CHAPTER 6

TWO RANDOM VARIABLES

6-1 BIVARIATE DISTRIBUTIONS

We are given two random variables $\mathbf{x}$ and $\mathbf{y}$, defined as in Sec. 4-1, and we wish to determine their joint statistics, that is, the probability that the point $(\mathbf{x}, \mathbf{y})$ is in a specified region[1] D in the xy plane. The distribution functions $F_x(x)$ and $F_y(y)$ of the given random variables determine their separate (marginal) statistics but not their joint statistics. In particular, the probability of the event

$$\{\mathbf{x} \le x\} \cap \{\mathbf{y} \le y\} = \{\mathbf{x} \le x, \mathbf{y} \le y\}$$

cannot be expressed in terms of $F_x(x)$ and $F_y(y)$. Here, we show that the joint statistics of the random variables $\mathbf{x}$ and $\mathbf{y}$ are completely determined if the probability of this event is known for every x and y.

Joint Distribution and Density

The joint (bivariate) distribution $F_{xy}(x, y)$ or, simply, $F(x, y)$ of two random variables $\mathbf{x}$ and $\mathbf{y}$ is the probability of the event

$$\{\mathbf{x} \le x, \mathbf{y} \le y\} = \{(\mathbf{x}, \mathbf{y}) \in D_1\}$$

where x and y are two arbitrary real numbers and D_1 is the quadrant shown in Fig. 6-1a:

$$F(x, y) = P\{\mathbf{x} \le x, \mathbf{y} \le y\} \tag{6-1}$$

[1]The region D is arbitrary subject only to the mild condition that it can be expressed as a countable union or intersection of rectangles.

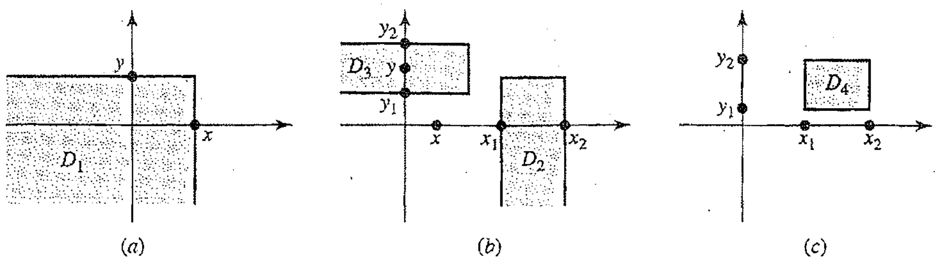


FIGURE 6-1

PROPERTIES

1. The function $F(x, y)$ is such that

$$F(-\infty, y) = 0, \quad F(x, -\infty) = 0, \quad F(\infty, \infty) = 1$$

Proof. As we know, $P\{\mathbf{x} = -\infty\} = P\{\mathbf{y} = -\infty\} = 0$. And since

$$\{\mathbf{x} = -\infty, \mathbf{y} \le y\} \subset \{\mathbf{x} = -\infty\} \qquad \{\mathbf{x} \le x, \mathbf{y} = -\infty\} \subset \{\mathbf{y} = -\infty\}$$

the first two equations follow. The last is a consequence of the identities

$$\{\mathbf{x} \le -\infty, \mathbf{y} \le -\infty\} = S \qquad P(S) = 1$$

2. The event $\{x_1 < \mathbf{x} \le x_2, \mathbf{y} \le y\}$ consists of all points $(\mathbf{x}, \mathbf{y})$ in the vertical half-strip D_2 and the event $\{\mathbf{x} \le x, y_1 < \mathbf{y} \le y_2\}$ consists of all points $(\mathbf{x}, \mathbf{y})$ in the horizontal half-strip D_3 of Fig. 6-1*b*. We maintain that

$$\{x_1 < \mathbf{x} \le x_2, \mathbf{y} \le y\} = F(x_2, y) - F(x_1, y) \tag{6-2}$$

$$\{\mathbf{x} \le x, y_1 < \mathbf{y} \le y_2\} = F(x, y_2) - F(x, y_1) \tag{6-3}$$

Proof. Clearly, for $x_2 > x_1$

$$\{\mathbf{x} \le x_2, \mathbf{y} \le y\} = \{\mathbf{x} \le x_1, \mathbf{y} \le y\} \cup \{x_1 < \mathbf{x} \le x_2, \mathbf{y} \le y\}$$

The last two events are mutually exclusive; hence [see (2-10)]

$$P\{\mathbf{x} \le x_2, \mathbf{y} \le y\} = P\{\mathbf{x} \le x_1, \mathbf{y} \le y\} + P\{x_1 < \mathbf{x} \le x_2, \mathbf{y} \le y\}$$

and (6-2) results. The proof of (6-3) is similar.

3. $$P\{x_1 < \mathbf{x} \le x_2, y_1 < \mathbf{y} \le y_2\} = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \tag{6-4}$$

This is the probability that $(\mathbf{x}, \mathbf{y})$ is in the rectangle D_4 of Fig. 6-1*c*.

Proof. It follows from (6-2) and (6-3) because

$$\{x_1 < \mathbf{x} \le x_2, \mathbf{y} \le y_2\} = \{x_1 < \mathbf{x} \le x_2, \mathbf{y} \le y_1\} \cup \{x_1 < \mathbf{x} \le x_2, y_1 < \mathbf{y} \le y_2\}$$

and the last two events are mutually exclusive.

JOINT DENSITY. The joint density of $\mathbf{x}$ and $\mathbf{y}$ is by definition the function

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} \tag{6-5}$$

From this and property 1 it follows that

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(\alpha, \beta)\, d\alpha\, d\beta \tag{6-6}$$

JOINT STATISTICS. We shall now show that the probability that the point $(\mathbf{x}, \mathbf{y})$ is in a region D of the xy plane equals the integral of $f(x, y)$ in D. In other words,

$$P\{(\mathbf{x}, \mathbf{y}) \in D\} = \int_D \int f(x, y)\, dx\, dy \tag{6-7}$$

where $\{(\mathbf{x}, \mathbf{y}) \in D\}$ is the event consisting of all outcomes ζ such that the point $[\mathbf{x}(\zeta), \mathbf{y}(\zeta)]$ is in D.

Proof. As we know, the ratio

$$\frac{F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - F(x + \Delta x, y) + F(x, y)}{\Delta x\, \Delta y}$$

tends to $\partial F(x, y)/\partial x \partial y$ as $\Delta x \to 0$ and $\Delta y \to 0$. Hence [see (6-4) and (6-5)]

$$P\{x < \mathbf{x} \le x + \Delta x, y < \mathbf{y} \le y + \Delta y\} \simeq f(x, y)\, \Delta x\, \Delta y \tag{6-8}$$

We have thus shown that the probability that $(\mathbf{x}, \mathbf{y})$ is in a differential rectangle equals $f(x, y)$ times the area $\Delta x\, \Delta y$ of the rectangle. This proves (6-7) because the region D can be written as the limit of the union of such rectangles.

MARGINAL STATISTICS. In the study of several random variables, the statistics of each are called marginal. Thus $F_x(x)$ is the *marginal distribution* and $f_x(x)$ the *marginal density* of $\mathbf{x}$. Here, we express the marginal statistics of $\mathbf{x}$ and $\mathbf{y}$ in terms of their joint statistics $F(x, y)$ and $f(x, y)$.

We maintain that

$$F_x(x) = F(x, \infty) \qquad F_y(y) = F(\infty, y) \tag{6-9}$$

$$f_x(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \qquad f_y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx \tag{6-10}$$

Proof. Clearly, $\{\mathbf{x} \le \infty\} = \{\mathbf{y} \le \infty\} = S$; hence

$$\{\mathbf{x} \le x\} = \{\mathbf{x} \le x, \mathbf{y} \le \infty\} \qquad \{\mathbf{y} \le y\} = \{\mathbf{x} \le \infty, \mathbf{y} \le y\}$$

The probabilistics of these two sides yield (6-9).

Differentiating (6-6), we obtain

$$\frac{\partial F(x, y)}{\partial x} = \int_{-\infty}^{y} f(x, \beta)\, d\beta \qquad \frac{\partial F(x, y)}{\partial y} = \int_{-\infty}^{y} f(\alpha, y)\, d\alpha \tag{6-11}$$

Setting $y = \infty$ in the first and $x = \infty$ in the second equation, we obtain (6-10) because [see (6-9)]

$$f_x(x) = \frac{\partial F(x, \infty)}{\partial x} \qquad f_y(x) = \frac{\partial F(\infty, y)}{\partial y}$$

EXISTENCE THEOREM. From properties 1 and 3 it follows that

$$F(-\infty, y) = 0 \qquad F(x, -\infty) = 0 \qquad F(\infty, \infty) = 1 \tag{6-12}$$

and

$$F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \geq 0 \tag{6-13}$$

for every $x_1 < x_2$, $y_1 < y_2$. Hence [see (6-6) and (6-8)]

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1 \qquad f(x, y) \geq 0 \tag{6-14}$$

Conversely, given $F(x, y)$ or $f(x, y)$ as before, we can find two random variables $\mathbf{x}$ and $\mathbf{y}$, defined in some space S, with distribution $F(x, y)$ or density $f(x, y)$. This can be done by extending the existence theorem of Sec. 4-3 to joint statistics.

Probability Masses

The probability that the point (x, y) is in a region D of the plane can be interpreted as the probability mass in this region. Thus the mass in the entire plane equals 1. The mass in the half-plane $\mathbf{x} \leq x$ to the left of the line L_x of Fig. 6-2 equals $F_x(x)$. The mass in the half-plane $\mathbf{y} \leq y$ below the line L_y equals $F_y(y)$. The mass in the doubly-shaded quadrant $\{\mathbf{x} \leq x, \mathbf{y} \leq y\}$ equals $F(x, y)$.

Finally, the mass in the clear quadrant $(\mathbf{x} > x, \mathbf{y} > y)$ equals

$$P\{\mathbf{x} > x, \mathbf{y} > y\} = 1 - F_x(x) - F_y(y) + F(x, y) \tag{6-15}$$

The probability mass in a region D equals the integral [see (6-7)]

$$\int_D \int f(x, y)\, dx\, dy$$

If, therefore, $f(x, y)$ is a bounded function, it can be interpreted as surface mass density.

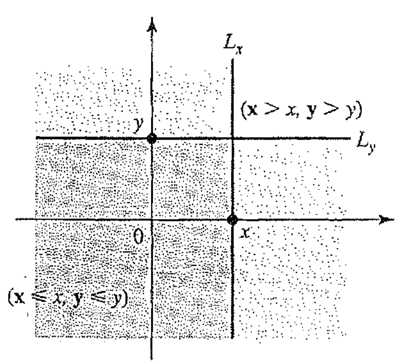


FIGURE 6-2

EXAMPLE 6-1

▶ Suppose that

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{6-16}$$

We shall find the mass m in the circle $x^2 + y^2 \le a^2$. Inserting (6-16) into (6-7) and using the transformation

$$x = r\cos\theta \qquad y = r\sin\theta$$

we obtain

$$m = \frac{1}{2\pi\sigma^2} \int_0^a \int_{-\pi}^{\pi} e^{-r^2/2\sigma^2} r\,dr\,d\theta = 1 - e^{-a^2/2\sigma^2} \tag{6-17}$$

◀

INDEPENDENCE

▶ Two random variables $\mathbf{x}$ and $\mathbf{y}$ are called (*statistically*) *independent* if the events $\{\mathbf{x} \in A\}$ and $\{\mathbf{y} \in B\}$ are independent [see (2-40)], that is, if

$$P\{\mathbf{x} \in A, \mathbf{y} \in B\} = P\{\mathbf{x} \in A\}P\{\mathbf{y} \in B\} \tag{6-18}$$

where A and B are two arbitrary sets on the x and y axes, respectively.

Applying this to the events $\{\mathbf{x} \le x\}$ and $\{\mathbf{y} \le y\}$, we conclude that, if the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent, then

$$F(x, y) = F_x(x)F_y(y) \tag{6-19}$$

Hence

$$f(x, y) = f_x(x)f_y(y) \tag{6-20}$$

It can be shown that, if (6-19) or (6-20) is true, then (6-18) is also true; that is, the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent [see (6-7)]. ◀

EXAMPLE 6-2

BUFFON'S NEEDLE

▶ A fine needle of length $2a$ is dropped at random on a board covered with parallel lines distance $2b$ apart where $b > a$ as in Fig. 6-3a. We shall show that the probability p that the needle intersects one of the lines equals $2a/\pi b$.

In terms of random variables the experiment just discussed can be phrased as: We denote by $\mathbf{x}$ the distance from the center of the needle to the nearest line and by $\boldsymbol{\theta}$ the angle between the needle and the direction perpendicular to the lines. We assume that the random variables $\mathbf{x}$ and $\boldsymbol{\theta}$ are independent, $\mathbf{x}$ is uniform in the interval $(0, b)$, and $\boldsymbol{\theta}$ is uniform in the interval $(0, \pi/2)$. From this it follows that

$$f(x, \theta) = f_x(x)f_\theta(\theta) = \frac{1}{b}\frac{2}{\pi} \qquad 0 \le x \le b \qquad 0 \le \theta \le \frac{\pi}{2}$$

and 0 elsewhere. Hence the probability that the point $(\mathbf{x}, \boldsymbol{\theta})$ is in a region D included in the rectangle R of Fig. 6-3b equals the area of D times $2/\pi b$.

The needle intersects the lines if $x < a\cos\theta$. Hence p equals the shaded area of Fig. 6-3b times $2/\pi b$:

$$p = P\{\mathbf{x} < a\cos\boldsymbol{\theta}\} = \frac{2}{\pi b}\int_0^{2/\pi} a\cos\theta\,d\theta = \frac{2a}{\pi b}$$

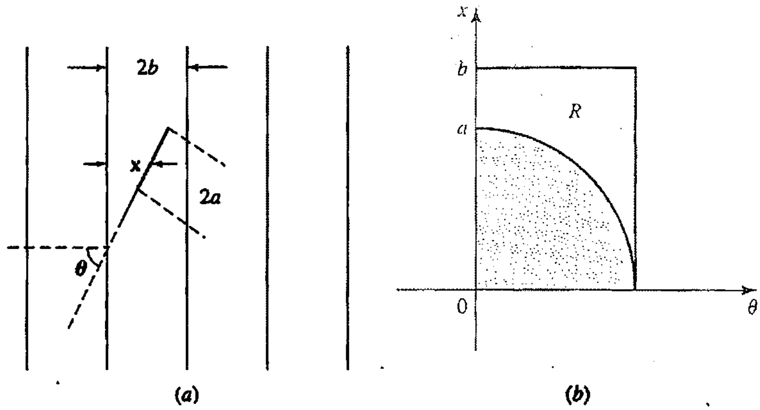


FIGURE 6-3

This can be used to determine experimentally the number π using the relative frequency interpretation of p: If the needle is dropped n times and it intersects the lines n_i times, then

$$\frac{n_i}{n} \simeq p = \frac{2a}{\pi b} \quad \text{hence} \quad \pi \simeq \frac{2an}{bn_i}$$

THEOREM 6-1

If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent, then the random variables

$$\mathbf{z} = g(\mathbf{x}) \qquad \mathbf{w} = h(\mathbf{y})$$

are also independent.

Proof. We denote by A_z the set of points on the x axis such that $g(x) \le z$ and by B_w the set of points on the y axis such that $h(y) \le w$. Clearly,

$$\{\mathbf{z} \le z\} = \{\mathbf{x} \in A_z\} \qquad \{\mathbf{w} \le w\} = \{\mathbf{y} \in B_w\} \tag{6-21}$$

Therefore the events $\{\mathbf{z} \le z\}$ and $\{\mathbf{w} \le w\}$ are independent because the events $\{\mathbf{x} \in A_z\}$ and $\{\mathbf{y} \in B_w\}$ are independent.

INDEPENDENT EXPERIMENTS. As in the case of events (Sec. 3-1), the concept of independence is important in the study of random variables defined on product spaces. Suppose that the random variable $\mathbf{x}$ is defined on a space S_1 consisting of the outcomes $\{\xi_1\}$ and the random variable $\mathbf{y}$ is defined on a space S_2 consisting of the outcomes $\{\xi_2\}$. In the combined experiment $S_1 \times S_2$ the random variables $\mathbf{x}$ and $\mathbf{y}$ are such that

$$\mathbf{x}(\xi_1\xi_2) = \mathbf{x}(\xi_1) \qquad \mathbf{y}(\xi_1\xi_2) = \mathbf{y}(\xi_2) \tag{6-22}$$

In other words, $\mathbf{x}$ depends on the outcomes of S_1 only, and $\mathbf{y}$ depends on the outcomes of S_2 only.

THEOREM 6-2

If the experiments S_1 and S_2 are independent, then the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent.

Proof. We denote by A_x the set $\{\mathbf{x} \le x\}$ in S_1 and by B_y the set $\{\mathbf{y} \le y\}$ in S_2. In the space $S_1 \times S_2$,

$$\{\mathbf{x} \le x\} = A_x \times S_2 \qquad \{\mathbf{y} \le y\} = S_1 \times B_y$$

From the independence of the two experiments, it follows that [see (3-4)] the events $A_x \times S_2$ and $S_1 \times B_y$ are independent. Hence the events $\{\mathbf{x} \le x\}$ and $\{\mathbf{y} \le y\}$ are also independent. ◀

JOINT NORMALITY

▶ We shall say that the random variables $\mathbf{x}$ and $\mathbf{y}$ are *jointly normal* if their joint density is given by

$$f(x, y) = A \exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{(x-\eta_1)^2}{\sigma_1^2} - 2r\frac{(x-\eta_1)(y-\eta_2)}{\sigma_1\sigma_2} + \frac{(y-\eta_2)^2}{\sigma_2^2}\right)\right\} \tag{6-23}$$

This function is positive and its integral equals 1 if

$$A = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \qquad |r| < 1 \tag{6-24}$$

Thus $f(x, y)$ is an exponential and its exponent is a negative quadratic because $|r| < 1$. The function $f(x, y)$ will be denoted by

$$N(\eta_1, \eta_2, \sigma_1^2, \sigma_2^2, r) \tag{6-25}$$

As we shall presently see, η_1 and η_2 are the expected values of $\mathbf{x}$ and $\mathbf{y}$, and σ_1^2 and σ_2^2 their variances. The significance of r will be given later in Example 6-30 (correlation coefficient).

We maintain that the marginal densities of $\mathbf{x}$ and $\mathbf{y}$ are given by

$$f_x(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-\eta_1)^2/2\sigma_1^2} \qquad f_y(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(y-\eta_2)^2/2\sigma_2^2} \tag{6-26}$$

Proof. To prove this, we must show that if (6-23) is inserted into (6-10), the result is (6-26). The bracket in (6-23) can be written in the form

$$(\cdots) = \left(\frac{x-\eta_1}{\sigma_1} - r\frac{y-\eta_2}{\sigma_2}\right)^2 + (1-r^2)\frac{(y-\eta_2)^2}{\sigma_2^2}$$

Hence

$$f_y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx = Ae^{-(y-\eta_2)^2/2\sigma_2^2} \int_{-\infty}^{\infty} e^{-(x-\eta)^2/2(1-r^2)\sigma_1^2}$$

where

$$\eta = \eta_1 + r\frac{(y-\eta_2)\sigma_1}{\sigma_2}$$

The last integral represents a normal random variable with mean μ and variance $(1-r^2)\sigma_1^2$. Therefore the last integral is a constant (independent of x and y) $B = \sqrt{2\pi(1-r^2)\sigma_1^2}$. Therefore

$$f_y(y) = ABe^{(y-\eta_2)^2/2\sigma_2^2}$$

And since $f_y(y)$ is a density, its area must equal 1. This yields $AB = 1/\sigma_2\sqrt{2\pi}$, from which we obtain $A = 1/2\pi\sigma_1\sigma_2\sqrt{1-r^2}$ proving (6-24), and the second equation in (6-26). The proof of the first equation is similar. ◀

Notes 1. From (6-26) it follows that if two random variables are jointly normal, they are also marginally normal. However, as the next example shows, *the converse is not true.*

2. Joint normality can be defined also as follows: Two random variables x and y are jointly normal if the sum $a\mathbf{x} + b\mathbf{y}$ is normal for every a and b [see (7-56)].

EXAMPLE 6-3 ▶ We shall construct two random variables **x** and **y** that are marginally but not jointly normal. Toward this, consider the function

$$f(x, y) = f_x(x) f_y(y)[1 + \rho\{2F_x(x) - 1\}\{2F_y(y) - 1\}] \qquad |\rho| < 1 \tag{6-27}$$

where $f_x(x)$ and $f_y(y)$ are two p.d.fs with respective distribution functions $F_x(x)$ and $F_y(y)$. It is easy to show that $f(x, y) \geq 0$ for all x, y, and

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\, dx\, dy = 1$$

which shows that (6-27) indeed represents a joint p.d.f. of two random variables **x** and **y**. Moreover, by direct integration

$$\int_{-\infty}^{+\infty} f(x, y)\, dy = f_x(x) + \rho(2F_x(x) - 1) f_x(x) \int_{-1}^{1} \frac{u\, du}{2} = f_x(x)$$

where we have made use of the substitution $u = 2F_y(y) - 1$. Similarly,

$$\int_{-\infty}^{+\infty} f(x, y)\, dx = f_y(y)$$

implying that $f_x(x)$ and $f_y(y)$ in (6-27) also represent the respective marginal p.d.f.s of **x** and **y**, respectively.

In particular, let $f_x(x)$ and $f_y(y)$ be normally distributed as in (6-26). In that case (6-27) represents a joint p.d.f. with normal marginals that is not jointly normal. ◀

Circular Symmetry

We say that the joint density of two random variables **x** and **y** is circularly symmetrical if it depends only on the distance from the origin, that is, if

$$f(x, y) = g(r) \qquad r = \sqrt{x^2 + y^2} \tag{6-28}$$

THEOREM 6-3 ▶ If the random variables **x** and **y** are circularly symmetrical and independent, then they are normal with zero mean and equal variance.

Proof. From (6-28) and (6-20) it follows that

$$g(\sqrt{x^2 + y^2}) = f_x(x) f_y(y) \tag{6-29}$$

Since

$$\frac{\partial g(r)}{\partial x} = \frac{dg(r)}{dr}\frac{\partial r}{\partial x} \qquad \text{and} \qquad \frac{\partial r}{\partial x} = \frac{x}{r}$$

we conclude, differentiating (6-29) with respect to x, that

$$\frac{x}{r}g'(r) = f'_x(x)f_y(y)$$

Dividing both sides by $xg(r) = xf_x(x)f_y(y)$, we obtain

$$\frac{1}{r}\frac{g'(r)}{g(r)} = \frac{1}{x}\frac{f'_x(x)}{f_x(x)} \tag{6-30}$$

The right side of (6-30) is independent of y and the left side is a function of $r = \sqrt{x^2+y^2}$. This shows that both sides are independent of x and y. Hence

$$\frac{1}{r}\frac{g'(r)}{g(r)} = \alpha = \text{constant}$$

and (6-28) yields

$$f(x,y) = g(\sqrt{x^2+y^2}) = Ae^{\alpha(x^2+y^2)/2} \tag{6-31}$$

Thus the random variables $\mathbf{x}$ and $\mathbf{y}$ are normal with zero mean and variance $\sigma^2 = -1/\alpha$. ◀

DISCRETE TYPE RANDOM VARIABLES. Suppose the random variables $\mathbf{x}$ and $\mathbf{y}$ are of discrete type taking the values of x_i and y_k with respective probabilities

$$P\{\mathbf{x} = x_i\} = p_i \qquad P\{\mathbf{y} = y_k\} = q_k \tag{6-32}$$

Their joint statistics are determined in terms of the *joint probabilities*

$$P\{\mathbf{x} = x_i, \mathbf{y} = y_k\} = p_{ik} \tag{6-33}$$

Clearly,

$$\sum_{i,k} p_{ik} = 1$$

because, as i and k take all possible values, the events $\{\mathbf{x} = x_i, \mathbf{y} = y_k\}$ are mutually exclusive, and their union equals the certain event.

We maintain that the *marginal probabilities* p_i and q_k can be expressed in terms of the joint probabilities p_{ik}:

$$p_i = \sum_k p_{ik} \qquad q_k = \sum_i p_{ik} \tag{6-34}$$

This is the discrete version of (6-10).

Proof. The events $\{\mathbf{y} = y_k\}$ form a partition of S. Hence as k ranges over all possible values, the events $\{\mathbf{x} = x_i, \mathbf{y} = y_k\}$ are mutually exclusive and their union equals $\{\mathbf{x} = x_i\}$. This yields the first equation in (6-34) [see (2-41)]. The proof of the second is similar.

POINT MASSES. If the random variables $\mathbf{x}$ and $\mathbf{y}$ are of discrete type taking the values x_i and y_k, then the probability masses are 0 everywhere except at the point (x_i, y_k). We have, thus, only point masses and the mass at each point equals p_{ik} [see (6-33)]. The probability $p_i = P\{\mathbf{x} = x_i\}$ equals the sum of all masses p_{ik} on the line $x = x_i$ in agreement with (6-34).

If $i = 1, \ldots, M$ and $k = 1, \ldots, N$, then the number of possible point masses on the plane equals MN. However, as Example 6-4 shows, some of these masses might be 0.

EXAMPLE 6-4

▶ (*a*) In the fair-die experiment, **x** equals the number of dots shown and **y** equals twice this number:

$$\mathbf{x}(f_i) = i \qquad \mathbf{y}(f_i) = 2i \quad i = 1, \ldots, 6$$

In other words, $x_i = i$, $y_k = 2k$, and

$$p_{ik} = P\{\mathbf{x} = i, \mathbf{y} = 2k\} = \begin{cases} \frac{1}{6} & i = k \\ 0 & i \neq k \end{cases}$$

Thus there are masses only on the six points $(i, 2i)$ and the mass of each point equals $1/6$ (Fig. 6-4*a*).

(*b*) We toss the die twice obtaining the 36 outcomes $f_i f_k$ and we define **x** and **y** such that **x** equals the first number that shows, and **y** the second

$$\mathbf{x}(f_i, f_k) = i \qquad \mathbf{y}(f_i f_k) = k \qquad i, k = 1, \ldots, 6$$

Thus $x_i = i$, $y_k = k$, and $p_{ik} = 1/36$. We have, therefore, 36 point masses (Fig. 6-4*b*) and the mass of each point equals $1/36$. On the line $x = i$ there are six points with total mass $1/6$.

(*c*) Again the die is tossed twice but now

$$\mathbf{x}(f_i f_k) = |i - k| \qquad \mathbf{y}(f_i f_k) = i + k$$

In this case, **x** takes the values $0, 1, \ldots, 5$ and **y** the values $2, 3, \ldots, 12$. The number of possible points $6 \times 11 = 66$; however, only 21 have positive masses (Fig. 6-4*c*). Specifically, if $\mathbf{x} = 0$, then $\mathbf{y} = 2$, or $4, \ldots,$ or 12 because if $\mathbf{x} = 0$, then $i = k$ and $\mathbf{y} = 2i$. There are, therefore, six mass points in this line and the mass of each point equals $1/36$. If $\mathbf{x} = 1$, then $\mathbf{y} = 3$, or $5, \ldots,$ or 11. Thus, there are, five mass points on the line $x = 1$ and the mass of each point equals $2/36$. For example, if $\mathbf{x} = 1$ and $\mathbf{y} = 7$, then $i = 3, k = 4$, or $i = 4, k = 3$; hence $P\{\mathbf{x} = 1, \mathbf{y} = 7\} = 2/36$. ◀

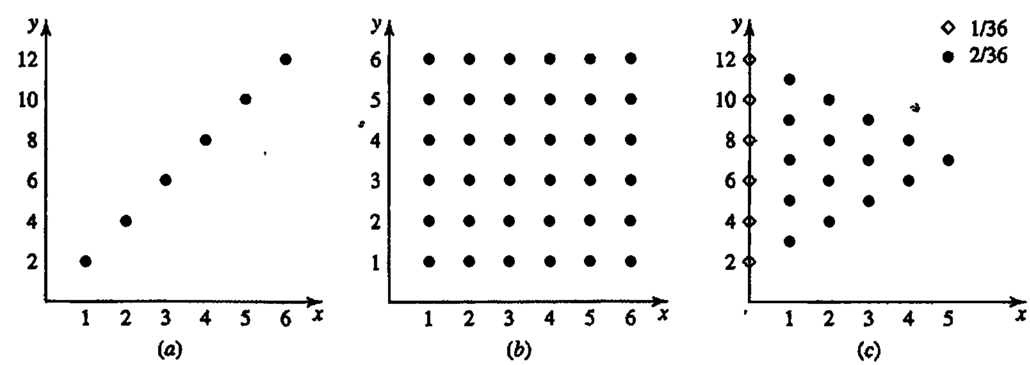
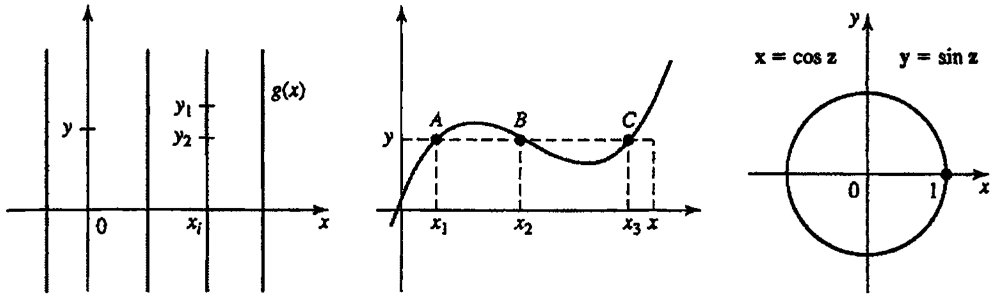


FIGURE 6-4



(a) (b) (c)

FIGURE 6-5

LINE MASSES. These cases lead to line masses:

1. If $\mathbf{x}$ is of discrete type taking the values x_i and $\mathbf{y}$ is of continuous type, then all probability masses are on the vertical lines $x = x_i$ (Fig. 6-5*a*). In particular, the mass between the point y_1 and y_2 on the line $x = x_i$ equals the probability of the event

$$\{\mathbf{x} = x_i, y_1 \le \mathbf{y} \le y_2\}$$

2. If $\mathbf{y} = g(\mathbf{x})$, then all the masses are on the curve $y = g(x)$. In this case, $F(x, y)$ can be expressed in terms of $F_x(x)$. For example, with x and y as in Fig. 6-5*b*, $F(x, y)$ equals the masses on the curve $y = g(x)$ to the left of the point A and between B and C equal $F_x(x_3) - F_x(x_2)$. Hence

$$F(x, y) = F_x(x_1) + F_x(x_3) - F_x(x_2) \qquad y = g(x_1) = g(x_2) = g(x_3)$$

3. If $\mathbf{x} = g(\mathbf{z})$ and $\mathbf{y} = h(\mathbf{z})$, then all probability masses are on the curve $x = g(z)$, $y = h(z)$ specified parametrically. For example, if $g(z) = \cos z$, $h(z) = \sin z$, then the curve is a circle (Fig. 6-5*c*). In this case, the joint statistics of $\mathbf{x}$ and $\mathbf{y}$ can be expressed in terms of $F_z(z)$.

If the random variables $\mathbf{x}$ and $\mathbf{y}$ are of discrete type as in (6-33) and *independent*, then

$$p_{ik} = p_i p_k \tag{6-35}$$

This follows if we apply (6-19) to the events $\{\mathbf{x} = x_i\}$ and $\{\mathbf{y} = y_k\}$. This is the discrete version of (6-20).

EXAMPLE 6-5 ▶ A die with $P\{f_i\} = p_i$ is tossed twice and the random variables $\mathbf{x}$ and $\mathbf{y}$ are such that

$$\mathbf{x}(f_i f_k) = i \qquad \mathbf{y}(f_i f_k) = k$$

Thus $\mathbf{x}$ equals the first number that shows and $\mathbf{y}$ equals the second; hence the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent. This leads to the conclusion that

$$p_{ik} = P\{\mathbf{x} = i, \mathbf{y} = k\} = p_i p_k$$ ◀

6-2 ONE FUNCTION OF TWO RANDOM VARIABLES

Given two random variables $\mathbf{x}$ and $\mathbf{y}$ and a function $g(x, y)$, we form a new random variable $\mathbf{z}$ as

$$\mathbf{z} = g(\mathbf{x}, \mathbf{y}) \tag{6-36}$$

Given the joint p.d.f. $f_{xy}(x, y)$, how does one obtain $f_z(z)$, the p.d.f. of $\mathbf{z}$? Problems of this type are of interest from a practical standpoint. For example, a received signal in a communication scene usually consists of the desired signal buried in noise, and this formulation in that case reduces to $\mathbf{z} = \mathbf{x} + \mathbf{y}$. It is important to know the statistics of the incoming signal for proper receiver design. In this context, we shall analyze problems of the type shown in Fig. 6-6. Referring to (6-36), to start with,

$$\begin{aligned} F_z(z) &= P\{\mathbf{z}(\xi) \le z\} = P\{g(\mathbf{x}, \mathbf{y}) \le z\} = P\{(\mathbf{x}, \mathbf{y}) \in D_z\} \\ &= \iint_{x,y \in D_z} f_{xy}(x, y)\, dx\, dy \end{aligned} \tag{6-37}$$

where D_z in the xy plane represents the region where the inequality $g(x, y) \le z$ is satisfied (Fig. 6-7).

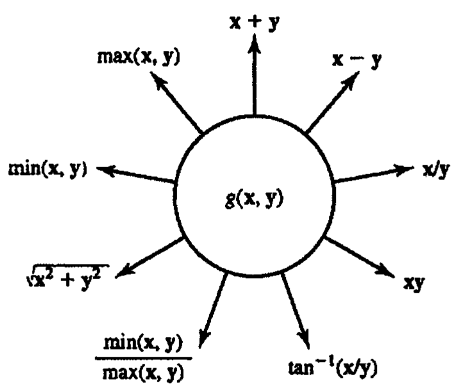


FIGURE 6-6

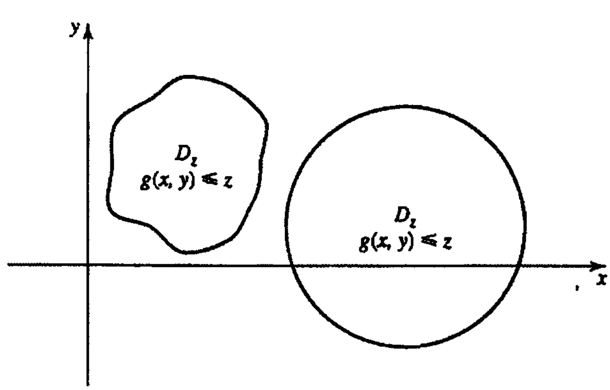


FIGURE 6-7

Note that D_z need not be simply connected. From (6-37), to determine $F_z(z)$ it is enough to find the region D_z for every z, and then evaluate the integral there.

We shall illustrate this method to determine the statistics of various functions of $\mathbf{x}$ and $\mathbf{y}$.

EXAMPLE 6-6

$\mathbf{z} = \mathbf{x} + \mathbf{y}$

▶ Let $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Determine the p.d.f. $f_z(z)$.

From (6-37),

$$F_z(z) = P\{\mathbf{x} + \mathbf{y} \le z\} = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{z-y} f_{xy}(x, y)\, dx\, dy \tag{6-38}$$

since the region D_z of the xy plane where $x + y \le z$ is the shaded area in Fig. 6-8 to the left of the line $x + y \le z$. Integrating over the horizontal strip along the x axis first (inner integral) followed by sliding that strip along the y axis from $-\infty$ to $+\infty$ (outer integral) we cover the entire shaded area.

We can find $f_z(z)$ by differentiating $F_z(z)$ directly. In this context it is useful to recall the differentiation rule due to Leibnitz. Suppose

$$F_z(z) = \int_{a(z)}^{b(z)} f(x, z)\, dx \tag{6-39}$$

Then

$$f_z(z) = \frac{dF_z(z)}{dz} = \frac{db(z)}{dz} f(b(z), z) - \frac{da(z)}{dz} f(a(z), z) + \int_{a(z)}^{b(z)} \frac{\partial f(x, z)}{\partial z}\, dx \tag{6-40}$$

Using (6-40) in (6-38) we get

$$\begin{aligned} f_z(z) &= \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial z} \int_{-\infty}^{z-y} f_{xy}(x, y)\, dx \right) dy \\ &= \int_{-\infty}^{\infty} \left(1 \cdot f_{xy}(z - y, y) - 0 + \int_{-\infty}^{z-y} \frac{\partial f_{xy}(x, y)}{\partial z} \right) dy \\ &= \int_{-\infty}^{\infty} f_{xy}(z - y, y)\, dy \end{aligned} \tag{6-41}$$

Alternatively, the integration in (6-38) can be carried out first along the y axis followed by the x axis as in Fig. 6-9 as well (see problem set).

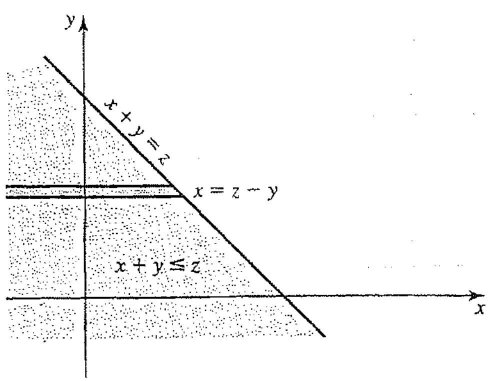


FIGURE 6-8

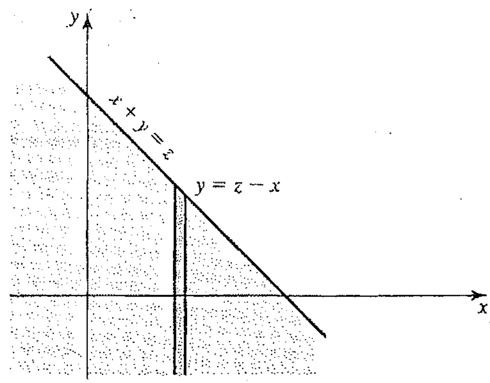


FIGURE 6-9

If x and y are independent, then

$$f_{xy}(x, y) = f_x(x) f_y(y) \tag{6-42}$$

and inserting (6-42) into (6-41) we get

$$f_z(z) = \int_{x=-\infty}^{\infty} f_x(z-y) f_y(y)\,dy = \int_{x=-\infty}^{\infty} f_x(x) f_y(z-x)\,dx \tag{6-43}$$

This integral is the convolution of the functions $f_x(z)$ and $f_y(z)$ expressed two different ways. We thus reach the following conclusion: If two random variables are *independent*, then the density of their sum equals the convolution of their densities.

As a special case, suppose that $f_x(x) = 0$ for $x < 0$ and $f_y(y) = 0$ for $y < 0$, then we can make use of Fig. 6-10 to determine the new limits for D_z.

In that case

$$F_z(z) = \int_{y=0}^{z} \int_{x=0}^{z-y} f_{xy}(x, y)\,dx\,dy$$

or

$$f_z(z) = \int_{y=0}^{z} \left(\frac{\partial}{\partial z} \int_{x=0}^{z-y} f_{xy}(x, y)\,dx \right) dy$$

$$= \begin{cases} \displaystyle\int_0^z f_{xy}(z-y, y)\,dy & z > 0 \\ 0 & z \le 0 \end{cases} \tag{6-44}$$

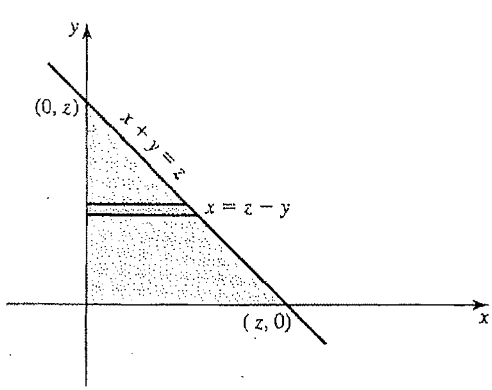


FIGURE 6-10

On the other hand, by considering vertical strips first in Fig. 6-10, we get

$$F_z(z) = \int_{x=0}^{z} \int_{y=0}^{z-x} f_{xy}(x, y)\, dy\, dx$$

or

$$f_z(z) = \int_{x=0}^{z} f_{xy}(x, z-x)\, dx$$

$$= \begin{cases} \int_0^z f_x(x) f_y(z-x)\, dx & z > 0 \\ 0 & z \le 0 \end{cases} \tag{6-45}$$

if **x** and **y** are independent random variables. ◀

EXAMPLE 6-7

▶ Suppose **x** and **y** are independent exponential random variables with common parameter λ. Then

$$f_x(x) = \lambda e^{-\lambda x} U(x) \qquad f_y(y) = \lambda e^{-\lambda y} U(y) \tag{6-46}$$

and we can make use of (6-45) to obtain the p.d.f. of $\mathbf{z} = \mathbf{x} + \mathbf{y}$.

$$f_z(z) = \int_0^z \lambda^2 e^{-\lambda x} e^{-\lambda(z-x)}\, dx = \lambda^2 e^{-\lambda z} \int_0^z dx$$

$$= z\lambda^2 e^{-\lambda z} U(z) \tag{6-47}$$

As Example 6-8 shows, care should be taken while using the convolution formula for random variables with finite range. ◀

EXAMPLE 6-8

▶ **x** and **y** are independent uniform random variables in the common interval (0, 1). Determine $f_z(z)$, where $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Clearly,

$$\mathbf{z} = \mathbf{x} + \mathbf{y} \Rightarrow 0 < z < 2$$

and as Fig. 6-11 shows there are two cases for which the shaded areas are quite different in shape, and they should be considered separately.

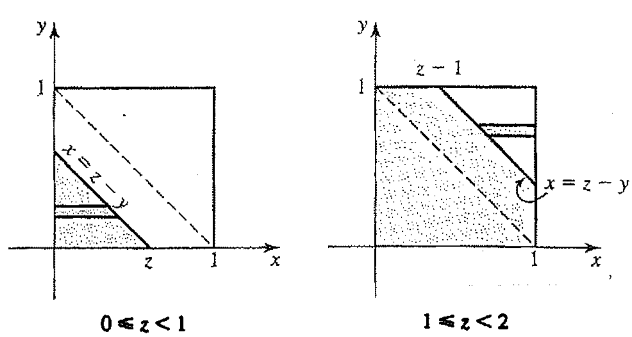


FIGURE 6-11

For $0 \le z < 1$,

$$F_z(z) = \int_{y=0}^{z} \int_{x=0}^{z-y} 1\, dx\, dy = \int_{y=0}^{z} (z - y)\, dy = \frac{z^2}{2} \qquad 0 < z < 1 \tag{6-48}$$

For $1 \le z < 2$, notice that it is easy to deal with the unshaded region. In that case,

$$F_z(z) = 1 - P\{\mathbf{z} > z\} = 1 - \int_{y=z-1}^{1} \int_{x=z-y}^{1} 1\, dx\, dy$$

$$= 1 - \int_{y=z-1}^{1} (1 - z + y)\, dy = 1 - \frac{(2-z)^2}{2} \qquad 1 \le z < 2 \tag{6-49}$$

Thus

$$f_z(z) = \frac{dF_z(z)}{dz} = \begin{cases} z & 0 \le z < 1 \\ 2 - z & 1 \le z < 2 \end{cases} \tag{6-50}$$

By direct convolution of $f_x(x)$ and $f_y(y)$, we obtain the same result as above. In fact, for $0 \le z < 1$ (Fig. 6-12a)

$$f_z(z) = \int f_x(z - x) f_y(x)\, dx = \int_0^z 1\, dx = z \tag{6-51}$$

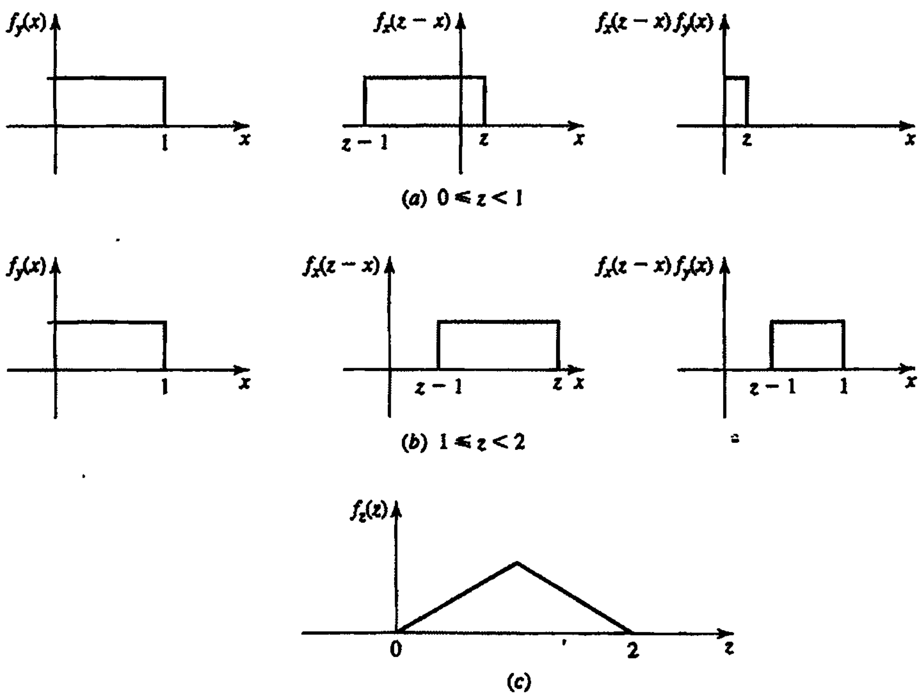


FIGURE 6-12

and for $1 \le z < 2$ (Fig. 6-12*b*)

$$f_z(z) = \int_{z-1}^{1} 1\,dx = 2 - z \tag{6-52}$$

Fig. 6-12*c* shows $f_z(z)$, which agrees with the convolution of two rectangular waveforms as well. ◀

EXAMPLE 6-9

$\mathbf{z} = \mathbf{x} - \mathbf{y}$

▶ Let $\mathbf{z} = \mathbf{x} - \mathbf{y}$. Determine $f_z(z)$.

From (6-37) and Fig. 6-13

$$F_z(z) = P\{\mathbf{x} - \mathbf{y} \le z\} = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{z+y} f_{xy}(x, y)\,dx\,dy$$

and hence

$$f_z(z) = \frac{dF_z(z)}{dz} = \int_{-\infty}^{\infty} f_{xy}(z + y, y)\,dy \tag{6-53}$$

If $\mathbf{x}$ and $\mathbf{y}$ are independent, then this formula reduces to

$$f_z(z) = \int_{-\infty}^{\infty} f_x(z + y) f_y(y)\,dy = f_x(-z) \otimes f_y(y) \tag{6-54}$$

which represents the convolution of $f_x(-z)$ with $f_y(z)$.

As a special case, suppose

$$f_x(x) = 0 \qquad x < 0, \qquad f_y(y) = 0 \qquad y < 0$$

In this case, $\mathbf{z}$ can be negative as well as positive, and that gives rise to two situations that should be analyzed separately, since the regions of integration for $z \ge 0$ and $z < 0$ are quite different.

For $z \ge 0$, from Fig. 6-14*a*

$$F_z(z) = \int_{y=0}^{\infty} \int_{x=0}^{z+y} f_{xy}(x, y)\,dx\,dy$$

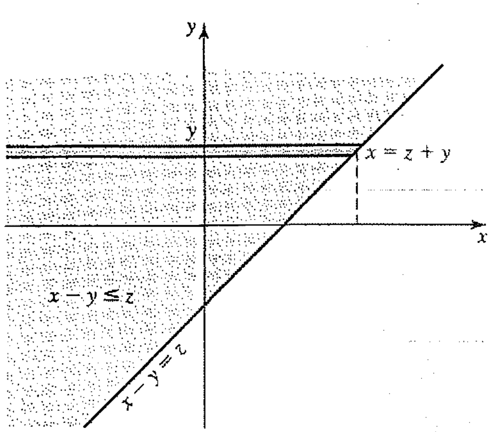


FIGURE 6-13

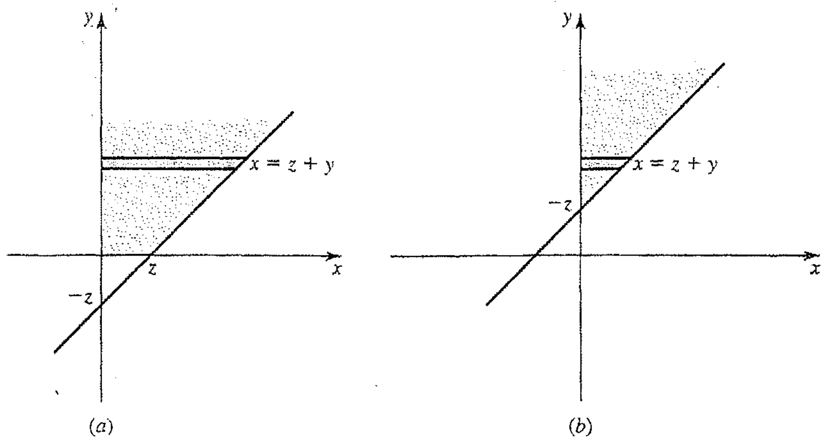


FIGURE 6-14

and for $z < 0$, from Fig. 6-14b

$$F_z(z) = \int_{y=-z}^{\infty} \int_{x=0}^{z+y} f_{xy}(x, y)\, dx\, dy$$

After differentiation, this gives

$$f_z(z) = \begin{cases} \displaystyle\int_0^{\infty} f_{xy}(z+y, y)\, dy & z \geq 0 \\ \displaystyle\int_{-z}^{\infty} f_{xy}(z+y, y)\, dy & z < 0 \end{cases} \tag{6-55}$$

◀

EXAMPLE 6-10

$\mathbf{z} = \mathbf{x}/\mathbf{y}$

▶ Let $\mathbf{z} = \mathbf{x}/\mathbf{y}$. Determine $f_z(z)$.

We have

$$F_z(z) = P\{\mathbf{x}/\mathbf{y} \leq z\} \tag{6-56}$$

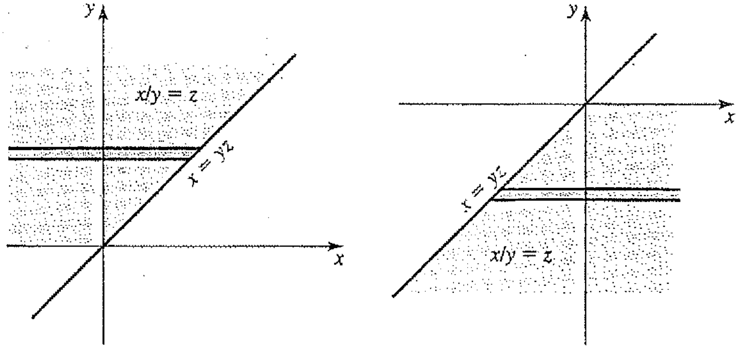
The inequality $x/y \leq z$ can be rewritten as $\mathbf{x} \leq \mathbf{y}z$ if $\mathbf{y} > 0$, and $\mathbf{x} \geq \mathbf{y}z$ if $\mathbf{y} < 0$. Hence the event $\{\mathbf{x}/\mathbf{y} \leq z\}$ in (6-56) needs to be conditioned by the event $A = \{\mathbf{y} > 0\}$ and its compliment $\overline{A}$. Since $A \cup \overline{A} = S$, by the partition theorem, we have

$$\begin{aligned} P\{\mathbf{x}/\mathbf{y} \leq z\} &= P\{\mathbf{x}/\mathbf{y} \leq z \cap (A \cup \overline{A})\} \\ &= P\{\mathbf{x}/\mathbf{y} \leq z, \mathbf{y} > 0\} + P\{\mathbf{x}/\mathbf{y} \leq z, \mathbf{y} < 0\} \\ &= P\{\mathbf{x} \leq \mathbf{y}z, \mathbf{y} > 0\} + P\{\mathbf{x} \geq \mathbf{y}z, \mathbf{y} < 0\} \end{aligned} \tag{6-57}$$

Fig. 6-15a shows the area corresponding to the first term, and Fig. 6-15b shows that corresponding to the second term in (6-57).

Integrating over these two regions, we get

$$F_z(z) = \int_{y=0}^{\infty} \int_{x=-\infty}^{yz} f_{xy}(x, y)\, dx\, dy + \int_{y=-\infty}^{0} \int_{x=yz}^{\infty} f_{xy}(x, y)\, dx\, dy \tag{6-58}$$



(a) $P(\mathbf{x} \leq \mathbf{y}z, \mathbf{y} > 0)$ (b) $P(\mathbf{x} \geq \mathbf{y}z, \mathbf{y} < 0)$

FIGURE 6-15

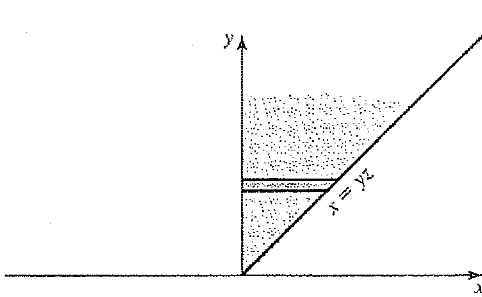


FIGURE 6-16

Differentiation gives

$$f_z(z) = \int_0^{\infty} y f_{xy}(yz, y)\,dy + \int_{-\infty}^{0} -y f_{xy}(yz, y)\,dy$$

$$= \int_{-\infty}^{\infty} |y| f_{xy}(yz, y)\,dy \tag{6-59}$$

Note that if $\mathbf{x}$ and $\mathbf{y}$ are non-negative random variables, then the area of integration reduces to that shown in Fig. 6-16.

This gives

$$F_z(z) = \int_{y=0}^{\infty} \int_{x=0}^{yz} f_{xy}(x, y)\,dx\,dy$$

or

$$f_z(z) = \int_{y=0}^{\infty} y f_{xy}(yz, y)\,dy \tag{6-60}$$

◀

EXAMPLE 6-11 ▶ $\mathbf{x}$ and $\mathbf{y}$ are jointly normal random variables with zero mean and

$$f_{xy}(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\left[\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right]} \tag{6-61}$$

Show that the ratio $\mathbf{z} = \mathbf{x}/\mathbf{y}$ has a Cauchy density centered at $r\sigma_1/\sigma_2$.

SOLUTION

Inserting (6-61) into (6-59) and using the fact that $f_{xy}(-x, -y) = f_{xy}(x, y)$, we obtain

$$f_z(z) = \frac{2}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int_0^\infty y e^{-y^2/2\sigma_0^2}\,dy = \frac{\sigma_0^2}{\pi\sigma_1\sigma_2\sqrt{1-r^2}}$$

where

$$\sigma_0^2 = \frac{1-r^2}{\left(z^2/\sigma_1^2\right) - (2rz/\sigma_1\sigma_2) + \left(1/\sigma_2^2\right)}$$

Thus

$$f_z(z) = \frac{\sigma_1\sigma_2\sqrt{1-r^2}/\pi}{\sigma_2^2(z - r\sigma_1/\sigma_2)^2 + \sigma_1^2(1-r^2)} \tag{6-62}$$

which represents a Cauchy random variable centered at $r\sigma_1/\sigma_2$. Integrating (6-62) from $-\infty$ to z, we obtain the corresponding distribution function to be

$$F_z(z) = \frac{1}{2} + \frac{1}{\pi}\arctan\frac{\sigma_2 z - r\sigma_1}{\sigma_1\sqrt{1-r^2}} \tag{6-63}$$

◀

As an application, we can use (6-63) to determine the probability masses m_1, m_2, m_3, and m_4 in the four quadrants of the xy plane for (6-61). From the spherical symmetry of (6-61), we have

$$m_1 = m_3 \qquad m_2 = m_4$$

But the second and fourth quadrants represent the region of the plane where $x/y < 0$. The probability that the point (x, y) is in that region equals, therefore, the probability that the random variable $\mathbf{z} = \mathbf{x}/\mathbf{y}$ is negative. Thus

$$m_2 + m_4 = P(\mathbf{z} \le 0) = F_z(0) = \frac{1}{2} - \frac{1}{\pi}\arctan\frac{r}{\sqrt{1-r^2}}$$

and

$$m_1 + m_3 = 1 - (m_2 + m_4) = \frac{1}{2} + \frac{1}{\pi}\arctan\frac{r}{\sqrt{1-r^2}}$$

If we define $\alpha = \text{arc tan}\, r/\sqrt{1-r^2}$, this gives

$$m_1 = m_3 = \frac{1}{4} + \frac{\alpha}{2\pi} \qquad m_2 = m_4 = \frac{1}{4} - \frac{\alpha}{2\pi} \tag{6-64}$$

Of course, we could have obtained this result by direct integration of (6-61) in each quadrant. However, this is simpler.

EXAMPLE 6-12

▶ Let $\mathbf{x}$ and $\mathbf{y}$ be independent gamma random variables with $\mathbf{x} \sim G(m, \alpha)$ and $\mathbf{y} \sim G(n, \alpha)$. Show that $\mathbf{z} = \mathbf{x}/(\mathbf{x}+\mathbf{y})$ has a beta distribution.

Proof.

$$f_{xy}(x, y) = f_x(x) f_y(y)$$

$$= \frac{1}{\alpha^{m+n}\Gamma(m)\Gamma(n)} x^{m-1} y^{n-1} e^{-(x+y)/\alpha} \qquad x > 0 \qquad y > 0 \tag{6-65}$$

Note that $0 < z < 1$, since $\mathbf{x}$ and $\mathbf{y}$ are non-negative random variables

$$F_z(z) = P\{\mathbf{z} \le z\} = P\left(\frac{\mathbf{x}}{\mathbf{x}+\mathbf{y}} \le z\right) = P\left(\mathbf{x} \le \mathbf{y}\frac{z}{1-z}\right)$$
$$= \int_0^\infty \int_0^{yz/(1-z)} f_{xy}(x, y)\,dx\,dy$$

where we have made use of Fig. 6-16. Differentiation with respect to z gives

$$f_z(z) = \int_0^\infty \frac{y}{(1-z)^2} f_{xy}(yz/(1-z), y)\,dy$$
$$= \int_0^\infty \frac{y}{(1-z)^2} \frac{1}{\alpha^{m+n}\Gamma(m)\Gamma(n)} \left(\frac{yz}{1-z}\right)^{m-1} y^{n-1} e^{-y/(1-z)\alpha}\,dy$$
$$= \frac{1}{\alpha^{m+n}\Gamma(m)\Gamma(n)} \frac{z^{m-1}}{(1-z)^{m+1}} \int_0^\infty y^{m+n-1} e^{-y/\alpha(1-z)}\,dy$$
$$= \frac{z^{m-1}(1-z)^{n-1}}{\Gamma(m)\Gamma(n)} \int_0^\infty u^{m+n-1} e^{-u}\,du = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} z^{m-1}(1-z)^{n-1}$$
$$= \begin{cases} \dfrac{1}{\beta(m,n)} z^{m-1}(1-z)^{n-1} & 0 < z < 1 \\ 0 & \text{otherwise} \end{cases} \tag{6-66}$$

which represents a beta distribution. ◀

EXAMPLE 6-13

$\mathbf{z} = \mathbf{x}^2 + \mathbf{y}^2$

▶ Let $\mathbf{z} = \mathbf{x}^2 + \mathbf{y}^2$. Determine $f_z(z)$.

We have

$$F_z(z) = P\{x^2 + y^2 \le z\} = \iint_{x^2+y^2 \le z} f_{xy}(x, y)\,dx\,dy$$

But, $\mathbf{x}^2 + \mathbf{y}^2 \le z$ represents the area of a circle with radius $\sqrt{z}$, and hence (see Fig. 6-17)

$$F_z(z) = \int_{y=-\sqrt{z}}^{\sqrt{z}} \int_{x=-\sqrt{z-y^2}}^{\sqrt{z-y^2}} f_{xy}(x, y)\,dx\,dy$$

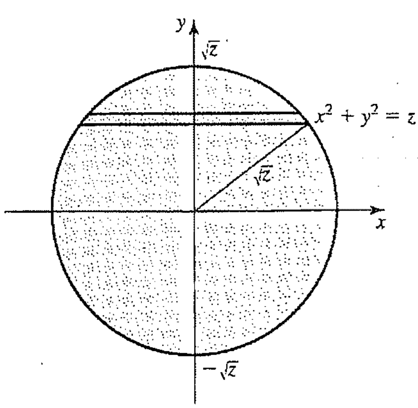


FIGURE 6-17

This gives

$$f_z(z) = \int_{-\sqrt{z}}^{\sqrt{z}} \frac{1}{2\sqrt{z-y^2}} \{f_{xy}(\sqrt{z-y^2}, y) + f_{xy}(-\sqrt{z-y^2}, y)\}\, dy \qquad (6\text{-}67)$$

◀

As an illustration, consider Example 6-14.

EXAMPLE 6-14

▶ $\mathbf{x}$ and $\mathbf{y}$ are independent normal random variables with zero mean and common variance σ^2. Determine $f_z(z)$ for $\mathbf{z} = \mathbf{x}^2 + \mathbf{y}^2$.

SOLUTION

Using (6-67), we get

$$\begin{aligned} f_z(z) &= \int_{-\sqrt{z}}^{\sqrt{z}} \frac{1}{2\sqrt{z-y^2}} \left(2 \cdot \frac{1}{2\pi\sigma^2} e^{(z-y^2+y^2)/2\sigma^2}\right) dy \\ &= \frac{e^{-z/2\sigma^2}}{\pi\sigma^2} \int_0^{\sqrt{z}} \frac{1}{\sqrt{z-y^2}}\, dy = \frac{e^{-z/2\sigma^2}}{\pi\sigma^2} \int_0^{\pi/2} \frac{\sqrt{z}\cos\theta}{\sqrt{z}\cos\theta}\, d\theta \\ &= \frac{1}{2\sigma^2} e^{-z/2\sigma^2} U(z) \end{aligned} \qquad (6\text{-}68)$$

where we have used the substitution $y = \sqrt{z}\sin\theta$. From (6-68), we have the following: If $\mathbf{x}$ and $\mathbf{y}$ are independent zero mean Gaussian random variables with common variance σ^2, then $\mathbf{x}^2 + \mathbf{y}^2$ is an exponential random variable with parameter $2\sigma^2$. ◀

EXAMPLE 6-15

$\mathbf{z} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$

▶ Let $\mathbf{z} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$. Find $f_z(z)$.

SOLUTION

From Fig. 6-17, the present case corresponds to a circle with radius z^2. Thus

$$F_z(z) = \int_{y=-z}^{z} \int_{x=-\sqrt{z^2-y^2}}^{\sqrt{z^2-y^2}} f_{xy}(x, y)\, dx\, dy$$

and by differentiation,

$$f_z(z) = \int_{-z}^{z} \frac{z}{\sqrt{z^2-y^2}} \{f_{xy}(\sqrt{z^2-y^2}, y) + f_{xy}(-\sqrt{z^2-y^2}, y)\}\, dy \qquad (6\text{-}69)$$

In particular, if $\mathbf{x}$ and $\mathbf{y}$ are zero mean independent Gaussian random variables as in the previous example, then

$$\begin{aligned} f_z(z) &= 2\int_0^{z} \frac{z}{\sqrt{z^2-y^2}} \frac{2}{2\pi\sigma^2} e^{-(z^2-y^2+y^2)/2\sigma^2}\, dy \\ &= \frac{2z}{\pi\sigma^2} e^{-z^2/2\sigma^2} \int_0^{z} \frac{1}{\sqrt{z^2-y^2}}\, dy = \frac{2z}{\pi\sigma^2} e^{-z^2/2\sigma^2} \int_0^{\pi/2} \frac{z\cos\theta}{z\cos\theta}\, d\theta \\ &= \frac{z}{\sigma^2} e^{-z^2/2\sigma^2} U(z) \end{aligned} \qquad (6\text{-}70)$$

which represents a Rayleigh distribution. Thus, if $\mathbf{w} = \mathbf{x} + i\mathbf{y}$, where $\mathbf{x}$ and $\mathbf{y}$ are real independent normal random variables with zero mean and equal variance, then the random variable $|\mathbf{w}| = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$ has a Rayleigh density. $\mathbf{w}$ is said to be a complex Gaussian random variable with zero mean, if its real and imaginary parts are independent. So far we have seen that the magnitude of a complex Gaussian random variable has Rayleigh distribution. What about its phase

$$\theta = \tan^{-1}\left(\frac{\mathbf{y}}{\mathbf{x}}\right) \tag{6-71}$$

Clearly, the principal value of θ lies in the interval $(-\pi/2, \pi/2)$. If we let $\mathbf{u} = \tan\theta = \mathbf{y}/\mathbf{x}$, then from Example 6-11, $\mathbf{u}$ has a Cauchy distribution (see (6-62) with $\sigma_1 = \sigma_2$, $r = 0$)

$$f_u(u) = \frac{1/\pi}{u^2+1} \qquad -\infty < u < \infty$$

As a result, the principal value of θ has the density function

$$f_\theta(\theta) = \frac{1}{|d\theta/du|} f_u(\tan\theta) = \frac{1}{(1/\sec^2\theta)} \frac{1/\pi}{\tan^2\theta + 1}$$

$$= \begin{cases} 1/\pi & -\pi/2 < \theta < \pi/2 \\ 0 & \text{otherwise} \end{cases} \tag{6-72}$$

However, in the representation $\mathbf{x} + j\mathbf{y} = \mathbf{r}e^{j\theta}$, the variable θ lies in the interval $(-\pi, \pi)$, and taking into account this scaling by a factor of two, we obtain

$$f_\theta(\theta) = \begin{cases} 1/2\pi & -\pi < \theta < \pi \\ 0 & \text{otherwise} \end{cases} \tag{6-73}$$

To summarize, the magnitude and phase of a zero mean complex Gaussian random variable have Rayleigh and uniform distributions respectively. Interestingly, as we will show later (Example 6-22), these two derived random variables are also statistically *independent* of each other! ◀

Let us reconsider Example 6-15 where $\mathbf{x}$ and $\mathbf{y}$ are independent Gaussian random variables with nonzero means μ_x and μ_y respectively. Then $\mathbf{z} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$ is said to be a Rician random variable. Such a scene arises in fading multipath situations where there is a dominant constant component (mean) in addition to a zero mean Gaussian random variable. The constant component may be the line of sight signal and the zero mean Gaussian random variable part could be due to random multipath components adding up incoherently. The envelope of such a signal is said to be Rician instead of Rayleigh.

EXAMPLE 6-16

▶ Redo Example 6-15, where $\mathbf{x}$ and $\mathbf{y}$ are independent Gaussian random variables with nonzero means μ_x and μ_y respectively.

SOLUTION

Since

$$f_{xy}(x, y) = \frac{1}{2\pi\sigma^2} e^{-[(x-\mu_x)^2+(y-\mu_y)^2]/2\sigma^2}$$

substituting this into (6-69) and letting $y = z\sin\theta$, $\mu = \sqrt{\mu_x^2 + \mu_y^2}$, $\mu_x = \mu\cos\phi$,

$\mu_y = \mu \sin\phi$, we get the Rician distribution to be

$$\begin{aligned} f_z(z) &= \frac{ze^{-(z^2+\mu^2)/2\sigma^2}}{2\pi\sigma^2} \int_{-\pi/2}^{\pi/2} \left(e^{z\mu\cos(\theta-\phi)/\sigma^2} + e^{-z\mu\cos(\theta+\phi)/\sigma^2}\right) d\theta \\ &= \frac{ze^{-(z^2+\mu^2)/2\sigma^2}}{2\pi\sigma^2} \left(\int_{-\pi/2}^{\pi/2} e^{z\mu\cos(\theta-\phi)/\sigma^2}\, d\theta + \int_{\pi/2}^{3\pi/2} e^{z\mu\cos(\theta-\phi)/\sigma^2}\, d\theta\right) \\ &= \frac{ze^{-(z^2+\mu^2)/2\sigma^2}}{\sigma^2} I_0\left(\frac{z\mu}{\sigma^2}\right) \end{aligned} \tag{6-74}$$

where

$$I_0(\eta) \triangleq \frac{1}{2\pi}\int_0^{2\pi} e^{\eta\cos(\theta-\phi)}\, d\theta = \frac{1}{\pi}\int_0^{\pi} e^{\eta\cos\theta}\, d\theta$$

is the modified Bessel function of the first kind and zeroth order. ◀

Order Statistics

In general, given any n-tuple $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, we can rearrange them in an increasing order of magnitude such that

$$\mathbf{x}_{(1)} \le \mathbf{x}_{(2)} \le \cdots \le \mathbf{x}_{(n)}$$

where $\mathbf{x}_{(1)} = \min(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$, and $\mathbf{x}_{(2)}$ is the second smallest value among $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, and finally $\mathbf{x}_{(n)} = \max(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. The functions *min* and *max* are nonlinear operators, and represent special cases of the more general order statistics. If $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ represent random variables, the function $\mathbf{x}_{(k)}$ that takes on the value $x_{(k)}$ in each possible sequence $(x_1, x_2, \ldots, x_n)$ is known as the kth-order statistic. $\{\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \ldots, \mathbf{x}_{(n)}\}$ represent the set of order statistics among n random variables. In this context

$$\mathbf{R} = \mathbf{x}_{(n)} - \mathbf{x}_{(1)} \tag{6-75}$$

represents the range, and when $n = 2$, we have the *max* and *min* statistics.

Order statistics is useful when relative magnitude of observations is of importance. When worst case scenarios have to be accounted for, then the function $max(\cdot)$ is quite useful. For example, let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ represent the recorded flood levels over the past n years at some location. If the objective is to construct a dam to prevent any more flooding, then the height H of the proposed dam should satisfy the inequality

$$H > \max(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \tag{6-76}$$

with some finite probability. In that case, the p.d.f. of the random variable on the right side of (6-76) can be used to compute the desired height. In another case, if a bulb manufacturer wants to determine the average time to failure (μ) of its bulbs based on a sample of size n, the sample mean $(\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n)/n$ can be used as an estimate for μ. On the other hand, an estimate based on the least time to failure has other attractive features. This estimate $\min(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ may not be as good as the sample mean in terms of their respective variances, but the $min(\cdot)$ can be computed as soon as the first bulb fuses, whereas to compute the sample mean one needs to wait till the last of the lot extinguishes.

EXAMPLE 6-17

$\mathbf{z} = \max(\mathbf{x}, \mathbf{y})$
$\mathbf{w} = \min(\mathbf{x}, \mathbf{y})$

▶ Let $\mathbf{z} = \max(\mathbf{x}, \mathbf{y})$ and $\mathbf{w} = \min(\mathbf{x}, \mathbf{y})$. Determine $f_z(z)$ and $f_w(w)$.

$$\mathbf{z} = \max(\mathbf{x}, \mathbf{y}) = \begin{cases} \mathbf{x} & \mathbf{x} > \mathbf{y} \\ \mathbf{y} & \mathbf{x} \le \mathbf{y} \end{cases} \tag{6-77}$$

we have [see (6-57)]

$$\begin{aligned} F_z(z) &= P\{\max(\mathbf{x}, \mathbf{y}) \le z\} \\ &= P\{(\mathbf{x} \le z, \mathbf{x} > \mathbf{y}) \cup (\mathbf{y} \le z, \mathbf{x} \le \mathbf{y})\} \\ &= P\{\mathbf{x} \le z, \mathbf{x} > \mathbf{y}\} + P\{\mathbf{y} \le z, \mathbf{x} \le \mathbf{y}\} \end{aligned}$$

since $\{\mathbf{x} > \mathbf{y}\}$ and $\{\mathbf{x} \le \mathbf{y}\}$ are mutually exclusive sets that form a partition. Figure 6-18*a* and 6-18*b* show the regions satisfying the corresponding inequalities in each term seen here.

Figure 6-18*c* represents the total region, and from there

$$F_z(z) = P\{\mathbf{x} \le z, \mathbf{y} \le z\} = F_{xy}(z, z) \tag{6-78}$$

If $\mathbf{x}$ and $\mathbf{y}$ are independent, then

$$F_z(z) = F_x(z) F_y(z)$$

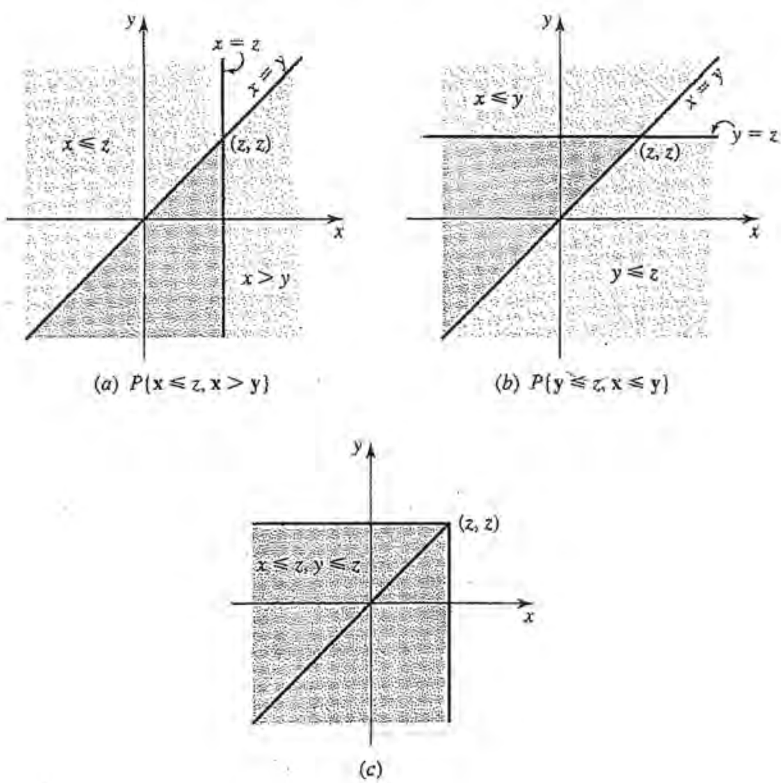


FIGURE 6-18

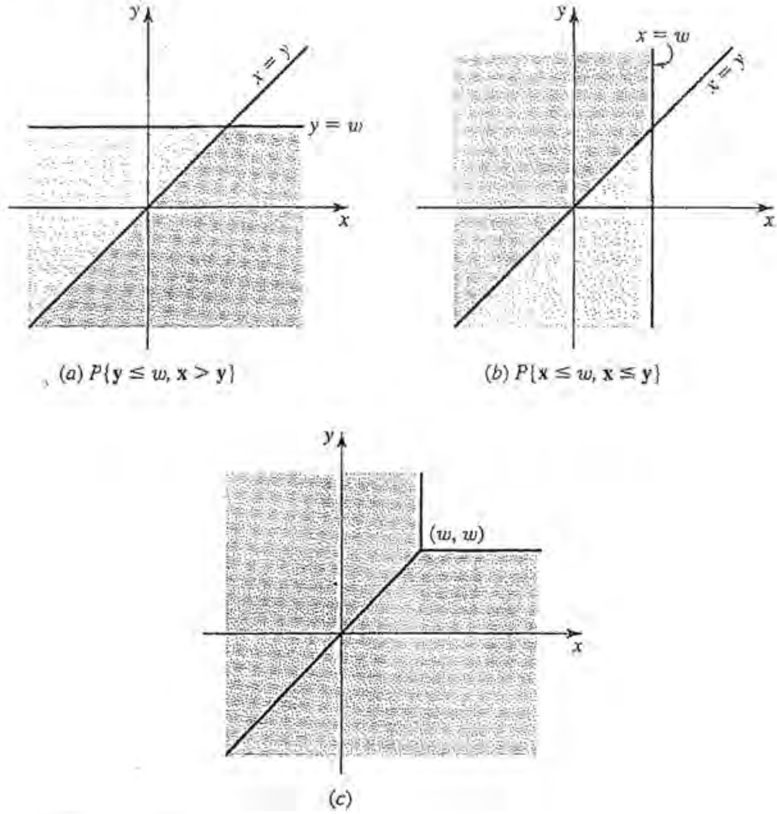


FIGURE 6-19

and hence

$$f_z(z) = F_x(z)f_y(z) + f_x(z)F_y(z) \tag{6-79}$$

Similarly,

$$\mathbf{w} = \min(\mathbf{x}, \mathbf{y}) = \begin{cases} \mathbf{y} & \mathbf{x} > \mathbf{y} \\ \mathbf{x} & \mathbf{x} \le \mathbf{y} \end{cases} \tag{6-80}$$

Thus,

$$\begin{aligned} F_w(w) &= P\{\min(\mathbf{x}, \mathbf{y}) \le w\} \\ &= P\{\mathbf{y} \le w, \mathbf{x} > \mathbf{y}\} + P\{\mathbf{x} \le w, \mathbf{x} \le \mathbf{y}\} \end{aligned}$$

Once again, the shaded areas in Fig. 6-19*a* and 6-19*b* show the regions satisfying these inequalities, and Fig. 6-19*c* shows them together.

From Fig. 6-19*c*,

$$\begin{aligned} F_w(w) &= 1 - P\{\mathbf{w} > w\} = 1 - P\{\mathbf{x} > w, \mathbf{y} > w\} \\ &= F_x(w) + F_y(w) - F_{xy}(w, w) \end{aligned} \tag{6-81}$$

where we have made use of (6-4) with $x_2 = y_2 = \infty$, and $x_1 = y_1 = w$. ◀

EXAMPLE 6-18

▶ Let x and y be independent exponential random variables with common parameter λ. Define $\mathbf{w} = \min(\mathbf{x}, \mathbf{y})$. Find $f_w(w)$.

SOLUTION

From (6-81)

$$F_w(w) = F_x(w) + F_y(w) - F_x(w)F_y(w)$$

and hence

$$f_w(w) = f_x(w) + f_y(w) - f_x(w)F_y(w) - F_x(w)f_y(w)$$

But $f_x(w) = f_y(w) = \lambda e^{-\lambda w}$, and $F_x(w) = F_y(w) = 1 - e^{-\lambda w}$, so that

$$f_w(w) = 2\lambda e^{\lambda w} - 2(1 - e^{-\lambda w})\lambda e^{-\lambda w} = 2\lambda e^{-2\lambda w}U(w) \tag{6-82}$$

Thus min(x, y) is also exponential with parameter 2λ. ◀

EXAMPLE 6-19

▶ Suppose x and y are as given in Example 6-18. Define

$$z = \frac{\min(\mathbf{x}, \mathbf{y})}{\max(\mathbf{x}, \mathbf{y})}$$

Although $\min(\cdot)/\max(\cdot)$ represents a complicated function, by partitioning the whole space as before, it is possible to simplify this function. In fact

$$\mathbf{z} = \begin{cases} \mathbf{x}/\mathbf{y} & \mathbf{x} \le \mathbf{y} \\ \mathbf{y}/\mathbf{x} & \mathbf{x} > \mathbf{y} \end{cases} \tag{6-83}$$

As before, this gives

$$\begin{aligned} F_z(z) &= P\{\mathbf{x}/\mathbf{y} \le z, \mathbf{x} \le \mathbf{y}\} + P\{\mathbf{y}/\mathbf{x} \le z, \mathbf{x} > \mathbf{y}\} \\ &= P\{\mathbf{x} \le \mathbf{y}z, \mathbf{x} \le \mathbf{y}\} + P\{\mathbf{y} \le \mathbf{x}z, \mathbf{x} > \mathbf{y}\} \end{aligned}$$

Since x and y are both positive random variables in this case, we have $0 < z < 1$. The shaded regions in Fig. 6-20*a* and 6-20*b* represent the two terms in this sum.

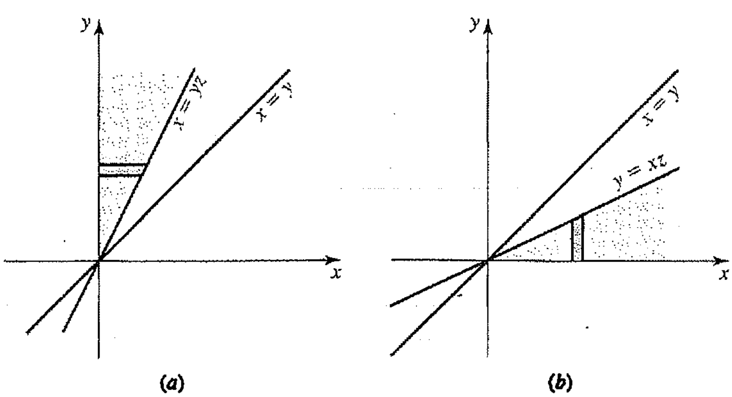


FIGURE 6-20

From Fig. 6-20,

$$F_z(z) = \int_0^{\infty}\int_{x=0}^{yz} f_{xy}(x, y)\, dx\, dy + \int_0^{\infty}\int_{y=0}^{xz} f_{xy}(x, y)\, dy\, dx$$

Hence

$$\begin{aligned} f_z(z) &= \int_0^{\infty} y f_{xy}(yz, y)\, dy + \int_0^{\infty} x f_{xy}(x, xz)\, dx \\ &= \int_0^{\infty} y\big(f_{xy}(yz, y) + f_{xy}(y, yz)\big)\, dy \\ &= \int_0^{\infty} y\lambda^2\big(e^{-\lambda(yz+y)} + e^{-\lambda(y+yz)}\big)\, dy \\ &= 2\lambda^2 \int_0^{\infty} y e^{-\lambda(1+z)y}\, dy = \frac{2}{(1+z)^2}\int_0^{\infty} u e^{-u}\, du \\ &= \begin{cases} \dfrac{2}{(1+z)^2} & 0 < z < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{6-84}$$

◀

EXAMPLE 6-20

DISCRETE CASE

▶ Let x and y be independent Poisson random variables with parameters λ_1 and λ_2, respectively. Let $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Determine the p.m.f. of z.

Since x and y both take values $\{0, 1, 2, \ldots\}$, the same is true for z. For any $n = 0, 1, 2, \ldots$, $\{\mathbf{x}+\mathbf{y} = n\}$ gives only a finite number of options for x and y. In fact, if $\mathbf{x} = 0$, then y must be n; if $\mathbf{x} = 1$, then y must be $n - 1$, and so on. Thus the event $\{\mathbf{x} + \mathbf{y} = n\}$ is the union of mutually exclusive events $A_k = \{\mathbf{x} = k, \mathbf{y} = n - k\}$, $k = 0 \to n$.

$$\begin{aligned} P\{\mathbf{z} = n\} = P\{\mathbf{x}+\mathbf{y} = n\} &= P\left(\bigcup_{k=0}^{n}\{\mathbf{x} = k, \mathbf{y} = n-k\}\right) \\ &= \sum_{k=0}^{n} P\{\mathbf{x} = k, \mathbf{y} = n - k\} \end{aligned} \tag{6-85}$$

If x and y are also independent, then

$$P\{\mathbf{x} = k, \mathbf{y} = n-k\} = P\{\mathbf{x} = k\}P\{\mathbf{y} = n - k\}$$

and hence

$$\begin{aligned} P\{\mathbf{z} = n\} &= \sum_{k=0}^{n} P\{\mathbf{x} = k\}P\{\mathbf{y} = n-k\} \\ &= \sum_{k=0}^{n} e^{-\lambda_1}\frac{\lambda_1^k}{k!}\, e^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!} = \frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k} \\ &= e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}, \qquad n = 0, 1, 2, \ldots, \infty \end{aligned} \tag{6-86}$$

Thus z represents a Poisson random variable with parameter $\lambda_1 + \lambda_2$, indicating that sum of independent Poisson random variables is a Poisson random variable whose parameter is the sum of the parameters of the original random variables.

As Example 6-20 indicates, this procedure is too tedious in the discrete case. As we shall see in Sec. 6-5, the joint characteristic function or the moment generating function can be used to solve problems of this type in a much easier manner. ◀

6-3 TWO FUNCTIONS OF TWO RANDOM VARIABLES

In the spirit of the previous section, let us look at an immediate generalization. Suppose $\mathbf{x}$ and $\mathbf{y}$ are two random variables with joint p.d.f. $f_{xy}(x, y)$. Given two functions $g(x, y)$ and $h(x, y)$, define two new random variables

$$\mathbf{z} = g(\mathbf{x}, \mathbf{y}) \tag{6-87}$$

$$\mathbf{w} = h(\mathbf{x}, \mathbf{y}) \tag{6-88}$$

How does one determine their joint p.d.f. $f_{zw}(z, w)$? Obviously with $f_{zw}(z, w)$ in hand, the marginal p.d.f.s $f_z(z)$ and $f_w(w)$ can be easily determined.

The procedure for determining $f_{zw}(z, w)$ is the same as that in (6-36). In fact for given numbers z and w,

$$\begin{aligned} F_{zw}(z, w) &= P\{\mathbf{z}(\xi) \le z, \mathbf{w}(\xi) \le w\} = P\{g(\mathbf{x}, \mathbf{y}) \le z, h(\mathbf{x}, \mathbf{y}) \le w\} \\ &= P\{(\mathbf{x}, \mathbf{y}) \in D_{z,w}\} = \iint_{(x,y)\in D_{z,w}} f_{xy}(x, y)\, dx\, dy \end{aligned} \tag{6-89}$$

where $D_{z,w}$ is the region in the xy plane such that the inequalities $g(x, y) \le z$ and $h(x, y) \le w$ are *simultaneously* satisfied in Fig. 6-21.

We illustrate this technique in Example 6-21.

EXAMPLE 6-21

▶ Suppose $\mathbf{x}$ and $\mathbf{y}$ are independent uniformly distributed random variables in the interval $(0, \theta)$. Define $\mathbf{z} = \min(\mathbf{x}, \mathbf{y})$, $\mathbf{w} = \max(\mathbf{x}, \mathbf{y})$. Determine $f_{zw}(z. w)$.

SOLUTION

Obviously both z and w vary in the interval $(0, \theta)$. Thus

$$F_{zw}(z, w) = 0 \qquad \text{if} \quad z < 0 \quad \text{or} \quad w < 0 \tag{6-90}$$

$$F_{zw}(z, w) = P\{\mathbf{z} \le z, \mathbf{w} \le w\} = P\{\min(\mathbf{x}, \mathbf{y}) \le z, \max(\mathbf{x}, \mathbf{y}) \le w\} \tag{6-91}$$

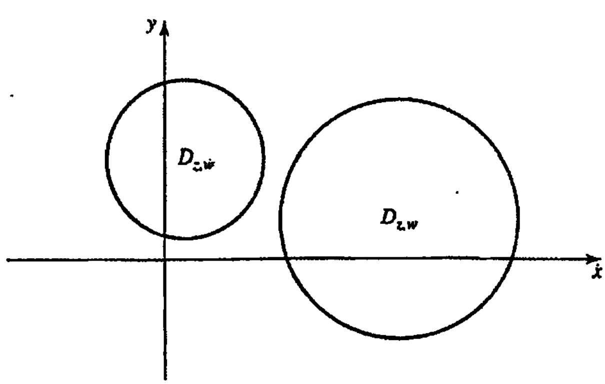


FIGURE 6-21

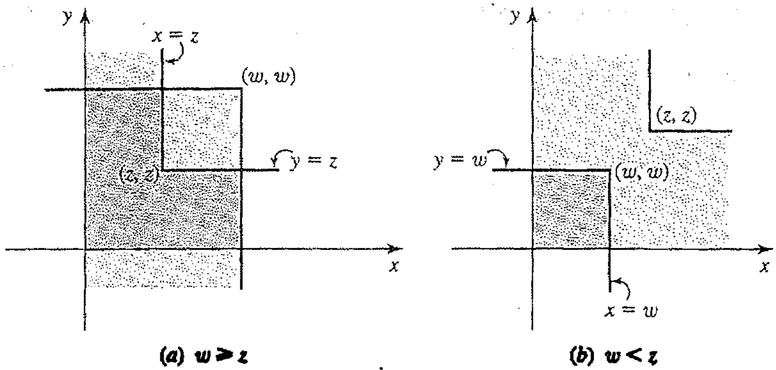


FIGURE 6-22

We must consider two cases: $w \geq z$ and $w < z$, since they give rise to different regions for $D_{z,w}$ (see Fig. 6-22*a* and 6-22*b*).

For $w \geq z$, from Fig. 6.22*a*, the region $D_{z,w}$ is represented by the doubly shaded area (see also Fig. 6-18*c* and Fig. 6-19*c*). Thus

$$F_{zw}(z, w) = F_{xy}(z, w) + F_{xy}(w, z) - F_{xy}(z, z) \qquad w \geq z \tag{6-92}$$

and for $w < z$, from Fig. 6.22*b*, we obtain

$$F_{zw}(z, w) = F_{xy}(w, w) \qquad w < z \tag{6-93}$$

with

$$F_{xy}(x, y) = F_x(x)F_y(y) = \frac{x}{\theta} \cdot \frac{y}{\theta} = \frac{xy}{\theta^2} \tag{6-94}$$

we obtain

$$F_{zw}(z, w) = \begin{cases} (2wz - z^2)/\theta^2 & 0 < z < w < \theta \\ w^2/\theta^2 & 0 < w < z < \theta \end{cases} \tag{6-95}$$

Thus

$$f_{zw}(z, w) = \begin{cases} 2/\theta^2 & 0 < z < w < \theta \\ 0 & \text{otherwise} \end{cases} \tag{6-96}$$

From (6-96), we also obtain

$$f_z(z) = \int_z^\theta f_{zw}(z, w)\, dw = \frac{2}{\theta}\left(1 - \frac{z}{\theta}\right) \qquad 0 < z < \theta \tag{6-97}$$

and

$$f_w(w) = \int_0^w f_{zw}(z, w)\, dz = \frac{2w}{\theta^2} \qquad 0 < w < \theta \tag{6-98}$$

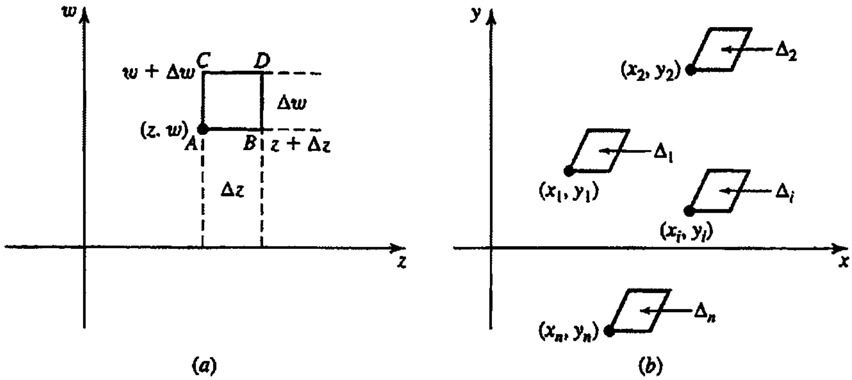


FIGURE 6-23

Joint Density

If $g(x, y)$ and $h(x, y)$ are continuous and differentiable functions, then, as in the case of one random variable [see (5-16)], it is possible to develop a formula to obtain the joint p.d.f. $f_{zw}(z, w)$ directly. Toward this, consider the equations

$$g(x, y) = z \qquad h(x, y) = w \tag{6-99}$$

For a given point (z, w), equation (6-99) can have many solutions. Let us say (x_1, y_1), $(x_2, y_2), \ldots, (x_n, y_n)$ represent these multiple solutions such that (see Fig. 6-23)

$$g(x_i, y_i) = z \qquad h(x_i, y_i) = w \tag{6-100}$$

Consider the problem of evaluating the probability

$$\begin{aligned} &P\{z < \mathbf{z} \le z + \Delta z, w < \mathbf{w} \le w + \Delta w\} \\ &\quad = P\{z < g(\mathbf{x}, \mathbf{y}) \le z + \Delta z, w < h(\mathbf{x}, \mathbf{y}) \le w + \Delta w\} \end{aligned} \tag{6-101}$$

Using (6-8) we can rewrite (6-101) as

$$P\{z < \mathbf{z} \le z + \Delta z, w < \mathbf{w} \le w + \Delta w\} = f_{zw}(z, w)\,\Delta z\,\Delta w \tag{6-102}$$

But to translate this probability in terms of $f_{xy}(x, y)$, we need to evaluate the equivalent region for $\Delta z\ \Delta w$ in the xy plane. Toward this, referring to Fig. 6-24, we observe that the point A with coordinates (z, w) gets mapped onto the point A' with coordinates (x_i, y_i) (as well as to other points as in Fig 6.23b). As z changes to $z + \Delta z$ to point B in Fig. 6.24a, let B' represent its image in the xy plane. Similarly, as w changes to $w + \Delta w$ to C, let C' represent its image in the xy plane.

Finally D goes to D', and $A'B'C'D'$ represents the equivalent parallelogram in the xy plane with area Δ_i. Referring to Fig. 6-23, because of the nonoverlapping nature of these regions the probability in (6-102) can be alternatively expressed as

$$\sum_i P\{(\mathbf{x}, \mathbf{y}) \in \Delta_i\} = \sum_i f_{xy}(x_i, y_i)\,\Delta_i \tag{6-103}$$

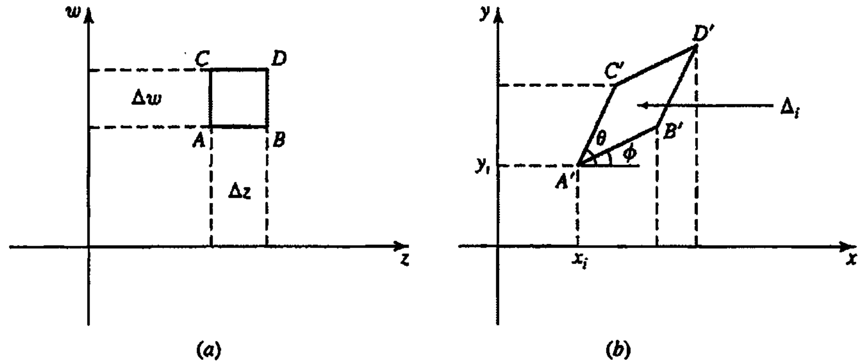


FIGURE 6-24

Equating (6-102) and (6-103) we obtain

$$f_{zw}(z, w) = \sum_i f_{xy}(x_i, y_i) \frac{\Delta_i}{\Delta z \, \Delta w} \tag{6-104}$$

To simplify (6-104), we need to evaluate the area Δ_i of the parallelograms in Fig. 6.24b in terms of $\Delta z \, \Delta w$. Toward this, let g_1 and h_1 denote the inverse transformation in (6-99), so that

$$x_i = g_1(z, w) \qquad y_i = h_1(z, w) \tag{6-105}$$

As the point (z, w) goes to $(x_i, y_i) \equiv A'$, the point $(z + \Delta z, w)$ goes to B', the point $(z, w + \Delta w)$ goes to C', and the point $(z + \Delta z, w + \Delta w)$ goes to D'. Hence the respective x and y coordinates of B' are given by

$$g_1(z + \Delta z, w) = g_1(z, w) + \frac{\partial g_1}{\partial z} \Delta z = x_i + \frac{\partial g_1}{\partial z} \Delta z \tag{6-106}$$

and

$$h_1(z + \Delta z, w) = h_1(z, w) + \frac{\partial h_1}{\partial z} \Delta z = y_i + \frac{\partial h_1}{\partial z} \Delta z \tag{6-107}$$

Similarly those of C' are given by

$$x_i + \frac{\partial g_1}{\partial w} \Delta w \qquad y_i + \frac{\partial h_1}{\partial w} \Delta w \tag{6-108}$$

The area of the parallelogram $A'B'C'D'$ in Fig. 6-24b is given by

$$\begin{aligned} \Delta_i &= (A'B')(A'C') \sin(\theta - \phi) \\ &= (A'B' \cos\phi)(A'C' \sin\theta) - (A'B' \sin\phi)(A'C' \cos\theta) \end{aligned} \tag{6-109}$$

But from Fig. 6-24b, and (6-106)–(6-108)

$$A'B' \cos\phi = \frac{\partial g_1}{\partial z} \Delta z \qquad A'C' \sin\theta = \frac{\partial h_1}{\partial w} \Delta w \tag{6-110}$$

$$A'B' \sin\phi = \frac{\partial h_1}{\partial z} \Delta z \qquad A'C' \cos\theta = \frac{\partial g_1}{\partial w} \Delta w \tag{6-111}$$

so that

$$\Delta_i = \left(\frac{\partial g_1}{\partial z}\frac{\partial h_1}{\partial w} - \frac{\partial g_1}{\partial w}\frac{\partial h_1}{\partial z}\right)\Delta z\,\Delta w \tag{6-112}$$

and

$$\frac{\Delta_i}{\Delta z\,\Delta w} = \left(\frac{\partial g_1}{\partial z}\frac{\partial h_1}{\partial w} - \frac{\partial g_1}{\partial w}\frac{\partial h_1}{\partial z}\right) = \begin{vmatrix} \dfrac{\partial g_1}{\partial z} & \dfrac{\partial g_1}{\partial w} \\ \dfrac{\partial h_1}{\partial z} & \dfrac{\partial h_1}{\partial w} \end{vmatrix} \tag{6-113}$$

The determinant on the right side of (6-113) represents the absolute value of the Jacobian $J(z, w)$ of the inverse transformation in (6-105). Thus

$$J(z, w) = \begin{vmatrix} \dfrac{\partial g_1}{\partial z} & \dfrac{\partial g_1}{\partial w} \\ \dfrac{\partial h_1}{\partial z} & \dfrac{\partial h_1}{\partial w} \end{vmatrix} \tag{6-114}$$

Substituting the absolute value of (6-114) into (6-104), we get

$$f_{zw}(z, w) = \sum_i |J(z, w)| f_{xy}(x_i, y_i) = \sum_i \frac{1}{|J(x_i, y_i)|} f_{xy}(x_i, y_i) \tag{6-115}$$

since

$$|J(z, w)| = \frac{1}{|J(x_i, y_i)|} \tag{6-116}$$

where the determinant $J(x_i, y_i)$ represents the Jacobian of the original transformation in (6-99) given by

$$J(x_i, y_i) = \begin{vmatrix} \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \\ \dfrac{\partial h}{\partial x} & \dfrac{\partial h}{\partial y} \end{vmatrix}_{x=x_i,\, y=y_i} \tag{6-117}$$

We shall illustrate the usefulness of the formulas in (6-115) through various examples.

Linear Transformation

$$z = ax + by \qquad w = cx + dy \tag{6-118}$$

If $ad - bc \neq 0$, then the system $ax + by = z$, $cx + dy = w$ has one and only one solution

$$x = Az + Bw \qquad y = Cz + Dw$$

Since $J(x, y) = ad - bc$, (6-115) yields

$$f_{zw}(z, w) = \frac{1}{|ad - bc|} f_{xy}(Az + Bw, Cz + Dw) \tag{6-119}$$

JOINT NORMALITY. From (6-119) it follows that if the random variables $\mathbf{x}$ and $\mathbf{y}$ are jointly normal as $N(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ and

$$\mathbf{z} = a\mathbf{x} + b\mathbf{y} \qquad \mathbf{w} = c\mathbf{x} + d\mathbf{y} \tag{6-120}$$

then $\mathbf{z}$ and $\mathbf{w}$ are also jointly normal since $f_{zw}(z, w)$ will be an exponential (similar to $f_{xy}(x, y)$) with a quadratic exponent in z and w. Using the notation in (6-25), $\mathbf{z}$ and $\mathbf{w}$ in (6-120) are jointly normal as $N(\mu_z, \mu_w, \sigma_z^2, \sigma_w^2, \rho_{zw})$, where by direct computation

$$\begin{aligned}
\mu_z &= a\mu_x + b\mu_y \\
\mu_w &= c\mu_x + d\mu_y \\
\sigma_z^2 &= a^2\sigma_x^2 + 2ab\rho\sigma_x\sigma_y + b^2\sigma_y^2 \\
\sigma_w^2 &= c^2\sigma_x^2 + 2cd\rho\sigma_x\sigma_y + d^2\sigma_y^2
\end{aligned} \tag{6-121}$$

and

$$\rho_{zw} = \frac{ac\sigma_x^2 + (ad + bc)\rho\sigma_x\sigma_y + bd\sigma_y^2}{\sigma_z\sigma_w}$$

In particular, *any* linear combination of two jointly normal random variables is normal.

EXAMPLE 6-22

▶ Suppose $\mathbf{x}$ and $\mathbf{y}$ are zero mean independent Gaussian random variables with common variance σ^2. Define $\mathbf{r} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$, $\boldsymbol{\theta} = \tan^{-1}(\mathbf{y}/\mathbf{x})$, where $|\theta| < \pi$. Obtain their joint density function.

SOLUTION
Here

$$f_{xy}(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{6-122}$$

Since

$$r = g(x, y) = \sqrt{x^2 + y^2} \qquad \theta = h(x, y) = \tan^{-1}(y/x) \tag{6-123}$$

and θ is known to vary in the interval $(-\pi, \pi)$, we have one solution pair given by

$$x_1 = r\cos\theta \qquad y_1 = r\sin\theta \tag{6-124}$$

We can use (6-124) to obtain $J(r, \theta)$. From (6-114)

$$J(r, \theta) = \begin{vmatrix} \dfrac{\partial x_1}{\partial r} & \dfrac{\partial x_1}{\partial \theta} \\ \dfrac{\partial y_1}{\partial r} & \dfrac{\partial y_1}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r \tag{6-125}$$

so that

$$|J(r, \theta)| = r \tag{6-126}$$

We can also compute $J(x, y)$ using (6-117). From (6-123),

$$J(x, y) = \begin{vmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{vmatrix} = \frac{1}{\sqrt{x^2 + y^2}} = \frac{1}{r} \tag{6-127}$$

Notice that $|J(r,\theta)| = 1/|J(x,y)|$, agreeing with (6-116). Substituting (6-122), (6-124) and (6-126) or (6-127) into (6-115), we get

$$f_{r\theta}(r,\theta) = r f_{xy}(x_1, y_1) = \frac{r}{2\pi\sigma^2} e^{-r^2/2\sigma^2} \qquad 0 < r < \infty \qquad |\theta| < \pi \tag{6-128}$$

Thus

$$f_r(r) = \int_{-\pi}^{\pi} f_{r\theta}(r,\theta)\, d\theta = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \qquad 0 < r < \infty \tag{6-129}$$

which represents a Rayleigh random variable with parameter σ^2, and

$$f_\theta(\theta) = \int_0^{\infty} f_{r\theta}(r,\theta)\, dr = \frac{1}{2\pi} \qquad |\theta| < \pi \tag{6-130}$$

which represents a uniform random variable in the interval $(-\pi, \pi)$. Moreover by direct computation

$$f_{r\theta}(r,\theta) = f_r(r) \cdot f_\theta(\theta) \tag{6-131}$$

implying that $\mathbf{r}$ and $\boldsymbol{\theta}$ are independent. We summarize these results in the following statement: If $\mathbf{x}$ and $\mathbf{y}$ are zero mean independent Gaussian random variables with common variance, then $\sqrt{\mathbf{x}^2 + \mathbf{y}^2}$ has a Rayleigh distribution, and $\tan^{-1}(\mathbf{y}/\mathbf{x})$ has a uniform distribution in $(-\pi, \pi)$ (see also Example 6-15). Moreover these two derived random variables are statistically independent. Alternatively, with $\mathbf{x}$ and $\mathbf{y}$ as independent zero mean random variables as in (6-122), $\mathbf{x} + j\mathbf{y}$ represents a complex Gaussian random variable. But

$$\mathbf{x} + j\mathbf{y} = \mathbf{r}e^{j\boldsymbol{\theta}} \tag{6-132}$$

with $\mathbf{r}$ and $\boldsymbol{\theta}$ as in (6-123), and hence we conclude that the magnitude and phase of a complex Gaussian random variable are independent with Rayleigh and uniform distributions respectively. The statistical independence of these derived random variables is an interesting observation. ◀

EXAMPLE 6-23

▶ Let $\mathbf{x}$ and $\mathbf{y}$ be independent exponential random variables with common parameter λ. Define $\mathbf{u} = \mathbf{x} + \mathbf{y}$, $\mathbf{v} = \mathbf{x} - \mathbf{y}$. Find the joint and marginal p.d.f. of $\mathbf{u}$ and $\mathbf{v}$.

SOLUTION

It is given that

$$f_{xy}(x,y) = \frac{1}{\lambda^2} e^{-(x+y)/\lambda} \qquad x > 0 \qquad y > 0 \tag{6-133}$$

Now since $u = x + y$, $v = x - y$, always $|v| < u$, and there is only one solution given by

$$x = \frac{u+v}{2} \qquad y = \frac{u-v}{2} \tag{6-134}$$

Moreover the Jacobian of the transformation is given by

$$J(x,y) = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$$

and hence

$$f_{uv}(u, v) = \frac{1}{2\lambda^2} e^{-u/\lambda} \qquad 0 < |v| < u < \infty \tag{6-135}$$

represents the joint p.d.f. of **u** and **v**. This gives

$$f_u(u) = \int_{-u}^{u} f_{uv}(u, v)\, dv = \frac{1}{2\lambda^2} \int_{-u}^{u} e^{-u/\lambda}\, dv = \frac{u}{\lambda^2} e^{-u/\lambda} \qquad 0 < u < \infty \tag{6-136}$$

and

$$f_v(v) = \int_{|v|}^{\infty} f_{uv}(u, v)\, du = \frac{1}{2\lambda^2} \int_{|v|}^{\infty} e^{-u/\lambda}\, du = \frac{1}{2\lambda} e^{-|v|/\lambda} \qquad -\infty < v < \infty \tag{6-137}$$

Notice that in this case $f_{uv}(u, v) \neq f_u(u) \cdot f_v(v)$, and the random variables **u** and **v** are not independent. ◀

As we show below, the general transformation formula in (6-115) making use of two functions can be made useful even when only one function is specified.

AUXILIARY VARIABLES. Suppose

$$\mathbf{z} = g(\mathbf{x}, \mathbf{y}) \tag{6-138}$$

where **x** and **y** are two random variables. To determine $f_z(z)$ by making use of the formulation in (6-115), we can define an auxiliary variable

$$\mathbf{w} = \mathbf{x} \quad \text{or} \quad \mathbf{w} = \mathbf{y} \tag{6-139}$$

and the p.d.f. of **z** can be obtained from $f_{zw}(z, w)$ by proper integration.

EXAMPLE 6-24

▶ Suppose $\mathbf{z} = \mathbf{x} + \mathbf{y}$ and let $\mathbf{w} = \mathbf{y}$ so that the transformation is one-to-one and the solution is given by $y_1 = w$, $x_1 = z - w$. The Jacobian of the transformation is given by

$$J(x, y) = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1$$

and hence

$$f_{zw}(x, y) = f_{xy}(x_1, y_1) = f_{xy}(z - w, w)$$

or

$$f_z(z) = \int f_{zw}(z, w)\, dw = \int_{-\infty}^{+\infty} f_{xy}(z - w, w)\, dw \tag{6-140}$$

which agrees with (6-41). Note that (6-140) reduces to the convolution of $f_x(z)$ and $f_y(z)$ if **x** and **y** are independent random variables. ◀

Next, we consider a less trivial example along these lines.

EXAMPLE 6-25

▶ Let $\mathbf{x} \sim U(0, 1)$ and $\mathbf{y} \sim U(0, 1)$ be independent random variables. Define

$$\mathbf{z} = (-2 \ln \mathbf{x})^{1/2} \cos(2\pi \mathbf{y}) \tag{6-141}$$

Find the density function of **z**.

SOLUTION

We can make use of the auxiliary variable $\mathbf{w} = \mathbf{y}$ in this case. This gives the only solution to be

$$x_1 = e^{-[z \sec(2\pi w)]^2/2} \tag{6-142}$$

$$y_1 = w \tag{6-143}$$

and using (6-114)

$$J(z, w) = \begin{vmatrix} \dfrac{\partial x_1}{\partial z} & \dfrac{\partial x_1}{\partial w} \\ \dfrac{\partial y_1}{\partial z} & \dfrac{\partial y_1}{\partial w} \end{vmatrix} = \begin{vmatrix} -z \sec^2(2\pi w) e^{-[z \sec(2\pi w)]^2/2} & \dfrac{\partial x_1}{\partial w} \\ 0 & 1 \end{vmatrix}$$

$$= -z \sec^2(2\pi w) e^{-[z \sec(2\pi w)]^2/2} \tag{6-144}$$

Substituting (6-142) and (6-144) into (6-115), we obtain

$$f_{zw}(z, w) = z \sec^2(2\pi w) e^{-[z \sec(2\pi w)]^2/2} \qquad -\infty < z < +\infty \quad 0 < w < 1 \tag{6-145}$$

and

$$f_z(z) = \int_0^1 f_{zw}(z, w)\, dw = e^{-z^2/2} \int_0^1 z \sec^2(2\pi w) e^{-[z \tan(2\pi w)]^2/2}\, dw \tag{6-146}$$

Let $u = z \tan(2\pi w)$ so that $du = 2\pi z \sec^2(2\pi w)\, dw$. Notice that as w varies from 0 to 1, u varies from $-\infty$ to $+\infty$. Using this in (6-146), we get

$$f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \underbrace{\int_{-\infty}^{\infty} e^{-u^2/2} \frac{du}{\sqrt{2\pi}}}_{1} = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad -\infty < z < \infty \tag{6-147}$$

which represents a zero mean Gaussian random variable with unit variance. Thus $\mathbf{z} \sim N(0, 1)$. Equation (6-141) can be used as a practical procedure to generate Gaussian random variables from two independent uniformly distributed random sequences. ◀

EXAMPLE 6-26

▶ Let $\mathbf{z} = \mathbf{xy}$. Then with $\mathbf{w} = \mathbf{x}$ the system $xy = z$, $x = w$ has a single solution: $x_1 = w$, $y_1 = z/w$. In this case, $J(x, y) = -w$ and (6-115) yields

$$f_{zw}(z, w) = \frac{1}{|w|} f_{xy}\left(w, \frac{z}{w}\right)$$

Hence the density of the random variable $\mathbf{z} = \mathbf{xy}$ is given by

$$f_z(z) = \int_{-\infty}^{\infty} \frac{1}{|w|} f_{xy}\left(w, \frac{z}{w}\right) dw \tag{6-148}$$

Special case: We now assume that the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent and each is uniform in the interval (0, 1). In this case, $z < w$ and

$$f_{xy}\left(w, \frac{z}{w}\right) = f_x(w) f_y\left(\frac{z}{w}\right) = 1$$

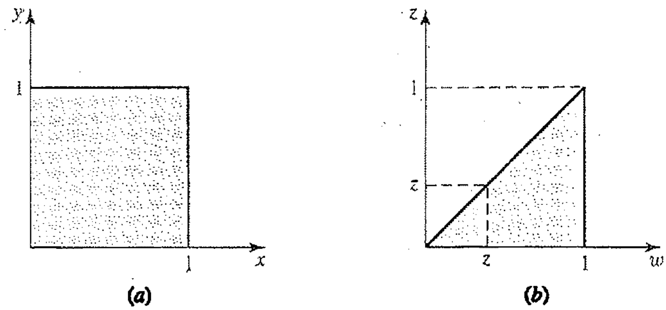


FIGURE 6-25

so that (see Fig. 6-25)

$$f_{zw}(z, w) = \begin{cases} 1/w & 0 < z < w < 1 \\ 0 & \text{otherwise} \end{cases} \tag{6-149}$$

Thus

$$f_z(z) = \int_z^1 \frac{1}{w}\, dw = \begin{cases} -\ln z & 0 < z < 1 \\ 0 & \text{elsewhere} \end{cases} \tag{6-150}$$

◀

EXAMPLE 6-27

▶ Let $\mathbf{x}$ and $\mathbf{y}$ be independent gamma random variables as in Example 6-12. Define $\mathbf{z} = \mathbf{x} + \mathbf{y}$ and $\mathbf{w} = \mathbf{x}/\mathbf{y}$. Show that $\mathbf{z}$ and $\mathbf{w}$ are independent random variables.

SOLUTION

Equations $\mathbf{z} = \mathbf{x} + \mathbf{y}$ and $\mathbf{w} = \mathbf{x}/\mathbf{y}$ generate one pair of solutions

$$x_1 = \frac{zw}{1+w} \qquad y_1 = \frac{z}{1+w}$$

Moreover

$$J(x, y) = \begin{vmatrix} 1 & 1 \\ 1/y & -x/y^2 \end{vmatrix} = -\frac{x+y}{y^2} = -\frac{(1+w)^2}{z}$$

Substituting these into (6-65) and (6-115) we get

$$\begin{aligned} f_{zw}(z, w) &= \frac{1}{\alpha^{m+n}\Gamma(m)\Gamma(n)} \frac{z}{(1+w)^2} \left(\frac{zw}{1+w}\right)^{m-1} \left(\frac{z}{1+w}\right)^{n-1} e^{-z/\alpha} \\ &= \frac{1}{\alpha^{m+n}} \frac{z^{m+n-1}}{\Gamma(m)\Gamma(n)} e^{-z/\alpha} \cdot \frac{w^{m-1}}{(1+w)^{m+n}} \\ &= \left(\frac{z^{m+n-1}}{\alpha^{m+n}\Gamma(m+n)} e^{-z/\alpha}\right) \cdot \left(\frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} \frac{w^{m-1}}{(1+w)^{m+n}}\right) \\ &= f_z(z) f_w(w) \qquad z > 0 \quad w > 0 \end{aligned} \tag{6-151}$$

showing that $\mathbf{z}$ and $\mathbf{w}$ are independent random variables. Notice that $\mathbf{z} \sim G(m+n, \alpha)$ and $\mathbf{w}$ represents the ratio of two independent gamma random variables. ◀

EXAMPLE 6-28

THE STUDENT t DISTRIBUTION

▶ A random variable $\mathbf{z}$ has a Student t distribution[2] $t(n)$ with n degrees of freedom if for $-\infty < z < \infty$

$$f_z(z) = \frac{\gamma_1}{\sqrt{(1+z^2/n)^{n+1}}} \qquad \gamma_1 = \frac{\Gamma((n+1)/2)}{\sqrt{\pi n}\,\Gamma(n/2)} \tag{6-152}$$

We shall show that if $\mathbf{x}$ and $\mathbf{y}$ are two independent random variables, $\mathbf{x}$ is $N(0, 1)$, and $\mathbf{y}$ is $\chi^2(n)$:

$$f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad f_y(y) = \frac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-y/2} U(y) \tag{6-153}$$

then the random variable

$$\mathbf{z} = \frac{\mathbf{x}}{\sqrt{\mathbf{y}/n}}$$

has a $t(n)$ distribution. Note that the Student t distribution represents the ratio of a normal random variable to the square root of an independent χ^2 random variable divided by its degrees of freedom.

SOLUTION

We introduce the random variable $\mathbf{w} = \mathbf{y}$ and use (6-115) with

$$x = z\sqrt{\frac{w}{n}} \qquad y = w \qquad J(z, w) = \sqrt{\frac{w}{n}} \quad \text{or} \quad J(x, y) = \sqrt{\frac{n}{w}}$$

This yields

$$\begin{aligned} f_{zw}(z, w) &= \sqrt{\frac{w}{n}} \frac{1}{\sqrt{2\pi}} e^{-z^2 w/2n} \frac{w^{n/2-1}}{2^{n/2}\Gamma(n/2)} e^{-w/2} U(w) \\ &= \frac{w^{(n-1)/2}}{\sqrt{2\pi n}\, 2^{n/2}\Gamma(n/2)} e^{-w(1+z^2/n)/2} U(w) \end{aligned}$$

Integrating with respect to w after replacing $w(1+z^2/n)/2 = u$, we obtain

$$\begin{aligned} f_z(z) &= \frac{1}{\sqrt{\pi n}\,\Gamma(n/2)} \frac{1}{(1+z^2/n)^{(n+1)/2}} \int_0^\infty u^{(n-1)/2} e^{-u}\, du \\ &= \frac{\Gamma((n+1)/2)}{\sqrt{\pi n}\,\Gamma(n/2)} \frac{1}{(1+z^2/n)^{(n+1)/2}} \\ &= \frac{1}{\sqrt{n}\beta(1/2, n/2)} \frac{1}{(1+z^2/n)^{(n+1)/2}} \qquad -\infty < z < \infty \end{aligned} \tag{6-154}$$

For $n = 1$, (6-154) represents a Cauchy random variable. Notice that for each n, (6-154) generates a different p.d.f. As n gets larger, the t distribution tends towards the normal distribution. In fact from (6-154)

$$(1+z^2/n)^{-(n+1)/2} \to e^{-z^2/2} \quad \text{as} \quad n \to \infty$$

For small n, the t distributions have "fatter tails" compared to the normal distribution because of its polynomial form. Like the normal distribution, Student t distribution is important in statistics and is often available in tabular form. ◀

---

[2] *Student* was the pseudonym of the English statistician W. S. Gosset, who first introduced this law in empirical form (*The probable error of a mean*, Biometrica, 1908.) The first rigorous proof of this result was published by R. A. Fisher.

EXAMPLE 6-29

THE F DISTRIBUTION

▶ Let x and y be independent random variables such that x has a chi-square distribution with m degrees of freedom and y has a chi-square distribution with n degrees of freedom. Then the random variable

$$\mathbf{F} = \frac{\mathbf{x}/m}{\mathbf{y}/n} \tag{6-155}$$

is said to have an *F distribution with* (m, n) degrees of freedom. Show that the p.d.f. of $\mathbf{z} = \mathbf{F}$ is given by

$$f_z(z) = \begin{cases} \dfrac{\Gamma((m+n)/2)m^{m/2}n^{n/2}}{\Gamma(m/2)\Gamma(n/2)} \dfrac{z^{m/2-1}}{(n+mz)^{(m+n)/2}} & z > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6-156}$$

SOLUTION

To compute the density of $\mathbf{F}$, using (6-153) we note that the density of $\mathbf{x}/m$ is given by

$$f_1(x) = \begin{cases} \dfrac{m(mx)^{m/2-1}e^{-mx/2}}{\Gamma(m/2)2^{m/2}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

and that of $\mathbf{y}/n$ by

$$f_2(y) = \begin{cases} \dfrac{n(ny)^{n/2-1}e^{-ny/2}}{\Gamma(n/2)2^{n/2}} & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

Using (6-60) from Example 6-10, the density of $\mathbf{z} = \mathbf{F}$ in (6-155) is given by

$$\begin{aligned} f_z(z) &= \int_0^\infty y \left(\frac{m(mzy)^{m/2-1}e^{-mzy/2}}{\Gamma(m/2)2^{m/2}} \right) \left(\frac{n(ny)^{n/2-1}e^{-ny/2}}{\Gamma(n/2)2^{n/2}} \right) dy \\ &= \frac{(m/2)^{m/2}(n/2)^{n/2}}{\Gamma(m/2)\Gamma(n/2)2^{(m+n)/2}} z^{m/2-1} \int_0^\infty y^{(m+n)/2-1} e^{y(n+mz)/2}\, dy \\ &= \frac{(m/2)^{m/2}(n/2)^{n/2}}{\Gamma(m/2)\Gamma(n/2)2^{(m+n)/2}} z^{m/2-1} \Gamma\left(\frac{m+n}{2}\right) \left(\frac{2}{n+mz}\right)^{(m+n)/2} \\ &= \frac{\Gamma((m+n)/2)m^{m/2}n^{n/2}}{\Gamma(m/2)\Gamma(n/2)} \frac{z^{m/2-1}}{(n+mz)^{(m+n)/2}} \\ &= \frac{(m/n)^{m/2}}{\beta(m/2, n/2)} z^{m/2-1}(1 + mz/n)^{-(m+n)/2} \qquad z > 0 \end{aligned} \tag{6-157}$$

and $f_z(z) = 0$ for $z \le 0$. The distribution in (6-157) is called Fisher's variance ration distribution. If $m = 1$ in (6-155), then from (6-154) and (6-157) we get $\mathbf{F} = [t(n)]^2$. Thus $F(1, n)$ and $t^2(n)$ have the same distribution. Moreover $F(1, 1) = t^2(1)$ represents the square of a Cauchy random variable. Both Student's t distribution and Fisher's F distribution play key roles in statistical tests of significance. ◀

6-4 JOINT MOMENTS

Given two random variables $\mathbf{x}$ and $\mathbf{y}$ and a function $g(x, y)$, we form the random variable $\mathbf{z} = g(\mathbf{x}, \mathbf{y})$. The expected value of this random variable is given by

$$E\{\mathbf{z}\} = \int_{-\infty}^{\infty} z f_z(z)\,dz \tag{6-158}$$

However, as the next theorem shows, $E\{\mathbf{z}\}$ can be expressed directly in terms of the function $g(x, y)$ and the joint density $f(x, y)$ of $\mathbf{x}$ and $\mathbf{y}$.

THEOREM 6-4

$$E\{g(\mathbf{x}, \mathbf{y})\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f(x, y)\,dx\,dy \tag{6-159}$$

Proof. The proof is similar to the proof of (5-55). We denote by ΔD_z the region of the xy plane such that $z < g(x, y) < z + dz$. Thus to each differential in (6-158) there corresponds a region ΔD_z in the xy plane. As dz covers the z axis, the regions ΔD_z are not overlapping and they cover the entire xy plane. Hence the integrals in (6-158) and (6-159) are equal.

We note that the expected value of $g(\mathbf{x})$ can be determined either from (6-159) or from (5-55) as a single integral

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x) f(x, y)\,dx\,dy = \int_{-\infty}^{\infty} g(x) f_x(x)\,dx$$

This is consistent with the relationship (6-10) between marginal and joint densities.

If the random variables $\mathbf{x}$ and $\mathbf{y}$ are of discrete type taking the values x_i and y_k with probability p_{ik} as in (6-33), then

$$E\{g(\mathbf{x}, \mathbf{y})\} = \sum_i \sum_k g(x_i, y_k) p_{ik} \tag{6-160}$$

Linearity From (6-159) it follows that

$$E\left\{\sum_{k=1}^{n} a_k g_k(\mathbf{x}, \mathbf{y})\right\} = \sum_{k=1}^{n} a_k E\{g_k(\mathbf{x}, \mathbf{y})\} \tag{6-161}$$

This fundamental result will be used extensively.

We note in particular that

$$E\{\mathbf{x} + \mathbf{y}\} = E\{\mathbf{x}\} + E\{\mathbf{y}\} \tag{6-162}$$

Thus the expected value of the sum of two random variables equals the sum of their expected values. We should stress, however, that, in general,

$$E\{\mathbf{x}\mathbf{y}\} \neq E\{\mathbf{x}\}E\{\mathbf{y}\}$$

Frequency interpretation As in (5-51)

$$\begin{aligned}
E\{\mathbf{x} + \mathbf{y}\} &\simeq \frac{\mathbf{x}(\xi_1) + \mathbf{y}(\xi_1) + \cdots + \mathbf{x}(\xi_n) + \mathbf{y}(\xi_n)}{n} \\
&= \frac{\mathbf{x}(\xi_1) + \cdots + \mathbf{x}(\xi_n)}{n} + \frac{\mathbf{y}(\xi_1) + \cdots + \mathbf{y}(\xi_n)}{n} \\
&\simeq E\{\mathbf{x}\} + E\{\mathbf{y}\}
\end{aligned}$$

However, in general,

$$E\{\mathbf{xy}\} \simeq \frac{\mathbf{x}(\xi_1)\mathbf{y}(\xi_i) + \cdots + \mathbf{x}(\xi_n)\mathbf{y}(\xi_n)}{n}$$
$$\neq \frac{\mathbf{x}(\xi_1) + \cdots + \mathbf{x}(\xi_n)}{n} \times \frac{\mathbf{y}(\xi_1) + \cdots + \mathbf{y}(\xi_n)}{n} \simeq E\{\mathbf{x}\}E\{\mathbf{y}\}$$

In the case of one random variable, we defined the parameters mean and variance to represent its average behavior. How does one parametrically represent similar cross behavior between two random variables? Toward this, we can generalize the variance definition as shown next.

COVARIANCE. The covariance C or C_{xy} of two random variables $\mathbf{x}$ and $\mathbf{y}$ is by definition the number

$$C_{xy} = E\{(\mathbf{x} - \eta_x)(\mathbf{y} - \eta_y)\} \tag{6-163}$$

where $E\{\mathbf{x}\} = \eta_x$ and $E\{\mathbf{y}\} = \eta_y$. Expanding the product in (6-163) and using (6-161) we obtain

$$C_{xy} = E\{\mathbf{xy}\} - E\{\mathbf{x}\}E\{\mathbf{y}\} \tag{6-164}$$

Correlation coefficient The correlation coefficient ρ or ρ_{xy} of the random variables $\mathbf{x}$ and $\mathbf{y}$ is by definition the ratio

$$\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \tag{6-165}$$

We maintain that

$$|\rho_{xy}| \leq 1 \qquad |C_{xy}| \leq \sigma_x \sigma_y \tag{6-166}$$

Proof. Clearly,

$$E\{[a(\mathbf{x} - \eta_x) + (\mathbf{y} - \eta_y)]^2\} = a^2\sigma_x^2 + 2aC_{xy} + \sigma_y^2 \tag{6-167}$$

Equation (6-167) is a positive quadratic for any a; hence its discriminant is negative. In other words,

$$C_{xy}^2 - \sigma_x^2\sigma_y^2 \leq 0 \tag{6-168}$$

and (6-166) results.

We note that the random variables $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{x} - \eta_x$, $\mathbf{y} - \eta_y$ have the same covariance and correlation coefficient.

EXAMPLE 6-30

▶ We shall show that the correlation coefficient of two jointly normal random variables is the parameter r in (6-23). It suffices to assume that $\eta_x = \eta_y = 0$ and to show that $E(\mathbf{xy}) = r\sigma_1\sigma_2$.

Since

$$\frac{x^2}{\sigma_1^2} - 2r\frac{xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2} = \left(\frac{x}{\sigma_1} - r\frac{y}{\sigma_2}\right)^2 + (1 - r^2)\frac{y^2}{\sigma_2^2}$$

we conclude with (6-23) that

$$E\{\mathbf{xy}\} = \frac{1}{\sigma_2\sqrt{2\pi}}\int_{-\infty}^{\infty} ye^{-y^2/2\sigma_2^2}\int_{-\infty}^{\infty}\frac{x}{\sigma_1\sqrt{2\pi(1-r^2)}}\exp\left(-\frac{(x-ry\sigma_1/\sigma_2)^2}{2\sigma_1^2(1-r^2)}\right)dx\,dy$$

The inner integral is a normal density with mean $ry\sigma_1/\sigma_2$ multiplied by x; hence it equals $ry\sigma_1/\sigma_2$. This yields

$$E\{\mathbf{xy}\} = r\sigma_1/\sigma_2\int_{-\infty}^{\infty}\frac{1}{\sigma_2\sqrt{2\pi}}y^2e^{-y^2/2\sigma_2^2}\,dy = r\sigma_1\sigma_2 \qquad \blacktriangleleft$$

Uncorrelatedness Two random variables are called uncorrelated if their covariance is 0. This can be phrased in the following equivalent forms

$$C_{xy} = 0 \qquad \rho_{xy} = 0 \qquad E\{\mathbf{xy}\} = E\{\mathbf{x}\}E\{\mathbf{y}\}$$

Orthogonality Two random variables are called orthogonal if

$$E\{\mathbf{xy}\} = 0$$

We shall use the notation

$$\mathbf{x} \perp \mathbf{y}$$

to indicate the random variables **x** and **y** are orthogonal.

Note (*a*) If **x** and **y** are uncorrelated, then $\mathbf{x} - \eta_x \perp \mathbf{y} - \eta_y$. (*b*) If **x** and **y** are uncorrelated and $\eta_x = 0$ or $\eta_y = 0$ then $\mathbf{x} \perp \mathbf{y}$.

Vector space of random variables. We shall find it convenient to interpret random variables as vectors in an abstract space. In this space, the second moment

$$E\{\mathbf{xy}\}$$

of the random variables **x** and **y** is by definition their *inner product* and $E\{\mathbf{x}^2\}$ and $E\{\mathbf{y}^2\}$ are the squares of their lengths. The ratio

$$\frac{E\{\mathbf{xy}\}}{\sqrt{E\{\mathbf{x}^2\}E\{\mathbf{y}^2\}}}$$

is the cosine of their angle.

We maintain that

$$E^2\{\mathbf{xy}\} \le E\{\mathbf{x}^2\}E\{\mathbf{y}^2\} \tag{6-169}$$

This is the *cosine inequality* and its proof is similar to the proof of (6-168): The quadratic

$$E\{(a\mathbf{x}-\mathbf{y})^2\} = a^2E\{\mathbf{x}^2\} - 2aE\{\mathbf{xy}\} + E\{\mathbf{y}^2\}$$

is positive for every a; hence its discriminant is negative and (6-169) results. If (6-169) is an equality, then the quadratic is 0 for some $a = a_0$, hence $\mathbf{y} = a_0\mathbf{x}$. This agrees with the geometric interpretation of random variables because, if (6-169) is an equality, then the vectors **x** and **y** are on the same line.

The following illustration is an example of the correspondence between vectors and random variables: Consider two random variables **x** and **y** such that $E\{\mathbf{x}^2\} = E\{\mathbf{y}^2\}$.

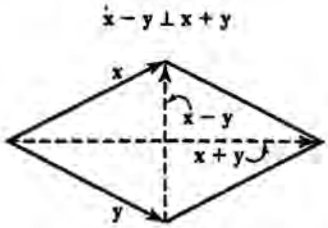


FIGURE 6-26

Geometrically, this means that the vectors $\mathbf{x}$ and $\mathbf{y}$ have the same length. If, therefore, we construct a parallelogram with sides $\mathbf{x}$ and $\mathbf{y}$, it will be a rhombus with diagonals $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ (Fig. 6-26). These diagonals are perpendicular because

$$E\{(\mathbf{x}+\mathbf{y})(\mathbf{x}-\mathbf{y})\} = E\{\mathbf{x}^2-\mathbf{y}^2\} = 0$$

THEOREM 6-5

▶ If two random variables are independent, that is, if [see also (6-20)]

$$f(x,y) = f_x(x)f_y(y) \tag{6-170}$$

then they are uncorrelated.

Proof. It suffices to show that

$$E\{\mathbf{xy}\} = E\{\mathbf{x}\}E\{\mathbf{y}\} \tag{6-171}$$

From (6-159) and (6-170) it follows that

$$E\{\mathbf{xy}\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_x(x)f_y(y)\,dx\,dy = \int_{-\infty}^{\infty} xf_x(x)\,dx \int_{-\infty}^{\infty} yf_y(y)\,dy$$

and (6-171) results.

If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent, then the random variables $g(\mathbf{x})$ and $h(\mathbf{y})$ are also independent [see (6-21)]. Hence

$$E\{g(\mathbf{x})h(\mathbf{y})\} = E\{g(\mathbf{x})\}E\{h(\mathbf{y})\} \tag{6-172}$$

This is not, in general, true if $\mathbf{x}$ and $\mathbf{y}$ are merely uncorrelated. ◀

As Example 6-31 shows if two random variables are uncorrelated they are not necessarily independent. However, for normal random variables uncorrelatedness is equivalent to independence. Indeed, if the random variables $\mathbf{x}$ and $\mathbf{y}$ are jointly normal and their correlation coefficient $r=0$, then [see (6-23)] $f_{xy}(x,y) = f_x(x)f_y(y)$.

EXAMPLE 6-31

▶ Let $\mathbf{x} \sim U(0,1)$, $\mathbf{y} \sim U(0,1)$. Suppose $\mathbf{x}$ and $\mathbf{y}$ are independent. Define $\mathbf{z}=\mathbf{x}+\mathbf{y}$, $\mathbf{w}=\mathbf{x}-\mathbf{y}$. Show that $\mathbf{z}$ and $\mathbf{w}$ are not independent, but uncorrelated random variables.

SOLUTION

$z=x+y$, $w=x-y$ gives the only solution set to be

$$x=\frac{z+w}{2} \qquad y=\frac{z-w}{2}$$

Moreover $0<z<2, -1<w<1, z+w\le 2, z-w\le 2, z>|w|$ and $|J(z,w)|=1/2$.

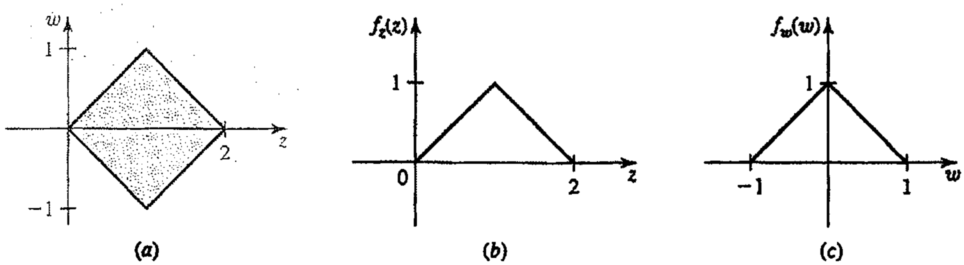


FIGURE 6-27

Thus (see the shaded region in Fig. 6-27)

$$f_{zw}(z,w) = \begin{cases} 1/2 & 0 < z < 2, -1 < w < 1, z + w \le 2, z - w \le 2, |w| < z \\ 0 & \text{otherwise} \end{cases} \tag{6-173}$$

and hence

$$f_z(z) = \int f_{zw}(z,w)\,dw = \begin{cases} \displaystyle\int_{-z}^{z} \frac{1}{2}\,dw = z & 0 < z < 1 \\ \displaystyle\int_{z-2}^{2-z} \frac{1}{2}\,dw = 2 - z & 1 < z < 2 \\ 0 & \text{otherwise} \end{cases} \tag{6-174}$$

and

$$f_w(w) = \int f_{zw}(z,w)\,dz = \int_{|w|}^{2-|w|} \frac{1}{2}\,dz = \begin{cases} 1 - |w| & -1 < w < 1 \\ 0 & \text{otherwise} \end{cases} \tag{6-175}$$

Clearly $f_{zw}(z,w) \neq f_z(z) f_w(w)$. Thus **z** and **w** are not independent. However,

$$E\{\mathbf{zw}\} = E\{(\mathbf{x}+\mathbf{y})(\mathbf{x}-\mathbf{y})\} = E\{\mathbf{x}^2\} - E\{\mathbf{y}^2\} = 0 \tag{6-176}$$

and

$$E\{\mathbf{w}\} = E\{\mathbf{x} - \mathbf{y}\} = 0 \tag{6-177}$$

and hence

$$\mathrm{Cov}\{\mathbf{z},\mathbf{w}\} = E\{\mathbf{zw}\} - E\{\mathbf{z}\}E\{\mathbf{w}\} = 0 \tag{6-178}$$

implying that **z** and **w** are uncorrelated random variables.

Variance of the sum of two random variables: If $\mathbf{z} = \mathbf{x} + \mathbf{y}$, then $\eta_z = \eta_x + \eta_y$; hence

$$\sigma_z^2 = E\{(\mathbf{z} - \eta_z)^2\} = E\{[(\mathbf{x} - \eta_x) + (\mathbf{y} - \eta_y)]^2\}$$

From this and (6-167) it follows that

$$\sigma_z^2 = \sigma_x^2 + 2\rho_{xy}\sigma_x\sigma_y + \sigma_y^2 \tag{6-179}$$

This leads to the conclusion that if $\rho_{xy} = 0$ then

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 \tag{6-180}$$

Thus, if two random variables are uncorrelated, then the variance of their sum equals the sum of their variances.

It follows from (6-171) that this is also true if $\mathbf{x}$ and $\mathbf{y}$ are independent. ◀

Moments

The mean

$$m_{kr} = E\{\mathbf{x}^k\mathbf{y}^r\} = \int_{-\infty}^{\infty} x^k y^r f_{xy}(x, y)\, dx\, dy \tag{6-181}$$

of the product $\mathbf{x}^k\mathbf{y}^r$ is by definition a joint moment of the random variables $\mathbf{x}$ and $\mathbf{y}$ of order $k + r = n$.

Thus $m_{10} = \eta_x$, $m_{01} = \eta_y$ are the first-order moments and

$$m_{20} = E\{\mathbf{x}^2\} \qquad m_{11} = E\{\mathbf{xy}\} \qquad m_{02} = E\{\mathbf{y}^2\}$$

are the second-order moments.

The joint central moments of $\mathbf{x}$ and $\mathbf{y}$ are the moments of $\mathbf{x} - \eta_x$ and $\mathbf{y} - \eta_y$:

$$\mu_{kr} = E\{(\mathbf{x} - \eta_x)^k(\mathbf{y} - \eta_y)^r\} = \int_{-\infty}^{\infty} (x - \eta_x)^k(y - \eta_y)^r f_{xy}(x, y)\, dx\, dy \tag{6-182}$$

Clearly, $\mu_{10} = \mu_{01} = 0$ and

$$\mu_{11} = C_{xy} \qquad \mu_{20} = \sigma_x^2 \qquad \mu_{02} = \sigma_y^2$$

Absolute and generalized moments are defined similarly [see (5-69) and (5-70)].

For the determination of the joint statistics of $\mathbf{x}$ and $\mathbf{y}$ knowledge of their joint density is required. However, in many applications, only the first- and second-moments are used. These moments are determined in terms of the five parameters

$$\eta_x \quad \eta_y \quad \sigma_x^2 \quad \sigma_y^2 \quad \rho_{xy}$$

If $\mathbf{x}$ and $\mathbf{y}$ are jointly normal, then [see (6-23)] these parameters determine uniquely $f_{xy}(x, y)$.

EXAMPLE 6-32

▶ The random variables $\mathbf{x}$ and $\mathbf{y}$ are jointly normal with

$$\eta_x = 10 \qquad \eta_y = 0 \qquad \sigma_x^2 = 4 \qquad \sigma_y^2 = 1 \qquad \rho_{xy} = 0.5$$

We shall find the joint density of the random variables

$$\mathbf{z} = \mathbf{x} + \mathbf{y} \qquad \mathbf{w} = \mathbf{x} - \mathbf{y}$$

Clearly,

$$\eta_z = \eta_x + \eta_y = 10 \quad \eta_w = \eta_x - \eta_y = 10$$

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2r_{xy}\sigma_x\sigma_y = 7 \qquad \sigma_w^2 = \sigma_x^2 + \sigma_y^2 - 2r_{xy}\sigma_x\sigma_y = 3$$

$$E(\mathbf{zw}) = E(\mathbf{x}^2 - \mathbf{y}^2) = (100 + 4) - 1 = 103$$

$$\rho_{zw} = \frac{E(\mathbf{zw}) - E(\mathbf{z})E(\mathbf{w})}{\sigma_z\sigma_w} = \frac{3}{\sqrt{7 \times 3}}$$

As we know [see (6-119)], the random variables **z** and **w** are jointly normal because they are linearly dependent on **x** and **y**. Hence their joint density is

$$N(10, 10, 7, 3, \sqrt{3/7})$$ ◀

ESTIMATE OF THE MEAN OF g(x, y). If the function $g(x, y)$ is sufficiently smooth near the point (η_x, η_y), then the mean η_g and variance σ_g^2 of $g(\mathbf{x}, \mathbf{y})$ can be estimated in terms of the mean, variance, and covariance of **x** and **y**:

$$\eta_g \simeq g + \frac{1}{2}\left(\frac{\partial^2 g}{\partial x^2}\sigma_x^2 + 2\frac{\partial^2 g}{\partial x \partial y}\rho_{xy}\sigma_x\sigma_y + \frac{\partial^2 g}{\partial y^2}\sigma_y^2\right) \tag{6-183}$$

$$\sigma_g^2 \simeq \left(\frac{\partial g}{\partial x}\right)^2 \sigma_x^2 + 2\left(\frac{\partial g}{\partial x}\right)\left(\frac{\partial g}{\partial y}\right)\rho_{xy}\sigma_x\sigma_y + \left(\frac{\partial g}{\partial y}\right)^2 \sigma_y^2 \tag{6-184}$$

where the function $g(x, y)$ and its derivatives are evaluated at $x = \eta_x$ and $y = \eta_y$.

Proof. We expand $g(x, y)$ into a series about the point (η_x, η_y):

$$g(x, y) = g(\eta_x, \eta_y) + (x - \eta_x)\frac{\partial g}{\partial x} + (y - \eta_y)\frac{\partial g}{\partial x} + \cdots \tag{6-185}$$

Inserting (6-185) into (6-159), we obtain the moment expansion of $E\{g(\mathbf{x}, \mathbf{y})\}$ in terms of the derivatives of $g(x, y)$ at (η_x, η_y) and the joint moments μ_{kr} of **x** and **y**. Using only the first five terms in (6-185), we obtain (6-183). Equation (6-184) follows if we apply (6-183) to the function $[g(x, y) - \eta_g]^2$ and neglect moments of order higher than 2.

6-5 JOINT CHARACTERISTIC FUNCTIONS

The *joint characteristic function* of the random variables **x** and **y** is by definition the integral

$$\Phi(\omega_1, \omega_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)e^{j(\omega_1 x + \omega_2 y)}\,dx\,dy \tag{6-186}$$

From this and the two-dimensional *inversion formula* for Fourier transforms, it follows that

$$f(x, y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi(\omega_1, \omega_2)e^{-j(\omega_1 x + \omega_2 y)}\,d\omega_1\,d\omega_2 \tag{6-187}$$

Clearly,

$$\Phi(\omega_1, \omega_2) = E\left\{e^{j(\omega_1 \mathbf{x} + \omega_2 \mathbf{y})}\right\} \tag{6-188}$$

The logarithm

$$\Psi(\omega_1, \omega_2) = \ln\,\Phi(\omega_1, \omega_2) \tag{6-189}$$

of $\Phi(\omega_1, \omega_2)$ is the joint logarithmic-characteristic function of **x** and **y**.

The *marginal* characteristic functions

$$\Phi_x(\omega) = E\{e^{j\omega\mathbf{x}}\} \qquad \Phi_y(\omega) = E\{e^{j\omega\mathbf{y}}\} \tag{6-190}$$

of $\mathbf{x}$ and $\mathbf{y}$ can be expressed in terms of their joint characteristic function $\Phi(\omega_1, \omega_2)$. From (6-188) and (6-190) it follows that

$$\Phi_x(\omega) = \Phi(\omega, 0) \qquad \Phi_y(\omega) = \Phi(0, \omega) \tag{6-191}$$

We note that, if $\mathbf{z} = a\mathbf{x} + b\mathbf{y}$ then

$$\Phi_z(\omega) = E\left\{e^{j(a\mathbf{x}+b\mathbf{y})\omega}\right\} = \Phi(a\omega, b\omega) \tag{6-192}$$

Hence $\Phi_z(1) = \Phi(a, b)$.

Cramér-Wold theorem The material just presented shows that if $\Phi_z(\omega)$ is known for every a and b, then $\Phi(\omega_1, \omega_2)$ is uniquely determined. In other words, if the density of $a\mathbf{x} + b\mathbf{y}$ is known for every a and b, then the joint density $f(x, y)$ of $\mathbf{x}$ and $\mathbf{y}$ is uniquely determined.

Independence and convolution

If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent, then [see (6-172)]

$$E\left\{e^{j(\omega_1\mathbf{x}+\omega_2\mathbf{y})}\right\} = E\{e^{j\omega_1\mathbf{x}}\}E\{e^{j\omega_2\mathbf{y}}\}$$

From this it follows that

$$\Phi(\omega_1, \omega_2) = \Phi_x(\omega_1)\Phi_y(\omega_2) \tag{6-193}$$

Conversely, if (6-193) is true, then the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent. Indeed, inserting (6-193) into the inversion formula (6-187) and using (5-102), we conclude that $f_xy(x, y) = f_x(x)f_y(y)$.

Convolution theorem If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent and $\mathbf{z} = \mathbf{x} + \mathbf{y}$, then

$$E\{e^{j\omega\mathbf{z}}\} = E\left\{e^{j\omega(\mathbf{x}+\mathbf{y})}\right\} = E\{e^{j\omega\mathbf{x}}\}E\{e^{j\omega\mathbf{y}}\}$$

Hence

$$\Phi_z(\omega) = \Phi_x(\omega)\Phi_y(\omega) \qquad \Psi_z(\omega) = \Psi_x(\omega) + \Psi_y(\omega) \tag{6-194}$$

As we know [see (6-43)], the density of $\mathbf{z}$ equals the convolution of $f_x(x)$ and $f_y(y)$. From this and (6-194) it follows that the characteristic function of the convolution of two densities equals the product of their characteristic functions.

EXAMPLE 6-33

► We shall show that if the random variables $\mathbf{x}$ and $\mathbf{y}$ are *independent* and Poisson distributed with parameters a and b, respectively, then their sum $\mathbf{z} = \mathbf{x} + \mathbf{y}$ is also Poisson distributed with parameter $a + b$.

Proof. As we know (see Example 5-31),

$$\Psi_x(\omega) = a(e^{j\omega} - 1) \qquad \Psi_y(\omega) = b(e^{j\omega} - 1)$$

Hence

$$\Psi_z(\omega) = \Psi_x(\omega) + \Psi_y(\omega) = (a + b)(e^{j\omega} - 1)$$

It can be shown that the converse is also true: If the random variables x and y are *independent* and their sum is Poisson distributed, then x and y are also Poisson distributed. The proof of this theorem is due to Raikov.[3] ◀

EXAMPLE 6-34

NORMAL RANDOM VARIABLES

▶ It was shown in Sec. 6-3 that if the random variables x and y are jointly normal, then the sum $a\mathbf{x} + b\mathbf{y}$ is also normal. Next we reestablish a special case of this result using (6-193): If x and y are *independent* and normal, then their sum $\mathbf{z} = \mathbf{x} + \mathbf{y}$ is also normal.

SOLUTION.
In this case [see (5-100)]

$$\Psi_x(\omega) = j\eta_x\omega - \frac{1}{2}\sigma_x^2\omega^2 \qquad \Psi_y(\omega) = j\eta_y\omega - \frac{1}{2}\sigma_y^2\omega^2$$

Hence

$$\Psi_z(\omega) = j(\eta_x + \eta_y)\omega - \frac{1}{2}(\sigma_x^2 + \sigma_y^2)\omega^2$$

It can be shown that the converse is also true (Cramér's theorem): If the random variables x and y are *independent* and their sum is *normal*, then they are also *normal*. The proof of this difficult theorem will not be given.[4]

In a similar manner, it is easy to show that if x and y are independent identically distributed normal random variables, then $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ are independent (and normal). Interestingly, in this case also, the converse is true (Bernstein's theorem): If x and y are independent and identically distributed and if $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ are also independent, then all random variables $(\mathbf{x}, \mathbf{y}, \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y})$ are normally distributed.

Darmois (1951) and Skitovitch (1954) have generalized this result as: If $\mathbf{x}_1$ and $\mathbf{x}_2$ are independent random variables and if two linear combinations $a_1\mathbf{x}_1 + a_2\mathbf{x}_2$ and $b_1\mathbf{x}_1 + b_2\mathbf{x}_2$ are also independent, where a_1, a_2, b_1, and b_2 represent nonzero coefficients, then *all* random variables are normally distributed. Thus if two nontrivial linear combinations of two independent random variables are also independent, then all of them represent normal random variables. ◀

More on Normal Random Variables

Let x and y be jointly Gaussian as $N(\eta_1, \eta_2, \sigma_1^2, \sigma_2^2, r)$ with p.d.f. as in (6-23) and (6-24). We shall show that the joint characteristic function of two jointly normal random variables is given by

$$\Phi(\omega_1, \omega_2) = e^{j(\eta_1\omega_1+\eta_2\omega_2)}e^{-(\omega_1^2\sigma_1^2+2r\sigma_1\sigma_2\omega_1\omega_2+\omega_2^2\sigma_2^2)/2} \tag{6-195}$$

Proof. This can be derived by inserting $f(x, y)$ into (6-186). The simpler proof presented here is based on the fact that the random variable $\mathbf{z} = \omega_1\mathbf{x} + \omega_2\mathbf{y}$ is normal and

$$\Phi_z(\omega) = e^{j\eta_z\omega-\sigma_z^2\omega^2/2} \tag{6-196}$$

[3] D. A. Raikov, "On the decomposition of Gauss and Poisson laws," *Izv. Akad. Nauk. SSSR*, Ser. Mat. 2, 1938, pp. 91–124.

[4] E. Lukacs, *Characteristic Functions*, Hafner Publishing Co., New York, 1960.

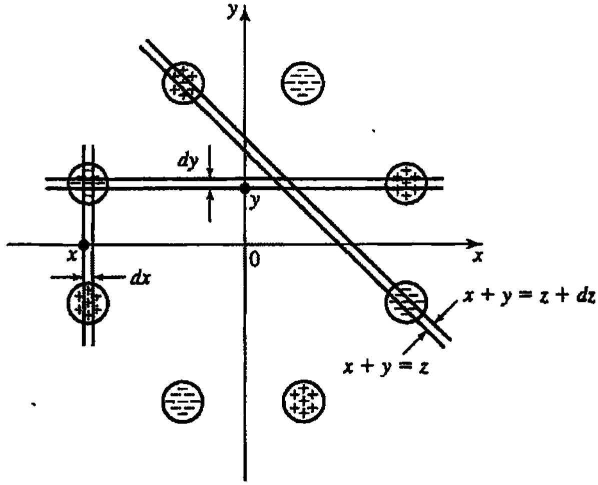


FIGURE 6-28

Since

$$\eta_z = \omega_1\eta_1 + \omega_2\eta_2 \qquad \sigma_z^2 = \omega_1^2\sigma_1^2 + 2r\omega_1\omega_2\sigma_1\sigma_2 + \omega_2^2\sigma_2^2$$

and $\Phi_z(\omega) = \Phi(\omega_1\omega, \omega_2\omega)$, (6-195) follows from (6-196) with $\omega = 1$.

This proof is based on the fact that the random variable $\mathbf{z} = \omega_1\mathbf{x} + \omega_2\mathbf{y}$ is normal for any ω_1 and ω_2; this leads to the conclusion: If it is known that the sum $a\mathbf{x} + b\mathbf{y}$ is normal for every a and b, then random variables $\mathbf{x}$ and $\mathbf{y}$ are jointly normal. We should stress, however, that this is not true if $a\mathbf{x} + b\mathbf{y}$ is normal for only a finite set of values of a and b. A counterexample can be formed by a simple extension of the construction in Fig. 6-28.

EXAMPLE 6-35

▶ We shall construct two random variables $\mathbf{x}_1$ and $\mathbf{x}_2$ with these properties: $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_1 + \mathbf{x}_2$ are normal but $\mathbf{x}_1$ and $\mathbf{x}_2$ are not jointly normal.

SOLUTION

Suppose that $\mathbf{x}$ and $\mathbf{y}$ are two jointly normal random variables with mass density $f(x, y)$. Adding and subtracting small masses in the region D of Fig. 6-28 consisting of eight circles as shown, we obtain a new function $f_1(x, y) = f(x, y) \pm \epsilon$ in D and $f_1(x, y) = f(x, y)$ everywhere else. The function $f_1(x, y)$ is a density; hence it defines two new random variables $\mathbf{x}_1$ and $\mathbf{y}_1$. These random variables are obviously not jointly normal. However, they are marginally normal because $\mathbf{x}$ and $\mathbf{y}$ are marginally normal and the masses in the vertical or horizontal strip have not changed. Furthermore, the random variable $\mathbf{z}_1 = \mathbf{x}_1 + \mathbf{y}_1$ is also normal because $\mathbf{z} = \mathbf{x} + \mathbf{y}$ is normal and the masses in any diagonal strip of the form $z \le x + y \le z + dz$ have not changed. ◀

THEOREM 6-6

MOMENT THEOREM

▶ The moment generating function of $\mathbf{x}$ and $\mathbf{y}$ is given by

$$\Phi(s_1, s_2) = E\{e^{s_1\mathbf{x}+s_2\mathbf{y}}\}.$$

Expanding the exponential and using the linearity of expected values, we obtain the series

$$\Phi(s_1, s_2) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{n=0}^{n} \binom{n}{k} E\{\mathbf{x}^k \mathbf{y}^{n-k}\} s_1^k s_2^{n-k}$$

$$= 1 + m_{10}s_1 + m_{01}s_2 + \frac{1}{2}\left(m_{20}s_1^2 + 2m_{11}s_2 + m_{02}s_2^2\right) + \cdots \quad (6\text{-}197)$$

From this it follows that

$$\frac{\partial^k \partial^r}{\partial s_1^k \partial s_2^r} \Phi(0, 0) = m_{kr} \quad (6\text{-}198)$$

The derivatives of the function $\Psi(s_1, s_2) = \ln \Phi(s_1, s_2)$ are by definition the joint cumulants λ_{kr} of x and y. It can be shown that

$$\lambda_{10} = m_{10} \qquad \lambda_{01} = m_{01} \qquad \lambda_{20} = \mu_{20} \qquad \lambda_{02} = \mu_{02} \qquad \lambda_{11} = \mu_{11}$$

Hence

$$\Psi(s_1, s_2) = \eta_1 s_1 + \eta_2 s_2 + \frac{1}{2}\left(\sigma_1^2 s_1^2 + 2r\sigma_1\sigma_2 s_1 s_2 + \sigma_2^2 s_2^2\right) + \cdots$$

◀

EXAMPLE 6-36

▶ Using (6-197), we shall show that if the random variables x and y are jointly normal with zero mean, then

$$E\{\mathbf{x}^2\mathbf{y}^2\} = E\{\mathbf{x}^2\}E\{\mathbf{y}^2\} + 2E^2\{\mathbf{xy}\} \quad (6\text{-}199)$$

SOLUTION

As we see from (6-195)

$$\Phi(s_1, s_2) = e^{-A} \qquad A = \frac{1}{2}\left(\sigma_1^2 s_1^2 + 2C s_1 s_2 + \sigma_2^2 s_2^2\right)$$

where $C = E\{\mathbf{xy}\} = r\sigma_1\sigma_2$. To prove (6-199), we shall equate the coefficient

$$\frac{1}{4!}\binom{4}{2} E\{\mathbf{x}^2\mathbf{y}^2\}$$

of $s_1^2 s_2^2$ in (6-197) with the corresponding coefficient of the expansion of e^{-A}. In this expansion, the factor $s_1^2 s_2^2$ appear only in the terms

$$\frac{A^2}{2} = \frac{1}{8}\left(\sigma_1^2 s_1^2 + 2C s_1 s_2 + \sigma_2^2 s_2^2\right)^2$$

Hence

$$\frac{1}{4!}\binom{4}{2} E\{\mathbf{x}^2\mathbf{y}^2\} = \frac{1}{8}\left(2\sigma_1^2\sigma_2^2 + 4C^2\right)$$

and (6-199) results. ◀

THEOREM 6-7

PRICE'S THEOREM[5]

▶ Given two jointly normal random variables x and y, we form the mean

$$I = E\{g(\mathbf{x}, \mathbf{y})\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx\, dy \quad (6\text{-}200)$$

[5] R. Price, "A Useful Theorem for Nonlinear Devices Having Gaussian Inputs," *IRE, PGIT*, Vol. IT-4, 1958. See also A. Papoulis, "On an Extension of Price's Theorem," *IEEE Transactions on Information Theory*, Vol. IT-11, 1965.

of some function $g(\mathbf{x}, \mathbf{y})$ of $(\mathbf{x}, \mathbf{y})$. The above integral is a function $I(\mu)$ of the covariance μ of the random variables $\mathbf{x}$ and $\mathbf{y}$ and of four parameters specifying the joint density $f(x, y)$ of $\mathbf{x}$ and $\mathbf{y}$. We shall show that if $g(x, y)f(x, y) \to 0$ as $(x, y) \to \infty$, then

$$\frac{\partial^n I(\mu)}{\partial \mu^n} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial^{2n} g(x, y)}{\partial x^n \partial y^n} f(x, y)\, dx\, dy = E\left(\frac{\partial^{2n} g(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}^n \partial \mathbf{y}^n}\right) \qquad (6\text{-}201)$$

Proof. Inserting (6-187) into (6-200) and differentiating with respect to μ, we obtain

$$\frac{\partial^n I(\mu)}{\partial \mu^n} = \frac{(-1)^n}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)$$
$$\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \omega_1^n \omega_2^n \Phi(\omega_1, \omega_2) e^{-j(\omega_1 x + \omega_2 y)}\, d\omega_1\, d\omega_2\, dx\, dy$$

From this and the derivative theorem, it follows that

$$\frac{\partial^n I(\mu)}{\partial \mu^n} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) \frac{\partial^{2n} f(x, y)}{\partial x^n \partial y^n}\, dx\, dy$$

After repeated integration by parts and using the condition at ∞, we obtain (6-201) (see also Prob. 5-48). ◁

EXAMPLE 6-37

▶ Using Price's theorem, we shall rederive (6-199). Setting $g(\mathbf{x}, \mathbf{y}) = \mathbf{x}^2\mathbf{y}^2$ into (6-201), we conclude with $n = 1$ that

$$\frac{\partial I(\mu)}{\partial \mu} = E\left(\frac{\partial^2 g(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}\, \partial \mathbf{y}}\right) = 4E\{\mathbf{xy}\} = 4\mu \qquad I(\mu) = \frac{4\mu^2}{2} + I(0)$$

If $\mu = 0$, the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent; hence $I(0) = E(\mathbf{x}^2\mathbf{y}^2) = E(\mathbf{x}^2)E(\mathbf{y}^2)$ and (6-199) results. ◀

6-6 CONDITIONAL DISTRIBUTIONS

As we have noted, the conditional distributions can be expressed as conditional probabilities:

$$F_z(z \mid M) = P\{\mathbf{z} \le z \mid M\} = \frac{P\{\mathbf{z} \le z, M\}}{P(M)}$$

$$F_{zw}(z, w \mid M) = P\{\mathbf{z} \le z, \mathbf{w} \le w \mid M\} = \frac{P\{\mathbf{z} \le z, \mathbf{w} \le w, M\}}{P(M)} \qquad (6\text{-}202)$$

The corresponding densities are obtained by appropriate differentiations. In this section, we evaluate these functions for various special cases.

EXAMPLE 6-38

▶ We shall first determine the conditional distribution $F_y(y \mid \mathbf{x} \le x)$ and density $f_y(y \mid \mathbf{x} \le x)$.

With $M = \{\mathbf{x} \le x\}$, (6-202) yields

$$F_y(y \mid \mathbf{x} \le x) = \frac{P\{\mathbf{x} \le x, \mathbf{y} \le y\}}{P\{\mathbf{x} \le x\}} = \frac{F(x, y)}{F_x(x)}$$

$$f_y(y \mid \mathbf{x} \le x) = \frac{\partial F(x, y)/\partial y}{F_x(x)}$$

EXAMPLE 6-39 We shall next determine the conditional distribution $F(x, y \mid M)$ for $M = \{x_1 < \mathbf{x} \le x_2\}$. In this case, $F(x, y \mid M)$ is given by

$$F(x, y \mid x_1 < \mathbf{x} \le x_2) = \frac{P\{\mathbf{x} \le x, \mathbf{y} \le y, x_1 < \mathbf{x} \le x_2\}}{P\{x_1 < \mathbf{x} \le x_2\}}$$

$$= \begin{cases} \dfrac{F(x_2, y) - F(x_1, y)}{F_x(x_2) - F_x(x_1)} & x > x_2 \\[2ex] \dfrac{F(x, y) - F(x_1, y)}{F_x(x_2) - F_x(x_1)} & x_1 < x \le x_2 \end{cases}$$

and it equals 0 for $x \le x_1$. Since $f = \partial^2 F/\partial x \partial y$, this yields

$$f(x, y \mid x_1 < \mathbf{x} \le x_2) = \frac{f(x, y)}{F_x(x_2) - F_x(x_1)} \qquad x_1 < x \le x_2 \tag{6-203}$$

and 0 otherwise.

The determination of the conditional density of $\mathbf{y}$ assuming $\mathbf{x} = x$ is of particular interest. This density cannot be derived directly from (6-202) because, in general, the event $\{\mathbf{x} = x\}$ has zero probability. It can, however, be defined as a limit. Suppose first that

$$M = \{x_1 < x \le x_2\}$$

In this case, (6-202) yields

$$F_y(y \mid x_1 < \mathbf{x} \le x_2) = \frac{P\{x_1 < \mathbf{x} \le x_2, \mathbf{y} \le y\}}{P\{x_1 < x \le x_2\}} = \frac{F(x_2, y) - F(x_1, y)}{F_x(x_2) - F_x(x_1)}$$

Differentiating with respect to y, we obtain

$$f_y(y \mid x_1 < \mathbf{x} \le x_2) = \frac{\int_{x_1}^{x_2} f(x, y)\, dx}{F_x(x_2) - F_x(x_1)} \tag{6-204}$$

because [see (6-6)]

$$\frac{\partial F(x, y)}{\partial y} = \int_{-\infty}^{x} f(\alpha, y)\, d\alpha$$

To determine $f_y(y \mid \mathbf{x} = x)$, we set $x_1 = x$ and $x_2 = x + \Delta x$ in (6-204). This yields

$$f_y(y \mid x < \mathbf{x} \le x + \Delta x) = \frac{\int_{x}^{x+\Delta x} f(\alpha, y)\, d\alpha}{F_x(x + \Delta x) - F_x(x)} \simeq \frac{f(x, y)\Delta x}{f_x(x)\Delta x}$$

Hence

$$f_y(y \mid \mathbf{x} = x) = \lim_{\Delta x \to 0} f_y(y \mid x < \mathbf{x} \le x + \Delta x) = \frac{f(x, y)}{f_x(x)}$$

If there is no fear of ambiguity, the function $f_y(y \mid \mathbf{x} = x) = f_{y|x}(y \mid x)$ will be written in the form $f(y \mid x)$. Defining $f(x \mid y)$ similarly, we obtain

$$f(y \mid x) = \frac{f(x, y)}{f(x)} \qquad f(x \mid y) = \frac{f(x, y)}{f(y)} \tag{6-205}$$

If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent, then

$$f(x, y) = f(x)f(y) \qquad f(y \mid x) = f(y) \qquad f(x \mid y) = f(x)$$ ◀

Next we shall illustrate the method of obtaining conditional p.d.f.s through an example.

EXAMPLE 6-40 ▶ Given

$$f_{xy}(x, y) = \begin{cases} k & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases} \tag{6-206}$$

determine $f_{x|y}(x \mid y)$ and $f_{y|x}(y \mid x)$.

SOLUTION

The joint p.d.f. is given to be a constant in the shaded region in Fig. 6-29. This gives

$$\iint f_{xy}(x, y)\, dx\, dy = \int_0^1 \int_0^y k\, dx\, dy = \int_0^1 ky\, dy = \frac{k}{2} = 1 \Rightarrow k = 2$$

Similarly

$$f_x(x) = \int f_{xy}(x, y)\, dy = \int_x^1 k\, dy = k(1 - x) \qquad 0 < x < 1 \tag{6-207}$$

and

$$f_y(y) = \int f_{xy}(x, y)\, dx = \int_0^y k\, dx = ky \qquad 0 < y < 1 \tag{6-208}$$

From (6-206)–(6-208), we get

$$f_{x|y}(x \mid y) = \frac{f_{xy}(x, y)}{f_y(y)} = \frac{1}{y} \qquad 0 < x < y < 1 \tag{6-209}$$

and

$$f_{y|x}(y \mid x) = \frac{f_{xy}(x, y)}{f_x(x)} = \frac{1}{1 - x} \qquad 0 < x < y < 1 \tag{6-210}$$

◀

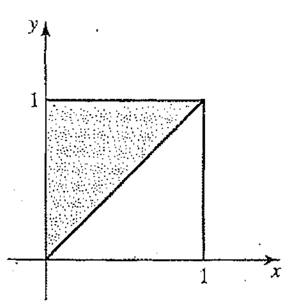


FIGURE 6-29

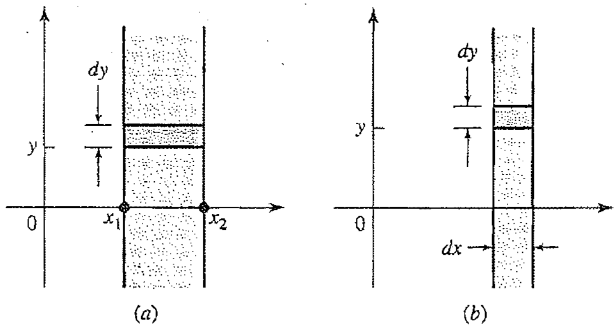


FIGURE 6-30

Notes 1. For a specific x, the function $f(x, y)$ is a *profile* of $f(x, y)$; that is, it equals the intersection of the surface $f(x, y)$ by the plane x = constant. The conditional density $f(y \mid x)$ is the equation of this curve normalized by the factor $1/f(x)$ so as to make its area 1. The function $f(x \mid y)$ has a similar interpretation: It is the normalized equation of the intersection of the surface $f(x, y)$ by the plane y = constant.

2. As we know, the product $f(y)dy$ equals the probability of the event $\{y < \mathbf{y} \le y + dy\}$. Extending this to conditional probabilities, we obtain

$$f_y(y \mid x_1 < \mathbf{x} \le x_2)\,dy = \frac{P\{x_1 < \mathbf{x} \le x_2, y < \mathbf{y} \le y + dy\}}{P\{x_1 < \mathbf{x} \le x_2\}}$$

This equals the mass in the rectangle of Fig. 6-30*a* divided by the mass in the vertical strip $x_1 < \mathbf{x} \le x_2$. Similarly, the product $f(y \mid x)dy$ equals the ratio of the mass in the differential rectangle $dx\,dy$ of Fig. 6-30*b* over the mass in the vertical strip $(x, x + dx)$.

3. The joint statistics of $\mathbf{x}$ and $\mathbf{y}$ are determined in terms of their joint density $f(x, y)$. Since

$$f(x, y) = f(y \mid x) f(x)$$

we conclude that they are also determined in terms of the marginal density $f(x)$ and the conditional density $f(y \mid x)$.

EXAMPLE 6-41

▶ We shall show that, if the random variables $\mathbf{x}$ and $\mathbf{y}$ are jointly normal with zero mean as in (6-61), then

$$f(y \mid x) = \frac{1}{\sigma_2\sqrt{2\pi(1 - r^2)}} \exp\left(-\frac{(y - r\sigma_2 x/\sigma_1)^2}{2\sigma_2^2(1 - r^2)}\right) \qquad (6\text{-}211)$$

Proof. The exponent in (6-61) equals

$$\frac{(y - r\sigma_2 x/\sigma_1)^2}{2\sigma_2^2(1 - r^2)} - \frac{x^2}{2\sigma_1^2}$$

Division by $f(x)$ removes the term $-x^2/2\sigma_1^2$ and (6-211) results.

The same reasoning leads to the conclusion that if $\mathbf{x}$ and $\mathbf{y}$ are jointly normal with $E\{\mathbf{x}\} = \eta_1$ and $E\{\mathbf{y}\} = \eta_2$, then $f(y \mid x)$ is given by (6-211) if y and x are replaced by $y - \eta_2$ and $x - \eta_1$, respectively. In other words, for a given x, $f(y \mid x)$ is a normal density with mean $\eta_2 + r\sigma_2(x - \eta_1)/\sigma_1$ and variance $\sigma_2^2(1 - r^2)$. ◀

BAYES' THEOREM AND TOTAL PROBABILITY. From (6-205) it follows that

$$f(x \mid y) = \frac{f(y \mid x) f(x)}{f(y)} \tag{6-212}$$

This is the density version of (2-43).

The denominator $f(y)$ can be expressed in terms of $f(y \mid x)$ and $f(x)$. Since

$$f(y) = \int_{-\infty}^{\infty} f(x, y)\, dx \quad \text{and} \quad f(x, y) = f(y \mid x) f(x)$$

we conclude that (total probability)

$$f(y) = \int_{-\infty}^{\infty} f(y \mid x) f(x)\, dx \tag{6-213}$$

Inserting into (6-212), we obtain Bayes' theorem for densities

$$f(x \mid y) = \frac{f(y \mid x) f(x)}{\int_{-\infty}^{\infty} f(y \mid x) f(x)\, dx} \tag{6-214}$$

Note As (6-213) shows, to remove the condition $\mathbf{x} = x$ from the conditional density $f(y \mid x)$, we multiply by the density $f(x)$ of $\mathbf{x}$ and integrate the product.

Equation (6-214) represents the p.d.f. version of Bayes' theorem. To appreciate the full significance of (6-214), we will look at a situation where observations are used to update our knowledge about unknown parameters. We shall illustrate this using the next example.

EXAMPLE 6-42

▶ An unknown random phase $\boldsymbol{\theta}$ is uniformly distributed in the interval $(0, 2\pi)$, and $\mathbf{r} = \boldsymbol{\theta} + \mathbf{n}$, where $\mathbf{n} \sim N(0, \sigma^2)$. Determine $f(\theta \mid r)$.

SOLUTION

Initially almost nothing about the random phase $\boldsymbol{\theta}$ is known, so that we assume its a priori p.d.f. to be uniform in the interval $(0, 2\pi)$. In the equation $\mathbf{r} = \boldsymbol{\theta} + \mathbf{n}$, we can think of $\mathbf{n}$ as the noise contribution and $\mathbf{r}$ as the observation. In practical situations, it is reasonable to assume that $\boldsymbol{\theta}$ and $\mathbf{n}$ are independent. If we assume so, then

$$f(r \mid \boldsymbol{\theta} = \theta) \sim N(\theta, \sigma^2) \tag{6-215}$$

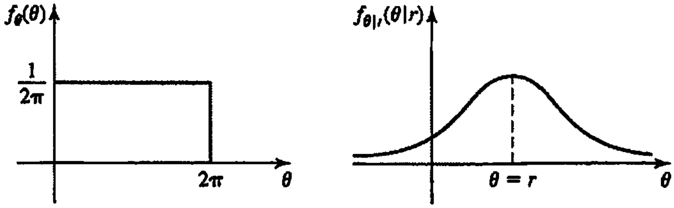
since it is given that $\boldsymbol{\theta}$ is a constant, and in that case $\mathbf{r} = \boldsymbol{\theta} + \mathbf{n}$ behaves like $\mathbf{n}$. Using (6-214), this gives the a posteriori p.d.f. of $\boldsymbol{\theta}$ given $\mathbf{r}$ to be (see Fig. 6-31b)

$$f(\theta \mid r) = \frac{f(r \mid \theta) f_\theta(\theta)}{\int_0^{2\pi} f(r \mid \theta) f_\theta(\theta)\, d\theta} = \frac{e^{-(r-\theta)^2/2\sigma^2}}{\int_0^{2\pi} e^{-(r-\theta)^2/2\sigma^2}\, d\theta}$$

$$= \varphi(r) e^{-(r-\theta)^2/2\sigma^2} \qquad 0 < \theta < 2\pi \tag{6-216}$$

where

$$\varphi(r) = \frac{1}{\int_0^{2\pi} e^{-(r-\theta)^2/2\sigma^2}\, d\theta}$$



(*a*) *a priori* p.d.f of θ

(*b*) *a posteriori* p.d.f of θ

FIGURE 6-31

Notice that knowledge about the observation $\mathbf{r}$ is reflected in the a posteriori p.d.f. of θ in Fig. 6.31*b*. It is no longer flat as the a priori p.d.f. in Fig. 6.31*a*, and it shows higher probabilities in the neighborhood of $\theta = r$.

Discrete Type Random Variables: Suppose that the random variables $\mathbf{x}$ and $\mathbf{y}$ are of discrete type

$$P\{\mathbf{x} = x_i\} = p_i \qquad P\{\mathbf{y} = y_k\} = q_k$$

$$P\{\mathbf{x} = x_i, \mathbf{y} = y_k\} = p_{ik} \qquad i = 1, \ldots, M \qquad k = 1, \ldots, N$$

where [see (6-34)]

$$p_i = \sum_k p_{ik} \qquad q_k = \sum_i p_{ik}$$

From the material just presented and (2-33) it follows that

$$P\{\mathbf{y} = y_k \mid \mathbf{x} = x_i\} = \frac{P\{\mathbf{x} = x_i, \mathbf{y} = y_k\}}{P\{\mathbf{x} = x_i\}} = \frac{p_{ik}}{p_i}$$

◀

MARKOV MATRIX. We denote by π_{ik} the above conditional probabilities

$$P\{\mathbf{y} = y_k \mid \mathbf{x} = x_i\} = \pi_{ik}$$

and by P the $M \times N$ matrix whose elements are π_{ik}. Clearly,

$$\pi_{ik} = \frac{p_{ik}}{p_i} \tag{6-217}$$

Hence

$$\pi_{ik} \geq 0, \quad \sum_k \pi_{ik} = 1 \tag{6-218}$$

Thus the elements of the matrix P are nonnegative and the sum on each row equals 1. Such a matrix is called *Markov* (see Chap. 15 for further details). The conditional probabilities

$$P\{\mathbf{x} = x_i \mid \mathbf{y} = y_k\} = \pi^{ki} = \frac{p_{ik}}{q_k}$$

are the elements of an $N \times M$ Markov matrix.

If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent, then

$$p_{ik} = p_i q_k \qquad \pi_{ik} = q_k \qquad \pi^{ki} = p_i$$

We note that

$$\pi^{ki} = \pi_{ik}\frac{p_i}{q_k} \qquad q_k = \sum_i \pi_{ik} p_i \tag{6-219}$$

These equations are the discrete versions of Eqs. (6-212) and (6-213).

Next we examine some interesting results involving conditional distributions and independent binomial/Poisson random variables.

EXAMPLE 6-43

▶ Suppose x and y are independent binomial random variables with parameters (m, p) and (n, p) respectively. Then $\mathbf{x}+\mathbf{y}$ is also binomial with parameter $(m+n, p)$, so that

$$P\{\mathbf{x}=x \mid \mathbf{x}+\mathbf{y}=x+y\} = \frac{P\{\mathbf{x}=x\}P\{\mathbf{y}=y\}}{P\{\mathbf{x}+\mathbf{y}=x+y\}} = \frac{\binom{m}{x}\binom{n}{y}}{\binom{m+n}{x+y}} \tag{6-220}$$

Thus the conditional distribution of x given $\mathbf{x}+\mathbf{y}$ is hypergeometric. Show that the converse of this result which states that if x and y are nonnegative independent random variables such that $P\{\mathbf{x}=0\} > 0$, $P\{\mathbf{y}=0\} > 0$ and the conditional distribution of x given $\mathbf{x}+\mathbf{y}$ is hypergeometric as in (6-220), then x and y are binomial random variables.

SOLUTION

From (6-220)

$$\frac{P\{\mathbf{x}=x\}}{\binom{m}{x}}\frac{P\{\mathbf{y}=y\}}{\binom{n}{y}} = \frac{P\{\mathbf{x}+\mathbf{y}=x+y\}}{\binom{m+n}{x+y}}$$

Let

$$\frac{P\{\mathbf{x}=x\}}{\binom{m}{x}} = f(x) \qquad \frac{P\{\mathbf{y}=y\}}{\binom{n}{y}} = g(y) \qquad \frac{P\{\mathbf{x}+\mathbf{y}=x+y\}}{\binom{m+n}{x+y}} = h(x+y)$$

Then

$$h(x+y) = f(x)\,g(y)$$

and hence

$$h(1) = f(1)g(0) = f(0)g(1)$$

$$h(2) = f(2)g(0) = f(1)g(1) = f(0)g(2)$$

$$\vdots$$

$$h(k) = f(k)g(0) = f(k-1)g(1) = \cdots$$

Thus

$$f(k) = f(k-1)\frac{g(1)}{g(0)} = f(0)\left(\frac{g(1)}{g(0)}\right)^k$$

or

$$P\{\mathbf{x}=k\} = \binom{m}{k}P\{\mathbf{x}=0\}a^k \qquad k = 0, 1, \ldots \tag{6-221}$$

where $a = g(1)/g(0) > 0$. But $\sum_{k=0}^{m} P\{\mathbf{x}=k\} = 1$ gives $P\{\mathbf{x}=0\}(1+a)^m = 1$, or $P\{\mathbf{x}=0\} = q^m$, where $q = 1/(1+a) < 1$. Hence $a = p/q$, where $p = 1-q > 0$

and from (6-221) we obtain

$$P\{\mathbf{x}=k\} = \binom{m}{k} p^k q^{m-k} \qquad k = 0, 1, 2, \ldots m$$

Similarly, it follows that

$$P\{\mathbf{y}=r\} = \binom{n}{r} p^r q^{n-r} \qquad r = 0, 1, 2, \ldots n$$

and the proof is complete.

Similarly if $\mathbf{x}$ and $\mathbf{y}$ are independent Poisson random variables with parameters λ and μ respectively, then their sum is also Poisson with parameter $\lambda + \mu$ [see (6-86)], and, moreover,

$$P\{\mathbf{x}=k \mid \mathbf{x}+\mathbf{y}=n\} = \frac{P\{\mathbf{x}=k\}P\{\mathbf{y}=n-k\}}{P\{\mathbf{x}+\mathbf{y}=n\}} = \frac{e^{-\lambda}\frac{\lambda^k}{k!}e^{-\mu}\frac{\mu^{n-k}}{(n-k)!}}{e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^n}{n!}}$$

$$= \binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{n-k} \qquad k = 0, 1, 2, \ldots n \qquad \text{(6-222)}$$

Thus if $\mathbf{x}$ and $\mathbf{y}$ are independent Poisson random variables, then the conditional density of $\mathbf{x}$ given $\mathbf{x}+\mathbf{y}$ is Binomial. Interestingly, the converse is also true, when $\mathbf{x}$ and $\mathbf{y}$ are independent random variables. The proof is left as an exercise.

Equivalently, this shows that if $\mathbf{y} = \sum_{i=1}^{\mathbf{n}} \mathbf{x}_i$ where $\mathbf{x}_i$ are independent Bernoulli random variables as in (4-55) and $\mathbf{n}$ is a Poisson random variable with parameter λ as in (4-57), then $\mathbf{y} \sim P(p\lambda)$ and $\mathbf{z} = \mathbf{n} - \mathbf{y} \sim P((1-p)\lambda)$. Further, $\mathbf{y}$ and $\mathbf{z}$ are independent random variables. Thus, for example, if the total number of eggs that a bird lays follows a Poisson random variable with parameter λ, and if each egg survives with probability p, then the number of baby birds that survive is also Poisson with parameter $p\lambda$. ◀

System Reliability

We shall use the term *system* to identify a physical device used to perform a certain function. The device might be a simple element, a lightbulb, for example, or a more complicated structure. We shall call the time interval from the moment the system is put into operation until it fails the *time to failure*. This interval is, in general, random. It specifies, therefore, a random variable $\mathbf{x} \geq 0$. The distribution $F(t) = P\{\mathbf{x} \leq t\}$ of this random variable is the probability that the system fails prior to time t where we assume that $t = 0$ is the moment the system is put into operation. The difference

$$R(t) = 1 - F(t) = P\{\mathbf{x} > t\}$$

is the *system reliability*. It equals the probability that the system functions at time t.

The *mean time to failure* of a system is the mean of $\mathbf{x}$. Since $F(x) = 0$ for $x < 0$, we conclude from (5-52) that

$$E\{\mathbf{x}\} = \int_0^\infty x f(x)\,dx = \int_0^\infty R(t)\,dt$$

The probability that a system functioning at time t fails prior to time $x > t$ equals

$$F(x \mid \mathbf{x} > t) = \frac{P\{\mathbf{x} \leq x, \mathbf{x} > t\}}{P\{\mathbf{x} > t\}} = \frac{F(x) - F(t)}{1 - F(t)}$$

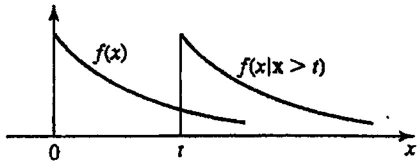


FIGURE 6-32

Differentiating with respect to x, we obtain

$$f(x \mid \mathbf{x} > t) = \frac{f(x)}{1 - F(t)} \qquad x > t \tag{6-223}$$

The product $f(x \mid \mathbf{x} > t)\,dx$ equals the probability that the system fails in the interval $(x, x + dx)$, assuming that it functions at time t.

EXAMPLE 6-44

▶ If $f(x) = ce^{-cx}$, then $F(t) = 1 - e^{-ct}$ and (6-223) yields

$$f(x \mid \mathbf{x} > t) = \frac{ce^{-cx}}{e^{-ct}} = f(x - t)$$

This shows that the probability that a system functioning at time t fails in the interval $(x, x + dx)$ depends only on the difference $x - t$ (Fig. 6-32). We show later that this is true only if $f(x)$ is an exponential density. ◀

CONDITIONAL FAILURE RATE. The conditional density $f(x \mid \mathbf{x} > t)$ is a function of x and t. Its value at $x = t$ is a function only of t. This function is denoted by $\beta(t)$ and is called the *conditional failure rate* or, the *hazard rate* of the system. From (6-223) and the definition of hazard rate it follows that

$$\beta(t) = f(t \mid \mathbf{x} > t) = \frac{f(t)}{1 - F(t)} \tag{6-224}$$

The product $\beta(t)\,dt$ is the probability that a system functioning at time t fails in the interval $(t, t + dt)$. In Sec. 7-1 (Example 7-3) we interpret the function $\beta(t)$ as the expected failure rate.

EXAMPLE 6-45

▶ (*a*) If $f(x) = ce^{-cx}$, then $F(t) = 1 - e^{-ct}$ and

$$\beta(t) = \frac{ce^{-ct}}{1 - (1 - e^{-ct})} = c$$

(*b*) If $f(x) = c^2xe^{-cx}$, then $F(x) = 1 - cxe^{-cx} - e^{-cx}$ and

$$\beta(t) = \frac{c^2te^{-ct}}{cte^{-ct} + e^{-ct}} = \frac{c^2t}{1 + ct}$$ ◀

From (6-224) it follows that

$$\beta(t) = \frac{F'(t)}{1 - F(t)} = -\frac{R'(t)}{R(t)}$$

We shall use this relationship to express the distribution of $\mathbf{x}$ in terms of the function

$\beta(t)$. Integrating from 0 to x and using the fact that $\ln R(0) = 0$, we obtain

$$-\int_0^x \beta(t)\,dt = \ln R(x)$$

Hence

$$R(x) = 1 - F(x) = \exp\left\{-\int_0^x \beta(t)\,dt\right\}$$

And since $f(x) = F'(x)$, this yields

$$f(x) = \beta(x)\exp\left\{-\int_0^x \beta(t)\,dt\right\} \tag{6-225}$$

EXAMPLE 6-46

MEMORYLESS SYSTEMS

▶ A system is called *memoryless* if the probability that it fails in an interval (t, x), assuming that it functions at time t, depends only on the length of this interval. In other words, if the system works a week, a month, or a year after it was put into operation, it is as good as new. This is equivalent to the assumption that $f(x \mid \mathbf{x} > t) = f(x - t)$ as in Fig. 6-32. From this and (6-224) it follows that with $x = t$:

$$\beta(t) = f(t \mid \mathbf{x} > t) = f(t - t) = f(0) = c$$

and (6-225) yields $f(x) = ce^{-cx}$. Thus a system is memoryless iff $\mathbf{x}$ has an exponential density. ◀

EXAMPLE 6-47

▶ A special form of $\beta(t)$ of particular interest in reliability theory is the function

$$\beta(t) = ct^{b-1}$$

This is a satisfactory approximation of a variety of failure rates, at least near the origin. The corresponding $f(x)$ is obtained from (6-225):

$$f(x) = cx^{b-1}\exp\left\{-\frac{cx^b}{b}\right\} \tag{6-226}$$

This function is called the *Weibull density*. (See (4-43) and Fig. 4-16.) ◀

We conclude with the observation that the function $\beta(t)$ equals the value of the conditional density $f(x \mid \mathbf{x} > t)$ for $x = t$; however, $\beta(t)$ is not a density because its area is not one. In fact its area is infinite. This follows from (6-224) because $R(\infty) = 1 - F(\infty) = 0$.

INTERCONNECTION OF SYSTEMS. We are given two systems S_1 and S_2 with times to failure $\mathbf{x}$ and $\mathbf{y}$, respectively, and we connect them in parallel or in series or in standby as in Fig. 6-33, forming a new system S. We shall express the properties of S in terms of the joint distribution of the random variables $\mathbf{x}$ and $\mathbf{y}$.

Parallel: We say that the two systems are connected in parallel if S fails when both systems fail. Denoting by $\mathbf{z}$ the time to failure of S, we conclude that $\mathbf{z} = t$ when the

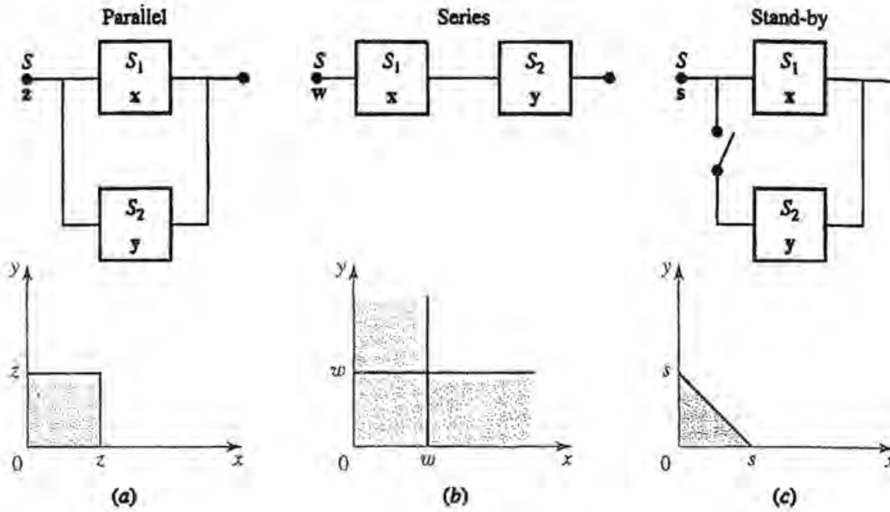


FIGURE 6-33

larger of the numbers **x** and **y** equals t. Hence [see (6-77)–(6-78)]

$$\mathbf{z} = \max(\mathbf{x}, \mathbf{y}) \qquad F_z(z) = F_{xy}(z, z)$$

If the random variables **x** and **y** are independent, $F_z(z) = F_x(z)F_y(z)$.

Series: We say that the two systems are connected in series if S fails when at least one of the two systems fails. Denoting by **w** the time to failure of S, we conclude that $\mathbf{w} = t$ when the smaller of the numbers **x** and **y** equals t. Hence [see (6-80)–(6-81)]

$$\mathbf{w} = \min(\mathbf{x}, \mathbf{y}) \qquad F_w(w) = F_x(w) + F_y(w) - F_{xy}(w, w)$$

If the random variables **x** and **y** are independent,

$$R_w(w) = R_x(w)R_y(w) \qquad \beta_w(t) = \beta_x(t) + \beta_y(t)$$

where $\beta_x(t)$, $\beta_y(t)$, and $\beta_w(t)$ are the conditional failure rates of systems S_1, S_2, and S, respectively.

Standby: We put system S_1 into operation, keeping S_2 in reserve. When S_1 fails, we put S_2 into operation. The system S so formed fails when S_2 fails. If t_1 and t_2 are the times of operation of S_1 and S_2, $t_1 + t_2$ is the time of operation of S. Denoting by **s** the time to failure of system S, we conclude that

$$\mathbf{s} = \mathbf{x} + \mathbf{y}$$

The distribution of **s** equals the probability that the point (**x**, **y**) is in the triangular shaded region of Fig. 6-33*c*. If the random variables **x** and **y** are independent, the density of **s** equals

$$f_s(s) = \int_0^s f_x(t)f_y(s-t)\,dt$$

as in (6-45).

6-7 CONDITIONAL EXPECTED VALUES

Applying theorem (5-55) to conditional densities, we obtain the conditional mean of $g(\mathbf{y})$:

$$E\{g(\mathbf{y}) \mid M\} = \int_{-\infty}^{\infty} g(y) f(y \mid M)\, dy \tag{6-227}$$

This can be used to define the conditional moments of $\mathbf{y}$.

Using a limit argument as in (6-205), we can also define the conditional mean $E\{g(\mathbf{y}) \mid x\}$. In particular,

$$\eta_{y|x} = E\{\mathbf{y} \mid x\} = \int_{-\infty}^{\infty} y f(y \mid x)\, dy \tag{6-228}$$

is the *conditional mean* of $\mathbf{y}$ assuming $\mathbf{x} = x$, and

$$\sigma_{y|x}^2 = E\{(\mathbf{y} - \eta_{y|x})^2 \mid x\} = \int_{-\infty}^{\infty} (y - \eta_{y|x})^2 f(y \mid x)\, dy \tag{6-229}$$

is its *conditional variance*.

We shall illustrate these calculations through an example.

EXAMPLE 6-48 ▶ Let

$$f_{xy}(x, y) = \begin{cases} 1 & 0 < |y| < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{6-230}$$

Determine $E\{\mathbf{x} \mid \mathbf{y}\}$ and $E\{\mathbf{y} \mid \mathbf{x}\}$.

SOLUTION

As Fig. 6-34 shows, $f_{xy}(x, y) = 1$ in the shaded area, and zero elsewhere. Hence

$$f_x(x) = \int_{-x}^{x} f_{xy}(x, y)\, dy = 2x \qquad 0 < x < 1$$

and

$$f_y(y) = \int_{-|y|}^{1} 1\, dx = 1 - |y| \qquad |y| < 1$$

This gives

$$f_{x|y}(x, y) = \frac{f_{xy}(x, y)}{f_y(y)} = \frac{1}{1 - |y|} \qquad 0 < |y| < x < 1 \tag{6-231}$$

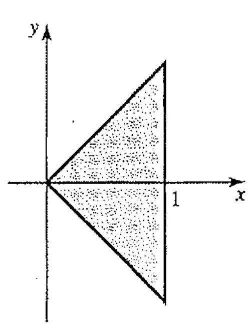


FIGURE 6-34

and

$$f_{y|x}(y \mid x) = \frac{f_{xy}(x, y)}{f_x(x)} = \frac{1}{2x} \qquad 0 < |y| < x < 1 \tag{6-232}$$

Hence

$$E\{\mathbf{x} \mid \mathbf{y}\} = \int x f_{x|y}(x \mid y)\, dx = \int_{|y|}^{1} \frac{x}{(1-|y|)}\, dx = \frac{1}{(1-|y|)} \frac{x^2}{2}\Big|_{|y|}^{1}$$

$$= \frac{1-|y|^2}{2(1-|y|)} = \frac{1+|y|}{2} \qquad |y| < 1 \tag{6-233}$$

$$E\{\mathbf{y} \mid \mathbf{x}\} = \int y f_{y|x}(y \mid x)\, dy = \int_{-x}^{x} \frac{y}{2x}\, dy = \frac{1}{2x} \frac{y^2}{2}\Big|_{-x}^{x} = 0 \qquad 0 < x < 1 \tag{6-234}$$

◀

For a given x, the integral in (6-228) is the center of gravity of the masses in the vertical strip $(x, x + dx)$. The locus of these points, as x varies from $-\infty$ to ∞, is the function

$$\varphi(x) = \int_{-\infty}^{\infty} y f(y \mid x)\, dy \tag{6-235}$$

known as the *regression line* (Fig. 6-35).

Note If the random variables $\mathbf{x}$ and $\mathbf{y}$ are functionally related, that is, if $\mathbf{y} = g(\mathbf{x})$, then the probability masses on the xy plane are on the line $y = g(x)$ (see Fig. 6-5b); hence $E\{\mathbf{y} \mid x\} = g(x)$.

Galton's law. The term *regression* has its origin in the following observation attributed to the geneticist Sir Francis Galton (1822–1911): "Population extremes *regress* toward their mean." This observation applied to parents and their adult children implies that children of tall (or short) parents are on the average shorter (or taller) than their parents. In statistical terms be phrased in terms of conditional expected values:

Suppose that the random variables $\mathbf{x}$ and $\mathbf{y}$ model the height of parents and their children respectively. These random variables have the same mean and variance, and they are positively correlated:

$$\eta_x = \eta_x = \eta \qquad \sigma_x = \sigma_y = \sigma \qquad r > 0$$

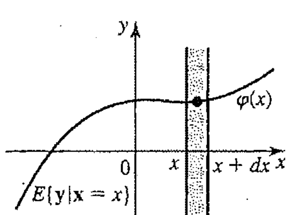


FIGURE 6-35

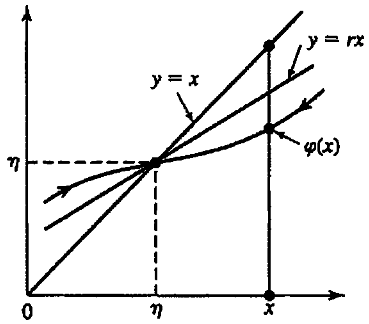


FIGURE 6-36

According to Galton's law, the conditional mean $E\{\mathbf{y} \mid x\}$ of the height of children whose parents height is x, is smaller (or larger) than x if $x > \eta$ (or $x < \eta$):

$$E\{\mathbf{y} \mid x\} = \varphi(x) \begin{cases} <x & \text{if} \quad x > \eta \\ >x & \text{if} \quad x < \eta \end{cases}$$

This shows that the regression line $\varphi(x)$ is below the line $y = x$ for $x > \eta$ and above this line if $x < \eta$ as in Fig. 6-36. If the random variables $\mathbf{x}$ and $\mathbf{y}$ are jointly normal, then [see (6-236) below] the regression line is the straight line $\varphi(x) = rx$. For arbitrary random variables, the function $\varphi(x)$ *does not* obey Galton's law. The term regression is used, however, to identify any conditional mean.

EXAMPLE 6-49

▶ If the random variables $\mathbf{x}$ and $\mathbf{y}$ are normal as in Example 6-41, then the function

$$E\{\mathbf{y} \mid x\} = \eta_2 + r\sigma_2 \frac{x - \eta_1}{\sigma_1} \tag{6-236}$$

is a straight line with slope $r\sigma_2/\sigma_1$ passing through the point (η_1, η_2). Since for normal random variables the conditional mean $E\{\mathbf{y} \mid x\}$ coincides with the maximum of $f(y \mid x)$, we conclude that the locus of the maxima of all profiles of $f(x, y)$ is the straight line (6-236).

From theorems (6-159) and (6-227) it follows that

$$E\{g(\mathbf{x}, \mathbf{y}) \mid M\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y \mid M)\, dx\, dy \tag{6-237}$$

◀

This expression can be used to determine $E\{g(\mathbf{x}, \mathbf{y}) \mid x\}$; however, the conditional density $f(x, y \mid x)$ consists of line masses on the line x-constant. To avoid dealing with line masses, we shall define $E\{g(\mathbf{x}, \mathbf{y}) \mid x\}$ as a limit:

As we have shown in Example 6-39, the conditional density $f(x, y \mid x < \mathbf{x} < x + \Delta x)$ is 0 outside the strip $(x, x + \Delta x)$ and in this strip it is given by (6-203) where $x_1 = x$ and $x_2 = x + \Delta x$. It follows, therefore, from (6-237) with $M = \{x < \mathbf{x} \le x + \Delta x\}$ that

$$E\{g(\mathbf{x}, \mathbf{y}) \mid x < \mathbf{x} \le x + \Delta x\} = \int_{-\infty}^{\infty} \int_{x}^{x+\Delta x} g(\alpha, y) \frac{f(\alpha, y)\, d\alpha}{F_x(x + \Delta x) - F_x(x)}\, dy$$

As $\Delta x \to 0$, the inner integral tends to $g(x, y) f(x, y)/f(x)$. Defining $E\{g(\mathbf{x}, \mathbf{y}) \mid x\}$ as the limit of the above integral, we obtain

$$E\{g(\mathbf{x}, \mathbf{y}) \mid x\} = \int_{-\infty}^{\infty} g(x, y) f(y \mid x)\, dy \tag{6-238}$$

We also note that

$$E\{g(x, \mathbf{y}) \mid x\} = \int_{-\infty}^{\infty} g(x, y) f(y \mid x)\, dy \tag{6-239}$$

because $g(x, \mathbf{y})$ is a function of the random variable $\mathbf{y}$, with x a parameter; hence its conditional expected value is given by (6-227). Thus

$$E\{g(\mathbf{x}, \mathbf{y}) \mid x\} = E\{g(x, \mathbf{y}) \mid x\} \tag{6-240}$$

One might be tempted from the above to conclude that (6-240) follows directly from (6-227); however, this is not so. The functions $g(\mathbf{x}, \mathbf{y})$ and $g(x, \mathbf{y})$ have the same expected value, assuming $\mathbf{x} = x$, but they are not equal. The first is a function $g(\mathbf{x}, \mathbf{y})$ of the random variables $\mathbf{x}$ and $\mathbf{y}$, and for a specific ζ it takes the value $g[\mathbf{x}(\zeta), \mathbf{y}(\zeta)]$. The second is a function $g(x, \mathbf{y})$ of the real variable x and the random variable $\mathbf{y}$, and for a specific ζ it takes the value $g[x, \mathbf{y}(\zeta)]$ where x is an arbitrary number.

Conditional Expected Values as Random Variables

The conditional mean of $\mathbf{y}$, assuming $\mathbf{x} = x$, is a function $\varphi(x) = E\{\mathbf{y} \mid x\}$ of x given by (6-235). Using this function, we can construct the random variable $\varphi(\mathbf{x}) = E\{\mathbf{y} \mid \mathbf{x}\}$ as in Sec. 5-1. As we see from (5-55), the mean of this random variable equals

$$E\{\varphi(\mathbf{x})\} = \int_{-\infty}^{\infty} \varphi(x) f(x)\, dx = \int_{-\infty}^{\infty} f(x) \int_{-\infty}^{\infty} y f(y \mid x)\, dy\, dx$$

Since $f(x, y) = f(x) f(y \mid x)$, the last equation yields

$$E\{E\{\mathbf{y} \mid \mathbf{x}\}\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y)\, dx\, dy = E\{\mathbf{y}\} \tag{6-241}$$

This basic result can be generalized: The conditional mean $E\{g(\mathbf{x}, \mathbf{y}) \mid x\}$ of $g(\mathbf{x}, \mathbf{y})$, assuming $\mathbf{x} = x$, is a function of the real variable x. It defines, therefore, the function $E\{g(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}\}$ of the random variable $\mathbf{x}$. As we see from (6-159) and (6-237), the mean of $E\{g(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}\}$ equals

$$\int_{-\infty}^{\infty} f(x) \int_{-\infty}^{\infty} g(x, y) f(y \mid x)\, dy\, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx\, dy$$

But the last integral equals $E\{g(\mathbf{x}, \mathbf{y})\}$; hence

$$E\{E\{g(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}\}\} = E\{g(\mathbf{x}, \mathbf{y})\} \tag{6-242}$$

We note, finally, that

$$E\{g_1(\mathbf{x}) g_2(\mathbf{y}) \mid x\} = E\{g_1(x) g_2(\mathbf{y}) \mid x\} = g_1(x) E\{g_2(\mathbf{y}) \mid x\} \tag{6-243}$$

$$E\{g_1(\mathbf{x}) g_2(\mathbf{y})\} = E\{E\{g_1(\mathbf{x}) g_2(\mathbf{y}) \mid \mathbf{x}\}\} = E\{g_1(\mathbf{x}) E\{g_2(\mathbf{y}) \mid \mathbf{x}\}\} \tag{6-244}$$

EXAMPLE 6-50 ▶ Suppose that the random variables x and y are $N(0, 0, \sigma_1^2, \sigma_2^2, r)$. As we know

$$E\{\mathbf{x}^2\} = \sigma_1^2 \qquad E\{\mathbf{x}^4\} = 3\sigma_1^4$$

Furthermore, $f(y \mid x)$ is a normal density with mean $r\sigma_2 x/\sigma_1$ and variance $\sigma_2\sqrt{1-r^2}$. Hence

$$E\{\mathbf{y}^2 \mid x\} = \eta_{y|x}^2 + \sigma_{y|x}^2 = \left(\frac{r\sigma_2 x}{\sigma_1}\right)^2 + \sigma_2^2(1-r^2) \tag{6-245}$$

Using (6-244), we shall show that

$$E\{\mathbf{xy}\} = r\sigma_1\sigma_2 \qquad E\{\mathbf{x}^2\mathbf{y}^2\} = E\{\mathbf{x}^2\}E\{\mathbf{y}^2\} + 2E^2\{\mathbf{xy}\} \tag{6-246}$$

Proof.

$$E\{\mathbf{xy}\} = E\{\mathbf{x}E\{\mathbf{y} \mid \mathbf{x}\}\} = E\left\{r\sigma_2\frac{\mathbf{x}^2}{\sigma_1}\right\} = r\sigma_2\frac{\sigma_1^2}{\sigma_1}$$

$$E\{\mathbf{x}^2\mathbf{y}^2\} = E\{\mathbf{x}^2E\{\mathbf{y}^2 \mid \mathbf{x}\}\} = E\left\{\mathbf{x}^2\left[r^2\sigma_2^2\frac{\mathbf{x}^2}{\sigma_1^2} + \sigma_2^2(1-r^2)\right]\right\}$$

$$= 3\sigma_1^4 r^2\frac{\sigma_2^2}{\sigma_1^2} + \sigma_1^2\sigma_2^2(1-r^2) = \sigma_1^2\sigma_2^2 + 2r^2\sigma_1^2\sigma_2^2$$

and the proof is complete [see also (6-199)]. ◀

PROBLEMS

6-1 x and y are independent, identically distributed (i.i.d.) random variables with common p.d.f.

$$f_x(x) = e^{-x}U(x) \qquad f_y(y) = e^{-y}U(y)$$

Find the p.d.f. of the following random variables (*a*) $\mathbf{x}+\mathbf{y}$, (*b*) $\mathbf{x}-\mathbf{y}$, (*c*) $\mathbf{xy}$, (*d*) $\mathbf{x}/\mathbf{y}$, (*e*) $\min(\mathbf{x}, \mathbf{y})$, (*f*) $\max(\mathbf{x}, \mathbf{y})$, (*g*) $\min(\mathbf{x}, \mathbf{y})/\max(\mathbf{x}, \mathbf{y})$.

6-2 x and y are independent and uniform in the interval $(0, a)$. Find the p.d.f. of (*a*) $\mathbf{x}/\mathbf{y}$, (*b*) $\mathbf{y}/(\mathbf{x}+\mathbf{y})$, (*c*) $|\mathbf{x}-\mathbf{y}|$.

6-3 The joint p.d.f. of the random variables x and y is given by

$$f_{xy}(x, y) = \begin{cases} 1 & \text{in the shaded area} \\ 0 & \text{otherwise} \end{cases}$$

Let $\mathbf{z} = \mathbf{x}+\mathbf{y}$. Find $F_z(z)$ and $f_z(z)$.

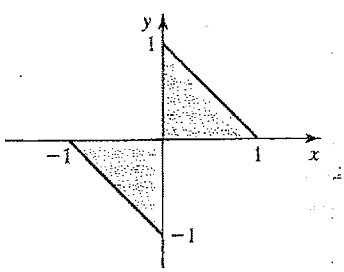


FIGURE P6-3

6-4 The joint p.d.f. of $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$f_{xy}(x,y) = \begin{cases} 6x & x \geq 0, y \geq 0, x+y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Define $\mathbf{z} = \mathbf{x} - \mathbf{y}$. Find the p.d.f. of $\mathbf{z}$.

6-5 $\mathbf{x}$ and $\mathbf{y}$ are independent identically distributed normal random variables with zero mean and common variance σ^2, that is, $\mathbf{x} \sim N(0, \sigma^2)$, $\mathbf{y} \sim N(0, \sigma^2)$ and $f_{xy}(x,y) = f_x(x) f_y(y)$. Find the p.d.f. of (*a*) $\mathbf{z} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$, (*b*) $\mathbf{w} = \mathbf{x}^2 + \mathbf{y}^2$, (*c*) $\mathbf{u} = \mathbf{x} - \mathbf{y}$.

6-6 The joint p.d.f. of $\mathbf{x}$ and $\mathbf{y}$ is given by

$$f_{xy}(x,y) = \begin{cases} 2(1-x) & 0 < x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Determine the probability density function of $\mathbf{z} = \mathbf{xy}$.

6-7 Given

$$f_{xy}(x,y) = \begin{cases} x+y & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Show that (*a*) $\mathbf{x} + \mathbf{y}$ has density $f_1(z) = z^2$, $0 < z < 1$, $f_1(z) = z(2-z)$, $1 < z < 2$, and 0 elsewhere. (*b*) $\mathbf{xy}$ has density $f_2(z) = 2(1-z)$, $0 < z < 1$, and 0 elsewhere. (*c*) $\mathbf{y}/\mathbf{x}$ has density $f_3(z) = (1+z)/3$, $0 < z < 1$, $f_3(z) = (1+z)/3z^3$, $z > 1$, and 0 elsewhere. (*d*) $\mathbf{y} - \mathbf{x}$ has density $f_4(z) = 1 - |z|$, $|z| < 1$, and 0 elsewhere.

6-8 Suppose $\mathbf{x}$ and $\mathbf{y}$ have joint density

$$f_{xy}(x,y) = \begin{cases} 1 & 0 \leq x \leq 2, 0 \leq y \leq 1, 2y \leq x \\ 0 & \text{otherwise} \end{cases}$$

Show that $\mathbf{z} = \mathbf{x} + \mathbf{y}$ has density

$$f_{xy}(x,y) = \begin{cases} (1/3)z & 0 < z < 2 \\ 2 - (2/3)z & 2 < z < 3 \\ 0 & \text{elsewhere} \end{cases}$$

6-9 $\mathbf{x}$ and $\mathbf{y}$ are uniformly distributed on the triangular region $0 \leq y \leq x \leq 1$. Show that (*a*) $\mathbf{z} = \mathbf{x}/\mathbf{y}$ has density $f_z(z) = 1/z^2$, $z \geq 1$, and $f_z(z) = 0$, otherwise. (*b*) Determine the density of $\mathbf{xy}$.

6-10 $\mathbf{x}$ and $\mathbf{y}$ are uniformly distributed on the triangular region $0 < x \leq y \leq x + y \leq 2$. Find the p.d.f. of $\mathbf{x} + \mathbf{y}$ and $\mathbf{x} - \mathbf{y}$.

6-11 $\mathbf{x}$ and $\mathbf{y}$ are independent Gamma random variables with common parameters α and β. Find the p.d.f. of (*a*) $\mathbf{x} + \mathbf{y}$, (*b*) $\mathbf{x}/\mathbf{y}$, (*c*) $\mathbf{x}/(\mathbf{x} + \mathbf{y})$.

6-12 $\mathbf{x}$ and $\mathbf{y}$ are independent uniformly distributed random variables on $(0, a)$. Find the joint p.d.f. of $\mathbf{x} + \mathbf{y}$ and $\mathbf{x} - \mathbf{y}$.

6-13 $\mathbf{x}$ and $\mathbf{y}$ are independent Rayleigh random variables with common parameter σ^2. Determine the density of $\mathbf{x}/\mathbf{y}$.

6-14 The random variables $\mathbf{x}$ and $\mathbf{y}$ are independent and $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Find $f_y(y)$ if

$$f_x(x) = ce^{-cx} U(x) \qquad f_z(z) = c^2 z e^{-cz} U(z)$$

6-15 The random variables $\mathbf{x}$ and $\mathbf{y}$ are independent and $\mathbf{y}$ is uniform in the interval (0, 1). Show that, if $\mathbf{z} = \mathbf{x} + \mathbf{y}$, then

$$f_z(z) = F_x(z) - F_x(z-1)$$

6-16 (*a*) The function $g(x)$ is monotone increasing and $\mathbf{y} = g(\mathbf{x})$. Show that

$$F_{xy}(x, y) = \begin{cases} F_x(x) & \text{if} \quad y > g(x) \\ F_y(y) & \text{if} \quad y < g(x) \end{cases}$$

(*b*) Find $F_{xy}(x, y)$ if $g(x)$ is monotone decreasing.

6-17 The random variables $\mathbf{x}$ and $\mathbf{y}$ are $N(0, 4)$ and independent. Find $f_z(z)$ and $F_z(z)$ if (*a*) $\mathbf{z} = 2\mathbf{x} + 3\mathbf{y}$, and (*b*) $\mathbf{z} = \mathbf{x}/\mathbf{y}$.

6-18 The random variables $\mathbf{x}$ and $\mathbf{y}$ are independent with

$$f_x(x) = \frac{x}{\alpha^2} e^{-x^2/2\alpha^2} U(x) \qquad f_y(y) = \begin{cases} 1/\pi\sqrt{1-y^2} & |y| < 1 \\ 0 & |y| > 1 \end{cases}$$

Show that the random variable $\mathbf{z} = \mathbf{xy}$ is $N(0, \alpha^2)$.

6-19 The random variables $\mathbf{x}$ and $\mathbf{y}$ are independent with Rayleigh densities

$$f_x(x) = \frac{x}{\alpha^2} e^{-x^2/2\alpha^2} U(x) \qquad f_y(y) = \frac{y}{\beta^2} e^{-y^2/2\beta^2} U(y)$$

(*a*) Show that if $\mathbf{z} = \mathbf{x}/\mathbf{y}$, then

$$f_z(z) = \frac{2\alpha^2}{\beta^2} \frac{z}{(z^2 + \alpha^2/\beta^2)^2} U(z) \tag{i}$$

(*b*) Using (i), show that for any $k > 0$,

$$P\{\mathbf{x} \le k\mathbf{y}\} = \frac{k^2}{k^2 + \alpha^2/\beta^2}$$

6-20 The random variables $\mathbf{x}$ and $\mathbf{y}$ are independent with exponential densities

$$f_x(x) = \alpha e^{-\alpha x} U(x) \qquad f_y(y) = \beta e^{-\beta y} U(y)$$

Find the densities of the following random variables:

(*a*) $2\mathbf{x} + \mathbf{y}$ (*b*) $\mathbf{x} - \mathbf{y}$ (*c*) $\dfrac{\mathbf{x}}{\mathbf{y}}$ (*d*) $\max(\mathbf{x}, \mathbf{y})$ (*e*) $\min(\mathbf{x}, \mathbf{y})$

6-21 The random variables $\mathbf{x}$ and $\mathbf{y}$ are independent and each is uniform in the interval $(0, a)$. Find the density of the random variable $\mathbf{z} = |\mathbf{x} - \mathbf{y}|$.

6-22 Show that (*a*) the convolution of two normal densities is a normal density, and (*b*) the convolution of two Cauchy densities is a Cauchy density.

6-23 The random variables $\mathbf{x}$ and $\mathbf{y}$ are independent with respective densities $\chi^2(m)$ and $\chi^2(n)$. Show that if (Example 6-29)

$$\mathbf{z} = \frac{\mathbf{x}/m}{\mathbf{y}/n} \qquad \text{then} \quad f_z(z) = \gamma \frac{z^{m/2-1}}{\sqrt{(1 + mz/n)^{m+n}}} U(z)$$

This distribution is denoted by $F(m, n)$ and is called the *Snedecor F* distribution. It is used in hypothesis testing (see Prob. 8-34).

6-24 Express $F_{zw}(z, w)$ in terms of $F_{xy}(x, y)$ if $\mathbf{z} = \max(\mathbf{x}, \mathbf{y})$, $\mathbf{w} = \min(\mathbf{x}, \mathbf{y})$.

6-25 Let $\mathbf{x}$ be the lifetime of a certain electric bulb, and $\mathbf{y}$ that of its replacement after the failure of the first bulb. Suppose $\mathbf{x}$ and $\mathbf{y}$ are independent with common exponential density function with parameter λ. Find the probability that the combined lifetime exceeds 2λ. What is the probability that the replacement outlasts the original component by λ?

6-26 $\mathbf{x}$ and $\mathbf{y}$ are independent uniformly distributed random variables in $(0, 1)$. Let

$$\mathbf{w} = \max(\mathbf{x}, \mathbf{y}) \qquad \mathbf{z} = \min(\mathbf{x}, \mathbf{y})$$

Find the p.d.f. of (*a*) $\mathbf{r} = \mathbf{w} - \mathbf{z}$, (*b*) $\mathbf{s} = \mathbf{w} + \mathbf{z}$.

6-27 Let x and y be independent identically distributed exponential random variables with common parameter λ. Find the p.d.f.s of (*a*) $\mathbf{z} = \mathbf{y}/\max(\mathbf{x}, \mathbf{y})$. (*b*) $\mathbf{w} = \mathbf{x}/\min(\mathbf{x}, 2\mathbf{y})$.

6-28 If x and y are independent exponential random variables with common parameter λ, show that $\mathbf{x}/(\mathbf{x}+\mathbf{y})$ is a uniformly distributed random variable in $(0, 1)$.

6-29 x and y are independent exponential random variables with common parameter λ. Show that

$$\mathbf{z} = \min(\mathbf{x}, \mathbf{y}) \quad \text{and} \quad \mathbf{w} = \max(\mathbf{x}, \mathbf{y}) - \min(\mathbf{x}, \mathbf{y})$$

are independent random variables.

6-30 Let x and y be independent random variables with common p.d.f. $f_x(x) = \beta^{-\alpha}\alpha x^{\alpha-1}$, $0 < x < \beta$, and zero otherwise ($\alpha \geq 1$). Let $\mathbf{z} = \min(\mathbf{x}, \mathbf{y})$ and $\mathbf{w} = \max(\mathbf{x}, \mathbf{y})$. (*a*) Find the p.d.f. of $\mathbf{x}+\mathbf{y}$. (*b*) Find the joint p.d.f. of z and w. (*c*) Show that $\mathbf{z}/\mathbf{w}$ and w are independent random variables.

6-31 Let x and y be independent gamma random variables with parameters (α_1, β) and (α_2, β), respectively. (*a*) Determine the p.d.f.s of the random variables $\mathbf{x}+\mathbf{y}$, $\mathbf{x}/\mathbf{y}$, and $\mathbf{x}/(\mathbf{x}+\mathbf{y})$. (*b*) Show that $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}/\mathbf{y}$ are independent random variables. (*c*) Show that $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}/(\mathbf{x}+\mathbf{y})$ are independent gamma and beta random variables, respectively. The converse to (*b*) due to Lukacs is also true. It states that with x and y representing nonnegative random variables, if $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}/\mathbf{y}$ are independent, then x and y are gamma random variables with common (second) parameter β.

6-32 Let x and y be independent normal random variables with zero mean and unit variances. (*a*) Find the p.d.f. of $\mathbf{x}/|\mathbf{y}|$ as well as that of $|\mathbf{x}|/|\mathbf{y}|$. (*b*) Let $\mathbf{u} = \mathbf{x}+\mathbf{y}$ and $\mathbf{v} = \mathbf{x}^2+\mathbf{y}^2$. Are u and v independent?

6-33 Let x and y be jointly normal random variables with parameters μ_x, μ_y, σ_x^2, σ_y^2, and r. Find a necessary and sufficient condition for $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ to be independent.

6-34 x and y are independent and identically distributed normal random variables with zero mean and variance σ^2. Define

$$\mathbf{u} = \frac{\mathbf{x}^2-\mathbf{y}^2}{\sqrt{\mathbf{x}^2+\mathbf{y}^2}} \qquad \mathbf{v} = \frac{\mathbf{x}\mathbf{y}}{\sqrt{\mathbf{x}^2+\mathbf{y}^2}}$$

(*a*) Find the joint p.d.f. $f_{uv}(u, v)$ of the random variables u and v. (*b*) Show that u and v are independent normal random variables. (*c*) Show that $[(\mathbf{x}-\mathbf{y})^2 - 2\mathbf{y}^2]/\sqrt{\mathbf{x}^2+\mathbf{y}^2}$ is also a normal random variable. Thus nonlinear functions of normal random variables *can* lead to normal random variables! (This result is due to Shepp.)

6-35 Suppose z has an F distribution with (m, n) degrees of freedom. (*a*) Show that $1/\mathbf{z}$ also has an F distribution with (n, m) degrees of freedom. (*b*) Show that $m\mathbf{z}/(m\mathbf{z}+n)$ has a beta distribution.

6-36 Let the joint p.d.f. of x and y be given by

$$f_{xy}(x, y) = \begin{cases} e^{-x} & 0 < y \leq x \leq \infty \\ 0 & \text{otherwise} \end{cases}$$

Define $\mathbf{z} = \mathbf{x}+\mathbf{y}$, $\mathbf{w} = \mathbf{x}-\mathbf{y}$. Find the joint p.d.f. of z and w. Show that z is an exponential random variable.

6-37 Let

$$f_{xy}(x, y) = \begin{cases} 2e^{-(x+y)} & 0 < x < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Define $\mathbf{z} = \mathbf{x}+\mathbf{y}$, $\mathbf{w} = \mathbf{y}/\mathbf{x}$. Determine the joint p.d.f. of z and w. Are z and w independent random variables?

6-38 The random variables x and θ are independent and θ is uniform in the interval $(-\pi, \pi)$. Show that if $z = x\cos(wt + \theta)$, then

$$f_z(z) = \frac{1}{\pi}\int_{-\infty}^{-|z|} \frac{f_x(y)}{\sqrt{y^2 - z^2}}\,dy + \frac{1}{\pi}\int_{|z|}^{\infty} \frac{f_x(y)}{\sqrt{y^2 - z^2}}\,dy$$

6-39 The random variables x and y are independent, x is $N(0, \sigma^2)$, and y is uniform in the interval $(0, \pi)$. Show that if $z = x + a\cos y$, then

$$f_z(z) = \frac{1}{\pi\sigma\sqrt{2\pi}}\int_0^{\pi} e^{-(z-a\cos y)^2/2\sigma^2}\,dy$$

6-40 The random variables x and y are of discrete type, independent, with $P\{x = n\} = a_n$, $P\{y = n\} = b_n$, $n = 0, 1, \ldots$. Show that, if $z = x + y$, then

$$P\{z = n\} = \sum_{k=0}^{n} a_k b_{n-k}, \qquad n = 0, 1, 2, \ldots$$

6-41 The random variable x is of discrete type taking the values x_n with $P\{x = x_n\} = p_n$ and the random variable y is of continuous type and independent of x. Show that if $z = x + y$ and $w = xy$, then

$$f_z(z) = \sum_n f_y(z - x_n)p_n \qquad f_w(w) = \sum_n \frac{1}{|x_n|} f_y\left(\frac{w}{x_n}\right) p_n$$

6-42 x and y are independent random variables with geometric p.m.f.

$$P\{x = k\} = pq^k \quad k = 0, 1, 2, \ldots \qquad P\{y = m\} = pq^m \quad m = 0, 1, 2, \ldots$$

Find the p.m.f. of (*a*) $x + y$ and (*b*) $x - y$.

6-43 Let x and y be independent identically distributed nonnegative discrete random variables with

$$P\{x = k\} = P\{y = k\} = p_k \qquad k = 0, 1, 2, \ldots$$

Suppose

$$P\{x = k \mid x + y = k\} = P\{x = k - 1 \mid x + y = k\} = \frac{1}{k+1} \qquad k \geq 0$$

Show that x and y are geometric random variables. (This result is due to Chatterji.)

6-44 x and y are independent, identically distributed binomial random variables with parameters n and p. Show that $z = x + y$ is also a binomial random variable. Find its parameters.

6-45 Let x and y be independent random variables with common p.m.f.

$$P(x = k) = pq^k \qquad k = 0, 1, 2, \ldots \qquad q = p - 1$$

(*a*) Show that $\min(x, y)$ and $x - y$ are independent random variables. (*b*) Show that $z = \min(x, y)$ and $w = \max(x, y) - \min(x, y)$ are independent random variables.

6-46 Let x and y be independent Poisson random variables with parameters λ_1 and λ_2, respectively. Show that the conditional density function of x given $x + y$ is binomial.

6-47 The random variables x_1 and x_2 are jointly normal with zero mean. Show that their density can be written in the form

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\Delta}}\exp\left\{-\frac{1}{2}XC^{-1}X^t\right\} \qquad C = \begin{bmatrix}\mu_{11} & \mu_{12}\\ \mu_{21} & \mu_{22}\end{bmatrix}$$

where $X: [x_1, x_2]$, $\mu_{ij} = E\{x_i x_j\}$, and $\Delta = \mu_{11}\mu_{22} - \mu_{12}^2$.

6-48 Show that if the random variables x and y are normal and independent, then

$$P\{\mathbf{xy} < 0\} = G\left(\frac{\eta_x}{\sigma_x}\right) + G\left(\frac{\eta_y}{\sigma_y}\right) - 2G\left(\frac{\eta_x}{\sigma_x}\right)G\left(\frac{\eta_y}{\sigma_y}\right)$$

6-49 The random variables x and y are $N(0; \sigma^2)$ and independent. Show that if $\mathbf{z} = |\mathbf{x} - \mathbf{y}|$, then $E\{\mathbf{z}\} = 2\sigma/\sqrt{\pi}$, $E\{\mathbf{z}^2\} = 2\sigma^2$.

6-50 Show that if x and y are two independent exponential random variables with $f_x(x) = e^{-x}U(x)$, $f_y(y) = e^{-y}U(y)$, and $\mathbf{z} = (\mathbf{x} - \mathbf{y})U(\mathbf{x} - \mathbf{y})$, then $E\{\mathbf{z}\} = 1/2$.

6-51 Show that for any x, y real or complex (*a*) $|E\{\mathbf{xy}\}|^2 \le E\{|\mathbf{x}|^2\}E\{|\mathbf{y}|^2\}$; (*b*) (*triangle inequality*) $\sqrt{E\{|\mathbf{x}+\mathbf{y}|^2\}} \le \sqrt{E\{|\mathbf{x}|^2\}} + \sqrt{E\{|\mathbf{y}|^2\}}$.

6-52 Show that, if the correlation coefficient $r_{xy} = 1$, then $\mathbf{y} = a\mathbf{x} + b$.

6-53 Show that, if $E\{\mathbf{x}^2\} = E\{\mathbf{y}^2\} = E\{\mathbf{xy}\}$, then $\mathbf{x} = \mathbf{y}$.

6-54 The random variable n is Poisson with parameter λ and the random variable x is independent of n. Show that, if $\mathbf{z} = \mathbf{nx}$ and

$$f_x(x) = \frac{\alpha}{\pi(\alpha^2 + x^2)} \qquad \text{then} \quad \Phi_z(\omega) = \exp\{\lambda e^{-\alpha|\omega|} - \lambda\}$$

6-55 Let x represent the number of successes and y the number of failures of n independent Bernoulli trials with p representing the probability of success in any one trial. Find the distribution of $\mathbf{z} = \mathbf{x} - \mathbf{y}$. Show that $E\{\mathbf{z}\} = n(2p - 1)$, $\text{Var}\{\mathbf{z}\} = 4np(1 - p)$.

6-56 x and y are zero mean independent random variables with variances σ_1^2 and σ_2^2, respectively, that is, $\mathbf{x} \sim N(0, \sigma_1^2)$, $\mathbf{y} \sim N(0, \sigma_2^2)$. Let

$$\mathbf{z} = a\mathbf{x} + b\mathbf{y} + c \qquad c \neq 0$$

(*a*) Find the characteristic function $\Phi_z(u)$ of z. (*b*) Using $\Phi_z(u)$ conclude that z is also a normal random variable. (*c*) Find the mean and variance of z.

6-57 Suppose the conditional distribution of x given $\mathbf{y} = n$ is binomial with parameters n and p_1. Further, y is a binomial random variable with parameters M and p_2. Show that the distribution of x is also binomial. Find its parameters.

6-58 The random variables x and y are jointly distributed over the region $0 < x < y < 1$ as

$$f_{xy}(x, y) = \begin{cases} kx & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

for some k. Determine k. Find the variances of x and y. What is the covariance between x and y?

6-59 x is a Poisson random variable with parameter λ and y is a normal random variable with mean μ and variance σ^2. Further x and y are given to be independent. (*a*) Find the joint characteristic function of x and y. (*b*) Define $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Find the characteristic function of z.

6-60 x and y are independent exponential random variables with common parameter λ. Find (*a*) $E[\min(\mathbf{x}, \mathbf{y})]$, (*b*) $E[\max(2\mathbf{x}, \mathbf{y})]$.

6-61 The joint p.d.f. of x and y is given by

$$f_{xy}(x, y) = \begin{cases} 6x & x > 0, y > 0, 0 < x + y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Define $\mathbf{z} = \mathbf{x} - \mathbf{y}$. (*a*) Find the p.d.f. of z. (*b*) Find the conditional p.d.f. of y given x. (*c*) Determine $\text{Var}\{\mathbf{x} + \mathbf{y}\}$.

6-62 Suppose x represents the inverse of a chi-square random variable with one degree of freedom, and the conditional p.d.f. of y given x is $N(0, x)$. Show that y has a Cauchy distribution.

6-63 For any two random variables x and y, let $\sigma_x^2 = \text{Var}\{\mathbf{x}\}, \sigma_y^2 = \text{Var}\{\mathbf{y}\}$ and $\sigma_{x+y}^2 = \text{Var}\{\mathbf{x}+\mathbf{y}\}$.
(*a*) Show that

$$\frac{\sigma_{x+y}}{\sigma_x + \sigma_y} \le 1$$

(*b*) More generally, show that for $p \ge 1$

$$\frac{\{E(|\mathbf{x}+\mathbf{y}|^p)\}^{1/p}}{\{E(|\mathbf{x}|^p)\}^{1/p} + \{E(|\mathbf{y}|^p)\}^{1/p}} \le 1$$

6-64 x and y are jointly normal with parameters $N(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho_{xy})$. Find (*a*) $E\{\mathbf{y} \mid \mathbf{x} = x\}$, and (*b*) $E\{\mathbf{x}^2 \mid \mathbf{y} = y\}$.

6-65 For any two random variables x and y with $E\{\mathbf{x}^2\} < \infty$, show that (*a*) $\text{Var}\{\mathbf{x}\} \ge E[\text{Var}\{\mathbf{x} \mid \mathbf{y}\}]$. (*b*) $\text{Var}\{\mathbf{x}\} = \text{Var}[E\{\mathbf{x} \mid \mathbf{y}\}] + E[\text{Var}\{\mathbf{x} \mid \mathbf{y}\}]$.

6-66 Let x and y be independent random variables with variances σ_1^2 and σ_2^2, respectively. Consider the sum

$$\mathbf{z} = a\mathbf{x} + (1-a)\mathbf{y} \qquad 0 \le a \le 1$$

Find a that minimizes the variance of z.

6-67 Show that, if the random variable x is of discrete type taking the values x_n with $P\{\mathbf{x} = x_n\} = p_n$ and $\mathbf{z} = g(\mathbf{x}, \mathbf{y})$, then

$$E\{\mathbf{z}\} = \sum_n E\{g(x_n, \mathbf{y})\}p_n \qquad f_z(z) = \sum_n f_z(z \mid x_n)p_n$$

6-68 Show that, if the random variables x and y are $N(0, 0, \sigma^2, \sigma^2, r)$, then

(*a*)
$$E\{f_y(\mathbf{y} \mid x)\} = \frac{1}{\sigma\sqrt{2\pi(2-r^2)}} \exp\left\{-\frac{r^2x^2}{2\sigma^2(2-r^2)}\right\}$$

(*b*)
$$E\{f_x(\mathbf{x})f_y(\mathbf{y})\} = \frac{1}{2\pi\sigma^2\sqrt{4-r^2}}$$

6-69 Show that if the random variables x and y are $N(0, 0, \sigma_1^2, \sigma_2^2, r)$ then

$$E\{|\mathbf{xy}|\} = \frac{2}{\pi}\int_0^c \arcsin\frac{\mu}{\sigma_1\sigma_2}\,d\mu + \frac{2\sigma_1\sigma_2}{\pi} = \frac{2\sigma_1\sigma_2}{\pi}(\cos\alpha + \alpha\sin\alpha)$$

where $r = \sin\alpha$ and $C = r\sigma_1\sigma_2$.
(*Hint:* Use (6-200) with $g(x, y) = |xy|$.)

6-70 The random variables x and y are $N(3, 4, 1, 4, 0.5)$. Find $f(y \mid x)$ and $f(x \mid y)$.

6-71 The random variables x and y are uniform in the interval $(-1, 1)$ and independent. Find the conditional density $f_r(r \mid M)$ of the random variable $\mathbf{r} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$, where $M = \{\mathbf{r} \le 1\}$.

6-72 Show that, if the random variables x and y are independent and $\mathbf{z} = \mathbf{x} + \mathbf{y}$, then $f_z(z \mid x) = f_y(z - x)$.

6-73 Show that, for any x and y, the random variables $\mathbf{z} = F_x(\mathbf{x})$ and $\mathbf{w} = F_y(\mathbf{y} \mid \mathbf{x})$ are independent and each is uniform in the interval (0, 1).

6-74 We have a pile of m coins. The probability of heads of the ith coin equals p_i. We select at random one of the coins, we toss it n times and heads shows k times. Show that the probability that we selected the rth coin equals

$$\frac{p_r^k(1-p_r)^{n-k}}{p_1^k(1-p_1)^{n-k} + \cdots + p_m^k(1-p_m)^{n-k}}$$

6-75 The random variable x has a Student t distribution $t(n)$. Show that $E\{\mathbf{x}^2\} = n/(n-2)$.

6-76 Show that if $\beta_x(t) = f_x(t \mid \mathbf{x} > t)$, $\beta_y(t \mid \mathbf{y} > t)$ and $\beta_x(t) = k\beta_y(t)$, then $1 - F_x(x)$ $[1 - F_y(x)]^k$.

6-77 Show that, for any $\mathbf{x}$, $\mathbf{y}$, and $\varepsilon > 0$,

$$P\{|\mathbf{x} - \mathbf{y}| > \varepsilon\} \leq \frac{1}{\varepsilon^2} E\{|\mathbf{x} - \mathbf{y}|^2\}$$

6-78 Show that the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent iff for any a and b:

$$E\{U(a - \mathbf{x})U(b - \mathbf{y})\} = E\{U(a - \mathbf{x})\}E\{U(b - \mathbf{y})\}$$

6-79 Show that

$$E\{\mathbf{y} \mid \mathbf{x} \leq 0\} = \frac{1}{F_x(0)} \int_{-\infty}^{0} E\{\mathbf{y} \mid \mathbf{x}\} f_x(x)\, dx$$