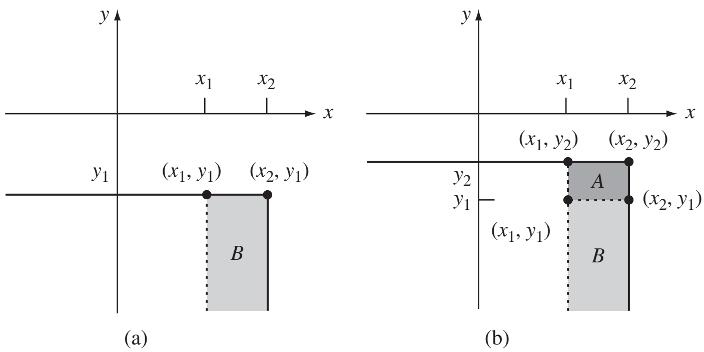


FIGURE 5.9
The joint cdf can be used to determine the probability of various events.

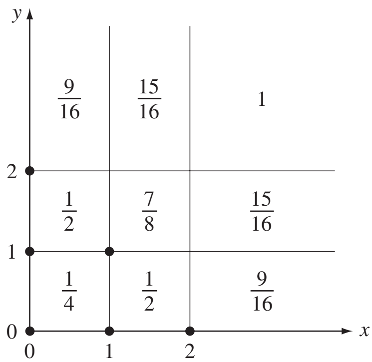


FIGURE 5.10
Joint cdf for packet switch example.

Example 5.10

Plot the joint cdf of X and Y from Example 5.6. Find the marginal cdf of X.

To find the cdf of $\mathbf{X}$, we identify the regions in the plane according to which points in $S_{X,Y}$ are included in the rectangular region defined by (x, y). For example,

- The regions outside the first quadrant do not include any of the points, so $F_{X,Y}(x, y) = 0$.
- The region $\{0 \le x < 1, 0 \le y < 1\}$ contains the point $(0,0)$, so $F_{X,Y}(x, y) = 1/4$.

Figure 5.10 shows the cdf after all possible regions are examined.

We need to consider several cases to find $F_X(x)$. For $x < 0$, we have $F_X(x) = 0$. For $0 \le x < 1$, we have $F_X(x) = F_{X,Y}(x, \infty) = 9/16$. For $1 \le x < 2$, we have $F_X(x) = F_{X,Y}(x, \infty) = 15/16$. Finally, for $x \ge 1$, we have $F_X(x) = F_{X,Y}(x, \infty) = 1$. Therefore $F_X(x)$ is a staircase function and X is a discrete random variable with $p_X(0) = 9/16$, $p_X(1) = 6/16$, and $p_X(2) = 1/16$.

Example 5.11

The joint cdf for the pair of random variables $\mathbf{X} = (X, Y)$ is given by

$$F_{X,Y}(x, y) = \begin{cases} 0 & x < 0 \text{ or } y < 0 \\ xy & 0 \le x \le 1, 0 \le y \le 1 \\ x & 0 \le x \le 1, y > 1 \\ y & 0 \le y \le 1, x > 1 \\ 1 & x \ge 1, y \ge 1. \end{cases} \tag{5.10}$$

Plot the joint cdf and find the marginal cdf of X.

Figure 5.11 shows a plot of the joint cdf of X and Y. $F_{X,Y}(x, y)$ is continuous for all points in the plane. $F_{X,Y}(x, y) = 1$ for all $x \ge 1$ and $y \ge 1$, which implies that X and Y each assume values less than or equal to one.

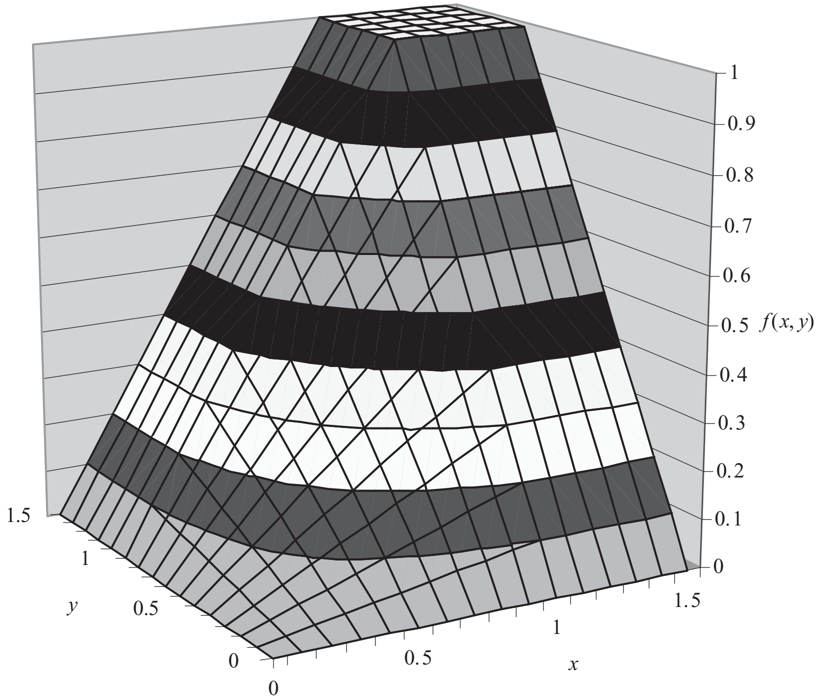


FIGURE 5.11
Joint cdf for two uniform random variables.

The marginal cdf of X is:

$$F_X(x) = F_{X,Y}(x, \infty) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x \ge 1. \end{cases}$$

X is uniformly distributed in the unit interval.

Example 5.12

The joint cdf for the vector of random variable $\mathbf{X} = (X, Y)$ is given by

$$F_{X,Y}(x, y) = \begin{cases} (1 - e^{-\alpha x})(1 - e^{-\beta y}) & x \ge 0, y \ge 0 \\ 0 & \text{elsewhere.} \end{cases}$$

Find the marginal cdf's.

The marginal cdf's are obtained by letting one of the variables approach infinity:

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x, y) = 1 - e^{-\alpha x} \quad x \ge 0$$

$$F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x, y) = 1 - e^{-\beta y} \quad y \ge 0.$$

X and Y individually have exponential distributions with parameters α and β, respectively.

Example 5.13

Find the probability of the events $A = \{X \le 1, Y \le 1\}$, $B = \{X > x, Y > y\}$, where $x > 0$ and $y > 0$, and $D = \{1 < X \le 2, 2 < Y \le 5\}$ in Example 5.12.

The probability of A is given directly by the cdf:

$$P[A] = P[X \le 1, Y \le 1] = F_{X,Y}(1, 1) = (1 - e^{-\alpha})(1 - e^{-\beta}).$$

The probability of B requires more work. By DeMorgan's rule:

$$B^c = (\{X > x\} \cap \{Y > y\})^c = \{X \le x\} \cup \{Y \le y\}.$$

Corollary 5 in Section 2.2 gives the probability of the union of two events:

$$\begin{aligned} P[B^c] &= P[X \le x] + P[Y \le y] - P[X \le x, Y \le y] \\ &= (1 - e^{-\alpha x}) + (1 - e^{-\beta y}) - (1 - e^{-\alpha x})(1 - e^{-\beta y}) \\ &= 1 - e^{-\alpha x}e^{-\beta y}. \end{aligned}$$

Finally we obtain the probability of B:

$$P[B] = 1 - P[B^c] = e^{-\alpha x}e^{-\beta y}.$$

You should sketch the region B on the plane and identify the events involved in the calculation of the probability of B^c.

The probability of event D is found by applying property (vi) of the joint cdf:

$$\begin{aligned} &P[1 < X \le 2, 2 < Y \le 5] \\ &\quad = F_{X,Y}(2, 5) - F_{X,Y}(2, 2) - F_{X,Y}(1, 5) + F_{X,Y}(1, 2) \\ &\quad = (1 - e^{-2\alpha})(1 - e^{-5\beta}) - (1 - e^{-2\alpha})(1 - e^{-2\beta}) \\ &\qquad -(1 - e^{-\alpha})(1 - e^{-5\beta}) + (1 - e^{-\alpha})(1 - e^{-2\beta}). \end{aligned}$$

5.3.1 Random Variables That Differ in Type

In some problems it is necessary to work with joint random variables that differ in type, that is, one is discrete and the other is continuous. Usually it is rather clumsy to work with the joint cdf, and so it is preferable to work with either $P[X = k, Y \le y]$ or $P[X = k, y_1 < Y \le y_2]$. These probabilities are sufficient to compute the joint cdf should we have to.

Example 5.14 Communication Channel with Discrete Input and Continuous Output

The input X to a communication channel is $+1$ volt or -1 volt with equal probability. The output Y of the channel is the input plus a noise voltage N that is uniformly distributed in the interval from -2 volts to $+2$ volts. Find $P[X = +1, Y \le 0]$.

This problem lends itself to the use of conditional probability:

$$P[X = +1, Y \le y] = P[Y \le y \mid X = +1]P[X = +1],$$

where $P[X = +1] = 1/2$. When the input $X = 1$, the output Y is uniformly distributed in the interval $[-1, 3]$; therefore

$$P[Y \le y \mid X = +1] = \frac{y+1}{4} \quad \text{for } -1 \le y \le 3.$$

Thus $P[X = +1, Y \le 0] = P[Y \le 0 \mid X = +1]P[X = +1] = (1/2)(1/4) = 1/8$.

5.4 THE JOINT PDF OF TWO CONTINUOUS RANDOM VARIABLES

The joint cdf allows us to compute the probability of events that correspond to "rectangular" shapes in the plane. To compute the probability of events corresponding to regions other than rectangles, we note that any reasonable shape (i.e., disk, polygon, or half-plane) can be approximated by the union of disjoint infinitesimal rectangles, $B_{j,k}$. For example, Fig. 5.12 shows how the events $A = \{X + Y \le 1\}$ and $B = \{X^2 + X^2 \le 1\}$ are approximated by rectangles of infinitesimal width. The probability of such events can therefore be approximated by the sum of the probabilities of infinitesimal rectangles, and if the cdf is sufficiently smooth, the probability of each rectangle can be expressed in terms of a density function:

$$P[B] \approx \sum_j \sum_k P[B_{j,k}] = \sum_{(x_j,\, y_k) \in B} \sum f_{X,Y}(x_j, y_k)\, \Delta x \Delta y.$$

As Δx and Δy approach zero, the above equation becomes an integral of a probability density function over the region B.

We say that the random variables **X and Y are jointly continuous** if the probabilities of events involving (X, Y) can be expressed as an integral of a probability density function. In other words, there is a nonnegative function $f_{X,Y}(x, y)$, called the **joint**

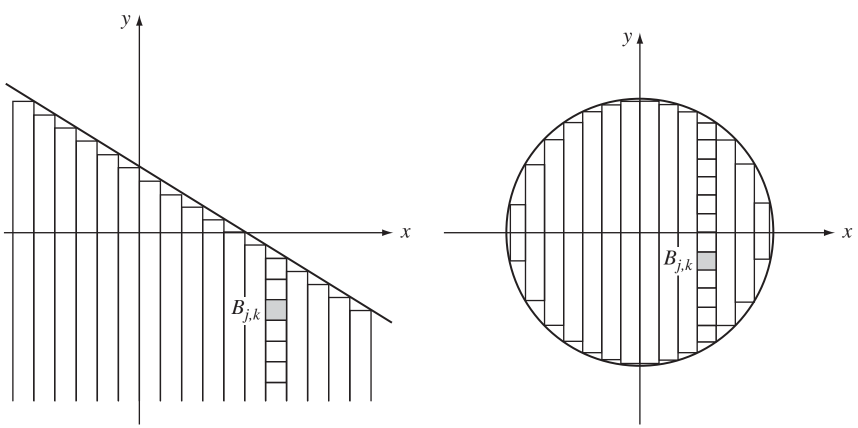


FIGURE 5.12
Some two-dimensional non-product form events.

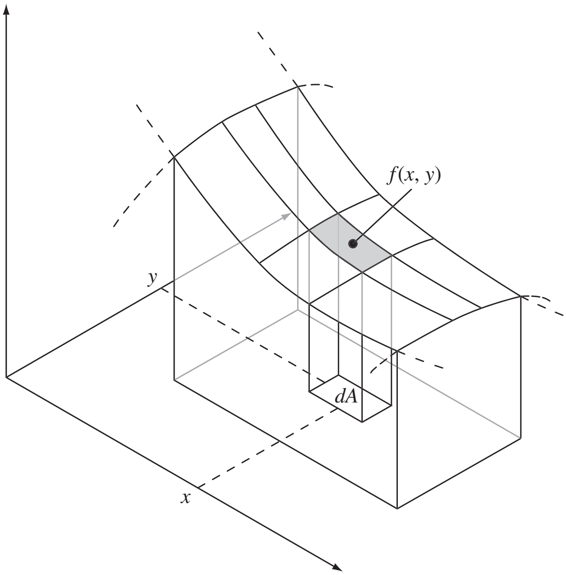


FIGURE 5.13
The probability of A is the integral of $f_{X,Y}(x, y)$ over the region defined by A.

probability density function, that is defined on the real plane such that for every event B, a subset of the plane,

$$P[\mathbf{X} \text{ in } B] = \int_B\int f_{X,Y}(x', y')\, dx'\, dy', \tag{5.11}$$

as shown in Fig. 5.13. Note the similarity to Eq. (5.5) for discrete random variables. When B is the entire plane, the integral must equal one:

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x', y')\, dx'\, dy'. \tag{5.12}$$

Equations (5.11) and (5.12) again suggest that the probability "mass" of an event is found by integrating the density of probability mass over the region corresponding to the event.

The joint cdf can be obtained in terms of the joint pdf of jointly continuous random variables by integrating over the semi-infinite rectangle defined by (x, y):

$$F_{X,Y}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(x', y')\, dx'\, dy'. \tag{5.13}$$

It then follows that *if* X and Y are jointly continuous random variables, then the pdf can be obtained from the cdf by differentiation:

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x\, \partial y}. \tag{5.14}$$

Note that if X and Y are not jointly continuous, then it is possible that the above partial derivative does not exist. In particular, if the $F_{X,Y}(x, y)$ is discontinuous or if its partial derivatives are discontinuous, then the joint pdf as defined by Eq. (5.14) will not exist.

The probability of a rectangular region is obtained by letting $B = \{(x, y): a_1 < x \leq b_1 \text{ and } a_2 < y \leq b_2\}$ in Eq. (5.11):

$$P[a_1 < X \leq b_1, a_2 < Y \leq b_2] = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f_{X,Y}(x', y')\, dx'\, dy'. \tag{5.15}$$

It then follows that the probability of an infinitesimal rectangle is the product of the pdf and the area of the rectangle:

$$\begin{aligned} P[x < X \leq x + dx, y < Y \leq y + dy] &= \int_{x}^{x+dx} \int_{y}^{y+dy} f_{X,Y}(x', y')\, dx'\, dy' \\ &\simeq f_{X,Y}(x, y)\, dx\, dy. \end{aligned} \tag{5.16}$$

Equation (5.16) can be interpreted as stating that the joint pdf specifies the probability of the product-form events

$$\{x < X \leq x + dx\} \cap \{y < Y \leq y + dy\}.$$

The **marginal pdf's** $f_X(x)$ and $f_Y(y)$ are obtained by taking the derivative of the corresponding marginal cdf's, $F_X(x) = F_{X,Y}(x,\infty)$ and $F_Y(y) = F_{X,Y}(\infty, y)$. Thus

$$\begin{aligned} f_X(x) &= \frac{d}{dx} \int_{-\infty}^{x} \left\{ \int_{-\infty}^{\infty} f_{X,Y}(x', y')\, dy' \right\} dx' \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x,y')\, dy'. \end{aligned} \tag{5.17a}$$

Similarly,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x', y)\, dx'. \tag{5.17b}$$

Thus the marginal pdf's are obtained by integrating out the variables that are not of interest.

Note that $f_X(x)\, dx \simeq P[x < X \leq x + dx, Y < \infty]$ is the probability of the infinitesimal strip shown in Fig. 5.14(a). This reminds us of the interpretation of the marginal pmf's as the probabilities of columns and rows in the case of discrete random variables. It is not surprising then that Eqs. (5.17a) and (5.17b) for the marginal pdf's and Eqs. (5.7a) and (5.7b) for the marginal pmf's are identical except for the fact that one contains an integral and the other a summation. As in the case of pmf's, we note that, in general, the joint pdf cannot be obtained from the marginal pdf's.

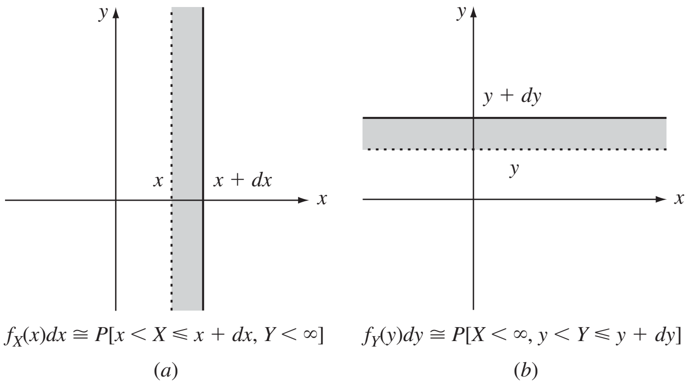


FIGURE 5.14
Interpretation of marginal pdf's.

Example 5.15 Jointly Uniform Random Variables

A randomly selected point (X, Y) in the unit square has the uniform joint pdf given by

$$f_{X,Y}(x, y) = \begin{cases} 1 & 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1 \\ 0 & \text{elsewhere.} \end{cases}$$

The scattergram in Fig. 5.3(a) corresponds to this pair of random variables. Find the joint cdf of X and Y.

The cdf is found by evaluating Eq. (5.13). You must be careful with the limits of the integral: The limits should define the region consisting of the intersection of the semi-infinite rectangle defined by (x, y) and the region where the pdf is nonzero. There are five cases in this problem, corresponding to the five regions shown in Fig. 5.15.

1. If $x < 0$ or $y < 0$, the pdf is zero and Eq. (5.14) implies

$$F_{X,Y}(x, y) = 0.$$

2. If (x, y) is inside the unit interval,

$$F_{X,Y}(x, y) = \int_0^x \int_0^y 1\, dx'\, dy' = xy.$$

3. If $0 \leq x \leq 1$ and $y > 1$,

$$F_{X,Y}(x, y) = \int_0^x \int_0^1 1\, dx'\, dy' = x.$$

4. Similarly, if $x > 1$ and $0 \leq y \leq 1$,

$$F_{X,Y}(x, y) = y.$$

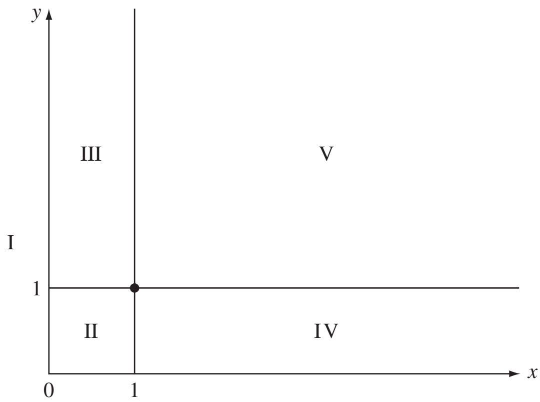


FIGURE 5.15
Regions that need to be considered separately in computing cdf in Example 5.15.

5. Finally, if $x > 1$ and $y > 1$,

$$F_{X,Y}(x, y) = \int_0^1 \int_0^1 1\, dx'\, dy' = 1.$$

We see that this is the joint cdf of Example 5.11.

Example 5.16

Find the normalization constant c and the marginal pdf's for the following joint pdf:

$$f_{X,Y}(x, y) = \begin{cases} ce^{-x}e^{-y} & 0 \le y \le x < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

The pdf is nonzero in the shaded region shown in Fig. 5.16(a). The constant c is found from the normalization condition specified by Eq. (5.12):

$$1 = \int_0^\infty \int_0^x ce^{-x}e^{-y}\, dy\, dx = \int_0^\infty ce^{-x}(1 - e^{-x})\, dx = \frac{c}{2}.$$

Therefore $c = 2$. The marginal pdf's are found by evaluating Eqs. (5.17a) and (5.17b):

$$f_X(x) = \int_0^\infty f_{X,Y}(x, y)\, dy = \int_0^x 2e^{-x}e^{-y}\, dy = 2e^{-x}(1 - e^{-x}) \qquad 0 \le x < \infty$$

and

$$f_Y(y) = \int_0^\infty f_{X,Y}(x, y)\, dx = \int_y^\infty 2e^{-x}e^{-y}\, dx = 2e^{-2y} \qquad 0 \le y < \infty.$$

You should fill in the steps in the evaluation of the integrals as well as verify that the marginal pdf's integrate to 1.

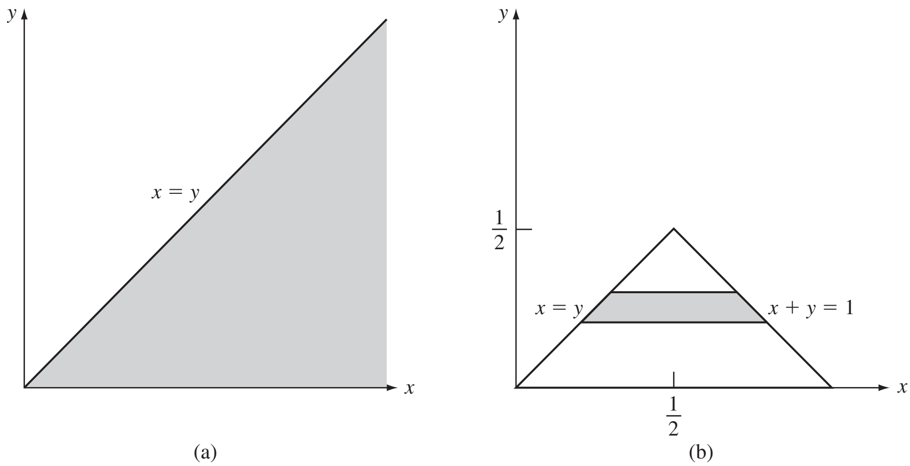


FIGURE 5.16
The random variables X and Y in Examples 5.16 and 5.17 have a pdf that is nonzero only in the shaded region shown in part (a).

Example 5.17

Find $P[X + Y \le 1]$ in Example 5.16.

Figure 5.16(b) shows the intersection of the event $\{X + Y \le 1\}$ and the region where the pdf is nonzero. We obtain the probability of the event by "adding" (actually integrating) infinitesimal rectangles of width dy as indicated in the figure:

$$P[X + Y \le 1] = \int_0^{.5} \int_y^{1-y} 2e^{-x}e^{-y}\,dx\,dy = \int_0^{.5} 2e^{-y}[e^{-y} - e^{-(1-y)}]\,dy$$
$$= 1 - 2e^{-1}.$$

Example 5.18 Jointly Gaussian Random Variables

The joint pdf of X and Y, shown in Fig. 5.17, is

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)} \qquad -\infty < x, y < \infty. \tag{5.18}$$

We say that X and Y are jointly Gaussian.[1] Find the marginal pdf's.

The marginal pdf of X is found by integrating $f_{X,Y}(x, y)$ over y:

$$f_X(x) = \frac{e^{-x^2/2(1-\rho^2)}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-(y^2-2\rho xy)/2(1-\rho^2)}\,dy.$$

[1]This is an important special case of jointly Gaussian random variables. The general case is discussed in Section 5.9.

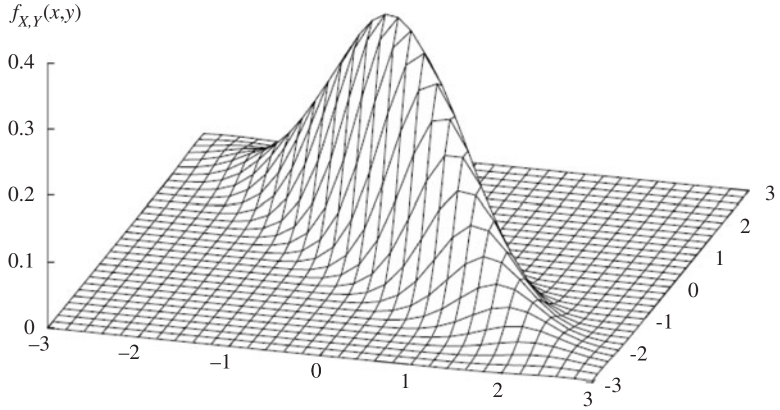


FIGURE 5.17
Joint pdf of two jointly Gaussian random variables.

We complete the square of the argument of the exponent by adding and subtracting $\rho^2 x^2$, that is, $y^2 - 2\rho xy + \rho^2 x^2 - \rho^2 x^2 = (y - \rho x)^2 - \rho^2 x^2$. Therefore

$$
\begin{aligned}
f_X(x) &= \frac{e^{-x^2/2(1-\rho^2)}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-[(y-\rho x)^2 - \rho^2 x^2]/2(1-\rho^2)}\, dy \\
&= \frac{e^{-x^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-(y-\rho x)^2/2(1-\rho^2)}}{\sqrt{2\pi(1-\rho^2)}}\, dy \\
&= \frac{e^{-x^2/2}}{\sqrt{2\pi}},
\end{aligned}
$$

where we have noted that the last integral equals one since its integrand is a Gaussian pdf with mean ρx and variance $1 - \rho^2$. The marginal pdf of X is therefore a one-dimensional Gaussian pdf with mean 0 and variance 1. From the symmetry of $f_{X,Y}(x, y)$ in x and y, we conclude that the marginal pdf of Y is also a one-dimensional Gaussian pdf with zero mean and unit variance.

5.5 INDEPENDENCE OF TWO RANDOM VARIABLES

***X* and *Y* are independent random variables** if *any* event A_1 defined in terms of X is independent of *any* event A_2 defined in terms of Y; that is,

$$P[X \text{ in } A_1, Y \text{ in } A_2] = P[X \text{ in } A_1]P[Y \text{ in } A_2]. \tag{5.19}$$

In this section we present a simple set of conditions for determining when X and Y are independent.

Suppose that X and Y are a pair of discrete random variables, and suppose we are interested in the probability of the event $A = A_1 \cap A_2$, where A_1 involves only X and A_2 involves only Y. In particular, if X and Y are independent, then A_1 and A_2 are independent events. If we let $A_1 = \{X = x_j\}$ and $A_2 = \{Y = y_k\}$, then the independence of X and Y implies that

$$\begin{aligned} p_{X,Y}(x_j, y_k) &= P[X = x_j, Y = y_k] \\ &= P[X = x_j]P[Y = y_k] \\ &= p_X(x_j)p_Y(y_k) \qquad \text{for all } x_j \text{ and } y_k. \end{aligned} \tag{5.20}$$

Therefore, *if X and Y are independent discrete random variables, then the joint pmf is equal to the product of the marginal pmf's.*

Now suppose that we don't know if X and Y are independent, but we do know that the pmf satisfies Eq. (5.20). Let $A = A_1 \cap A_2$ be a product-form event as above, then

$$\begin{aligned} P[A] &= \sum_{x_j \text{ in } A_1} \sum_{y_k \text{ in } A_2} p_{X,Y}(x_j, y_k) \\ &= \sum_{x_j \text{ in } A_1} \sum_{y_k \text{ in } A_2} p_X(x_j)p_Y(y_k) \\ &= \sum_{x_j \text{ in } A_1} p_X(x_j) \sum_{y_k \text{ in } A_2} p_Y(y_k) \\ &= P[A_1]P[A_2], \end{aligned} \tag{5.21}$$

which implies that A_1 and A_2 are independent events. Therefore, *if the joint pmf of X and Y equals the product of the marginal pmf's, then X and Y are independent.* We have just proved that the statement "X and Y are independent" is equivalent to the statement "the joint pmf is equal to the product of the marginal pmf's." In mathematical language, we say, the "*discrete random variables X and Y are independent if and only if the joint pmf is equal to the product of the marginal pmf's for all x_j, y_k.*"

Example 5.19

Is the pmf in Example 5.6 consistent with an experiment that consists of the independent tosses of two fair dice?

The probability of each face in a toss of a fair die is 1/6. If two fair dice are tossed and if the tosses are independent, then the probability of any pair of faces, say j and k, is:

$$P[X = j, Y = k] = P[X = j]P[Y = k] = \frac{1}{36}.$$

Thus all possible pairs of outcomes should be equiprobable. This is not the case for the joint pmf given in Example 5.6. Therefore the tosses in Example 5.6 are not independent.

Example 5.20

Are Q and R in Example 5.9 independent? From Example 5.9 we have

$$\begin{aligned} P[Q = q]P[R = r] &= (1 - p^M)(p^M)^q \frac{(1 - p)}{1 - p^M} p^r \\ &= (1 - p)p^{Mq+r} \end{aligned}$$

$$= P[Q = q, R = r] \qquad \text{for all } q = 0, 1, \ldots \quad r = 0, \ldots, M - 1.$$

Therefore Q and R are independent.

In general, it can be shown that the random variables *X and Y are independent if and only if their joint cdf is equal to the product of its marginal cdf's*:

$$F_{X,Y}(x, y) = F_X(x)F_Y(y) \qquad \text{for all } x \text{ and } y. \tag{5.22}$$

Similarly, if X and Y are jointly continuous, then *X and Y are independent if and only if their joint pdf is equal to the product of the marginal pdf's*:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) \qquad \text{for all } x \text{ and } y. \tag{5.23}$$

Equation (5.23) is obtained from Eq. (5.22) by differentiation. Conversely, Eq. (5.22) is obtained from Eq. (5.23) by integration.

Example 5.21

Are the random variables X and Y in Example 5.16 independent?

Note that $f_X(x)$ and $f_Y(y)$ are nonzero for all $x > 0$ and all $y > 0$. Hence $f_X(x)f_Y(y)$ is nonzero in the entire positive quadrant. However $f_{X,Y}(x, y)$ is nonzero only in the region $y < x$ inside the positive quadrant. Hence Eq. (5.23) does not hold for all x, y and the random variables are not independent. You should note that in this example the joint pdf appears to factor, but nevertheless it is not the product of the marginal pdf's.

Example 5.22

Are the random variables X and Y in Example 5.18 independent? The product of the marginal pdf's of X and Y in Example 5.18 is

$$f_X(x)f_Y(y) = \frac{1}{2\pi}e^{-(x^2+y^2)/2} \qquad -\infty < x, y < \infty.$$

By comparing to Eq. (5.18) we see that the product of the marginals is equal to the joint pdf if and only if $\rho = 0$. Therefore the jointly Gaussian random variables X and Y are independent if and only if $\rho = 0$. We see in a later section that ρ is the *correlation coefficient* between X and Y.

Example 5.23

Are the random variables X and Y independent in Example 5.12? If we multiply the marginal cdf's found in Example 5.12 we find

$$F_X(x)F_Y(y) = (1 - e^{-\alpha x})(1 - e^{-\beta y}) = F_{X,Y}(x, y) \qquad \text{for all } x \text{ and } y.$$

Therefore Eq. (5.22) is satisfied so X and Y are independent.

If X and Y are independent random variables, then the random variables defined by any pair of functions $g(X)$ and $h(Y)$ are also independent. To show this, consider the one-dimensional events A and B. Let A' be the set of all values of x such that if x is in A' then $g(x)$ is in A, and let B' be the set of all values of y such that if y is in B' then $h(y)$ is in B. (In Chapter 3 we called A' and B' the equivalent events of A and B.) Then

$$\begin{aligned} P[g(X) \text{ in } A, h(Y) \text{ in } B] &= P[X \text{ in } A', Y \text{ in } B'] \\ &= P[X \text{ in } A']P[Y \text{ in } B'] \\ &= P[g(X) \text{ in } A]P[h(Y) \text{ in } B]. \end{aligned} \tag{5.24}$$

The first and third equalities follow from the fact that A and A' and B and B' are equivalent events. The second equality follows from the independence of X and Y. Thus $g(X)$ and $h(Y)$ are independent random variables.

5.6 JOINT MOMENTS AND EXPECTED VALUES OF A FUNCTION OF TWO RANDOM VARIABLES

The expected value of X identifies the center of mass of the distribution of X. The variance, which is defined as the expected value of $(X - m)^2$, provides a measure of the spread of the distribution. In the case of two random variables we are interested in how X and Y vary together. In particular, we are interested in whether the variation of X and Y are correlated. For example, if X increases does Y tend to increase or to decrease? The joint moments of X and Y, which are defined as expected values of functions of X and Y, provide this information.

5.6.1 Expected Value of a Function of Two Random Variables

The problem of finding the expected value of a function of two or more random variables is similar to that of finding the expected value of a function of a single random variable. It can be shown that the expected value of $Z = g(X, Y)$ can be found using the following expressions:

$$E[Z] = \begin{cases} \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)\,dx\,dy & X, Y \text{ jointly continuous} \\ \displaystyle\sum_i \sum_n g(x_i, y_n) p_{X,Y}(x_i, y_n) & X, Y \text{ discrete.} \end{cases} \tag{5.25}$$

Example 5.24 Sum of Random Variables

Let $Z = X + Y$. Find $E[Z]$.

$$\begin{aligned} E[Z] &= E[X + Y] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x' + y') f_{X,Y}(x', y')\,dx'\,dy' \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x' f_{X,Y}(x', y')\,dy'\,dx' + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y' f_{X,Y}(x', y')\,dx'\,dy' \\ &= \int_{-\infty}^{\infty} x' f_X(x')\,dx' + \int_{-\infty}^{\infty} y' f_Y(y')\,dy' = E[X] + E[Y]. \end{aligned} \tag{5.26}$$

Thus the expected value of the sum of two random variables is equal to the sum of the individual expected values. Note that X and Y need not be independent.

The result in Example 5.24 and a simple induction argument show that *the expected value of a sum of n random variables is equal to the sum of the expected values:*

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + \cdots + E[X_n]. \tag{5.27}$$

Note that the random variables do not have to be independent.

Example 5.25 Product of Functions of Independent Random Variables

Suppose that X and Y are independent random variables, and let $g(X, Y) = g_1(X)g_2(Y)$. Find $E[g(X, Y)] = E[g_1(X)g_2(Y)]$.

$$\begin{aligned} E[g_1(X)g_2(Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(x')g_2(y')f_X(x')f_Y(y')\,dx'\,dy' \\ &= \left\{\int_{-\infty}^{\infty} g_1(x')f_X(x')\,dx'\right\}\left\{\int_{-\infty}^{\infty} g_2(y')f_Y(y')\,dy'\right\} \\ &= E[g_1(X)]E[g_2(Y)]. \end{aligned}$$

5.6.2 Joint Moments, Correlation, and Covariance

The joint moments of two random variables X and Y summarize information about their joint behavior. The jkth **joint moment of X and Y** is defined by

$$E[X^jY^k] = \begin{cases} \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^jy^kf_{X,Y}(x, y)\,dx\,dy & X, Y \text{ jointly continuous} \\ \displaystyle\sum_i\sum_n x_i^jy_n^kp_{X,Y}(x_i, y_n) & X, Y \text{ discrete.} \end{cases} \tag{5.28}$$

If $j = 0$, we obtain the moments of Y, and if $k = 0$, we obtain the moments of X. In electrical engineering, it is customary to call the $j = 1$ $k = 1$ moment, $E[XY]$, the **correlation of X and Y**. If $E[XY] = 0$, then we say that **X and Y are orthogonal**.

The jkth **central moment of X and Y** is defined as the joint moment of the centered random variables, $X - E[X]$ and $Y - E[Y]$:

$$E[(X - E[X])^j(Y - E[Y])^k].$$

Note that $j = 2$ $k = 0$ gives VAR(X) and $j = 0$ $k = 2$ gives VAR(Y).

The **covariance of X and Y** is defined as the $j = k = 1$ central moment:

$$\text{COV}(X, Y) = E[(X - E[X])(Y - E[Y])]. \tag{5.29}$$

The following form for COV(X, Y) is sometimes more convenient to work with:

$$\text{COV}(X, Y) = E[XY - XE[Y] - YE[X] + E[X]E[Y]]$$

$$= E[XY] - 2E[X]E[Y] + E[X]E[Y]$$
$$= E[XY] - E[X]E[Y]. \tag{5.30}$$

Note that $\text{COV}(X, Y) = E[XY]$ if either of the random variables has mean zero.

Example 5.26 Covariance of Independent Random Variables

Let X and Y be independent random variables. Find their covariance.

$$\begin{aligned}\text{COV}(X, Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[X - E[X]]E[Y - E[Y]] \\ &= 0,\end{aligned}$$

where the second equality follows from the fact that X and Y are independent, and the third equality follows from $E[X - E[X]] = E[X] - E[X] = 0$. Therefore *pairs of independent random variables have covariance zero*.

Let's see how the covariance measures the correlation between X and Y. The covariance measures the deviation from $m_X = E[X]$ and $m_Y = E[Y]$. If a positive value of $(X - m_X)$ tends to be accompanied by a positive values of $(Y - m_Y)$, and negative $(X - m_X)$ tend to be accompanied by negative $(Y - m_Y)$; then $(X - m_X)(Y - m_Y)$ will tend to be a positive value, and its expected value, COV(X, Y), will be positive. This is the case for the scattergram in Fig. 5.3(d) where the observed points tend to cluster along a line of positive slope. On the other hand, if $(X - m_X)$ and $(Y - m_Y)$ tend to have opposite signs, then COV(X, Y) will be negative. A scattergram for this case would have observation points cluster along a line of negative slope. Finally if $(X - m_X)$ and $(Y - m_Y)$ sometimes have the same sign and sometimes have opposite signs, then COV(X, Y) will be close to zero. The three scattergrams in Figs. 5.3(a), (b), and (c) fall into this category.

Multiplying either X or Y by a large number will increase the covariance, so we need to normalize the covariance to measure the correlation in an absolute scale. The **correlation coefficient of X and Y** is defined by

$$\rho_{X,Y} = \frac{\text{COV}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[XY] - E[X]E[Y]}{\sigma_X \sigma_Y}, \tag{5.31}$$

where $\sigma_X = \sqrt{\text{VAR}(X)}$ and $\sigma_Y = \sqrt{\text{VAR}(Y)}$ are the standard deviations of X and Y, respectively.

The correlation coefficient is a number that is at most 1 in magnitude:

$$-1 \le \rho_{X,Y} \le 1. \tag{5.32}$$

To show Eq. (5.32), we begin with an inequality that results from the fact that the expected value of the square of a random variable is nonnegative:

$$0 \le E\left\{\left(\frac{X - E[X]}{\sigma_X} \pm \frac{Y - E[Y]}{\sigma_Y}\right)^2\right\} \rightarrow \frac{(x - m_x)\sigma_y \pm (y - m_y)\sigma_x}{\sigma_x \sigma_y}$$

FALOPA

ESO AL cuad $\left|\frac{1}{\sigma_x \sigma_y}\right|^2 \left[(x-m_x)\sigma_y \pm (y-m_y)\sigma_y\right]^2$ → cuadrado del binomio y tal

$1+\rho > 0 \Rightarrow \rho > -1$

ó $1-\rho > 0 \Rightarrow \rho < 1$

$$= 1 \pm 2\rho_{X,Y} + 1$$
$$= 2(1 \pm \rho_{X,Y}). \geq 0$$

The last equation implies Eq. (5.32).

The extreme values of $\rho_{X,Y}$ are achieved when X and Y are related linearly, $Y = aX + b$; $\rho_{X,Y} = 1$ if $a > 0$ and $\rho_{X,Y} = -1$ if $a < 0$. In Section 6.5 we show that $\rho_{X,Y}$ can be viewed as a statistical measure of the extent to which Y can be predicted by a linear function of X.

X and Y are said to be **uncorrelated** if $\rho_{X,Y} = 0$. If X and Y are independent, then $\text{COV}(X, Y) = 0$, so $\rho_{X,Y} = 0$. Thus *if X and Y are independent, then X and Y are uncorrelated.* In Example 5.22, we saw that if *X and Y are jointly Gaussian and $\rho_{X,Y} = 0$, then X and Y are independent random variables.* Example 5.27 shows that this is not always true for non-Gaussian random variables: It is possible for X and Y to be uncorrelated but not independent.

Mejor: Ahora $Cov > 0 \rightarrow 0 < Cov < \sigma_x \sigma_y = \sqrt{V_x V_y}$ → ¿Media geométrica?

Example 5.27 **Uncorrelated but Dependent Random Variables**

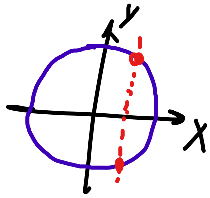
Let Θ be uniformly distributed in the interval $(0, 2\pi)$. Let

$$X = \cos\Theta \qquad \text{and} \qquad Y = \sin\Theta.$$

Se debe cumplir $x^2 + y^2 = 1$

The point (X, Y) then corresponds to the point on the unit circle specified by the angle Θ, as shown in Fig. 5.18. In Example 4.36, we saw that the marginal pdf's of X and Y are arcsine pdf's, which are nonzero in the interval $(-1, 1)$. The product of the marginals is nonzero in the square defined by $-1 \leq x \leq 1$ and $-1 \leq y \leq 1$, so if X and Y were independent the point (X, Y) would assume all values in this square. This is not the case, so X and Y are dependent.

We now show that X and Y are uncorrelated:



Si conozco un valor de X, Y solo puede tomar 2 valores posibles, no son indep

$$E[XY] = E[\sin\Theta\cos\Theta] = \frac{1}{2\pi}\int_0^{2\pi} \sin\phi\cos\phi \, d\phi$$
$$= \frac{1}{4\pi}\int_0^{2\pi} \sin 2\phi \, d\phi = 0. \rightarrow Cov = 0$$

Since $E[X] = E[Y] = 0$, Eq. (5.30) then implies that X and Y are uncorrelated.

Example 5.28

Let X and Y be the random variables discussed in Example 5.16. Find $E[XY]$, $\text{COV}(X, Y)$, and $\rho_{X,Y}$.

Equations (5.30) and (5.31) require that we find the mean, variance, and correlation of X and Y. From the marginal pdf's of X and Y obtained in Example 5.16, we find that $E[X] = 3/2$ and $\text{VAR}[X] = 5/4$, and that $E[Y] = 1/2$ and $\text{VAR}[Y] = 1/4$. The correlation of X and Y is

$$E[XY] = \int_0^{\infty}\int_0^{x} xy2e^{-x}e^{-y} \, dy \, dx$$
$$= \int_0^{\infty} 2xe^{-x}(1 - e^{-x} - xe^{-x}) \, dx = 1.$$

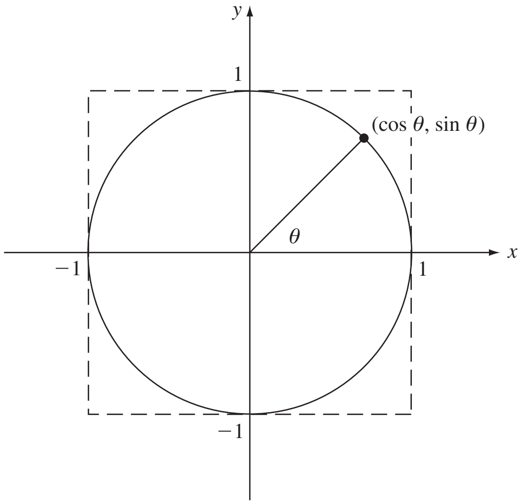


FIGURE 5.18
(X, Y) is a point selected at random *on* the unit circle. X and Y are uncorrelated but not independent.

Thus the correlation coefficient is given by

$$\rho_{X,Y} = \frac{1 - \frac{3}{2}\frac{1}{2}}{\sqrt{\frac{5}{4}}\sqrt{\frac{1}{4}}} = \frac{1}{\sqrt{5}}.$$

5.7 CONDITIONAL PROBABILITY AND CONDITIONAL EXPECTATION

Many random variables of practical interest are not independent: The output Y of a communication channel must depend on the input X in order to convey information; consecutive samples of a waveform that varies slowly are likely to be close in value and hence are not independent. In this section we are interested in computing the probability of events concerning the random variable Y given that we know $X = x$. We are also interested in the expected value of Y given $X = x$. We show that the notions of conditional probability and conditional expectation are extremely useful tools in solving problems, even in situations where we are only concerned with one of the random variables.

5.7.1 Conditional Probability

The definition of conditional probability in Section 2.4 allows us to compute the probability that Y is in A given that we know that $X = x$:

$$P[Y \text{ in } A \,|\, X = x] = \frac{P[Y \text{ in } A, X = x]}{P[X = x]} \quad \text{for } P[X = x] > 0. \tag{5.33}$$

Case 1: *X* Is a Discrete Random Variable

For X and Y discrete random variables, the **conditional pmf of *Y* given $X = x$** is defined by:

$$p_Y(y \mid x) = P[Y = y \mid X = x] = \frac{P[X = x, Y = y]}{P[X = x]} = \frac{p_{X,Y}(x, y)}{p_X(x)} \tag{5.34}$$

for x such that $P[X = x] > 0$. We define $p_Y(y \mid x) = 0$ for x such that $P[X = x] = 0$. Note that $p_Y(y \mid x)$ is a function of y over the real line, and that $p_Y(y \mid x) > 0$ only for y in a discrete set $\{y_1, y_2, \dots\}$.

The conditional pmf satisfies all the properties of a pmf, that is, it assigns non-negative values to every y and these values add to 1. Note from Eq. (5.34) that $p_Y(y \mid x_k)$ is simply the cross section of $p_{X,Y}(x_k, y)$ along the $X = x_k$ column in Fig. 5.6, but normalized by the probability $p_X(x_k)$.

The probability of an event A given $X = x_k$ is found by adding the pmf values of the outcomes in A:

$$P[Y \text{ in } A \mid X = x_k] = \sum_{y_j \text{ in } A} p_Y(y_j \mid x_k). \tag{5.35}$$

If *X and Y are independent*, then using Eq (5.20)

$$p_Y(y_j \mid x_k) = \frac{P[X = x_k, Y = y_j]}{P[X = x_k]} = P[Y = y_j] = p_Y(y_j). \tag{5.36}$$

In other words, knowledge that $X = x_k$ does not affect the probability of events A involving Y.

Equation (5.34) implies that the joint pmf $p_{X,Y}(x, y)$ can be expressed as the product of a conditional pmf and a marginal pmf:

$$p_{X,Y}(x_k, y_j) = p_Y(y_j \mid x_k) p_X(x_k) \text{ and } p_{X,Y}(x_k, y_j) = p_X(x_k \mid y_j) p_Y(y_j). \tag{5.37}$$

This expression is very useful when we can view the pair (X, Y) as being generated sequentially, e.g., first X, and then Y given $X = x$. We find the probability that Y is in A as follows:

$$\begin{aligned} P[Y \text{ in } A] &= \sum_{\text{all } x_k} \sum_{y_j \text{ in } A} p_{X,Y}(x_k, y_j) \\ &= \sum_{\text{all } x_k} \sum_{y_j \text{ in } A} p_Y(y_j \mid x_k) p_X(x_k) \\ &= \sum_{\text{all } x_k} p_X(x_k) \sum_{y_j \text{ in } A} p_Y(y_j \mid x_k) \\ &= \sum_{\text{all } x_k} P[Y \text{ in } A \mid X = x_k] p_X(x_k). \end{aligned} \tag{5.38}$$

Equation (5.38) is simply a restatement of the theorem on total probability discussed in Chapter 2. In other words, to compute $P[Y \text{ in } A]$ we can first compute $P[Y \text{ in } A \mid X = x_k]$ and then "average" over X_k.

Example 5.29 Loaded Dice

Find $p_Y(y \mid 5)$ in the loaded dice experiment considered in Examples 5.6 and 5.8.

In Example 5.8 we found that $p_X(5) = 1/6$. Therefore:

$$p_Y(y \mid 5) = \frac{p_{X,Y}(5, y)}{p_X(5)} \text{ and so } p_Y(5 \mid 5) = 2/7 \text{ and}$$

$$p_Y(1 \mid 5) = p_Y(2 \mid 5) = p_Y(3 \mid 5) = p_Y(4 \mid 5) = p_Y(6 \mid 5) = 1/7.$$

Clearly this die is loaded.

Example 5.30 Number of Defects in a Region; Random Splitting of Poisson Counts

The total number of defects X on a chip is a Poisson random variable with mean α. Each defect has a probability p of falling in a specific region R and the location of each defect is independent of the locations of other defects. Find the pmf of the number of defects Y that fall in the region R.

We can imagine performing a Bernoulli trial each time a defect occurs with a "success" occurring when the defect falls in the region R. If the total number of defects is $X = k$, then Y is a binomial random variable with parameters k and p:

$$p_Y(j \mid k) = \begin{cases} 0 & j > k \\ \binom{k}{j} p^j (1 - p)^{k-j} & 0 \le j \le k. \end{cases}$$

From Eq. (5.38) and noting that $k \ge j$, we have

$$\begin{aligned} p_Y(j) &= \sum_{k=0}^{\infty} p_Y(j \mid k) p_X(k) = \sum_{k=j}^{\infty} \frac{k!}{j!(k - j)!} p^j (1 - p)^{k-j} \frac{\alpha^k}{k!} e^{-\alpha} \\ &= \frac{(\alpha p)^j e^{-\alpha}}{j!} \sum_{k=j}^{\infty} \frac{\{(1 - p)\alpha\}^{k-j}}{(k - j)!} \\ &= \frac{(\alpha p)^j e^{-\alpha}}{j!} e^{(1-p)\alpha} = \frac{(\alpha p)^j}{j!} e^{-\alpha p}. \end{aligned}$$

Thus Y is a Poisson random variable with mean αp.

Suppose Y is a continuous random variable. Eq. (5.33) can be used to define the **conditional cdf of Y given $X = x_k$**:

$$F_Y(y \mid x_k) = \frac{P[Y \le y, X = x_k]}{P[X = x_k]}, \qquad \text{for } P[X = x_k] > 0. \tag{5.39}$$

It is easy to show that $F_Y(y \mid x_k)$ satisfies all the properties of a cdf. The **conditional pdf of Y given $X = x_k$**, if the derivative exists, is given by

$$f_Y(y \mid x_k) = \frac{d}{dy} F_Y(y \mid x_k). \tag{5.40}$$

If X and Y are independent, $P[Y \le y, X = X_k] = P[Y \le y]P[X = X_k]$ so $F_Y(y \mid x) = F_Y(y)$ and $f_Y(y \mid x) = f_Y(y)$. The probability of event A given $X = x_k$ is obtained by integrating the conditional pdf:

$$P[Y \text{ in } A \mid X = x_k] = \int_{y \text{ in } A} f_Y(y \mid x_k)\, dy. \tag{5.41}$$

We obtain $P[Y \text{ in } A]$ using Eq. (5.38).

Example 5.31 Binary Communications System

The input X to a communication channel assumes the values $+1$ or -1 with probabilities 1/3 and 2/3. The output Y of the channel is given by $Y = X + N$, where N is a zero-mean, unit variance Gaussian random variable. Find the conditional pdf of Y given $X = +1$, and given $X = -1$. Find $P[X = +1 \mid Y > 0]$.

The conditional cdf of Y given $X = +1$ is:

$$\begin{aligned} F_Y(y \mid +1) &= P[Y \le y \mid X = +1] = P[N + 1 \le y] \\ &= P[N \le y - 1] = \int_{-\infty}^{y-1} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \end{aligned}$$

where we noted that if $X = +1$, then $Y = N + 1$ and Y depends only on N. Thus, if $X = +1$, then Y is a Gaussian random variable with mean 1 and unit variance. Similarly, if $X = -1$, then Y is Gaussian with mean -1 and unit variance.

The probabilities that $Y > 0$ given $X = +1$ and $X = -1$ is:

$$P[Y > 0 \mid X = +1] = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x-1)^2/2}\, dx = \int_{-1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt = 1 - Q(1) = 0.841.$$

$$P[Y > 0 \mid X = -1] = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x+1)^2/2}\, dx = \int_{1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt = Q(1) = 0.159.$$

Applying Eq. (5.38), we obtain:

$$P[Y > 0] = P[Y > 0 \mid X = +1]\frac{1}{3} + P[Y > 0 \mid X = -1]\frac{2}{3} = 0.386.$$

From Bayes' theorem we find:

$$P[X = +1 \mid Y > 0] = \frac{P[Y > 0 \mid X = +1]P[X = +1]}{P[Y > 0]} = \frac{(1 - Q(1))/3}{(1 + Q(1))/3} = 0.726.$$

We conclude that if $Y > 0$, then $X = +1$ is more likely than $X = -1$. Therefore the receiver should decide that the input is $X = +1$ when it observes $Y > 0$.

In the previous example, we made an interesting step that is worth elaborating on because it comes up quite frequently: $P[Y \le y \mid X = +1] = P[N + 1 \le y]$, where $Y = X + N$. Let's take a closer look:

$$P[Y \leq z \mid X = x] = \frac{P[\{X + N \leq z\} \cap \{X = x\}]}{P[X = x]} = \frac{P[\{x + N \leq z\} \cap \{X = x\}]}{P[X = x]}$$
$$= P[x + N \leq z \mid X = x] = P[N \leq z - x \mid X = x].$$

In the first line, the events $\{X + N \leq z\}$ and $\{x + N \leq z\}$ are quite different. The first involves the two random variables X and N, whereas the second only involves N and consequently is much simpler. We can then apply an expression such as Eq. (5.38) to obtain $P[Y \leq z]$. The step we made in the example, however, is even more interesting. Since X and N are independent random variables, we can take the expression one step further:

$$P[Y \leq z \mid X = x] = P[N \leq z - x \mid X = x] = P[N \leq z - x].$$

The independence of X and N allows us to dispense with the conditioning on x altogether!

Case 2: *X* Is a Continuous Random Variable

If X is a continuous random variable, then $P[X = x] = 0$ so Eq. (5.33) is undefined for all x. If X and Y have a joint pdf that is continuous and nonzero over some region of the plane, we define the **conditional cdf of Y given $X = x$** by the following limiting procedure:

$$F_Y(y \mid x) = \lim_{h \to 0} F_Y(y \mid x < X \leq x + h). \tag{5.42}$$

The conditional cdf on the right side of Eq. (5.42) is:

$$F_Y(y \mid x < X \leq x + h) = \frac{P[Y \leq y, x < X \leq x + h]}{P[x < X \leq x + h]}$$
$$= \frac{\displaystyle\int_{-\infty}^{y} \int_{x}^{x+h} f_{X,Y}(x', y')\, dx'\, dy'}{\displaystyle\int_{x}^{x+h} f_X(x')\, dx'} = \frac{\displaystyle\int_{-\infty}^{y} f_{X,Y}(x, y')\, dy' h}{f_X(x) h}. \tag{5.43}$$

As we let h approach zero, Eqs. (5.42) and (5.43) imply that

$$F_Y(y \mid x) = \frac{\displaystyle\int_{-\infty}^{y} f_{X,Y}(x, y')\, dy'}{f_X(x)}. \tag{5.44}$$

The **conditional pdf of Y given $X = x$** is then:

$$f_Y(y \mid x) = \frac{d}{dy} F_Y(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)}. \tag{5.45}$$

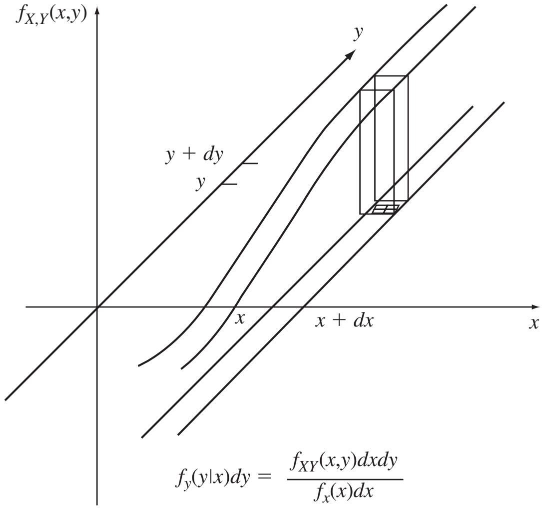


FIGURE 5.19
Interpretation of conditional pdf.

It is easy to show that $f_Y(y \mid x)$ satisfies the properties of a pdf. We can interpret $f_Y(y \mid x)\,dy$ as the probability that Y is in the infinitesimal strip defined by $(y, y + dy)$ given that X is in the infinitesimal strip defined by $(x, x + dx)$, as shown in Fig. 5.19.

The probability of event A given $X = x$ is obtained as follows:

$$P[Y \text{ in } A \mid X = x] = \int_{y \text{ in } A} f_Y(y \mid x)\, dy. \tag{5.46}$$

There is a strong resemblance between Eq. (5.34) for the discrete case and Eq. (5.45) for the continuous case. Indeed many of the same properties hold. For example, we obtain the multiplication rule from Eq. (5.45):

$$f_{X,Y}(x, y) = f_Y(y \mid x) f_X(x) \text{ and } f_{X,Y}(x, y) = f_X(x \mid y) f_Y(y). \tag{5.47}$$

If X and Y are independent, then $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ and $f_Y(y \mid x) = f_Y(y)$, $f_X(x \mid y) = f_X(\mathrm{x})$, $F_Y(y \mid x) = F_Y(y)$, and $F_X(x \mid y) = F_X(x)$.

By combining Eqs. (5.46) and (5.47), we can show that:

$$P[Y \text{ in } A] = \int_{-\infty}^{\infty} P[Y \text{ in } A \mid X = x] f_X(x)\, dx. \tag{5.48}$$

You can think of Eq. (5.48) as the "continuous" version of the theorem on total probability. The following examples show the usefulness of the above results in calculating the probabilities of complicated events.

Example 5.32

Let X and Y be the random variables in Example 5.8. Find $f_X(x \mid y)$ and $f_Y(y \mid x)$.

Using the marginal pdf's obtained in Example 5.8, we have

$$f_X(y \mid x) = \frac{2e^{-x}e^{-y}}{2e^{-2y}} = e^{-(x-y)} \qquad \text{for } x \geq y$$

$$f_Y(y \mid x) = \frac{2e^{-x}e^{-y}}{2e^{-x}(1 - e^{-x})} = \frac{e^{-y}}{1 - e^{-x}} \qquad \text{for } 0 < y < x.$$

The conditional pdf of X is an exponential pdf shifted by y to the right. The conditional pdf of Y is an exponential pdf that has been truncated to the interval $[0, x]$.

Example 5.33 Number of Arrivals During a Customer's Service Time

The number N of customers that arrive at a service station during a time t is a Poisson random variable with parameter βt. The time T required to service each customer is an exponential random variable with parameter α. Find the pmf for the number N that arrive during the service time T of a specific customer. Assume that the customer arrivals are independent of the customer service time.

Equation (5.48) holds even if Y is a discrete random variable, thus

$$\begin{aligned} P[N = k] &= \int_0^\infty P[N = k \mid T = t] f_T(t)\, dt \\ &= \int_0^\infty \frac{(\beta t)^k}{k!} e^{-\beta t} \alpha e^{-\alpha t}\, dt \\ &= \frac{\alpha \beta^k}{k!} \int_0^\infty t^k e^{-(\alpha+\beta)t}\, dt. \end{aligned}$$

Let $r = (\alpha + \beta)t$, then

$$\begin{aligned} P[N = k] &= \frac{\alpha \beta^k}{k!(\alpha + \beta)^{k+1}} \int_0^\infty r^k e^{-r}\, dr \\ &= \frac{\alpha \beta^k}{(\alpha + \beta)^{k+1}} = \left(\frac{\alpha}{(\alpha + \beta)}\right)\left(\frac{\beta}{(\alpha + \beta)}\right)^k, \end{aligned}$$

where we have used the fact that the last integral is a gamma function and is equal to $k!$. Thus N is a geometric random variable with probability of "success" $\alpha/(\alpha + \beta)$. Each time a customer arrives we can imagine that a new Bernoulli trial begins where "success" occurs if the customer's service time is completed before the next arrival.

Example 5.34

X is selected at random from the unit interval; Y is then selected at random from the interval$(0, X)$. Find the cdf of Y.

When $X = x$, Y is uniformly distributed in $(0, x)$ so the conditional cdf given $X = x$ is

$$P[Y \le y \,|\, X = k] = \begin{cases} y/x & 0 \le y \le x \\ 1 & x < y. \end{cases}$$

Equation (5.48) and the above conditional cdf yield:

$$F_Y(y) = P[Y \le y] = \int_0^1 P[Y \le y \,|\, X = x] f_X(x)\, dx =$$
$$= \int_0^y 1\, dx' + \int_y^1 \frac{y}{x'}\, dx' = y - y \ln y.$$

The corresponding pdf is obtained by taking the derivative of the cdf:

$$f_Y(y) = -\ln y \quad 0 \le y \le 1.$$

Example 5.35 Maximum A Posteriori Receiver

For the communications system in Example 5.31, find the probability that the input was $X = +1$ given that the output of the channel is $Y = y$.

This is a tricky version of Bayes' rule. Condition on the event $\{y < Y \le y + \Delta\}$ instead of $\{Y = y\}$:

$$P[X = +1 \,|\, y < Y < y + \Delta] = \frac{P[y < Y < y + \Delta \,|\, X = +1] P[X = +1]}{P[y < Y < y + \Delta]}$$

$$= \frac{f_Y(y \,|\, +1)\Delta(1/3)}{f_Y(y \,|\, +1)\Delta(1/3) + f_Y(y \,|\, -1)\Delta(2/3)}$$

$$= \frac{\frac{1}{\sqrt{2\pi}} e^{-(y-1)^2/2}(1/3)}{\frac{1}{\sqrt{2\pi}} e^{-(y-1)^2/2}(1/3) + \frac{1}{\sqrt{2\pi}} e^{-(y+1)^2/2}(2/3)}$$

$$= \frac{e^{-(y-1)^2/2}}{e^{-(y-1)^2/2} + 2e^{-(y+1)^2/2}} = \frac{1}{1 + 2e^{-2y}}.$$

The above expression is equal to 1/2 when $y_T = 0.3466$. For $y > y_T$, $X = +1$ is more likely, and for $y < y_T$, $X = -1$ is more likely. A receiver that selects the input X that is more likely given $Y = y$ is called a *maximum a posteriori receiver*.

5.7.2 Conditional Expectation

The **conditional expectation of Y given $X = x$** is defined by

$$E[Y \,|\, x] = \int_{-\infty}^{\infty} y f_Y(y \,|\, x)\, dy. \tag{5.49a}$$

In the special case where X and Y are both discrete random variables we have:

$$E[Y \mid x_k] = \sum_{y_j} y_j p_Y(y_j \mid x_k). \tag{5.49b}$$

Clearly, $E[Y \mid x]$ is simply the center of mass associated with the conditional pdf or pmf.

The conditional expectation $E[Y \mid x]$ can be viewed as defining a function of x: $g(x) = E[Y \mid x]$. It therefore makes sense to talk about the random variable $g(X) = E[Y \mid X]$. We can imagine that a random experiment is performed and a value for X is obtained, say $X = x_0$, and then the value $g(x_0) = E[Y \mid x_0]$ is produced. We are interested in $E[g(X)] = E[E[Y \mid X]]$. In particular, we now show that

$$E[Y] = E[E[Y \mid X]], \tag{5.50}$$

where the right-hand side is

$$E[E[Y \mid X]] = \int_{-\infty}^{\infty} E[Y \mid x] f_X(x)\, dx \qquad X \text{ continuous} \tag{5.51a}$$

$$E[E[Y \mid X]] = \sum_{x_k} E[Y \mid x_k] p_X(x_k) \qquad X \text{ discrete.} \tag{5.51b}$$

We prove Eq. (5.50) for the case where X and Y are jointly continuous random variables, then

$$\begin{aligned}
E[E[Y \mid X]] &= \int_{-\infty}^{\infty} E[Y \mid x] f_X(x)\, dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_Y(y \mid x)\, dy\, f_X(x)\, dx \\
&= \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} y f_Y(y)\, dy = E[Y].
\end{aligned}$$

The above result also holds for the expected value of a function of Y:

$$E[h(Y)] = E[E[h(Y) \mid X]].$$

In particular, the kth moment of Y is given by

$$E[Y^k] = E[E[Y^k \mid X]].$$

Example 5.36 Average Number of Defects in a Region

Find the mean of Y in Example 5.30 using conditional expectation.

$$E[Y] = \sum_{k=0}^{\infty} E[Y \mid X = k] P[X = k] = \sum_{k=0}^{\infty} k p P[X = k] = pE[X] = p\alpha.$$

The second equality uses the fact that $E[Y \mid X = k] = kp$ since Y is binomial with parameters k and p. Note that the second to the last equality holds for *any* pmf of X. The fact that X is Poisson with mean α is not used until the last equality.

Example 5.37 Binary Communications Channel

Find the mean of the output Y in the communications channel in Example 5.31.

Since Y is a Gaussian random variable with mean $+1$ when $X = +1$, and -1 when $X = -1$, the conditional expected values of Y given X are:

$$E[Y \mid +1] = 1 \qquad \text{and} \quad E[Y \mid -1] = -1.$$

Equation (5.38b) implies

$$E[Y] = \sum_{k=0}^{\infty} E[Y \mid X = k]P[X = k] = +1(1/3) - 1(2/3) = -1/3.$$

The mean is negative because the $X = -1$ inputs occur twice as often as $X = +1$.

Example 5.38 Average Number of Arrivals in a Service Time

Find the mean and variance of the number of customer arrivals N during the service time T of a specific customer in Example (5.33).

N is a Poisson random variable with parameter βt when $T = t$ is given, so the first two conditional moments are:

$$E[N \mid T = t] = \beta t \qquad E[N^2 \mid T = t] = (\beta t) + (\beta t)^2.$$

The first two moments of N are obtained from Eq. (5.50):

$$E[N] = \int_0^{\infty} E[N \mid T = t] f_T(t)\, dt = \int_0^{\infty} \beta t f_T(t)\, dt = \beta E[T]$$

$$E[N^2] = \int_0^{\infty} E[N^2 \mid T = t] f_T(t)\, dt = \int_0^{\infty} \{\beta t + \beta^2 t^2\} f_T(t)\, dt$$

$$= \beta E[T] + \beta^2 E[T^2].$$

The variance of N is then

$$\begin{aligned} \text{VAR}[N] &= E[N^2] - (E[N])^2 \\ &= \beta^2 E[T^2] + \beta E[T] - \beta^2 (E[T])^2 \\ &= \beta^2\, \text{VAR}[T] + \beta E[T]. \end{aligned}$$

Note that if T is not random (i.e., $E[T] = \text{constant}$ and $\text{VAR}[T] = 0$) then the mean and variance of N are those of a Poisson random variable with parameter $\beta E[T]$. When T is random, the mean of N remains the same but the variance of N increases by the term $\beta^2\, \text{VAR}[T]$, that is, the variability of T causes greater variability in N. Up to this point, we have intentionally avoided using the fact that T has an exponential distribution to emphasize that the above results hold for *any* service time distribution $f_T(t)$. If *T is exponential with parameter* α, then $E[T] = 1/\alpha$ and $\text{VAR}[T] = 1/\alpha^2$, so

$$E[N] = \frac{\beta}{\alpha} \qquad \text{and} \qquad \text{VAR}[N] = \frac{\beta^2}{\alpha^2} + \frac{\beta}{\alpha}.$$

5.8 FUNCTIONS OF TWO RANDOM VARIABLES

Quite often we are interested in one or more functions of the random variables associated with some experiment. For example, if we make repeated measurements of the same random quantity, we might be interested in the maximum and minimum value in the set, as well as the sample mean and sample variance. In this section we present methods of determining the probabilities of events involving functions of two random variables.

5.8.1 One Function of Two Random Variables

Let the random variable Z be defined as a function of two random variables:

$$Z = g(X, Y). \tag{5.52}$$

The cdf of Z is found by first finding the equivalent event of $\{Z \le z\}$, that is, the set $R_z = \{\mathbf{x} = (x, y) \text{ such that } g(\mathbf{x}) \le z\}$, then

$$F_z(z) = P[\mathbf{X} \text{ in } R_z] = \iint_{(x, y) \in R_z} f_{X,Y}(x', y')\, dx'\, dy'. \tag{5.53}$$

The pdf of Z is then found by taking the derivative of $F_z(z)$.

Example 5.39 Sum of Two Random Variables

Let $Z = X + Y$. Find $F_Z(z)$ and $f_Z(z)$ in terms of the joint pdf of X and Y.

The cdf of Z is found by integrating the joint pdf of X and Y over the region of the plane corresponding to the event $\{Z \le z\}$, as shown in Fig. 5.20.

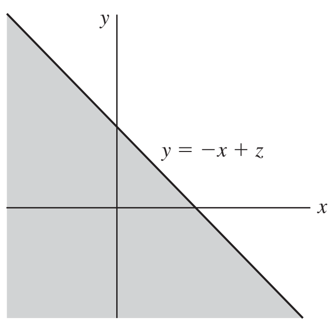


FIGURE 5.20
$P[Z \le z] = P[X + Y \le z]$.

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x'} f_{X,Y}(x', y')\, dy'\, dx'.$$

The pdf of Z is

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x', z - x')\, dx'. \tag{5.54}$$

Thus the pdf for the sum of two random variables is given by a *superposition* integral.

If X and Y are independent random variables, then by Eq. (5.23) the pdf is given by the *convolution integral* of the marginal pdf's of X and Y:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x') f_Y(z - x')\, dx'. \tag{5.55}$$

In Chapter 7 we show how transform methods are used to evaluate convolution integrals such as Eq. (5.55).

Example 5.40 Sum of Nonindependent Gaussian Random Variables

Find the pdf of the sum $Z = X + Y$ of two zero-mean, unit-variance Gaussian random variables with correlation coefficient $\rho = -1/2$.

The joint pdf for this pair of random variables was given in Example 5.18. The pdf of Z is obtained by substituting the pdf for the joint Gaussian random variables into the superposition integral found in Example 5.39:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{X,Y}(x', z - x')\, dx' \\ &= \frac{1}{2\pi(1 - \rho^2)^{1/2}} \int_{-\infty}^{\infty} e^{-[x'^2 - 2\rho x'(z-x') + (z-x')^2]/2(1-\rho^2)}\, dx' \\ &= \frac{1}{2\pi(3/4)^{1/2}} \int_{-\infty}^{\infty} e^{-(x'^2 - x'z + z^2)/2(3/4)}\, dx'. \end{aligned}$$

After completing the square of the argument in the exponent we obtain

$$f_Z(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}.$$

Thus the sum of these two nonindependent Gaussian random variables is also a zero-mean, unit-variance Gaussian random variable.

Example 5.41 A System with Standby Redundancy

A system with standby redundancy has a single key component in operation and a duplicate of that component in standby mode. When the first component fails, the second component is put into operation. Find the pdf of the lifetime of the standby system if the components have independent exponentially distributed lifetimes with the same mean.

Let T_1 and T_2 be the lifetimes of the two components, then the system lifetime is $T = T_1 + T_2$, and the pdf of T is given by Eq. (5.55). The terms in the integrand are

$$f_{T_1}(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$f_{T_2}(z - x) = \begin{cases} \lambda e^{-\lambda(z-x)} & z - x \geq 0 \\ 0 & x > z. \end{cases}$$

Note that the first equation sets the lower limit of integration to 0 and the second equation sets the upper limit to z. Equation (5.55) becomes

$$f_T(z) = \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)}\, dx$$

$$= \lambda^2 e^{-\lambda z} \int_0^z dx = \lambda^2 z e^{-\lambda z}.$$

Thus T is an Erlang random variable with parameter $m = 2$.

The conditional pdf can be used to find the pdf of a function of several random variables. Let $Z = g(X, Y)$, and suppose we are given that $Y = y$, then $Z = g(X, y)$ is a function of one random variable. Therefore we can use the methods developed in Section 4.5 for single random variables to find the pdf of Z *given* $Y = y$: $f_Z(z \mid Y = y)$. The pdf of Z is then found from

$$f_Z(z) = \int_{-\infty}^{\infty} f_Z(z \mid y') f_Y(y')\, dy'.$$

Example 5.42

Parecido al 6 de la guía

Let $Z = X/Y$. Find the pdf of Z if X and Y are independent and both exponentially distributed with mean one.

Assume $Y = y$, then $Z = X/y$ is simply a scaled version of X. Therefore from Example 4.31

$$f_Z(z \mid y) = |y| f_X(yz \mid y).$$

The pdf of Z is therefore

$$f_Z(z) = \int_{-\infty}^{\infty} |y'| f_X(y'z \mid y') f_Y(y')\, dy' = \int_{-\infty}^{\infty} |y'| f_{X,Y}(y'z, y')\, dy'.$$

We now use the fact that X and Y are independent and exponentially distributed with mean one:

$$f_Z(z) = \int_0^{\infty} y' f_X(y'z) f_Y(y')\, dy' \qquad z > 0$$

$$= \int_0^{\infty} y' e^{-y'z} e^{-y'}\, dy' \quad ¿¿?$$

$$= \frac{1}{(1 + z)^2} \qquad z > 0.$$

5.8.2 Transformations of Two Random Variables

Let X and Y be random variables associated with some experiment, and let the random variables Z_1 and Z_2 be defined by two functions of $\mathbf{X} = (X, Y)$:

$\vec{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} g_1(x,y) \\ g_2(x,y) \end{bmatrix}$

$$Z_1 = g_1(\mathbf{X}) \qquad \text{and} \qquad Z_2 = g_2(\mathbf{X}).$$

We now consider the problem of finding the joint cdf and pdf of Z_1 and Z_2.

The joint cdf of Z_1 and Z_2 at the point $\mathbf{z} = (z_1, z_2)$ is equal to the probability of the region of $\mathbf{x}$ where $g_k(\mathbf{x}) \leq z_k$ for $k = 1, 2$:

$$F_{z_1, z_2}(z_1, z_2) = P[g_1(\mathbf{X}) \leq z_1, g_2(\mathbf{X}) \leq z_2]. \tag{5.56a}$$

If X, Y have a joint pdf, then

$$F_{z_1, z_2}(z_1, z_2) = \iint\limits_{\mathbf{x}': \mathbf{g_k}(\mathbf{x}') \leq \mathbf{z_k}} f_{X,Y}(x', y')\, dx'\, dy'. \tag{5.56b}$$

Example 5.43

Let the random variables W and Z be defined by

$$W = \min(X, Y) \qquad \text{and} \qquad Z = \max(X, Y).$$

Find the joint cdf of W and Z in terms of the joint cdf of X and Y.

Equation (5.56a) implies that

$$F_{W,Z}(w\ z) = P[\{\min(X, Y) \leq w\} \cap \{\max(X, Y) \leq z\}].$$

The region corresponding to this event is shown in Fig. 5.21. From the figure it is clear that if $z > w$, the above probability is the probability of the semi-infinite rectangle defined by the

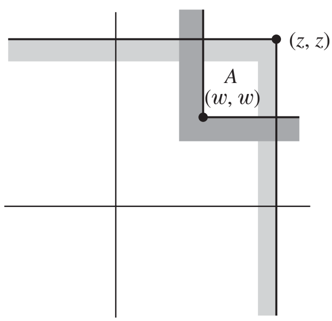


FIGURE 5.21
$\{\min(X, Y) \leq w = \{X \leq w\} \cup \{Y \leq w\}$ and
$\{\max(X, Y) \leq z = \{X \leq z\} \cap \{Y \leq z\}$.

point (z, z) minus the square region denoted by A. Thus if $z > w$,

$$\begin{aligned} F_{W,Z}(w, z) &= F_{X,Y}(z, z) - P[A] \\ &= F_{X,Y}(z, z) \\ &\quad - \{F_{X,Y}(z, z) - F_{X,Y}(w, z) - F_{X,Y}(z, w) + F_{X,Y}(w, w)\} \\ &= F_{X,Y}(w, z) + F_{X,Y}(z, w) - F_{X,Y}(w, w). \end{aligned}$$

If $z < w$ then

$$F_{W,Z}(w, z) = F_{X,Y}(z, z).$$

Box MULLER

Box Muller

Example 5.44 Radius and Angle of Independent Gaussian Random Variables

Let X and Y be zero-mean, unit-variance independent Gaussian random variables. Find the joint cdf and pdf of R and Θ, the radius and angle of the point (X, Y):

$$R = (X^2 + Y^2)^{1/2} \qquad \Theta = \tan^{-1}(Y/X).$$

The joint cdf of R and Θ is:

$$F_{R,\Theta}(r_0, \theta_0) = P[R \le r_0, \Theta \le \theta_0] = \iint_{(x,y)\in R_{(r_0,\theta_0)}} \frac{e^{-(x^2+y^2)/2}}{2\pi}\,dx\,dy$$

where

$$R_{(r_0,\theta_0)} = \{(x, y): \sqrt{x^2 + y^2} \le r_0, 0 < \tan^{-1}(Y/X) \le \theta_0\}.$$

The region R_{r_0,θ_0} is the pie-shaped region in Fig. 5.22. We change variables from Cartesian to polar coordinates to obtain:

$$\begin{aligned} F_{R,\Theta}(r_0, \theta_0) = P[R \le r_0, \Theta \le \theta_0] &= \int_0^{r_0}\int_0^{\theta_0} \frac{e^{-r^2/2}}{2\pi} r\,dr\,d\theta \\ &= \frac{\theta_0}{2\pi}\left(1 - e^{-r_0^2/2}\right), \quad 0 < \theta_0 < 2\pi \quad 0 < r_0 < \infty. \end{aligned} \tag{5.57}$$

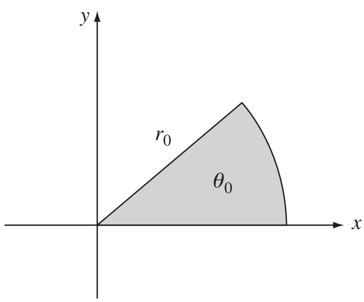


FIGURE 5.22
Region of integration R_{r_0,θ_0} in Example 5.44.

R and Θ are independent random variables, where R has a Rayleigh distribution and Θ is uniformly distributed in $(0, 2\pi)$. The joint pdf is obtained by taking partial derivatives with respect to r and θ:

$$
\begin{aligned}
f_{R,\Theta}(r, \theta) &= \frac{\partial^2}{\partial r \partial \theta} \frac{\theta}{2\pi}(1 - e^{-r^2/2}) \\
&= \frac{1}{2\pi}\left(re^{-r^2/2}\right), \quad 0 < \theta < 2\pi \quad 0 < r < \infty.
\end{aligned}
$$

This transformation maps every point in the plane from Cartesian coordinates to polar coordinates. We can also go *backwards* from polar to Cartesian coordinates. First we generate independent Rayleigh R and uniform Θ random variables. We then transform R and Θ into Cartesian coordinates to obtain an independent pair of zero-mean, unit-variance Gaussians. Neat!

5.8.3 pdf of Linear Transformations

The joint pdf of $\mathbf{Z}$ can be found directly in terms of the joint pdf of $\mathbf{X}$ by finding the equivalent events of infinitesimal rectangles. We consider the **linear transformation** of two random variables:

¿Estará ahí la clave del Jacobiano?

$$
\begin{aligned}
V &= aX + bY \\
W &= cX + eY
\end{aligned}
\qquad \text{or} \qquad
\begin{bmatrix} V \\ W \end{bmatrix} = \begin{bmatrix} a & b \\ c & e \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}.
$$

Denote the above matrix by A. We will assume that A has an inverse, that is, it has determinant $|ae - bc| \neq 0$, so each point (v, w) has a unique corresponding point (x, y) obtained from

$$
\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1} \begin{bmatrix} v \\ w \end{bmatrix}. \tag{5.58}
$$

Consider the infinitesimal rectangle shown in Fig. 5.23. The points in this rectangle are mapped into the parallelogram shown in the figure. The infinitesimal rectangle and the parallelogram are equivalent events, so their probabilities must be equal. Thus

$$
f_{X,Y}(x, y) dx\, dy \simeq f_{V,W}(v, w)\, dP
$$

where dP is the area of the parallelogram. The joint pdf of V and W is thus given by

$$
f_{V,W}(v, w) = \frac{f_{X,Y}(x, y)}{\left| \dfrac{dP}{dx\, dy} \right|}, \tag{5.59}
$$

where x and y are related to (v, w) by Eq. (5.58). Equation (5.59) states that the joint pdf of V and W at (v, w) is the pdf of X and Y at the corresponding point (x, y), but rescaled by the "stretch factor" $dP/dx\, dy$. It can be shown that $dP = (|ae - bc|)\, dx\, dy$, so the "stretch factor" is

$$
\left| \frac{dP}{dx\, dy} \right| = \frac{|ae - bc|(dx\, dy)}{(dx\, dy)} = |ae - bc| = |A|,
$$

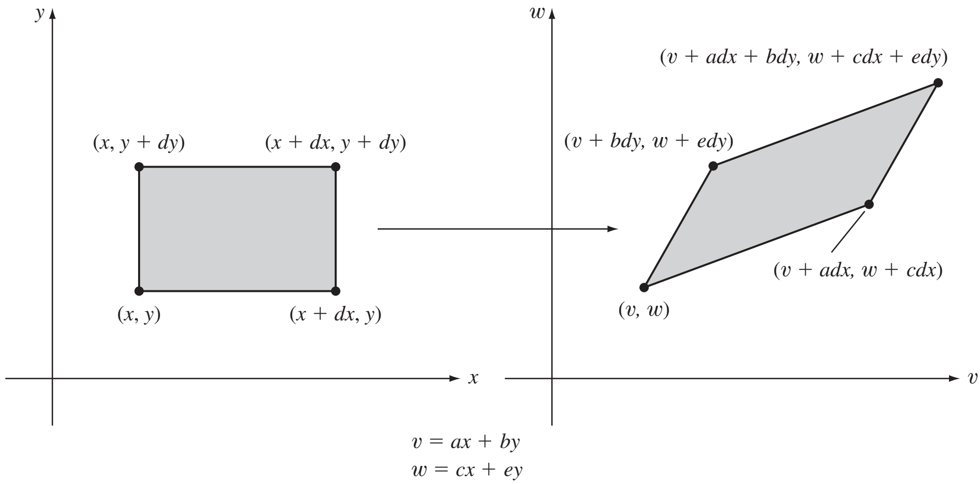


FIGURE 5.23
Image of an infinitesimal rectangle under a linear transformation.

where $|A|$ is the determinant of A.

The above result can be written compactly using matrix notation. Let the vector $\mathbf{Z}$ be

$$\mathbf{Z} = A\mathbf{X},$$

where A is an $n \times n$ invertible matrix. The joint pdf of $\mathbf{Z}$ is then

$$f_{\mathbf{z}}(\mathbf{z}) = \frac{f_{\mathbf{x}}(A^{-1}\mathbf{z}).}{|A|} \tag{5.60}$$

Example 5.45 Linear Transformation of Jointly Gaussian Random Variables

Let X and Y be the jointly Gaussian random variables introduced in Example 5.18. Let V and W be obtained from (X, Y) by

$$\begin{bmatrix} V \\ W \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix} = A\begin{bmatrix} X \\ Y \end{bmatrix}.$$

Find the joint pdf of V and W.

The determinant of the matrix is $|A| = 1$, and the inverse mapping is given by

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} V \\ W \end{bmatrix},$$

so $X = (V - W)/\sqrt{2}$ and $Y = (V + W)/\sqrt{2}$. Therefore the pdf of V and W is

$$f_{V,W}(v, w) = f_{X,Y}\left(\frac{v - w}{\sqrt{2}}, \frac{v + w}{\sqrt{2}}\right),$$

where

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)}.$$

By substituting for x and y, the argument of the exponent becomes

$$\frac{(v - w)^2/2 - 2\rho(v - w)(v + w)/2 + (v + w)^2/2}{2(1 - \rho^2)}$$
$$= \frac{v^2}{2(1 + \rho)} + \frac{w^2}{2(1 - \rho)}.$$

Thus

$$f_{V,W}(v, w) = \frac{1}{2\pi(1 - \rho^2)^{1/2}} e^{-\{[v^2/2(1+\rho)]+[w^2/2(1-\rho)]\}}.$$

It can be seen that the transformed variables V and W are independent, zero-mean Gaussian random variables with variance $1 + \rho$ and $1 - \rho$, respectively. Figure 5.24 shows contours of equal value of the joint pdf of (X, Y). It can be seen that the pdf has elliptical symmetry about the origin with principal axes at 45° with respect to the axes of the plane. In Section 5.9 we show that the above linear transformation corresponds to a rotation of the coordinate system so that the axes of the plane are aligned with the axes of the ellipse.

5.9 PAIRS OF JOINTLY GAUSSIAN RANDOM VARIABLES

The jointly Gaussian random variables appear in numerous applications in electrical engineering. They are frequently used to model signals in signal processing applications, and they are the most important model used in communication systems that involve dealing with signals in the presence of noise. They also play a central role in many statistical methods.

The random variables X and Y are said to be **jointly Gaussian** if their joint pdf has the form

$$f_{X,Y}(x, y) = \frac{\exp\left\{\dfrac{-1}{2(1-\rho_{X,Y}^2)}\left[\left(\dfrac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{X,Y}\left(\dfrac{x-m_1}{\sigma_1}\right)\left(\dfrac{y-m_2}{\sigma_2}\right) + \left(\dfrac{y-m_2}{\sigma_2}\right)^2\right]\right\}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{X,Y}^2}} \tag{5.61a}$$

for $-\infty < x < \infty$ and $-\infty < y < \infty$.

The pdf is centered at the point (m_1, m_2), and it has a bell shape that depends on the values of σ_1, σ_2, and $\rho_{X,Y}$ as shown in Fig. 5.25. As shown in the figure, the pdf is constant for values x and y for which the argument of the exponent is constant:

$$\left[\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{X,Y}\left(\frac{x-m_1}{\sigma_1}\right)\left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2\right] = \text{constant}. \tag{5.61b}$$

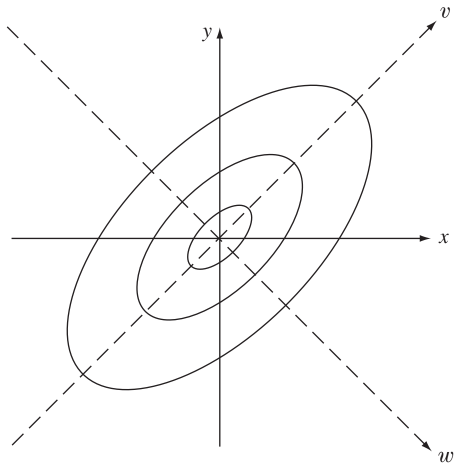
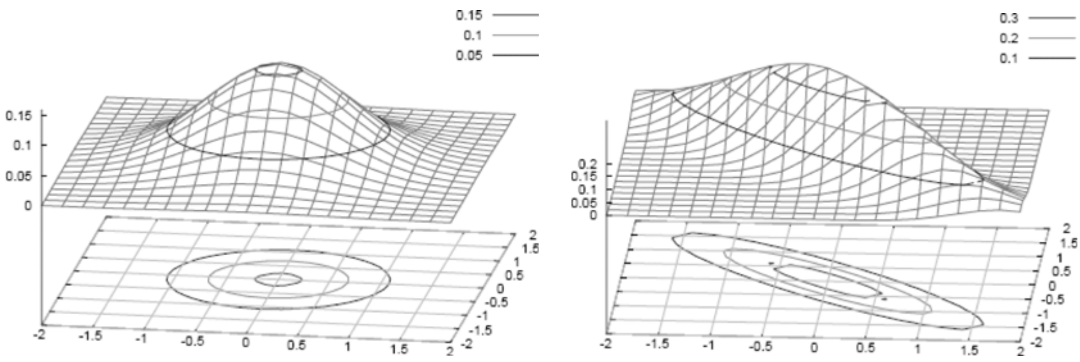


FIGURE 5.24
Contours of equal value of joint Gaussian pdf discussed in Example 5.45.



(a) (b)

FIGURE 5.25
Jointly Gaussian pdf (a) $\rho = 0$ (b) $\rho = -0.9$.

Figure 5.26 shows the orientation of these elliptical contours for various values of σ_1, σ_2, and $\rho_{X,Y}$. When $\rho_{X,Y} = 0$, that is, when X and Y are independent, the equal-pdf contour is an ellipse with principal axes aligned with the x- and y-axes. When $\rho_{X,Y} \neq 0$, the major axis of the ellipse is oriented along the angle [Edwards and Penney, pp. 570–571]

$$\theta = \frac{1}{2} \arctan^{-1} \tan\left(\frac{2\rho_{X,Y}\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}\right). \tag{5.62}$$

Note that the angle is 45° when the variances are equal.

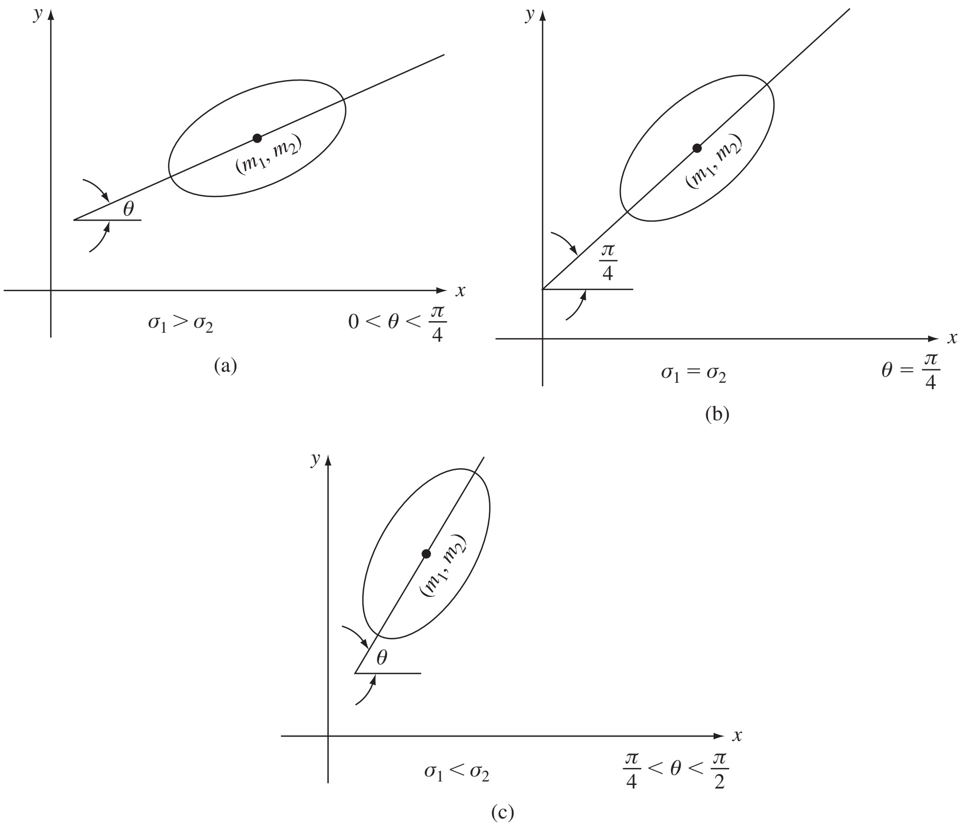


FIGURE 5.26
Orientation of contours of equal value of joint Gaussian pdf for $\rho_{X,Y} > 0$.

The marginal pdf of X is found by integrating $f_{X,Y}(x, y)$ over all y. The integration is carried out by completing the square in the exponent as was done in Example 5.18. The result is that the marginal pdf of X is

$$f_X(x) = \frac{e^{-(x-m_1)^2/2\sigma_1^2}}{\sqrt{2\pi}\sigma_1}, \tag{5.63}$$

that is, X is a Gaussian random variable with mean m_1 and variance σ_1^2. Similarly, the marginal pdf for Y is found to be Gaussian with pdf mean m_2 and variance σ_2^2.

The conditional pdf's $f_X(x \mid y)$ and $f_Y(y \mid x)$ give us information about the interrelation between X and Y. The conditional pdf of X given $Y = y$ is

$$\begin{aligned} f_X(x \mid y) &= \frac{f_{X,Y}(x, y)}{f_Y(y)} \\ &= \frac{\exp\left\{\dfrac{-1}{2(1-\rho_{X,Y}^2)\sigma_1^2}\left[x - \rho_{X,Y}\dfrac{\sigma_1}{\sigma_2}(y - m_2) - m_1\right]^2\right\}}{\sqrt{2\pi\sigma_1^2(1-\rho_{X,Y}^2)}}. \end{aligned} \tag{5.64}$$

Equation (5.64) shows that the conditional pdf of X given $Y = y$ is also Gaussian but with conditional mean $m_1 + \rho_{X,Y}(\sigma_1/\sigma_2)(y - m_2)$ and conditional variance $\sigma_1^2(1 - \rho_{X,Y}^2)$. Note that when $\rho_{X,Y} = 0$, the conditional pdf of X given $Y = y$ equals the marginal pdf of X. This is consistent with the fact that X and Y are independent when $\rho_{X,Y} = 0$. On the other hand, as $|\rho_{X,Y}| \to 1$ the variance of X about the conditional mean approaches zero, so the conditional pdf approaches a delta function at the conditional mean. Thus when $|\rho_{X,Y}| = 1$, the conditional variance is zero and X is equal to the conditional mean with probability one. We note that similarly $f_Y(y \mid x)$ is Gaussian with conditional mean $m_2 + \rho_{X,Y}(\sigma_2/\sigma_1)(x - m_1)$ and conditional variance $\sigma_2^2(1-\rho_{X,Y}^2)$.

We now show that the $\rho_{X,Y}$ in Eq. (5.61a) is indeed the correlation coefficient between X and Y. The covariance between X and Y is defined by

$$\begin{aligned} \text{COV}(X, Y) &= E[(X - m_1)(Y - m_2)] \\ &= E[E[(X - m_1)(Y - m_2) \mid Y]]. \end{aligned}$$

Now the conditional expectation of $(X - m_1)(Y - m_2)$ given $Y = y$ is

$$\begin{aligned} E[(X - m_1)(Y - m_2) \mid Y = y] &= (y - m_2)E[X - m_1 \mid Y = y] \\ &= (y - m_2)(E[X \mid Y = y] - m_1) \\ &= (y - m_2)\left(\rho_{X,Y}\frac{\sigma_1}{\sigma_2}(y - m_2)\right), \end{aligned}$$

where we have used the fact that the conditional mean of X given $Y = y$ is $m_1 + \rho_{X,Y}(\sigma_1/\sigma_2)(y - m_2)$. Therefore

$$E[(X - m_1)(Y - m_2) \mid Y] = \rho_{X,Y}\frac{\sigma_1}{\sigma_2}(Y - m_2)^2$$

and

$$\begin{aligned} \text{COV}(X, Y) &= E[E[(X - m_1)(Y - m_2) \mid Y]] = \rho_{X,Y}\frac{\sigma_1}{\sigma_2}E[(Y - m_2)^2] \\ &= \rho_{X,Y}\sigma_1\sigma_2. \end{aligned}$$

The above equation is consistent with the definition of the correlation coefficient, $\rho_{X,Y} = \text{COV}(X, Y)/\sigma_1\sigma_2$. Thus the $\rho_{X,Y}$ in Eq. (5.61a) is indeed the correlation coefficient between X and Y.

Example 5.46

The amount of yearly rainfall in city 1 and in city 2 is modeled by a pair of jointly Gaussian random variables, X and Y, with pdf given by Eq. (5.61a). Find the most likely value of X given that we know $Y = y$.

The most likely value of X given $Y = y$ is the value of x for which $f_X(x \mid y)$ is maximum. The conditional pdf of X given $Y = y$ is given by Eq. (5.64), which is maximum at the conditional mean

$$E[X \mid y] = m_1 + \rho_{X,Y}\frac{\sigma_1}{\sigma_2}(y - m_2).$$

Note that this "maximum likelihood" estimate is a linear function of the observation y.

Example 5.47 Estimation of Signal in Noise

Let $Y = X + N$ where X (the "signal") and N (the "noise') are independent zero-mean Gaussian random variables with different variances. Find the correlation coefficient between the observed signal Y and the desired signal X. Find the value of x that maximizes $f_X(x \mid y)$.

The mean and variance of Y and the covariance of X and Y are:

$$E[Y] = E[X] + E[N] = 0$$

$$\sigma_Y^2 = E[Y^2] = E[(X+N)^2] = E[X^2 + 2XN + N^2] = E[X^2] + E[N^2] = \sigma_X^2 + \sigma_N^2.$$

$$\text{COV}(X,Y) = E[(X - E[X])(E(Y - E[Y])] = E[XY] = E[X(X+N)] = \sigma_X^2.$$

Therefore, the correlation coefficient is:

$$\rho_{X,Y} = \frac{\text{COV}(X,Y)}{\sigma_X \sigma_Y} = \frac{\sigma_X}{\sigma_Y} = \frac{\sigma_X}{(\sigma_X^2 + \sigma_N^2)^{1/2}} = \frac{1}{\left(1 + \dfrac{\sigma_N^2}{\sigma_X^2}\right)^{1/2}}.$$

Note that $\rho_{X,Y}^2 = \sigma_X^2/\sigma_Y^2 = 1 - \sigma_N^2/\sigma_Y^2$.

To find the joint pdf of X and Y consider the following linear transformation:

$$\begin{aligned} X &= X \\ Y &= X + N \end{aligned} \qquad \text{which has inverse} \qquad \begin{aligned} X &= X \\ N &= -X + Y. \end{aligned}$$

From Eq. (5.52) we have:

$$f_{X,Y}(x,y) = \left.\frac{f_{X,N}(x,y)}{\det A}\right|_{x=x,\, n=y-x} = \left.\frac{e^{-x^2/2\sigma_X^2}}{\sqrt{2\pi}\sigma_X} \frac{e^{-n^2/2\sigma_N^2}}{\sqrt{2\pi}\sigma_N}\right|_{x=x,\, n=y-x}$$

$$= \frac{e^{-x^2/2\sigma_X^2}}{\sqrt{2\pi}\sigma_X} \frac{e^{-(y-x)^2/2\sigma_N^2}}{\sqrt{2\pi}\sigma_N}.$$

The conditional pdf of the signal X given the observation Y is then:

$$f_X(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{e^{-x^2/2\sigma_X^2}}{\sqrt{2\pi}\sigma_X} \frac{e^{-(y-x)^2/2\sigma_N^2}}{\sqrt{2\pi}\sigma_N} \frac{\sqrt{2\pi}\sigma_Y}{e^{-y^2/2\sigma_Y^2}}$$

$$= \frac{\exp\left\{-\frac{1}{2}\left(\left(\frac{x}{\sigma_X}\right)^2 + \left(\frac{y-x}{\sigma_N}\right)^2 - \left(\frac{y}{\sigma_Y}\right)^2\right)\right\}}{\sqrt{2\pi}\sigma_N \sigma_X/\sigma_Y} = \frac{\exp\left\{-\frac{1}{2}\frac{\sigma_Y^2}{\sigma_X^2\sigma_N^2}\left(x - \frac{\sigma_X^2}{\sigma_Y^2}y\right)^2\right\}}{\sqrt{2\pi}\sigma_N\sigma_X/\sigma_Y}$$

$$= \frac{\exp\left\{-\frac{1}{2(1-\rho_{X,Y}^2)\sigma_X^2}\left(x - \left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_X^2}\right)y\right)^2\right\}}{\sqrt{1-\rho_{X,Y}^2}\,\sigma_X}.$$

This pdf has its maximum value, when the argument of the exponent is zero, that is,

$$x = \left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_N^2}\right)y = \left(\frac{1}{1 + \dfrac{\sigma_N^2}{\sigma_X^2}}\right)y.$$

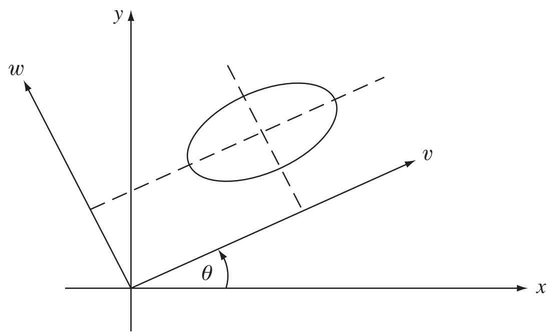


FIGURE 5.27
A rotation of the coordinate system transforms a pair of dependent Gaussian random variables into a pair of independent Gaussian random variables.

The signal-to-noise ratio (SNR) is defined as the ratio of the variance of X and the variance of N. At high SNRs this estimator gives $x \approx y$, and at very low signal-to-noise ratios, it gives $x \approx 0$.

Example 5.48 Rotation of Jointly Gaussian Random Variables

The ellipse corresponding to an arbitrary two-dimensional Gaussian vector forms an angle

$$\theta = \frac{1}{2}\arctan\left(\frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}\right)$$

relative to the x-axis. Suppose we define a new coordinate system whose axes are aligned with those of the ellipse as shown in Fig. 5.27. This is accomplished by using the following rotation matrix:

$$\begin{bmatrix} V \\ W \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix}.$$

To show that the new random variables are independent it suffices to show that they have covariance zero:

$$\begin{aligned} \text{COV}(V, W) &= E[(V - E[V])(W - E[W])] \\ &= E[\{(X - m_1)\cos\theta + (Y - m_2)\sin\theta\} \\ &\quad \times \{-(X - m_1)\sin\theta + (Y - m_2)\cos\theta\}] \\ &= -\sigma_1^2 \sin\theta\cos\theta + \text{COV}(X, Y)\cos^2\theta \\ &\quad -\text{COV}(X, Y)\sin^2\theta + \sigma_2^2\sin\theta\cos\theta \\ &= \frac{(\sigma_2^2 - \sigma_1^2)\sin 2\theta + 2\,\text{COV}(X, Y)\cos 2\theta}{2} \\ &= \frac{\cos 2\theta[(\sigma_2^2 - \sigma_1^2)\tan 2\theta + 2\,\text{COV}(X, Y)]}{2}. \end{aligned}$$

If we let the angle of rotation θ be such that

$$\tan 2\theta = \frac{2\,\text{COV}(X, Y)}{\sigma_1^2 - \sigma_2^2},$$

then the covariance of V and W is zero as required.

---

*5.10 GENERATING INDEPENDENT GAUSSIAN RANDOM VARIABLES

BOX MULLER

We now present a method for generating unit-variance, uncorrelated (and hence independent) jointly Gaussian random variables. Suppose that X and Y are two independent zero-mean, unit-variance jointly Gaussian random variables with pdf:

$$f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}.$$

In Example 5.44 we saw that the transformation

$$R = \sqrt{X^2 + Y^2} \qquad \text{and} \qquad \Theta = \tan^{-1} Y/X$$

leads to the pair of independent random variables

$$f_{R,\Theta}(r, \theta) = \frac{1}{2\pi} r e^{-r^2/2} = f_R(r) f_\Theta(\theta),$$

where R is a Rayleigh random variable and Θ is a uniform random variable. The above transformation is invertible. Therefore we can also start with independent Rayleigh and uniform random variables and produce zero-mean, unit-variance independent Gaussian random variables through the transformation:

$$X = R\cos\Theta \quad \text{and} \quad Y = R\sin\Theta. \tag{5.65}$$

Consider $W = R^2$ where R is a Rayleigh random variable. From Example 5.41 we then have that: W has pdf

$$f_W(w) = \frac{f_R(\sqrt{w})}{2\sqrt{w}} = \frac{\sqrt{w}e^{-\sqrt{w}^2/2}}{2\sqrt{w}} = \frac{1}{2}e^{-w/2}.$$

$W = R^2$ has an exponential distribution with $\lambda = 1/2$.

Therefore we can generate R^2 by generating an exponential random variable with parameter 1/2, and we can generate Θ by generating a random variable that is uniformly distributed in the interval $(0, 2\pi)$. If we substitute these random variables into Eq. (5.65), we then obtain a pair of independent zero-mean, unit-variance Gaussian random variables. The above discussion thus leads to the following algorithm:

1. Generate U_1 and U_2, two independent random variables uniformly distributed in the unit interval.
2. Let $R^2 = -2\log U_1$ and $\Theta = 2\pi U_2$.
3. Let $X = R\cos\Theta = (-2\log U_1)^{1/2}\cos 2\pi U_2$ and $Y = R\sin\Theta = (-2\log U_1)^{1/2}$ $\sin 2\pi U_2$.

Then X and Y are independent, zero-mean, unit-variance Gaussian random variables. By repeating the above procedure we can generate any number of such random variables.

Example 5.49

Use Octave or MATLAB to generate 1000 independent zero-mean, unit-variance Gaussian random variables. Compare a histogram of the observed values with the pdf of a zero-mean unit-variance random variable.

The Octave commands below show the steps for generating the Gaussian random variables. A set of histogram range values K from -4 to 4 is created and used to build a normalized histogram Z. The points in Z are then plotted and compared to the value predicted to fall in each interval by the Gaussian pdf. These plots are shown in Fig. 5.28, which shows excellent agreement.

```
> U1=rand(1000,1);          % Create a 1000-element vector U1 (step 1).
> U2=rand(1000,1);          % Create a 1000-element vector U2 (step 1).
> R2=-2*log(U1);            % Find R^2 (step 2).
> TH=2*pi*U2;               % Find θ (step 2).
> X=sqrt(R2).*sin(TH);      % Generate X (step 3).
```

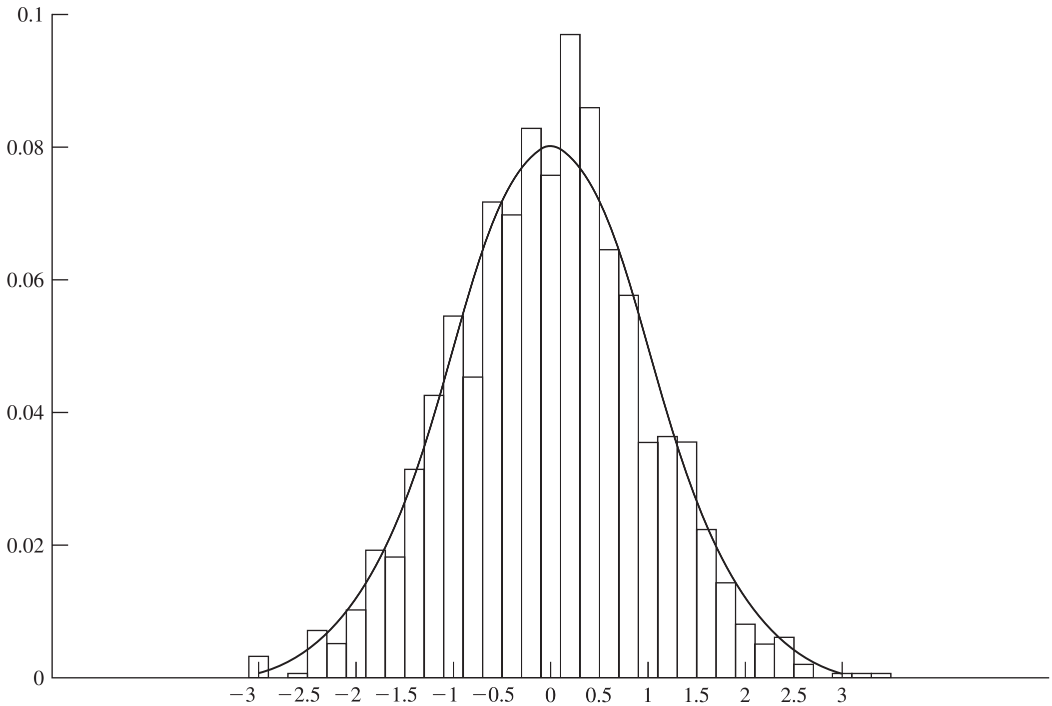


FIGURE 5.28
Histogram of 1000 observations of a Gaussian random variable.

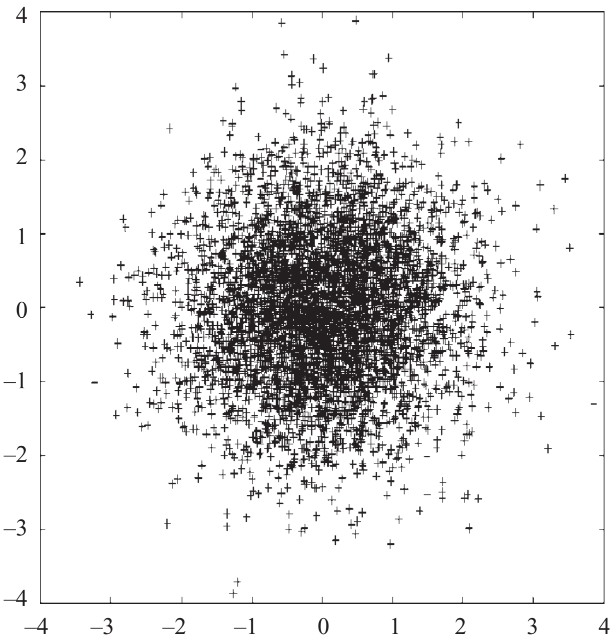


FIGURE 5.29
Scattergram of 5000 pairs of jointly Gaussian random variables.

```
> Y=sqrt(R2).*cos(TH);          % Generate Y (step 3).
> K=-4:.2:4;                    % Create histogram range values K.
> Z=hist(X,K)/1000              % Create normalized histogram Z based on K.
> bar(K,Z)                      % Plot Z.
> hold on
> stem(K,.2*normal_pdf(K,0,1))  % Compare to values predicted by pdf.
```

We also plotted the X values vs. the Y values for 5000 pairs of generated random variables in a scattergram as shown in Fig. 5.29. Good agreement with the circular symmetry of the jointly Gaussian pdf of zero-mean, unit-variance pairs is observed.

In the next chapter we will show how to generate a vector of jointly Gaussian random variables with an arbitrary covariance matrix.

SUMMARY

- The joint statistical behavior of a pair of random variables X and Y is specified by the joint cumulative distribution function, the joint probability mass function, or the joint probability density function. The probability of any event involving the joint behavior of these random variables can be computed from these functions.

- The statistical behavior of individual random variables from **X** is specified by the marginal cdf, marginal pdf, or marginal pmf that can be obtained from the joint cdf, joint pdf, or joint pmf of **X**.
- Two random variables are independent if the probability of a product-form event is equal to the product of the probabilities of the component events. Equivalent conditions for the independence of a set of random variables are that the joint cdf, joint pdf, or joint pmf factors into the product of the corresponding marginal functions.
- The covariance and the correlation coefficient of two random variables are measures of the linear dependence between the random variables.
- If **X** and **Y** are independent, then **X** and **Y** are uncorrelated, but not vice versa. If **X** and **Y** are jointly Gaussian and uncorrelated, then they are independent.
- The statistical behavior of **X**, given the exact values of X or Y, is specified by the conditional cdf, conditional pmf, or conditional pdf. Many problems lend themselves to a solution that involves conditioning on the value of one of the random variables. In these problems, the expected value of random variables can be obtained by conditional expectation.
- The joint pdf of a pair of jointly Gaussian random variables is determined by the means, variances, and covariance. All marginal pdf's and conditional pdf's are also Gaussian pdf's.
- Independent Gaussian random variables can be generated by a transformation of uniform random variables.

CHECKLIST OF IMPORTANT TERMS

Central moments of X and Y
Conditional cdf
Conditional expectation
Conditional pdf
Conditional pmf
Correlation of X and Y
Covariance X and Y
Independent random variables
Joint cdf
Joint moments of X and Y
Joint pdf
Joint pmf
Jointly continuous random variables
Jointly Gaussian random variables
Linear transformation
Marginal cdf
Marginal pdf
Marginal pmf
Orthogonal random variables
Product-form event
Uncorrelated random variables

ANNOTATED REFERENCES

Papoulis [1] is the standard reference for electrical engineers for the material on random variables. References [2] and [3] present many interesting examples involving multiple random variables. The book by Jayant and Noll [4] gives numerous applications of probability concepts to the digital coding of waveforms.

1. A. Papoulis and S. Pillai, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 2002.

2. L. Breiman, *Probability and Stochastic Processes*, Houghton Mifflin, Boston, 1969.

3. H. J. Larson and B. O. Shubert, *Probabilistic Models in Engineering Sciences*, vol. 1, Wiley, New York, 1979.

4. N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice Hall, Englewood Cliffs, N.J., 1984.

5. N. Johnson et al., *Continuous Multivariate Distributions*, Wiley, New York, 2000.

6. H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, Prentice Hall, Englewood Cliffs, N.J., 1986.

7. H. Anton, *Elementary Linear Algebra*, 9th ed., Wiley, New York, 2005.

8. C. H. Edwards, Jr., and D. E. Penney, *Calculus and Analytic Geometry*, 4th ed., Prentice Hall, Englewood Cliffs, N.J., 1994.

PROBLEMS

Section 5.1: Two Random Variables

5.1. Let X be the maximum and let Y be the minimum of the number of heads obtained when Carlos and Michael each flip a fair coin twice.

- **(a)** Describe the underlying space S of this random experiment and show the mapping from S to S_{XY}, the range of the pair (X, Y).
- **(b)** Find the probabilities for all values of (X, Y).
- **(c)** Find $P[X = Y]$.
- **(d)** Repeat parts b and c if Carlos uses a biased coin with $P[\text{heads}] = 3/4$.

5.2. Let X be the difference and let Y be the sum of the number of heads obtained when Carlos and Michael each flip a fair coin twice.

- **(a)** Describe the underlying space S of this random experiment and show the mapping from S to S_{XY}, the range of the pair (X, Y).
- **(b)** Find the probabilities for all values of (X, Y).
- **(c)** Find $P[X + Y = 1]$, $P[X + Y = 2]$.

5.3. The input X to a communication channel is "-1"or "1", with respective probabilities 1/4 and 3/4. The output of the channel Y is equal to: the corresponding input X with probability $1 - p - p_e$; $-X$ with probability p; 0 with probability p_e.

- **(a)** Describe the underlying space S of this random experiment and show the mapping from S to S_{XY}, the range of the pair (X, Y).
- **(b)** Find the probabilities for all values of (X, Y).
- **(c)** Find $P[X \neq Y]$, $P[Y = 0]$.

5.4.
- **(a)** Specify the range of the pair (N_1, N_2) in Example 5.2.
- **(b)** Specify and sketch the event "more revenue comes from type 1 requests than type 2 requests."

5.5.
- **(a)** Specify the range of the pair (Q, R) in Example 5.3.
- **(b)** Specify and sketch the event "last packet is more than half full."

5.6. Let the pair of random variables H and W be the height and weight in Example 5.1. The body mass index is a measure of body fat and is defined by $\text{BMI} = W/H^2$ where W is in kilograms and H is in meters. Determine and sketch on the plane the following events: $A = \{\text{"obese," BMI} \geq 30\}$; $B = \{\text{"overweight," } 25 \leq \text{BMI} < 30\}$; $C = \{\text{"normal," } 18.5 \leq \text{BMI} < 25\}$; and $D = \{\text{"underweight," BMI} < 18.5\}$.

5.7. Let (X, Y) be the two-dimensional noise signal in Example 5.4. Specify and sketch the events:

(a) "Maximum noise magnitude is greater than 5."

(b) "The noise power $X^2 + Y^2$ is greater than 4."

(c) "The noise power $X^2 + Y^2$ is greater than 4 and less than 9."

5.8. For the pair of random variables (X, Y) sketch the region of the plane corresponding to the following events. Identify which events are of product form.

(a) $\{X + Y > 3\}$.

(b) $\{e^X > Ye^3\}$.

(c) $\{\min(X, Y) > 0\} \cup \{\max\{X, Y) < 0\}$.

(d) $\{|X - Y| \geq 1\}$.

(e) $\{|X/Y| > 2\}$.

(f) $\{X/Y < 2\}$.

(g) $\{X^3 > Y\}$.

(h) $\{XY < 0\}$.

(i) $\{\max(|X|, Y) < 3\}$.

Section 5.2: Pairs of Discrete Random Variables

5.9. **(a)** Find and sketch $p_{X,Y}(x, y)$ in Problem 5.1 when using a fair coin.

(b) Find $p_X(x)$ and $p_Y(y)$.

(c) Repeat parts a and b if Carlos uses a biased coin with $P[\text{heads}] = 3/4$.

5.10. **(a)** Find and sketch $p_{X,Y}(x, y)$ in Problem 5.2 when using a fair coin.

(b) Find $p_X(x)$ and $p_Y(y)$.

(c) Repeat parts a and b if Carlos uses a biased coin with $P[\text{heads}] = 3/4$.

5.11. **(a)** Find the marginal pmf's for the pairs of random variables with the indicated joint pmf.

(i)				(ii)				(iii)			
X/Y	−1	0	1	X/Y	−1	0	1	X/Y	−1	0	1
−1	1/6	1/6	0	−1	1/9	1/9	1/9	−1	1/3	0	0
0	0	0	1/3	0	1/9	1/9	1/9	0	0	1/3	0
1	1/6	1/6	0	1	1/9	1/9	1/9	1	0	0	1/3

(b) Find the probability of the events $A = \{X > 0\}, B = \{X \geq Y\}$, and $C = \{X = -Y\}$ for the above joint pmf's.

5.12. A modem transmits a two-dimensional signal (X, Y) given by:

$$X = r\cos(2\pi\Theta/8) \quad \text{and} \quad Y = r\sin(2\pi\Theta/8)$$

where Θ is a discrete uniform random variable in the set $\{0, 1, 2, \ldots, 7\}$.

(a) Show the mapping from S to S_{XY}, the range of the pair (X, Y).

(b) Find the joint pmf of X and Y.

(c) Find the marginal pmf of X and of Y.

(d) Find the probability of the following events: $A = \{X = 0\}, B = \{Y \leq r/\sqrt{2}\}$, $C = \{X \geq r/\sqrt{2}, Y \geq r/\sqrt{2}\}, D = \{X < -r/\sqrt{2}\}$.

5.13. Let N_1 be the number of Web page requests arriving at a server in a 100-ms period and let N_2 be the number of Web page requests arriving at a server in the next 100-ms period. Assume that in a 1-ms interval either zero or one page request takes place with respective probabilities $1 - p = 0.95$ and $p = 0.05$, and that the requests in different 1-ms intervals are independent of each other.

(a) Describe the underlying space S of this random experiment and show the mapping from S to S_{XY}, the range of the pair (X, Y).

(b) Find the joint pmf of X and Y.

(c) Find the marginal pmf for X and for Y.

(d) Find the probability of the events $A = \{X \geq Y\}$, $B = \{X = Y = 0\}$, $C = \{X > 5, Y > 3\}$.

(e) Find the probability of the event $D = \{X + Y = 10\}$.

5.14. Let N_1 be the number of Web page requests arriving at a server in the period $(0, 100)$ ms and let N_2 be the *total* combined number of Web page requests arriving at a server in the period $(0, 200)$ ms. Assume arrivals occur as in Problem 5.13.

(a) Describe the underlying space S of this random experiment and show the mapping from S to S_{XY}, the range of the pair (X, Y).

(b) Find the joint pmf of N_1 and N_2.

(c) Find the marginal pmf for N_1 and N_2.

(d) Find the probability of the events $A = \{N_1 < N_2\}$, $B = \{N_2 = 0\}$, $C = \{N_1 > 5, N_2 > 3\}$, $D = \{|N_2 - 2N_1| < 2\}$.

5.15. At even time instants, a robot moves either $+\Delta$ cm or $-\Delta$ cm in the x-direction according to the outcome of a coin flip; at odd time instants, a robot moves similarly according to another coin flip in the y-direction. Assuming that the robot begins at the origin, let X and Y be the coordinates of the location of the robot after $2n$ time instants.

(a) Describe the underlying space S of this random experiment and show the mapping from S to S_{XY}, the range of the pair (X, Y).

(b) Find the marginal pmf of the coordinates X and Y.

(c) Find the probability that the robot is within distance $\sqrt{2}$ of the origin after $2n$ time instants.

Section 5.3: The Joint cdf of x and y

5.16. **(a)** Sketch the joint cdf for the pair (X, Y) in Problem 5.1 and verify that the properties of the joint cdf are satisfied. You may find it helpful to first divide the plane into regions where the cdf is constant.

(b) Find the marginal cdf of X and of Y.

5.17. A point (X, Y) is selected at random inside a triangle defined by $\{(x, y): 0 \leq y \leq x \leq 1\}$. Assume the point is equally likely to fall anywhere in the triangle.

(a) Find the joint cdf of X and Y.

(b) Find the marginal cdf of X and of Y.

(c) Find the probabilities of the following events in terms of the joint cdf: $A = \{X \leq 1/2, Y \leq 3/4\}$; $B = \{1/4 < X \leq 3/4, 1/4 < Y \leq 3/4\}$.

5.18. A dart is equally likely to land at any point (X_1, X_2) inside a circular target of unit radius. Let R and Θ be the radius and angle of the point (X_1, X_2).

(a) Find the joint cdf of R and Θ.

(b) Find the marginal cdf of R and Θ.

(c) Use the joint cdf to find the probability that the point is in the first quadrant of the real plane and that the radius is greater than 0.5.

5.19. Find an expression for the probability of the events in Problem 5.8 parts c, h, and i in terms of the joint cdf of X and Y.

5.20. The pair (X, Y) has joint cdf given by:

$$F_{X,Y}(x, y) = \begin{cases} (1 - 1/x^2)(1 - 1/y^2) & \text{for } x > 1, y > 1 \\ 0 & \text{elsewhere.} \end{cases}$$

(a) Sketch the joint cdf.

(b) Find the marginal cdf of X and of Y.

(c) Find the probability of the following events: $\{X < 3, Y \le 5\}$, $\{X > 4, Y > 3\}$.

5.21. Is the following a valid cdf? Why?

$$F_{X,Y}(x, y) = \begin{cases} (1 - 1/x^2y^2) & \text{for } x > 1, y > 1 \\ 0 & \text{elsewhere.} \end{cases}$$

5.22. Let $F_X(x)$ and $F_Y(y)$ be valid one-dimensional cdf's. Show that $F_{X,Y}(x, y) = F_X(x)F_Y(y)$ satisfies the properties of a two-dimensional cdf.

5.23. The number of users logged onto a system N and the time T until the next user logs off have joint probability given by:

$$P[N = n, X \le t] = (1 - \rho)\rho^{n-1}(1 - e^{-n\lambda t}) \qquad \text{for } n = 1, 2, \ldots \quad t > 0.$$

(a) Sketch the above joint probability.

(b) Find the marginal pmf of N.

(c) Find the marginal cdf of X.

(d) Find $P[N \le 3, X > 3/\lambda]$.

5.24. A factory has n machines of a certain type. Let p be the probability that a machine is working on any given day, and let N be the total number of machines working on a certain day. The time T required to manufacture an item is an exponentially distributed random variable with rate $k\alpha$ if k machines are working. Find and $P[T \le t]$. Find $P[T \le t]$ as $t \to \infty$ and explain the result.

Section 5.4: The Joint pdf of Two Continuous Random Variables

5.25. The amplitudes of two signals X and Y have joint pdf:

$$f_{X,Y}(x, y) = e^{-x/2}ye^{-y^2} \qquad \text{for } x > 0, y > 0.$$

(a) Find the joint cdf.

(b) Find $P[X^{1/2} > Y]$.

(c) Find the marginal pdfs.

5.26. Let X and Y have joint pdf:

$$f_{X,Y}(x, y) = k(x + y) \qquad \text{for } 0 \le x \le 1, 0 \le y \le 1.$$

(a) Find k.

(b) Find the joint cdf of (X, Y).

(c) Find the marginal pdf of X and of Y.

(d) Find $P[X < Y]$, $P[Y < X^2]$, $P[X + Y > 0.5]$.

5.27. Let X and Y have joint pdf:

$$f_{X,Y}(x, y) = kx(1 - x)y \qquad \text{for } 0 < x < 1, 0 < y < 1.$$

(a) Find k.
(b) Find the joint cdf of (X, Y).
(c) Find the marginal pdf of X and of Y.
(d) Find $P[Y < X^{1/2}]$, $P[X < Y]$.

5.28. The random vector (X, Y) is uniformly distributed (i.e., $f(x, y) = k$) in the regions shown in Fig. P5.1 and zero elsewhere.

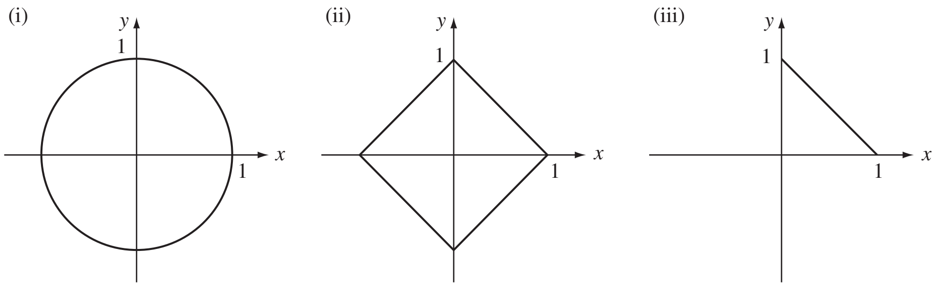


FIGURE P5.1

(a) Find the value of k in each case.
(b) Find the marginal pdf for X and for Y in each case.
(c) Find $P[X > 0, Y > 0]$.

5.29. **(a)** Find the joint cdf for the vector random variable introduced in Example 5.16.
(b) Use the result of part a to find the marginal cdf of X and of Y.

5.30. Let X and Y have the joint pdf:

$$f_{X,Y}(x, y) = ye^{-y(1+x)} \qquad \text{for } x > 0, y > 0.$$

Find the marginal pdf of X and of Y.

5.31. Let X and Y be the pair of random variables in Problem 5.17.
(a) Find the joint pdf of X and Y.
(b) Find the marginal pdf of X and of Y.
(c) Find $P[Y < X^2]$.

5.32. Let R and Θ be the pair of random variables in Problem 5.18.
(a) Find the joint pdf of R and Θ.
(b) Find the marginal pdf of R and of Θ.

5.33. Let (X, Y) be the jointly Gaussian random variables discussed in Example 5.18. Find $P[X^2 + Y^2 > r^2]$ when $\rho = 0$. *Hint:* Use polar coordinates to compute the integral.

5.34. The general form of the joint pdf for two jointly Gaussian random variables is given by Eq. (5.61a). Show that X and Y have marginal pdfs that correspond to Gaussian random variables with means m_1 and m_2 and variances σ_1^2 and σ_2^2 respectively.

5.35. The input X to a communication channel is +1 or –1 with probability p and $1 - p$, respectively. The received signal Y is the sum of X and noise N which has a Gaussian distribution with zero mean and variance $\sigma^2 = 0.25$.

(a) Find the joint probability $P[X = j, Y \leq y]$.

(b) Find the marginal pmf of X and the marginal pdf of Y.

(c) Suppose we are given that $Y > 0$. Which is more likely, $X = 1$ or $X = -1$?

5.36. A modem sends a two-dimensional signal $\mathbf{X}$ from the set $\{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$. The channel adds a noise signal (N_1, N_2), so the received signal is $\mathbf{Y} = \mathbf{X} + \mathbf{N} = (X_1 + N_1, X_2 + N_2)$. Assume that (N_1, N_2) have the jointly Gaussian pdf in Example 5.18 with $\rho = 0$. Let the distance between $\mathbf{X}$ and $\mathbf{Y}$ be $d(\mathbf{X}, \mathbf{Y}) = \{(X_1 - Y_1)^2 + (X_2 - Y_2)^2\}^{1/2}$.

(a) Suppose that $\mathbf{X} = (1, 1)$. Find and sketch region for the event $\{\mathbf{Y}$ is closer to $(1, 1)$ than to the other possible values of $\mathbf{X}\}$. Evaluate the probability of this event.

(b) Suppose that $\mathbf{X} = (1, 1)$. Find and sketch region for the event $\{\mathbf{Y}$ is closer to $(1, -1)$ than to the other possible values of $\mathbf{X}\}$. Evaluate the probability of this event.

(c) Suppose that $\mathbf{X} = (1, 1)$. Find and sketch region for the event $\{d(\mathbf{X}, \mathbf{Y}) > 1\}$. Evaluate the probability of this event. Explain why this probability is an upper bound on the probability that $\mathbf{Y}$ is closer to a signal other than $\mathbf{X} = (1, 1)$.

Section 5.5: Independence of Two Random Variables

5.37. Let X be the number of full pairs and let Y be the remainder of the number of dots observed in a toss of a fair die. Are X and Y independent random variables?

5.38. Let X and Y be the coordinates of the robot in Problem 5.15 after $2n$ time instants. Determine whether X and Y are independent random variables.

5.39. Let X and Y be the coordinates of the two-dimensional modem signal (X, Y) in Problem 5.12.

(a) Determine if X and Y are independent random variables.

(b) Repeat part a if even values of Θ are twice as likely as odd values.

5.40. Determine which of the joint pmfs in Problem 5.11 correspond to independent pairs of random variables.

5.41. Michael takes the 7:30 bus every morning. The arrival time of the bus at the stop is uniformly distributed in the interval [7:27, 7:37]. Michael's arrival time at the stop is also uniformly distributed in the interval [7:25, 7:40]. Assume that Michael's and the bus's arrival times are independent random variables.

(a) What is the probability that Michael arrives more than 5 minutes before the bus?

(b) What is the probability that Michael misses the bus?

5.42. Are R and Θ independent in Problem 5.18?

5.43. Are X and Y independent in Problem 5.20?

5.44. Are the signal amplitudes X and Y independent in Problem 5.25?

5.45. Are X and Y independent in Problem 5.26?

5.46. Are X and Y independent in Problem 5.27?

5.47. Let X and Y be independent random variables. Find an expression for the probability of the following events in terms of $F_X(x)$ and $F_Y(y)$.

(a) $\{a < X \le b\} \cap \{Y > d\}$.

(b) $\{a < X \le b\} \cap \{c \le Y < d\}$.

(c) $\{|X| < a\} \cap \{c \le Y \le d\}$.

5.48. Let X and Y be independent random variables that are uniformly distributed in $[-1, 1]$. Find the probability of the following events:

(a) $P[X^2 < 1/2, |Y| < 1/2]$.

(b) $P[4X < 1, Y < 0]$.

(c) $P[XY < 1/2]$.

(d) $P[\max(X, Y) < 1/3]$.

5.49. Let X and Y be random variables that take on values from the set $\{-1, 0, 1\}$.

(a) Find a joint pmf for which X and Y are independent.

(b) Are X^2 and Y^2 independent random variables for the pmf in part a?

(c) Find a joint pmf for which X and Y are not independent, but for which X^2 and Y^2 are independent.

5.50. Let X and Y be the jointly Gaussian random variables introduced in Problem 5.34.

(a) Show that X and Y are independent random variables if and only if $\rho = 0$.

(b) Suppose $\rho = 0$, find $P[XY < 0]$.

5.51. Two fair dice are tossed repeatedly until a pair occurs. Let K be the number of tosses required and let X be the number showing up in the pair. Find the joint pmf of K and X and determine whether K and X are independent.

5.52. The number of devices L produced in a day is geometric distributed with probability of success p. Let N be the number of working devices and let M be the number of defective devices produced in a day.

(a) Are N and M independent random variables?

(b) Find the joint pmf of N and M.

(c) Find the marginal pmfs of N and M. (See hint in Problem 5.87b.)

(d) Are L and M independent random variables?

5.53. Let N_1 be the number of Web page requests arriving at a server in a 100-ms period and let N_2 be the number of Web page requests arriving at a server in the next 100-ms period. Use the result of Problem 5.13 parts a and b to develop a model where N_1 and N_2 are independent Poisson random variables.

5.54. **(a)** Show that Eq. (5.22) implies Eq. (5.21).

(b) Show that Eq. (5.21) implies Eq. (5.22).

5.55. Verify that Eqs. (5.22) and (5.23) can be obtained from each other.

Section 5.6: Joint Moments and Expected Values of a Function of Two Random Variables

5.56. **(a)** Find $E[(X + Y)^2]$.

(b) Find the variance of $X + Y$.

(c) Under what condition is the variance of the sum equal to the sum of the individual variances?

5.57. Find $E[|X - Y|]$ if X and Y are independent exponential random variables with parameters $\lambda_1 = 1$ and $\lambda_2 = 2$, respectively.

5.58. Find $E[X^2 e^Y]$ where X and Y are independent random variables, X is a zero-mean, unit-variance Gaussian random variable, and Y is a uniform random variable in the interval $[0, 3]$.

5.59. For the discrete random variables X and Y in Problem 5.1, find the correlation and covariance, and indicate whether the random variables are independent, orthogonal, or uncorrelated.

5.60. For the discrete random variables X and Y in Problem 5.2, find the correlation and covariance, and indicate whether the random variables are independent, orthogonal, or uncorrelated.

5.61. For the three pairs of discrete random variables in Problem 5.11, find the correlation and covariance of X and Y, and indicate whether the random variables are independent, orthogonal, or uncorrelated.

5.62. Let N_1 and N_2 be the number of Web page requests in Problem 5.13. Find the correlation and covariance of N_1 and N_2, and indicate whether the random variables are independent, orthogonal, or uncorrelated.

5.63. Repeat Problem 5.62 for N_1 and N_2, the number of Web page requests in Problem 5.14.

5.64. Let N and T be the number of users logged on and the time till the next logoff in Problem 5.23. Find the correlation and covariance of N and T, and indicate whether the random variables are independent, orthogonal, or uncorrelated.

5.65. Find the correlation and covariance of X and Y in Problem 5.26. Determine whether X and Y are independent, orthogonal, or uncorrelated.

5.66. Repeat Problem 5.65 for X and Y in Problem 5.27.

5.67. For the three pairs of continuous random variables X and Y in Problem 5.28, find the correlation and covariance, and indicate whether the random variables are independent, orthogonal, or uncorrelated.

5.68. Find the correlation coefficient between X and $Y = aX + b$. Does the answer depend on the sign of a?

5.69. Propose a method for estimating the covariance of two random variables.

5.70. **(a)** Complete the calculations for the correlation coefficient in Example 5.28.

(b) Repeat the calculations if X and Y have the pdf:

$$f_{X,Y}(x, y) = e^{-(x+|y|)} \qquad \text{for } x > 0, -x < y < x.$$

5.71. The output of a channel $Y = X + N$, where the input X and the noise N are independent, zero-mean random variables.

(a) Find the correlation coefficient between the input X and the output Y.

(b) Suppose we estimate the input X by a linear function $g(Y) = aY$. Find the value of a that minimizes the mean squared error $E[(X - aY)^2]$.

(c) Express the resulting mean-square error in terms of σ_X/σ_N.

5.72. In Example 5.27 let $X = \cos \Theta/4$ and $Y = \sin \Theta/4$. Are X and Y uncorrelated?

5.73. **(a)** Show that $\text{COV}(X, E[Y | X]) = \text{COV}(X, Y)$.

(b) Show that $E[Y | X = x] = E[Y]$, for all x, implies that X and Y are uncorrelated.

5.74. Use the fact that $E[(tX + Y)^2] \geq 0$ for all t to prove the Cauchy-Schwarz inequality:

$$(E[XY])^2 \leq E[X^2]E[Y^2].$$

Hint: Consider the discriminant of the quadratic equation in t that results from the above inequality.

Section 5.7: Conditional Probability and Conditional Expectation

5.75. **(a)** Find $p_Y(y \mid x)$ and $p_X(x \mid y)$ in Problem 5.1 assuming fair coins are used.

(b) Find $p_Y(y \mid x)$ and $p_X(x \mid y)$ in Problem 5.1 assuming Carlos uses a coin with $p = 3/4$.

(c) What is the effect on $p_X(x \mid y)$ of Carlos using a biased coin?

(d) Find $E[Y \mid X = x]$ and $E[X \mid Y = y]$ in part a; then find $E[X]$ and $E[Y]$.

(e) Find $E[Y \mid X = x]$ and $E[X \mid Y = y]$ in part b; then find $E[X]$ and $E[Y]$.

5.76. **(a)** Find $p_X(x \mid y)$ for the communication channel in Problem 5.3.

(b) For each value of y, find the value of x that maximizes $p_X(x \mid y)$. State any assumptions about p and p_e.

(c) Find the probability of error if a receiver uses the decision rule from part b.

5.77. **(a)** In Problem 5.11(i), which conditional pmf given X provides the most information about Y: $p_Y(y \mid -1)$, $p_Y(y \mid 0)$, or $p_Y(y \mid +1)$? Explain why.

(b) Compare the conditional pmfs in Problems 5.11(ii) and (iii) and explain which of these two cases is "more random."

(c) Find $E[Y \mid X = x]$ and $E[X \mid Y = y]$ in Problems 5.11(i), (ii), (iii); then find $E[X]$ and $E[Y]$.

(d) Find $E[Y^2 \mid X = x]$ and $E[X^2 \mid Y = y]$ in Problems 5.11(i), (ii), (iii); then find VAR[X] and VAR[Y].

5.78. **(a)** Find the conditional pmf of N_1 given N_2 in Problem 5.14.

(b) Find $P[N_1 = k \mid N_2 = 2k]$ for $k = 5, 10, 20$. *Hint:* Use Stirling's fromula.

(c) Find $E[N_1 \mid N_2 = k]$, then find $E[N_1]$.

5.79. In Example 5.30, let Y be the number of defects inside the region R and let Z be the number of defects outside the region.

(a) Find the pmf of Z given Y.

(b) Find the joint pmf of Y and Z.

(c) Are Y and Z independent random variables? Is the result intuitive?

5.80. **(a)** Find $f_Y(y \mid x)$ in Problem 5.26.

(b) Find $P[Y > X \mid x]$.

(c) Find $P[Y > X]$ using part b.

(d) Find $E[Y \mid X = x]$.

5.81. **(a)** Find $f_Y(y \mid x)$ in Problem 5.28(i).

(b) Find $E[Y|X = x]$ and $E[Y]$.

(c) Repeat parts a and b of Problem 5.28(ii).

(d) Repeat parts a and b of Problem 5.28(iii).

5.82. **(a)** Find $f_Y(y \mid x)$ in Example 5.27.

(b) Find $E[Y| X = x]$.

(c) Find $E[Y]$.

(d) Find $E[XY \mid X = x]$.

(e) Find $E[XY]$.

5.83. Find $f_Y(y \mid x)$ and $f_X(x \mid y)$ for the jointly Gaussian pdf in Problem 5.34.

5.84. **(a)** Find $f_X(t \mid N = n)$ in Problem 5.23.

(b) Find $E[X^t \mid N = n]$.

(c) Find the value of n that maximizes $P[N = n \mid t < X < t + dt]$.

5.85. **(a)** Find $p_Y(y \mid x)$ and $p_X(x \mid y)$ in Problem 5.12.

(b) Find $E[Y \mid X = x]$.

(c) Find $E[XY \mid X = x]$ and $E[XY]$.

5.86. A customer enters a store and is equally likely to be served by one of three clerks. The time taken by clerk 1 is a constant random variable with mean two minutes; the time for clerk 2 is exponentially distributed with mean two minutes; and the time for clerk 3 is Pareto distributed with mean two minutes and $\alpha = 2.5$.

(a) Find the pdf of T, the time taken to service a customer.

(b) Find $E[T]$ and VAR$[T]$.

5.87. A message requires N time units to be transmitted, where N is a geometric random variable with pmf $p_i = (1 - a)a^{i-1}, i = 1, 2, \ldots$. A single new message arrives during a time unit with probability p, and no messages arrive with probability $1 - p$. Let K be the number of new messages that arrive during the transmission of a single message.

(a) Find $E[K]$ and VAR$[K]$ using conditional expectation.

(b) Find the pmf of K. *Hint:* $(1 - \beta)^{-(k+1)} = \sum_{n=k}^{\infty} \binom{n}{k} \beta^{n-k}$.

(c) Find the conditional pmf of N given $K = k$.

(d) Find the value of n that maximizes $P[N = n \mid X = k]$.

5.88. The number of defects in a VLSI chip is a Poisson random variable with rate r. However, r is itself a gamma random variable with parameters α and λ.

(a) Use conditional expectation to find $E[N]$ and VAR$[N]$.

(b) Find the pmf for N, the number of defects.

5.89. **(a)** In Problem 5.35, find the conditional pmf of the input X of the communication channel given that the output is in the interval $y < Y \leq y + dy$.

(b) Find the value of X that is more probable given $y < Y \leq y + dy$.

(c) Find an expression for the probability of error if we use the result of part b to decide what the input to the channel was.

Section 5.8: Functions of Two Random Variables

5.90. Two toys are started at the same time each with a different battery. The first battery has a lifetime that is exponentially distributed with mean 100 minutes; the second battery has a Rayleigh-distributed lifetime with mean 100 minutes.

(a) Find the cdf to the time T until the battery in a toy first runs out.

(b) Suppose that both toys are still operating after 100 minutes. Find the cdf of the time T_2 that subsequently elapses until the battery in a toy first runs out.

(c) In part b, find the cdf of the total time that elapses until a battery first fails.

5.91. **(a)** Find the cdf of the time that elapses until both batteries run out in Problem 5.90a.

(b) Find the cdf of the remaining time until both batteries run out in Problem 5.90b.

5.92. Let K and N be independent random variables with nonnegative integer values.

(a) Find an expression for the pmf of $M = K + N$.

(b) Find the pmf of M if K and N are binomial random variables with parameters (k, p) and (n, p).

(c) Find the pmf of M if K and N are Poisson random variables with parameters α_1 and α_2, respectively.

5.93. The number X of goals the Bulldogs score against the Flames has a geometric distribution with mean 2; the number of goals Y that the Flames score against the Bulldogs is also geometrically distributed but with mean 4.

(a) Find the pmf of the $Z = X - Y$. Assume X and Y are independent.

(b) What is the probability that the Bulldogs beat the Flames? Tie the Flames?

(c) Find $E[Z]$.

5.94. Passengers arrive at an airport taxi stand every minute according to a Bernoulli random variable. A taxi will not leave until it has two passengers.

(a) Find the pmf until the time T when the taxi has two passengers.

(b) Find the pmf for the time that the first customer waits.

5.95. Let X and Y be independent random variables that are uniformly distributed in the interval $[0, 1]$. Find the pdf of $Z = XY$.

5.96. Let X_1, X_2, and X_3 be independent and uniformly distributed in $[-1, 1]$.

(a) Find the cdf and pdf of $Y = X_1 + X_2$.

(b) Find the cdf of $Z = Y + X_3$.

5.97. Let X and Y be independent random variables with gamma distributions and parameters (α_1, λ) and (α_2, λ), respectively. Show that $Z = X + Y$ is gamma-distributed with parameters $(\alpha_1 + \alpha_2, \lambda)$. *Hint:* See Eq. (4.59).

5.98. Signals X and Y are independent. X is exponentially distributed with mean 1 and Y is exponentially distributed with mean 1.

(a) Find the cdf of $Z = |X - Y|$.

(b) Use the result of part a to find $E[Z]$.

5.99. The random variables X and Y have the joint pdf

$$f_{X,Y}(x, y) = e^{-(x+y)} \qquad \text{for } 0 < y < x < 1.$$

Find the pdf of $Z = X + Y$.

5.100. Let X and Y be independent Rayleigh random variables with parameters $\alpha = \beta = 1$. Find the pdf of $Z = X/Y$.

5.101. Let X and Y be independent Gaussian random variables that are zero mean and unit variance. Show that $Z = X/Y$ is a Cauchy random variable.

5.102. Find the joint cdf of $W = \min(X, Y)$ and $Z = \max(X, Y)$ if X and Y are independent and X is uniformly distributed in $[0, 1]$ and Y is uniformly distributed in $[0, 1]$.

5.103. Find the joint cdf of $W = \min(X, Y)$ and $Z = \max(X, Y)$ if X and Y are independent exponential random variables with the same mean.

5.104. Find the joint cdf of $W = \min(X, Y)$ and $Z = \max(X, Y)$ if X and Y are the independent Pareto random variables with the same distribution.

5.105. Let $W = X + Y$ and $Z = X - Y$.

(a) Find an expression for the joint pdf of W and Z.

(b) Find $f_{W,Z}(z, w)$ if X and Y are independent exponential random variables with parameter $\lambda = 1$.

(c) Find $f_{W,Z}(z, w)$ if X and Y are independent Pareto random variables with the same distribution.

5.106. The pair (X, Y) is uniformly distributed in a ring centered about the origin and inner and outer radii $r_1 < r_2$. Let R and Θ be the radius and angle corresponding to (X, Y). Find the joint pdf of R and Θ.

5.107. Let X and Y be independent, zero-mean, unit-variance Gaussian random variables. Let $V = aX + bY$ and $W = cX + eY$.

(a) Find the joint pdf of V and W, assuming the transformation matrix A is invertible.

(b) Suppose A is not invertible. What is the joint pdf of V and W?

5.108. Let X and Y be independent Gaussian random variables that are zero mean and unit variance. Let $W = X^2 + Y^2$ and let $\Theta = \tan^{-1}(Y/X)$. Find the joint pdf of W and Θ.

5.109. Let X and Y be the random variables introduced in Example 5.4. Let $R = (X^2 + Y^2)^{1/2}$ and let $\Theta = \tan^{-1}(Y/X)$.

(a) Find the joint pdf of R and Θ.

(b) What is the joint pdf of X and Y?

Section 5.9: Pairs of Jointly Gaussian Variables

5.110. Let X and Y be jointly Gaussian random variables with pdf

$$f_{X,Y}(x, y) = \frac{\exp\{-2x^2 - y^2/2\}}{2\pi c} \quad \text{for all } x, y.$$

Find VAR[X], VAR[Y], and COV(X, Y).

5.111. Let X and Y be jointly Gaussian random variables with pdf

$$f_{X,Y}(x, y) = \frac{\exp\left\{\frac{-1}{2}[x^2 + 4y^2 - 3xy + 3y - 2x + 1]\right\}}{2\pi} \quad \text{for all } x, y.$$

Find $E[X]$, $E[Y]$, VAR[X], VAR[Y], and COV(X, Y).

5.112. Let X and Y be jointly Gaussian random variables with $E[Y] = 0$, $\sigma_1 = 1$, $\sigma_2 = 2$, and $E[X \mid Y] = Y/4 + 1$. Find the joint pdf of X and Y.

5.113. Let X and Y be zero-mean, independent Gaussian random variables with $\sigma^2 = 1$.

(a) Find the value of r for which the probability that (X, Y) falls inside a circle of radius r is 1/2.

(b) Find the conditional pdf of (X, Y) given that (X, Y) is not inside a ring with inner radius r_1 and outer radius r_2.

5.114. Use a plotting program (as provided by Octave or MATLAB) to show the pdf for jointly Gaussian zero-mean random variables with the following parameters:

(a) $\sigma_1 = 1, \sigma_2 = 1, \rho = 0$.

(b) $\sigma_1 = 1, \sigma_2 = 1, \rho = 0.8$.

(c) $\sigma_1 = 1, \sigma_2 = 1, \rho = -0.8$.

(d) $\sigma_1 = 1, \sigma_2 = 2, \rho = 0$.

(e) $\sigma_1 = 1, \sigma_2 = 2, \rho = 0.8$.

(f) $\sigma_1 = 1, \sigma_2 = 10, \rho = 0.8$.

5.115. Let X and Y be zero-mean, jointly Gaussian random variables with $\sigma_1 = 1, \sigma_2 = 2$, and correlation coefficient ρ.

(a) Plot the principal axes of the constant-pdf ellipse of (X, Y).

(b) Plot the conditional expectation of Y given $X = x$.

(c) Are the plots in parts a and b the same or different? Why?

5.116. Let X and Y be zero-mean, unit-variance jointly Gaussian random variables for which $\rho = 1$. Sketch the joint cdf of X and Y. Does a joint pdf exist?

5.117. Let $h(x, y)$ be a joint Gaussian pdf for zero-mean, unit-variance Gaussian random variables with correlation coefficient ρ_1. Let $g(x, y)$ be a joint Gaussian pdf for zero-mean, unit-variance Gaussian random variables with correlation coefficient $\rho_2 \neq \rho_1$. Suppose the random variables X and Y have joint pdf

$$f_{X,Y}(x, y) = \{h(x, y) + g(x, y)\}/2.$$

(a) Find the marginal pdf for X and for Y.

(b) Explain why X and Y are not jointly Gaussian random variables.

5.118. Use conditional expectation to show that for X and Y zero-mean, jointly Gaussian random variables, $E[X^2Y^2] = E[X^2]E[Y^2] + 2E[XY]^2$.

5.119. Let $\mathbf{X} = (X, Y)$ be the zero-mean jointly Gaussian random variables in Problem 5.110. Find a transformation A such that $\mathbf{Z} = A\mathbf{X}$ has components that are zero-mean, unit-variance Gaussian random variables.

5.120. In Example 5. 47, suppose we estimate the value of the signal X from the noisy observation Y by:

$$\hat{X} = \frac{1}{1 + \sigma_N^2/\sigma_X^2} Y.$$

(a) Evaluate the mean square estimation error: $E[(X - \hat{X})^2]$.

(b) How does the estimation error in part a vary with signal-to-noise ratio σ_X/σ_N?

Section 5.10: Generating Independent Gaussian Random Variables

5.121. Find the inverse of the cdf of the Rayleigh random variable to derive the transformation method for generating Rayleigh random variables. Show that this method leads to the same algorithm that was presented in Section 5.10.

5.122. Reproduce the results presented in Example 5.49.

5.123. Consider the two-dimensional modem in Problem 5.36.

(a) Generate 10,000 discrete random variables uniformly distributed in the set $\{1, 2, 3, 4\}$. Assign each outcome in this set to one of the signals $\{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$. The sequence of discrete random variables then produces a sequence of 10,000 signal points $\mathbf{X}$.

(b) Generate 10,000 noise pairs $\mathbf{N}$ of independent zero-mean, unit-variance jointly Gaussian random variables.

(c) Form the sequence of 10,000 received signals $\mathbf{Y} = (Y_1, Y_2) = \mathbf{X} + \mathbf{N}$.

(d) Plot the scattergram of received signal vectors. Is the plot what you expected?

(e) Estimate the transmitted signal by the quadrant that $\mathbf{Y}$ falls in: $\hat{X} = (\text{sgn}(Y_1), \text{sgn}(Y_2))$.

(f) Compare the estimates with the actually transmitted signals to estimate the probability of error.

5.124. Generate a sequence of 1000 pairs of independent zero-mean Gaussian random variables, where X has variance 2 and N has variance 1. Let $Y = X + N$ be the noisy signal from Example 5.47.

(a) Estimate X using the estimator in Problem 5.120, and calculate the sequence of estimation errors.

(b) What is the pdf of the estimation error?

(c) Compare the mean, variance, and relative frequencies of the estimation error with the result from part b.

5.125. Let $X_1, X_2, \ldots, X_{1000}$ be a sequence of zero-mean, unit-variance independent Gaussian random variables. Suppose that the sequence is "smoothed" as follows:

$$Y_n = (X_n + X_{N-1})/2 \text{ where } X_0 = 0.$$

(a) Find the pdf of (Y_n, Y_{n+1}).

(b) Generate the sequence of X_n and the corresponding sequence Y_n. Plot the scattergram of (Y_n, Y_{n+1}). Does it agree with the result from part a?

(c) Repeat parts a and b for $Z_n = (X_n - X_{N-1})/2$.

5.126. Let X and Y be independent, zero-mean, unit-variance Gaussian random variables. Find the linear transformation to generate jointly Gaussian random variables with means m_1, m_2, variances σ_1^2, σ_2^2, and correlation coefficient ρ. *Hint:* Use the conditional pdf in Eq. (5.64).

5.127. (a) Use the method developed in Problem 5.126 to generate 1000 pairs of jointly Gaussian random variables with $m_1 = 1, m_2 = -1$, variances $\sigma_1^2 = 1, \sigma_2^2 = 2$, and correlation coefficient $\rho = -1/2$.

(b) Plot a two-dimensional scattergram of the 1000 pairs and compare to equal-pdf contour lines for the theoretical pdf.

5.128. Let H and W be the height and weight of adult males. Studies have shown that H (in cm) and $V = \ln W$ (W in kg) are jointly Gaussian with parameters $m_H = 174$ cm, $m_V = 4.4$, $\sigma_H^2 = 42.36, \sigma_V^2 = 0.021$, and $\text{COV}(H, V) = 0.458$.

(a) Use the method in part a to generate 1000 pairs (H, V). Plot a scattergram to check the joint pdf.

(b) Convert the (H, V) pairs into (H, W) pairs.

(c) Calculate the body mass index for each outcome, and estimate the proportion of the population that is underweight, normal, overweight, or obese. (See Problem 5.6.)

Problems Requiring Cumulative Knowledge

5.129. The random variables X and Y have joint pdf:

$$f_{X,Y}(x, y) = c \sin(x + y) \qquad 0 \le x \le \pi/2, 0 \le y \le \pi/2.$$

(a) Find the value of the constant c.

(b) Find the joint cdf of X and Y.

(c) Find the marginal pdf's of X and of Y.

(d) Find the mean, variance, and covariance of X and Y.

5.130. An inspector selects an item for inspection according to the outcome of a coin flip: The item is inspected if the outcome is heads. Suppose that the time between item arrivals is an exponential random variable with mean one. Assume the time to inspect an item is a constant value t.

(a) Find the pmf for the number of item arrivals between consecutive inspections.

(b) Find the pdf for the time X between item inspections. *Hint:* Use conditional expectation.

(c) Find the value of p, so that with a probability of 90% an inspection is completed before the next item is selected for inspection.

5.131. The lifetime X of a device is an exponential random variable with mean $= 1/R$. Suppose that due to irregularities in the production process, the parameter R is random and has a gamma distribution.

(a) Find the joint pdf of X and R.

(b) Find the pdf of X.

(c) Find the mean and variance of X.

5.132. Let X and Y be samples of a random signal at two time instants. Suppose that X and Y are independent zero-mean Gaussian random variables with the same variance. When signal "0" is present the variance is σ_0^2, and when signal "1" is present the variance is $\sigma_1^2 > \sigma_0^2$. Suppose signals 0 and 1 occur with probabilities p and $1 - p$, respectively. Let $R^2 = X^2 + Y^2$ be the total energy of the two observations.

(a) Find the pdf of R^2 when signal 0 is present; when signal 1 is present. Find the pdf of R^2.

(b) Suppose we use the following "signal detection" rule: If $R^2 > T$, then we decide signal 1 is present; otherwise, we decide signal 0 is present. Find an expression for the probability of error in terms of T.

(c) Find the value of T that minimizes the probability of error.

5.133. Let $U_0, U_1, \ldots$ be a sequence of independent zero-mean, unit-variance Gaussian random variables. A "low-pass filter" takes the sequence U_i and produces the output sequence $X_n = (U_n + U_{n-1})/2$, and a "high-pass filter" produces the output sequence $Y_n = (U_n - U_{n-1})/2$.

(a) Find the joint pdf of X_n and X_{n-1}; of X_n and X_{n+m}, $m > 1$.

(b) Repeat part a for Y_n.

(c) Find the joint pdf of X_n and Y_m.