CHAPTER 5

FUNCTIONS OF ONE RANDOM VARIABLE

5-1 THE RANDOM VARIABLE $g(\mathbf{x})$

Suppose that $\mathbf{x}$ is a random variable and $g(x)$ is a function of the real variable x. The expression

$$\mathbf{y} = g(\mathbf{x})$$

is a new random variable defined as follows: For a given ζ, $\mathbf{x}(\zeta)$ is a number and $g[\mathbf{x}(\zeta)]$ is another number specified in terms of $\mathbf{x}(\zeta)$ and $g(x)$. This number is the value $\mathbf{y}(\zeta) = g[\mathbf{x}(\zeta)]$ assigned to the random variable $\mathbf{y}$. Thus a function of a random variable $\mathbf{x}$ is a composite function $\mathbf{y} = g(\mathbf{x}) = g[\mathbf{x}(\zeta)]$ with domain the set S of experimental outcomes.

The distribution function $F_y(y)$ of the random variable so formed is the probability of the event $\{\mathbf{y} \le y\}$ consisting of all outcomes ζ such that $\mathbf{y}(\zeta) = g[\mathbf{x}(\zeta)] \le y$. Thus

$$F_y(y) = P\{\mathbf{y} \le y\} = P\{g(\mathbf{x}) \le y\} \tag{5-1}$$

For a specific y, the values of x such that $g(x) \le y$ form a set on the x axis denoted by R_y. Clearly, $g[\mathbf{x}(\zeta)] \le y$ if $\mathbf{x}(\zeta)$ is a number in the set R_y. Hence

$$F_y(y) = P\{\mathbf{x} \in R_y\} \tag{5-2}$$

This discussion leads to the conclusion that for $g(\mathbf{x})$ to be a random variable, the function $g(x)$ must have these properties:

1. Its domain must include the range of the random variable $\mathbf{x}$.
2. It must be a *Borel* function, that is, for every y, the set R_y such that $g(x) \le y$ must consist of the union and intersection of a countable number of intervals. Only then $\{\mathbf{y} \le y\}$ is an event.
3. The events $\{g(\mathbf{x}) = \pm\infty\}$ must have zero probability.

5-2 THE DISTRIBUTION OF $g(\mathbf{x})$

We shall express the distribution function $F_y(y)$ of the random variable $\mathbf{y} = g(\mathbf{x})$ in terms of the distribution function $F_x(x)$ of the random variable $\mathbf{x}$ and the function $g(x)$. For this purpose, we must determine the set R_y of the x axis such that $g(x) \le y$, and the probability that $\mathbf{x}$ is in this set. The method will be illustrated with several examples. Unless otherwise stated, it will be assumed that $F_x(x)$ is continuous.

1. We start with the function $g(x)$ in Fig. 5-1. As we see from the figure, $g(x)$ is between a and b for any x. This leads to the conclusion that if $y \ge b$, then $g(x) \le y$ for every x, hence $P\{\mathbf{y} \le y\} = 1$; if $y < a$, then there is no x such that $g(x) \le y$, hence $P\{\mathbf{y} \le y\} = 0$. Thus

$$F_y(y) = \begin{cases} 1 & y \ge b \\ 0 & y < a \end{cases}$$

With x_1 and $y_1 = g(x_1)$ as shown, we observe that $g(x) \le y_1$ for $x \le x_1$. Hence

$$F_y(y_1) = P\{\mathbf{x} \le x_1\} = F_x(x_1)$$

We finally note that

$$g(x) \le y_2 \qquad \text{if} \quad x \le x_2' \text{ or if } x_2'' \le x \le x_2'''$$

Hence

$$F_y(y_2) = P\{\mathbf{x} \le x_2'\} + P\{x_2'' \le \mathbf{x} \le x_2'''\} = F_x(x_2') + F_x(x_2''') - F_x(x_2'')$$

because the events $\{\mathbf{x} \le x_2'\}$ and $\{x_2'' \le \mathbf{x} \le x_2'''\}$ are mutually exclusive.

EXAMPLE 5-1 ▶

$$\mathbf{y} = a\mathbf{x} + b \tag{5-3}$$

To find $F_y(y)$, we must find the values of x such that $ax + b \le y$.
(*a*) If $a > 0$, then $ax + b \le y$ for $x \le (y - b)/a$ (Fig. 5-2*a*). Hence

$$F_y(y) = P\left\{\mathbf{x} \le \frac{y-b}{a}\right\} = F_x\left(\frac{y-b}{a}\right) \qquad a > 0$$

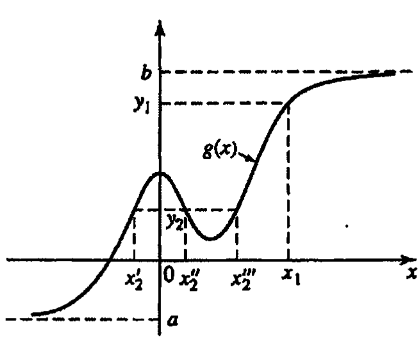


FIGURE 5-1

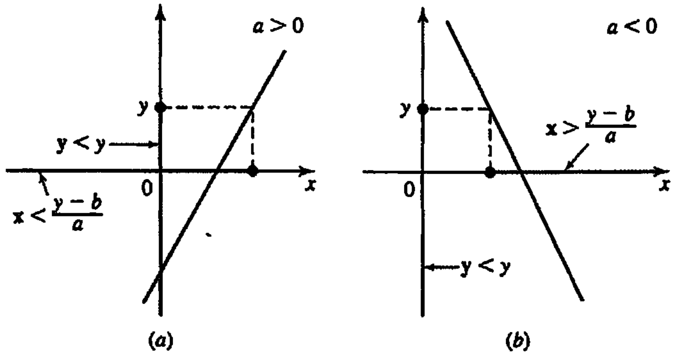


FIGURE 5-2

(b) If $a < 0$, then $ax + b \le y$ for $x > (y - b)/a$ (Fig. 5-2b). Hence [see also (5-17)–(5-18)]

$$F_y(y) = P\left\{\mathbf{x} \ge \frac{y-b}{a}\right\} = 1 - F_x\left(\frac{y-b}{a}\right) \qquad a < 0$$ ◀

EXAMPLE 5-2 ▶

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

If $y \ge 0$, then $x^2 \le y$ for $-\sqrt{y} \le x \le \sqrt{y}$ (Fig. 5-3a). Hence

$$F_y(y) = P\{-\sqrt{y} \le \mathbf{x} \le \sqrt{y}\} = F_x(\sqrt{y}) - F_x(-\sqrt{y}) \qquad y > 0$$

If $y < 0$, then there are no values of x such that $x^2 < y$. Hence

$$F_y(y) = P\{\emptyset\} = 0 \qquad y < 0$$

By direct differentiation of $F_y(y)$, we get

$$f_y(y) = \begin{cases} \dfrac{1}{2\sqrt{y}}\left(f_x(\sqrt{y}) + f_x(-\sqrt{y})\right) & y > 0 \\ 0 & \text{otherwise} \end{cases} \tag{5-4}$$

If $f_x(x)$ represents an even function, then (5-4) reduces to

$$f_y(y) = \frac{1}{\sqrt{y}} f_x(\sqrt{y}) U(y) \tag{5-5}$$

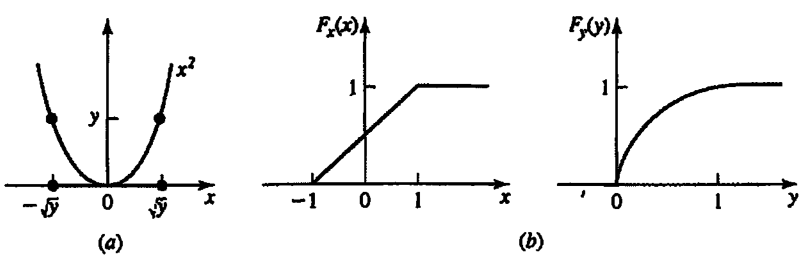


FIGURE 5-3

In particular if $\mathbf{x} \sim N(0, 1)$, so that

$$f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{5-6}$$

and substituting this into (5-5), we obtain the p.d.f. of $\mathbf{y} = \mathbf{x}^2$ to be

$$f_y(y) = \frac{1}{\sqrt{2\pi y}} e^{-y/2} U(y) \tag{5-7}$$

On comparing this with (4-39), we notice that (5-7) represents a chi-square random variable with $n = 1$, since $\Gamma(1/2) = \sqrt{\pi}$. Thus, if $\mathbf{x}$ is a Gaussian random variable with $\mu = 0$, then $\mathbf{y} = \mathbf{x}^2$ represents a chi-square random variable with one degree of freedom. ◀

Special case If $\mathbf{x}$ is uniform in the interval $(-1, 1)$, then

$$F_x(x) = \frac{1}{2} + \frac{x}{2} \qquad |x| < 1$$

(Fig. 5-3*b*). Hence

$$F_y(y) = \sqrt{y} \quad \text{for} \quad 0 \le y \le 1 \quad \text{and} \quad F_y(y) = \begin{cases} 1 & y > 1 \\ 0 & y < 0 \end{cases}$$

2. Suppose now that the function $g(x)$ is constant in an interval (x_0, x_1):

$$g(x) = y_1 \qquad x_0 < x \le x_1 \tag{5-8}$$

In this case

$$P\{\mathbf{y} = y_1\} = P\{x_0 < \mathbf{x} \le x_1\} = F_x(x_1) - F_x(x_0) \tag{5-9}$$

Hence $F_y(y)$ is discontinuous at $\mathbf{y} = y_1$ and its discontinuity equals $F_x(x_1) - F_x(x_0)$.

EXAMPLE 5-3

▶ Consider the function (Fig. 5-4)

$$g(x) = 0 \quad \text{for} \quad -c \le x \le c \quad \text{and} \quad g(x) = \begin{cases} x - c & x > c \\ x + c & x < -c \end{cases} \tag{5-10}$$

In this case, $F_y(y)$ is discontinuous for $y = 0$ and its discontinuity equals $F_x(c) - F_x(-c)$. Furthermore,

$$\begin{array}{lll} \text{If } y \ge 0 & \text{then} & P\{\mathbf{y} \le y\} = P\{\mathbf{x} \le y + c\} = F_x(y + c) \\ \text{If } y < 0 & \text{then} & P\{\mathbf{y} \le y\} = P\{\mathbf{x} \le y - c\} = F_x(y - c) \end{array}$$ ◀

EXAMPLE 5-4

LIMITER

▶ The curve $g(x)$ of Fig. 5-5 is constant for $x \le -b$ and $x \ge b$ and in the interval $(-b, b)$ it is a straight line. With $\mathbf{y} = g(\mathbf{x})$, it follows that $F_y(y)$ is discontinuous for $y = g(-b) = -b$ and $y = g(b) = b$, respectively. Furthermore,

$$\begin{array}{llllll} \text{If} & y \ge b & \text{then} & g(x) \le y & \text{for every } x; & \text{hence} \quad F_y(y) = 1 \\ \text{If} & -b \le y < b & \text{then} & g(x) \le y & \text{for } x \le y; & \text{hence} \quad F_y(y) = F_x(y) \\ \text{If} & y < -b & \text{then} & g(x) \le y & \text{for no } x; & \text{hence} \quad F_y(y) = 0 \end{array}$$ ◀

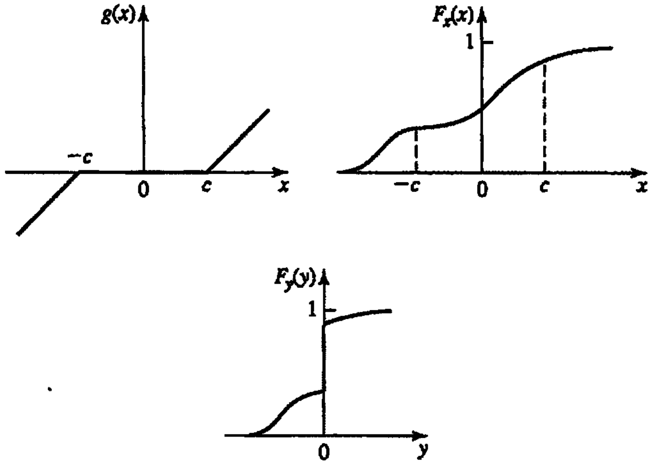


FIGURE 5-4

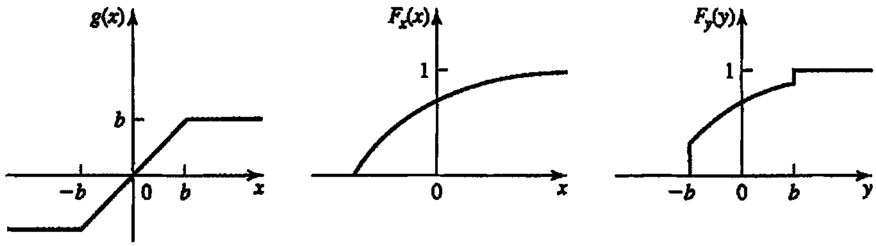


FIGURE 5-5

3. We assume next that $g(x)$ is a staircase function

$$g(x) = g(x_i) = y_i \qquad x_{i-1} < x \leq x_i$$

In this case, the random variable $\mathbf{y} = g(\mathbf{x})$ is of discrete type taking the values y_i with

$$P\{\mathbf{y} = y_i\} = P\{x_{i-1} < \mathbf{x} \leq x_i\} = F_x(x_i) - F_x(x_{i-1})$$

EXAMPLE 5-5

HARD LIMITER

▶ If

$$g(x) = \begin{cases} 1 & x > 0 \\ -1 & x \leq 0 \end{cases} \tag{5-11}$$

then $\mathbf{y}$ takes the values ± 1 with

$$P\{\mathbf{y} = -1\} = P\{\mathbf{x} \leq 0\} = F_x(0)$$

$$P\{\mathbf{y} = 1\} = P\{\mathbf{x} > 0\} = 1 - F_x(0)$$

Hence $F_y(y)$ is a staircase function as in Fig. 5-6. ◀

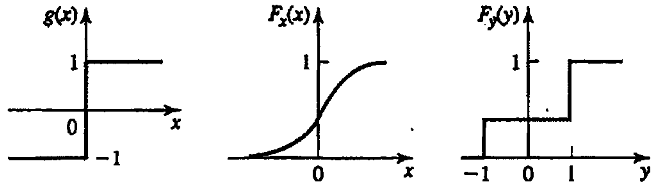


FIGURE 5-6

EXAMPLE 5-6

QUANTIZATION

▶ If

$$g(x) = ns \qquad (n-1)s < x \le ns \tag{5-12}$$

then $\mathbf{y}$ takes the values $y_n = ns$ with

$$P\{\mathbf{y} = ns\} = P\{(n-1)s < \mathbf{x} \le ns\} = F_x(ns) - F_x(ns - s) \tag{5-13}$$

◀

4. We assume, finally, that the function $g(x)$ is discontinuous at $x = x_0$ and such that

$$g(x) < g(x_0^-) \qquad \text{for} \quad x < x_0 \qquad g(x) > g(x_0^+) \qquad \text{for} \quad x > x_0$$

In this case, if y is between $g(x_0^-)$ and $g(x_0^+)$, then $g(x) < y$ for $x \le x_0$. Hence

$$F_y(y) = P\{\mathbf{x} \le x_0\} = F_x(x_0) \qquad g(x_0^-) \le y \le g(x_0^+)$$

EXAMPLE 5-7

▶ Suppose that

$$g(x) = \begin{cases} x + c & x \ge 0 \\ x - c & x < 0 \end{cases} \tag{5-14}$$

is discontinuous (Fig. 5-7). Thus $g(x)$ is discontinuous for $x = 0$ with $g(0^-) = -c$ and

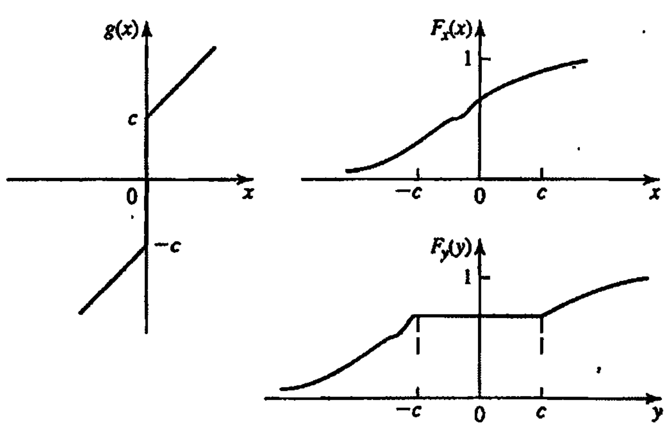


FIGURE 5-7

$g(0^+) = c$. Hence $F_y(y) = F_x(0)$ for $|y| \le c$. Furthermore,

$$
\begin{array}{llllll}
\text{If} & y \ge c & \text{then} & g(x) \le y & \text{for } x \le y - c; & \text{hence} \quad F_y(y) = F_x(y-c) \\
\text{If} & -c \le y \le c & \text{then} & g(x) \le y & \text{for } x \le 0; & \text{hence} \quad F_y(y) = F_x(0) \\
\text{If} & y \le -c & \text{then} & g(x) \le y & \text{for } x \le y + c; & \text{hence} \quad F_y(y) = F_x(y+c)
\end{array}
$$

◀

EXAMPLE 5-8

▶ The function $g(x)$ in Fig. 5-8 equals 0 in the interval $(-c, c)$ and it is discontinuous for $x = \pm c$ with $g(c^+) = c$, $g(c^-) = 0$, $g(-c^-) = -c$, $g(-c^+) = 0$. Hence $F_y(y)$ is discontinuous for $y = 0$ and it is constant for $0 \le y \le c$ and $-c \le y \le 0$. Thus

$$
\begin{array}{llllll}
\text{If} & y \ge c & \text{then} & g(x) \le y & \text{for } x \le y; & \text{hence} \quad F_y(y) = F_x(y) \\
\text{If} & 0 \le y < c & \text{then} & g(x) \le y & \text{for } x < c; & \text{hence} \quad F_y(y) = F_x(c) \\
\text{If} & -c \le y < c & \text{then} & g(x) \le y & \text{for } x \le -c; & \text{hence} \quad F_y(y) = F_x(-c) \\
\text{If} & y < -c & \text{then} & g(x) \le y & \text{for } x \le y; & \text{hence} \quad F_y(y) = F_x(y)
\end{array}
$$

◀

5. We now assume that the random variable $\mathbf{x}$ is of discrete type taking the values x_k with probability p_k. In this case, the random variable $\mathbf{y} = g(\mathbf{x})$ is also of discrete type taking the values $y_k = g(x_k)$.

If $y_k = g(x)$ for only one $x = x_k$, then

$$P\{\mathbf{y} = y_k\} = P\{\mathbf{x} = x_k\} = p_k$$

If, however, $y_k = g(x)$ for $x = x_k$ and $x = x_l$, then

$$P\{\mathbf{y} = y_k\} = P\{\mathbf{x} = x_k\} + P\{\mathbf{x} = x_l\} = p_k + p_l$$

EXAMPLE 5-9

▶
$$\mathbf{y} = \mathbf{x}^2$$

(a) If $\mathbf{x}$ takes the values $1, 2, \ldots, 6$ with probability $1/6$, then $\mathbf{y}$ takes the values $1^2, 2^2, \ldots, 6^2$ with probability $1/6$.

(b) If, however, $\mathbf{x}$ takes the values $-2, -1, 0, 1, 2, 3$ with probability $1/6$, then $\mathbf{y}$ takes the values $0, 1, 4, 9$ with probabilities $1/6, 2/6, 2/6, 1/6$, respectively. ◀

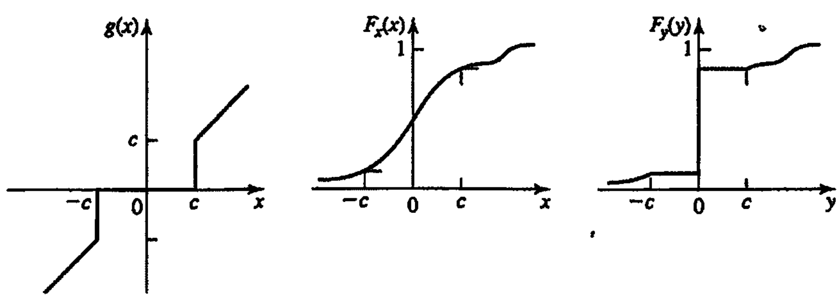


FIGURE 5-8

Determination of $f_y(y)$

We wish to determine the density of $\mathbf{y} = g(\mathbf{x})$ in terms of the density of $\mathbf{x}$. Suppose, first, that the set R of the y axis is not in the range of the function $g(x)$, that is, that $g(x)$ is not a point of R for any x. In this case, the probability that $g(\mathbf{x})$ is in R equals 0. Hence $f_y(y) = 0$ for $y \in R$. It suffices, therefore, to consider the values of y such that for some x, $g(x) = y$.

FUNDAMENTAL THEOREM. To find $f_y(y)$ for a specific y, we solve the equation $y = g(x)$. Denoting its real roots by x_n,

$$y = g(x_1) = \cdots = g(x_n) = \cdots \tag{5-15}$$

we shall show that

$$f_y(y) = \frac{f_x(x_1)}{|g'(x_1)|} + \cdots + \frac{f_x(x_n)}{|g'(x_n)|} + \cdots \tag{5-16}$$

where $g'(x)$ is the derivative of $g(x)$.

Proof. To avoid generalities, we assume that the equation $y = g(x)$ has three roots as in Fig. 5-9. As we know

$$f_y(y)\,dy = P\{y < \mathbf{y} \le y + dy\}$$

It suffices, therefore, to find the set of values x such that $y < g(x) \le y + dy$ and the probability that $\mathbf{x}$ is in this set. As we see from the figure, this set consists of the following three intervals

$$x_1 < x < x_1 + dx_1 \qquad x_2 + dx_2 < x < x_2 \qquad x_3 < x < x_3 + dx_3$$

where $dx_1 > 0$, $dx_3 > 0$ but $dx_2 < 0$. From this it follows that

$$\begin{aligned} P\{y < \mathbf{y} < y + dy\} &= P\{x_1 < \mathbf{x} < x_1 + dx_1\} \\ &\quad + P\{x_2 + dx_2 < \mathbf{x} < x_2\} + P\{x_3 < \mathbf{x} < x_3 + dx_3\} \end{aligned}$$

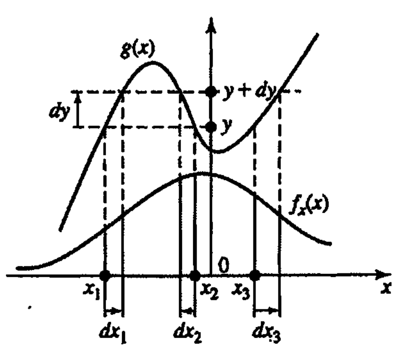


FIGURE 5-9

The right side equals the shaded area in Fig. 5-9. Since

$$P\{x_1 < \mathbf{x} < x_1 + dx_1\} = f_x(x_1)\,dx_1 \qquad dx_1 = dy/g'(x_1)$$

$$P\{x_2 + dx_2 < \mathbf{x} < x_2\} = f_x(x_2)\,|dx_2| \quad dx_2 = dy/g'(x_2)$$

$$P\{x_3 < \mathbf{x} < x_3 + dx_3\} = f_x(x_3)\,dx_3 \qquad dx_3 = dy/g'(x_3)$$

we conclude that

$$f_y(y)\,dy = \frac{f_x(x_1)}{g'(x_1)}\,dy + \frac{f_x(x_2)}{|g'(x_2)|}\,dy + \frac{f_x(x_3)}{g'(x_3)}\,dy$$

and (5-16) results.

We note, finally, that if $g(x) = y_1 =$ constant for every x in the interval (x_0, x_1), then [see (5-9)] $F_y(y)$ is discontinuous for $y = y_1$. Hence $f_y(y)$ contains an impulse $\delta(y - y_1)$ of area $F_x(x_1) - F_x(x_0)$.

Conditional density The conditional density $f_y(y \mid M)$ of the random variable $\mathbf{y} = g(\mathbf{x})$ assuming an event M is given by (5-5) if on the right side we replace the terms $f_x(x_i)$ by $f_x(x_i \mid M)$ (see, for example, Prob. 5-21).

Illustrations

We give next several applications of (5-2) and (5-16).

1. $$\mathbf{y} = a\mathbf{x} + b \qquad g'(x) = a \tag{5-17}$$

The equation $y = ax + b$ has a single solution $x = (y - b)/a$ for every y. Hence

$$f_y(y) = \frac{1}{|a|} f_x\left(\frac{y-b}{a}\right) \tag{5-18}$$

Special case If $\mathbf{x}$ is uniform in the interval (x_1, x_2), then $\mathbf{y}$ is uniform in the interval $(ax_1 + b, ax_2 + b)$.

EXAMPLE 5-10

▶ Suppose that the voltage $\mathbf{v}$ is a random variable given by

$$\mathbf{v} = i(\mathbf{r} + r_0)$$

where $i = 0.01$ A and $r_0 = 1000\,\Omega$. If the resistance $\mathbf{r}$ is a random variable uniform between 900 and 1100 Ω, then $\mathbf{v}$ is uniform between 19 and 21 V. ◀

2. $$\mathbf{y} = \frac{1}{\mathbf{x}} \qquad g'(x) = -\frac{1}{x^2} \tag{5-19}$$

The equation $y = 1/x$ has a single solution $x = 1/y$. Hence

$$f_y(y) = \frac{1}{y^2} f_x\left(\frac{1}{y}\right) \tag{5-20}$$

Cauchy density: If $\mathbf{x}$ has a *Cauchy density* with parameter α,

$$f_x(x) = \frac{\alpha/\pi}{x^2 + \alpha^2} \qquad \text{then} \quad f_y(y) = \frac{1/\alpha\pi}{y^2 + 1/\alpha^2}$$

in (5-19) is also a Cauchy density with parameter $1/\alpha$.

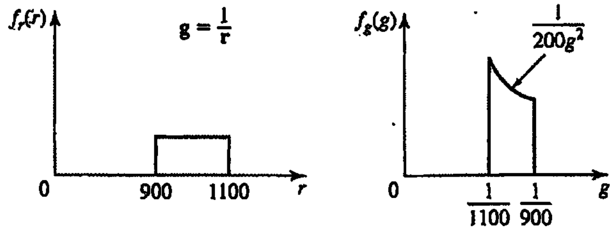


FIGURE 5-10

EXAMPLE 5-11 ▶ Suppose that the resistance **r** is uniform between 900 and 1100 Ω as in Fig. 5-10. We shall determine the density of the corresponding conductance

$$\mathbf{g} = 1/\mathbf{r}$$

Since $f_r(r) = 1/200$ S for r between 900 and 1100 it follows from (5-20) that

$$f_g(g) = \frac{1}{200g^2} \qquad \text{for} \quad \frac{1}{1100} < g < \frac{1}{900}$$

and 0 elsewhere. ◀

3. $$y = a\mathbf{x}^2 \quad a > 0 \quad g'(x) = 2ax \tag{5-21}$$

If $y \le 0$, then the equation $y = ax^2$ has no real solutions; hence $f_y(y) = 0$. If $y > 0$, then it has two solutions

$$x_1 = \sqrt{\frac{y}{a}} \qquad x_2 = -\sqrt{\frac{y}{a}}$$

and (5-16) yields [see also (5-4)]

$$f_y(y) = \frac{1}{2a\sqrt{y/a}}\left[f_x\left(\sqrt{\frac{y}{a}}\right) + f_x\left(-\sqrt{\frac{y}{a}}\right)\right] \qquad y > 0 \tag{5-22}$$

We note that $F_y(y) = 0$ for $y < 0$ and

$$F_y(y) = P\left\{-\sqrt{\frac{y}{a}} \le \mathbf{x} \le \sqrt{\frac{y}{a}}\right\} = F_x\left(\sqrt{\frac{y}{a}}\right) - F_x\left(-\sqrt{\frac{y}{a}}\right) \qquad y > 0$$

EXAMPLE 5-12 ▶ The voltage across a resistor is a random variable **e** uniform between 5 and 10 V. We shall determine the density of the power

$$\mathbf{w} = \frac{\mathbf{e}^2}{r} \qquad r = 1000\,\Omega$$

dissipated in r.

Since $f_e(e) = 1/5$ for e between 5 and 10 and 0 elsewhere, we conclude from (5-8) with $a = 1/r$ that

$$f_w(w) = \sqrt{\frac{10}{w}} \qquad \frac{1}{40} < w < \frac{1}{10}$$

and 0 elsewhere. ◀

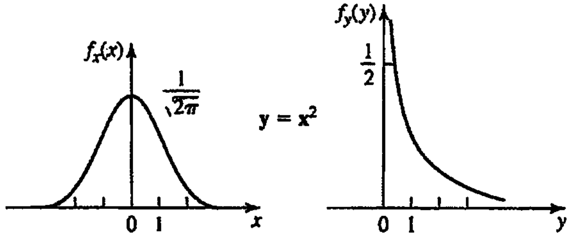


FIGURE 5-11

Special case Suppose that

$$f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad \mathbf{y} = \mathbf{x}^2$$

With $a = 1$, it follows from (5-22) and the evenness of $f_x(x)$ that (Fig. 5-11)

$$f_y(y) = \frac{1}{\sqrt{y}} f_x(\sqrt{y}) = \frac{1}{\sqrt{2\pi y}} e^{-y/2} U(y)$$

We have thus shown that if $\mathbf{x}$ is an $N(0, 1)$ random variable, the random variable $\mathbf{y} = \mathbf{x}^2$ has a chi-square distribution with one degree of freedom [see (4-39) and also (5-7)].

4.
$$\mathbf{y} = \sqrt{\mathbf{x}} \quad g'(x) = \frac{1}{2\sqrt{x}} \tag{5-23}$$

The equation $y = \sqrt{x}$ has a single solution $x = y^2$ for $y > 0$ and no solution for $y < 0$. Hence

$$f_y(y) = 2y f_x(y^2) U(y) \tag{5-24}$$

The chi density Suppose that $\mathbf{x}$ has a chi-square density as in (4-39),

$$f_x(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2} U(y)$$

and $y = \sqrt{x}$. In this case, (5-24) yields

$$f_y(y) = \frac{2}{2^{n/2}\Gamma(n/2)} y^{n-1} e^{-y^2/2} U(y) \tag{5-25}$$

This function is called the **chi density** with n degrees of freedom. The following cases are of special interest.

Maxwell For $n = 3$, (5-25) yields the Maxwell density [see also (4-54)]

$$f_y(y) = \sqrt{2/\pi}\, y^2 e^{-y^2/2}$$

Rayleigh For $n = 2$, we obtain the Rayleigh density $f_y(y) = y e^{-y^2/2} U(y)$.

5.
$$\mathbf{y} = \mathbf{x} U(\mathbf{x}) \qquad g'(x) = U(x) \tag{5-26}$$

Clearly, $f_y(y) = 0$ and $F_y(y) = 0$ for $y < 0$ (Fig. 5-12). If $y > 0$, then the equation $y = xU(x)$ has a single solution $x_1 = y$. Hence

$$f_y(y) = f_x(y) \qquad F_y(y) = F_x(y) \qquad y > 0 \tag{5-27}$$

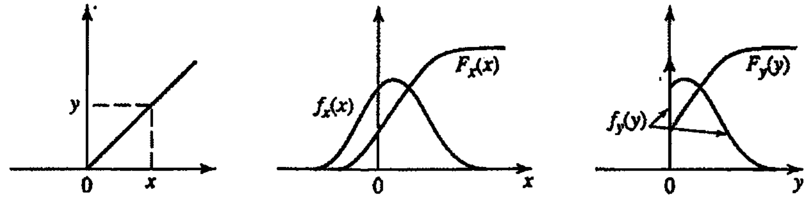


FIGURE 5-12
Half wave rectifier.

Thus $F_y(y)$ is discontinuous at $y=0$ with discontinuity $F_y(0^+) - F_y(0^-) = F_x(0)$. Hence

$$f_y(y) = f_x(y)U(y) + F_x(0)\delta(y)$$

6. $$y = e^x \qquad g'(x) = e^x \tag{5-28}$$

If $y > 0$, then the equation $y = e^x$ has the single solution $x = \ln y$. Hence

$$f_y(y) = \frac{1}{y} f_x(\ln y) \qquad y > 0 \tag{5-29}$$

If $y < 0$, then $f_y(y) = 0$.

lognormal: If x is $N(\eta; \sigma)$, then

$$f_y(y) = \frac{1}{\sigma y \sqrt{2\pi}} e^{-(\ln y - \eta)^2 / 2\sigma^2} \tag{5-30}$$

This density is called *lognormal.*

7. $$y = a\sin(x + \theta) \qquad a > 0 \tag{5-31}$$

If $|y| > a$, then the equation $y = a\sin(x+\theta)$ has no solutions; hence $f_y(y) = 0$. If $|y| < a$, then it has infinitely many solutions (Fig. 5-13*a*)

$$x_n = \arcsin\frac{y}{a} - \theta \qquad n = -\cdots -1, 0, 1, \ldots$$

Since $g'(x_n) = a\cos(x_n + \theta) = \sqrt{a^2 - y^2}$, (5-5) yields

$$f_y(y) = \frac{1}{\sqrt{a^2 - y^2}} \sum_{n=-\infty}^{\infty} f_x(x_n) \qquad |y| < a \tag{5-32}$$

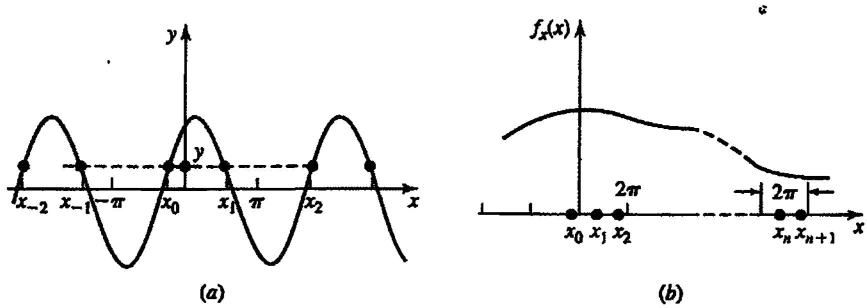


FIGURE 5-13

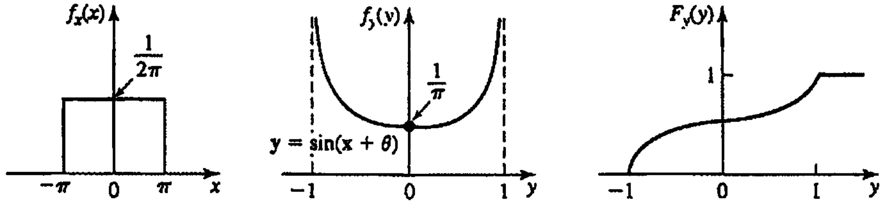


FIGURE 5-14

Special case: Suppose that $\mathbf{x}$ is uniform in the interval $(-\pi, \pi)$. In this case, the equation $y = a\sin(x+\theta)$ has exactly two solutions in the interval $(-\pi, \pi)$ for any θ (Fig. 5-14). The function $f_x(x)$ equals $1/2\pi$ for these two values and it equals 0 for any x_n outside the interval $(-\pi, \pi)$. Retaining the two nonzero terms in (5-32), we obtain

$$f_y(y) = \frac{2}{2\pi\sqrt{a^2-y^2}} = \frac{1}{\pi\sqrt{a^2-y^2}} \qquad |y| < a \tag{5-33}$$

To find $F_y(y)$, we observe that $\mathbf{y} \le y$ if $\mathbf{x}$ is either between $-\pi$ and x_0 or between x_1 and π (Fig. 5-13a). Since the total length of the two intervals equals $\pi + 2x_0 + 2\theta$, we conclude, dividing by 2π, that

$$F_y(y) = \frac{1}{2} + \frac{1}{\pi}\arcsin\frac{y}{a} \qquad |y| < a \tag{5-34}$$

We note that although $f_y(\pm a) = \infty$, the probability that $\mathbf{y} = \pm a$ is 0.

Smooth phase If the density $f_x(x)$ of $\mathbf{x}$ is sufficiently smooth so that it can be approximated by a constant in any interval of length 2π (see Fig. 5-13b), then

$$\pi\sum_{n=-\infty}^{\infty} f_x(x_n) \simeq \int_{-\infty}^{\infty} f_x(x)\,dx = 1$$

because in each interval of length 2π this sum has two terms. Inserting into (5-32), we conclude that the density of $\mathbf{y}$ is given approximately by (5-33).

EXAMPLE 5-13

▶ A particle leaves the origin under the influence of the force of gravity and its initial velocity v forms an angle φ with the horizontal axis. The path of the particle reaches the ground at a distance

$$\mathbf{d} = \frac{v^2}{g}\sin 2\varphi$$

from the origin (Fig. 5-15). Assuming that φ is a random variable uniform between 0 and $\pi/2$, we shall determine: (a) the density of $\mathbf{d}$ and (b) the probability that $\mathbf{d} \le d_0$.

SOLUTION

(a) Clearly,

$$\mathbf{d} = a\sin\mathbf{x} \qquad a = v^2/g$$

where the random variable $\mathbf{x} = 2\varphi$ is uniform between 0 and π. If $0 < d < a$, then the equation $d = a\sin x$ has exactly two solutions in the interval $(0, \pi)$. Reasoning as in

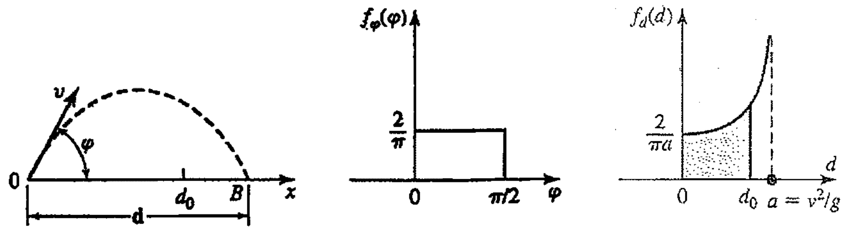


FIGURE 5-15

(5-33), we obtain

$$f_d(d) = \frac{2}{\pi\sqrt{a^2 - d^2}} \qquad 0 < d < a$$

and 0 otherwise.

(*b*) The probability that $\mathbf{d} \le d_0$ equals the shaded area in Fig. 5-15:

$$P\{\mathbf{d} \le d_0\} = F_d(d_0) = \frac{2}{\pi} \arcsin \frac{d_0}{a}$$

◀

8. $$\mathbf{y} = \tan \mathbf{x} \tag{5-35}$$

The equation $y = \tan x$ has infinitely many solutions for any y (Fig. 5-16*a*)

$$x_n = \arctan y \qquad n = \ldots, -1, 0, 1, \ldots$$

Since $g'(x) = 1/\cos^2 x = 1 + y^2$, Eq. (5-16) yields

$$f_y(y) = \frac{1}{1+y^2} \sum_{n=-\infty}^{\infty} f_x(x_n) \tag{5-36}$$

Special case If $\mathbf{x}$ is uniform in the interval $(-\pi/2, \pi/2)$, then the term $f_x(x_1)$ in (5-36) equals $1/\pi$ and all others are 0 (Fig. 5-16*b*). Hence $\mathbf{y}$ has a *Cauchy* density given by

$$f_y(y) = \frac{1/\pi}{1+y^2} \tag{5-37}$$

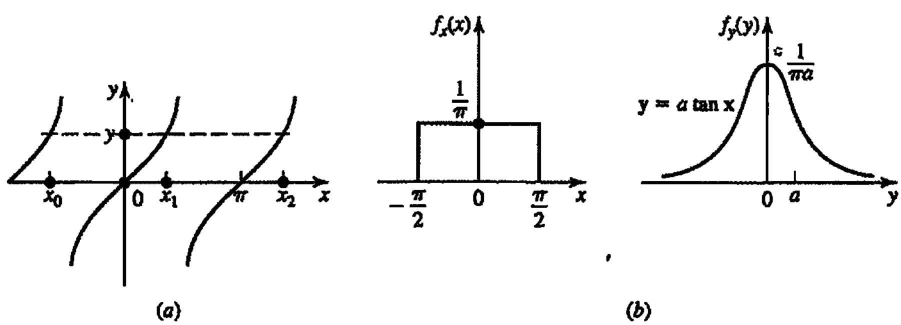


FIGURE 5-16

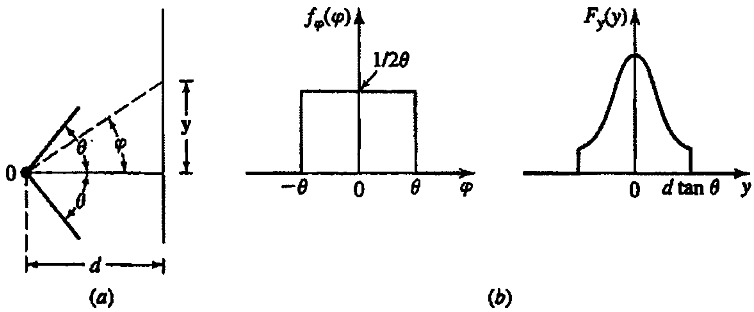


FIGURE 5-17

As we see from the figure, $\mathbf{y} \le y$ if $\mathbf{x}$ is between $-\pi/2$ and x_1. Since the length of this interval equals $x_1 + \pi/2$, we conclude, dividing by π, that

$$F_y(y) = \frac{1}{\pi}\left(x_1 + \frac{\pi}{2}\right) = \frac{1}{2} + \frac{1}{\pi}\text{arc tan } y \tag{5-38}$$

EXAMPLE 5-14

▶ A particle leaves the origin in a free motion as in Fig. 5-17 crossing the vertical line $x = d$ at

$$\mathbf{y} = d \tan \boldsymbol{\varphi}$$

Assuming that the angle $\boldsymbol{\varphi}$ is uniform in the interval $(-\theta, \theta)$, we conclude as in (5-37) that

$$f_y(y) = \frac{d/2\theta}{d^2 + y^2}. \qquad \text{for} \quad |y| < d \tan \theta$$

and 0 otherwise. ◀

EXAMPLE 5-15

▶ Suppose $f_x(x) = 2x/\pi^2$, $0 < x < \pi$, and $\mathbf{y} = \sin \mathbf{x}$. Determine $f_y(y)$.

SOLUTION

Since $\mathbf{x}$ has zero probability of falling outside the interval $(0, \pi)$, $y = \sin x$ has zero probability of falling outside the interval $(0, 1)$ and $f_y(y) = 0$ outside this interval. For any $0 < y < 1$, from Fig. 5.18*b*, the equation $y = \sin x$ has an infinite number of solutions $\ldots, x_1, x_2, x_3, \ldots$, where $x_1 = \sin^{-1} y$ is the principal solution. Moreover, using the symmetry we also get $x_2 = \pi - x_1$ and so on. Further,

$$\frac{dy}{dx} = \cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2}$$

so that

$$\left|\frac{dy}{dx}\right|_{x=x_i} = \sqrt{1 - y^2}$$

Using this in (5-16), we obtain

$$f_y(y) = \sum_{i=-\infty}^{+\infty} \frac{1}{\sqrt{1 - y^2}} f_x(x_i) \qquad 0 < y < 1 \tag{5-39}$$

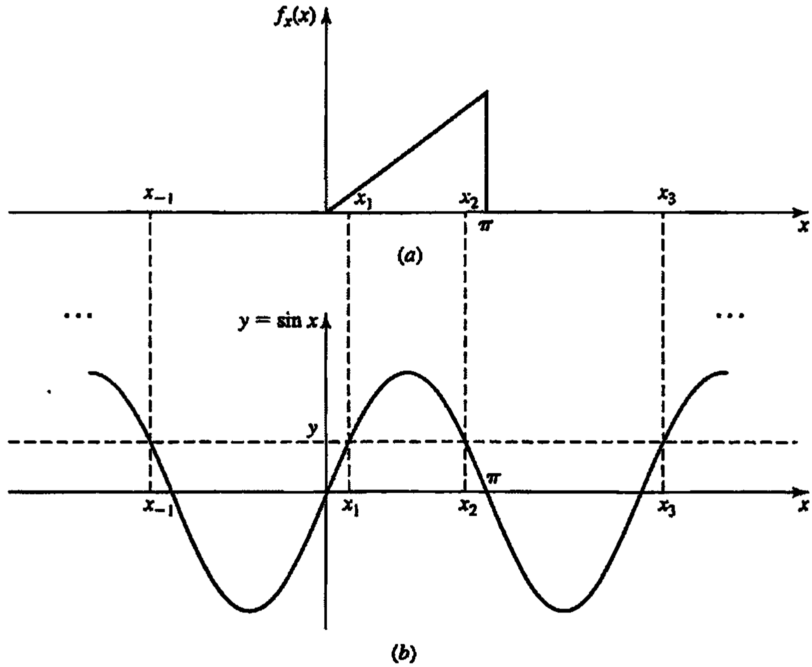


FIGURE 5-18

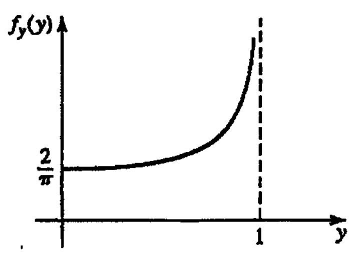


FIGURE 5-19

But from Fig. 5.18a, here $f_x(x_{-1}) = f_x(x_3) = f_x(x_4) = \cdots = 0$ (except for $f_x(x_1)$ and $f_x(x_2)$ the rest are all zeros). Thus (Fig. 5-19)

$$f_y(y) = \frac{1}{\sqrt{1-y^2}}(f_x(x_1) + f_x(x_2)) = \frac{1}{\sqrt{1-y^2}}\left(\frac{2x_1}{\pi^2} + \frac{2x_2}{\pi^2}\right)$$

$$= \frac{2(x_1 + \pi - x_1)}{\pi^2\sqrt{1-y^2}} = \begin{cases} \dfrac{2}{\pi\sqrt{1-y^2}} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases} \tag{5-40}$$

◀

THE INVERSE PROBLEM. In the preceding discussion, we were given a random variable $\mathbf{x}$ with known distribution $F_x(x)$ and a function $g(x)$ and we determined the distribution $F_y(y)$ of the random variable $\mathbf{y} = g(\mathbf{x})$. We consider now the inverse problem: We are given the distribution of $\mathbf{x}$ and we wish to find a function $g(x)$ such that the distribution of the random variable $\mathbf{y} = g(\mathbf{x})$ equals a specified function $F_y(y)$. This topic is developed further in Sec. 7-5. We start with two special cases.

From $F_x(x)$ to a uniform distribution. Given a random variable $\mathbf{x}$ with distribution $F_x(x)$, we wish to find a function $g(x)$ such that the random variable $\mathbf{u} = g(\mathbf{x})$ is uniformly distributed in the interval (0, 1). We maintain that $g(x) = F_x(x)$, that is, if

$$\mathbf{u} = F_x(\mathbf{x}) \quad \text{then } F_u(u) = u \text{ for } 0 \le u \le 1 \tag{5-41}$$

Proof. Suppose that x is an arbitrary number and $u = F_x(x)$. From the monotonicity of $F_x(x)$ it follows that $\mathbf{u} \le u$ iff $\mathbf{x} \le x$. Hence

$$F_u(u) = P\{\mathbf{u} \le u\} = P\{\mathbf{x} \le x\} = F_x(x) = u$$

and (5-41) results.

The random variable $\mathbf{u}$ can be considered as the output of a nonlinear memoryless system (Fig. 5-20) with input $\mathbf{x}$ and transfer characteristic $F_x(x)$. Therefore if we use $\mathbf{u}$ as the input to another system with transfer characteristic the inverse $F_x^{(-1)}(u)$ of the function $u = F_x(x)$, the resulting output will equal $\mathbf{x}$:

$$\text{If} \quad \mathbf{x} = F_x^{(-1)}(\mathbf{u}) \qquad \text{then} \quad P\{\mathbf{x} \le x\} = F_x(x)$$

From uniform to $F_y(y)$. Given a random variable $\mathbf{u}$ with uniform distribution in the interval (0, 1), we wish to find a function $g(u)$ such that the distribution of the random variable $\mathbf{y} = g(\mathbf{u})$ is a specified function $F_y(y)$. We maintain that $g(u)$ is the inverse of the function $u = F_y(y)$:

$$\text{If} \quad y = F_y^{(-1)}(\mathbf{u}) \qquad \text{then} \quad P\{\mathbf{y} \le y\} = F_y(y) \tag{5-42}$$

Proof. The random variable $\mathbf{u}$ in (5-41) is uniform and the function $F_x(x)$ is arbitrary. Replacing $F_x(x)$ by $F_y(y)$, we obtain (5-42) (see also Fig. 5-20).

From $F_x(x)$ to $F_y(y)$. We consider, finally, the general case: Given $F_x(x)$ and $F_y(y)$, find $g(x)$ such that the distribution of $\mathbf{y} = g(\mathbf{x})$ equals $F_y(y)$. To solve this problem, we form the random variable $\mathbf{u} = F_x(\mathbf{x})$ as in (5-41) and the random variable $\mathbf{y} = F^{(-1)}(\mathbf{u})$

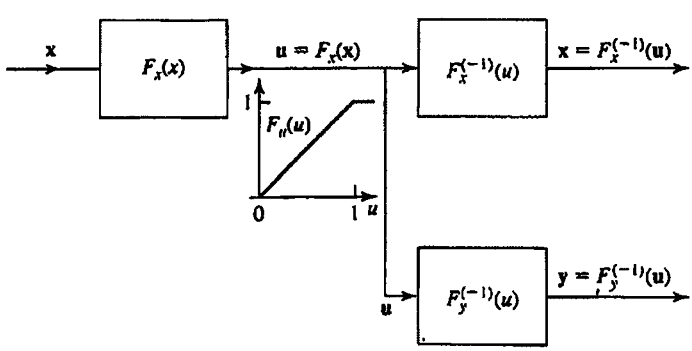


FIGURE 5-20

as in (5-42). Combining the two, we conclude:

$$\text{If}\quad \mathbf{y} = F_y^{(-1)}(F_x(\mathbf{x})) \qquad \text{then}\quad P\{\mathbf{y} \le y\} = F_y(y) \tag{5-43}$$

5-3 MEAN AND VARIANCE

The *expected value* or *mean* of a random variable $\mathbf{x}$ is by definition the integral

$$E\{\mathbf{x}\} = \int_{-\infty}^{\infty} x f(x)\,dx \tag{5-44}$$

This number will also be denoted by η_x or η.

EXAMPLE 5-16

▶ If $\mathbf{x}$ is uniform in the interval (x_1, x_2), then $f(x) = 1/(x_2 - x_1)$ in this interval. Hence

$$E\{\mathbf{x}\} = \frac{1}{x_2 - x_1}\int_{x_1}^{x_2} x\,dx = \frac{x_1 + x_2}{2}$$ ◀

We note that, if the vertical line $x = a$ is an axis of symmetry of $f(x)$ then $E\{\mathbf{x}\} = a$; in particular, if $f(-x) = f(x)$, then $E\{\mathbf{x}\} = 0$. In Example 5-16, $f(x)$ is symmetrical about the line $x = (x_1 + x_2)/2$.

Discrete type For discrete type random variables the integral in (5-44) can be written as a sum. Indeed, suppose that $\mathbf{x}$ takes the values x_i with probability p_i. In this case [see (4-15)]

$$f(x) = \sum_i p_i \delta(x - x_i) \tag{5-45}$$

Inserting into (5-44) and using the identity

$$\int_{-\infty}^{\infty} x\delta(x - x_i)\,dx = x_i$$

we obtain

$$E\{\mathbf{x}\} = \sum_i p_i x_i \qquad p_i = P\{\mathbf{x} = x_i\} \tag{5-46}$$

EXAMPLE 5-17

▶ If $\mathbf{x}$ takes the values $1, 2, \ldots, 6$ with probability $1/6$, then

$$E\{\mathbf{x}\} = \tfrac{1}{6}(1 + 2 + \cdots + 6) = 3.5$$ ◀

Conditional mean The conditional mean of a random variable $\mathbf{x}$ assuming an event M is given by the integral in (5-44) if $f(x)$ is replaced by the conditional density $f(x \mid M)$:

$$E\{\mathbf{x} \mid M\} = \int_{-\infty}^{\infty} x f(x \mid M)\,dx \tag{5-47}$$

For discrete-type random variables (5-47) yields

$$E\{\mathbf{x} \mid M\} = \sum_i x_i P\{\mathbf{x} = x_i \mid M\} \tag{5-48}$$

EXAMPLE 5-18 ▶ With $M = \{x \ge a\}$, it follows from (5-47) that

$$E\{\mathbf{x} \mid \mathbf{x} \ge a\} = \int_{-\infty}^{\infty} x f(x \mid \mathbf{x} \ge a)\, dx = \frac{\int_a^{\infty} x f(x)\, dx}{\int_a^{\infty} f(x)\, dx}$$ ◀

Lebesgue integral. The mean of a random variable can be interpreted as a Lebesgue integral. This interpretation is important in mathematics but it will not be used in our development. We make, therefore, only a passing reference.

We divide the x axis into intervals (x_k, x_{k+1}) of length Δx as in Fig. 5-21a. If Δx is small, then the Riemann integral in (5-44) can be approximated by a sum

$$\int_{-\infty}^{\infty} x f(x)\, dx \simeq \sum_{k=-\infty}^{\infty} x_k f(x_k)\, \Delta x \tag{5-49}$$

And since $f(x_k)\,\Delta x \simeq P\{x_k < \mathbf{x} < x_k + \Delta x\}$, we conclude that

$$E\{\mathbf{x}\} \simeq \sum_{k=-\infty}^{\infty} x_k P\{x_k < \mathbf{x} < x_k + \Delta x\}$$

Here, the sets $\{x_k < \mathbf{x} < x_k + \Delta x\}$ are differential events specified in terms of the random variable $\mathbf{x}$, and their union is the space S (Fig. 5-21b). Hence, to find $E\{\mathbf{x}\}$, we multiply the probability of each differential event by the corresponding value of $\mathbf{x}$ and sum over all k. The resulting limit as $\Delta x \to 0$ is written in the form

$$E\{\mathbf{x}\} = \int_s \mathbf{x}\, dP \tag{5-50}$$

and is called the *Lebesgue integral* of $\mathbf{x}$.

Frequency interpretation We maintain that the arithmetic average $\bar{x}$ of the observed values x_i of $\mathbf{x}$ tends to the integral in (5-44) as $n \to \infty$:

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} \to E\{\mathbf{x}\} \tag{5-51}$$

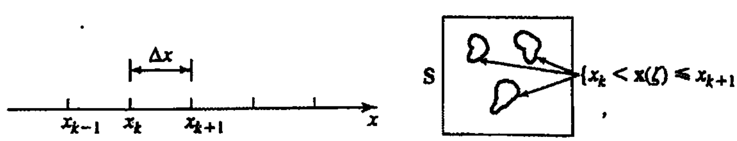
Proof. We denote by Δn_k the number of x_i's that are between z_k and $z_k + \Delta x = z_{k+1}$. From this it follows that

$$x_1 + \cdots + x_n \simeq \sum z_k\, \Delta n_k$$

and since $f(z_k)\,\Delta x \simeq \Delta n_k/n$ [see (4-21)] we conclude that

$$\bar{x} \simeq \frac{1}{n} \sum z_k\, \Delta n_k = \sum z_k f(z_k)\, \Delta x \simeq \int_{-\infty}^{\infty} x f(x)\, dx$$

and (5-51) results.



(a) (b)

FIGURE 5-21

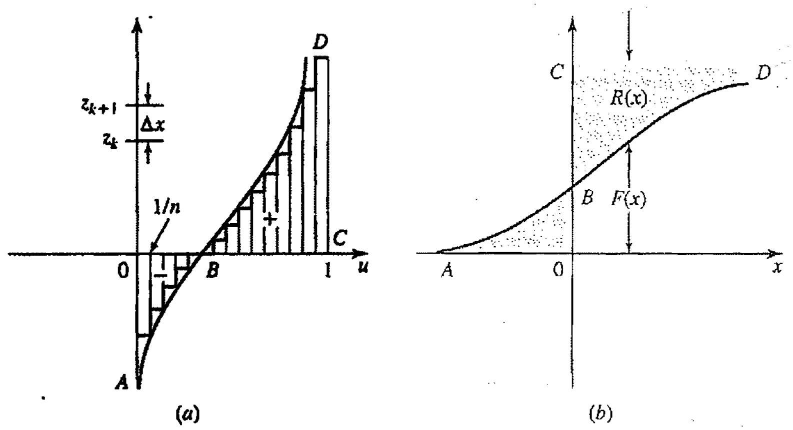


FIGURE 5-22

We shall use the above frequency interpretation to express the mean of $\mathbf{x}$ in terms of its distribution. From the construction of Fig. 5-22a it readily follows that $\bar{x}$ equals the area under the empirical percentile curve of $\mathbf{x}$. Thus

$$\bar{x} = (BCD) - (OAB)$$

where (BCD) and (OAB) are the shaded areas above and below the u axis, respectively. These areas equal the corresponding areas of Fig. 5-22b; hence

$$\bar{x} = \int_0^{\infty} [1 - F_n(x)]\,dx - \int_{-\infty}^{0} F_n(x)\,dx$$

where $F_n(x)$ is the empirical distribution of $\mathbf{x}$. With $n \to \infty$ this yields

$$E\{\mathbf{x}\} = \int_0^{\infty} R(x)\,dx - \int_{-\infty}^{0} F(x)\,dx, \qquad R(x) = 1 - F(x) = P\{\mathbf{x} > x\} \tag{5-52}$$

In particular, for a random variable that takes only nonnegative values, we also obtain

$$E\{\mathbf{x}\} = \int_0^{\infty} R(x)\,dx \tag{5-53}$$

Mean of $g(\mathbf{x})$. Given a random variable $\mathbf{x}$ and a function $g(x)$, we form the random variable $\mathbf{y} = g(\mathbf{x})$. As we see from (5-44), the mean of this random variable is given by

$$E\{\mathbf{y}\} = \int_{-\infty}^{\infty} y f_y(y)\,dy \tag{5-54}$$

It appears, therefore, that to determine the mean of $\mathbf{y}$, we must find its density $f_y(y)$. This, however, is not necessary. As the next basic theorem shows, $E\{\mathbf{y}\}$ can be expressed directly in terms of the function $g(x)$ and the density $f_x(x)$ of $\mathbf{x}$.

THEOREM 5-1

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g(x) f_x(x)\,dx \tag{5-55}$$

Proof. We shall sketch a proof using the curve $g(x)$ of Fig. 5-23. With $y = g(x_1) = g(x_2) = g(x_3)$ as in the figure, we see that

$$f_y(y)\,dy = f_x(x_1)\,dx_1 + f_x(x_2)\,dx_2 + f_x(x_3)\,dx_3$$

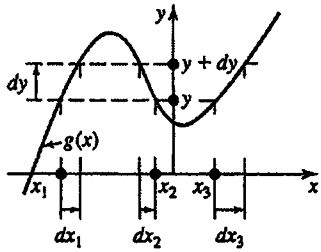


FIGURE 5-23

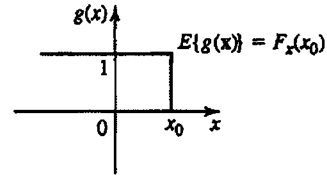


FIGURE 5-24

Multiplying by y, we obtain

$$y f_y(y)\,dy = g(x_1) f_x(x_1)\,dx_1 + g(x_2) f_x(x_2)\,dx_2 + g(x_3) f_x(x_3)\,dx_3$$

Thus to each differential in (5-54) there correspond one or more differentials in (5-55). As dy covers the y axis, the corresponding dx's are nonoverlapping and they cover the entire x axis. Hence the integrals in (5-54) and (5-55) are equal.

If $\mathbf{x}$ is of discrete type as in (5-45), then (5-55) yields

$$E\{g(\mathbf{x})\} = \sum_i g(x_i) P\{\mathbf{x} = x_i\} \tag{5-56}$$

◀

EXAMPLE 5-19

▶ With x_0 an arbitrary number and $g(x)$ as in Fig. 5-24, (5-55) yields

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{x_0} f_x(x)\,dx = F_x(x_0)$$

This shows that the distribution function of a random variable can be expressed as expected value. ◀

EXAMPLE 5-20

▶ In this example, we show that the probability of any event A can be expressed as expected value. For this purpose we form the zero–one random variable $\mathbf{x}_A$ associated with the event A:

$$\mathbf{x}_A(\zeta) = \begin{cases} 1 & \zeta \in A \\ 0 & \zeta \notin A \end{cases}$$

Since this random variable takes the values 1 and 0 with respective probabilities $P(A)$ and $P(\overline{A})$, yields

$$E\{\mathbf{x}_A\} = 1 \times P(A) + 0 \times P(\overline{A}) = P(A)$$ ◀

Linearity: From (5-55) it follows that

$$E\{a_1 g_1(\mathbf{x}) + \cdots + a_n g_n(\mathbf{x})\} = a_1 E\{g_1(\mathbf{x})\} + \cdots + a_n E\{g_n(\mathbf{x})\} \tag{5-57}$$

In particular, $E\{a\mathbf{x} + b\} = aE\{\mathbf{x}\} + b$

Complex random variables: If $\mathbf{z} = \mathbf{x} + j\mathbf{y}$ is a complex random variable, then its expected value is by definition

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

From this and (5-55) it follows that if

$$g(\mathbf{x}) = g_1(\mathbf{x}) + jg_2(\mathbf{x})$$

is a complex function of the real random variable $\mathbf{x}$ then

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g_1(x)f(x)\,dx + j\int_{-\infty}^{\infty} g_2(x)f(x)\,dx = \int_{-\infty}^{\infty} g(x)f(x)\,dx \tag{5-58}$$

In other words, (5-55) holds even if $g(x)$ is complex.

Variance

Mean alone will not be able to truly represent the p.d.f. of any random variable. To illustrate this, consider two Gaussian random variables $\mathbf{x}_1 \sim N(0, 1)$ and $\mathbf{x}_2 \sim N(0, 3)$. Both of them have the same mean $\mu = 0$. However, as Fig. 5-25 shows, their p.d.fs are quite different. Here $\mathbf{x}_1$ is more concentrated around the mean, whereas $\mathbf{x}_2$ has a wider spread. Clearly, we need at least an additional parameter to measure this spread around the mean!

For a random variable $\mathbf{x}$ with mean μ, $\mathbf{x} - \mu$ represents the deviation of the random variable from its mean. Since this deviation can be either positive or negative, consider the quantity $(\mathbf{x} - \mu)^2$, and its average value $E[(\mathbf{x} - \mu)^2]$ represents the average square deviation of $\mathbf{x}$ around its mean. Define

$$\sigma_x^2 \triangleq E[(\mathbf{x} - \mu)^2] > 0 \tag{5-59}$$

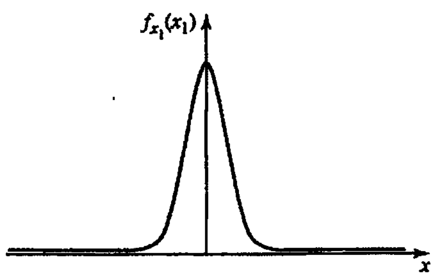
With $g(\mathbf{x}) = (\mathbf{x} - \mu)^2$ and using (5-55) we get

$$\sigma_x^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f_x(x)\,dx > 0 \tag{5-60}$$

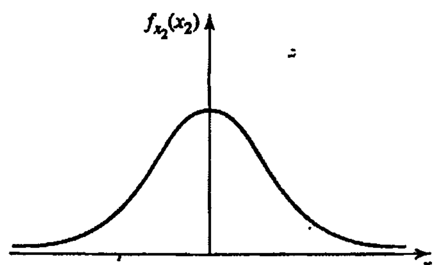
The positive constant σ_x^2 is known as the *variance* of the random variable $\mathbf{x}$, and its positive square root $\sigma_x = \sqrt{E(\mathbf{x} - \mu)^2}$ is known as the standard deviation of $\mathbf{x}$. Note that the standard deviation represents the root mean square value of the random variable $\mathbf{x}$ around its mean μ.

From the definition it follows that σ^2 is the mean of the random variable $(\mathbf{x} - \eta)^2$. Thus

$$\mathrm{Var}\{\mathbf{x}\} = \sigma^2 = E\{(\mathbf{x} - \eta)^2\} = E\{\mathbf{x}^2 - 2\mathbf{x}\eta + \eta^2\} = E\{\mathbf{x}^2\} - 2\eta E\{\mathbf{x}\} + \eta^2$$



(a) $\sigma^2 = 1$



(b) $\sigma^2 = 3$

FIGURE 5-25

Hence

$$\sigma^2 = E\{\mathbf{x}^2\} - \eta^2 = E\{\mathbf{x}^2\} - (E\{\mathbf{x}\})^2 \tag{5-61}$$

or, for any random variable

$$E\{\mathbf{x}^2\} \geq (E\{\mathbf{x}\})^2$$

EXAMPLE 5-21 ▶ If $\mathbf{x}$ is uniform in the interval $(-c, c)$, then $\eta = 0$ and

$$\sigma^2 = E\{\mathbf{x}^2\} = \frac{1}{2c}\int_{-c}^{c} x^2\,dx = \frac{c^2}{3}$$ ◀

EXAMPLE 5-22 ▶ We have written the density of a normal random variable in the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\eta)^2/2\sigma^2}$$

where up to now η and σ^2 were two arbitrary constants. We show next that η is indeed the mean of $\mathbf{x}$ and σ^2 its variance.

Proof. Clearly, $f(x)$ is symmetrical about the line $x = \eta$; hence $E\{\mathbf{x}\} = \eta$. Furthermore,

$$\int_{-\infty}^{\infty} e^{-(x-\eta)^2/2\sigma^2}\,dx = \sigma\sqrt{2\pi}$$

because the area of $f(x)$ equals 1. Differentiating with respect to σ, we obtain

$$\int_{-\infty}^{\infty} \frac{(x-\eta)^2}{\sigma^3} e^{-(x-\eta)^2/2\sigma^2}\,dx = \sqrt{2\pi}$$

Multiplying both sides by $\sigma^2/\sqrt{2\pi}$, we conclude that $E(\mathbf{x}-\eta)^2 = \sigma^2$ and the proof is complete. ◀

Discrete type. If the random variable $\mathbf{x}$ is of discrete type as in (5-45), then

$$\sigma^2 = \sum_i p_i(x_i - \eta)^2 \qquad p_i = P\{\mathbf{x} = x_i\} \tag{5-62}$$

EXAMPLE 5-23 ▶ The random variable $\mathbf{x}$ takes the values 1 and 0 with probabilities p and $q = 1 - p$ respectively. In this case

$$E\{\mathbf{x}\} = 1 \times p + 0 \times q = p$$
$$E\{\mathbf{x}^2\} = 1^2 \times p + 0^2 \times q = p$$

Hence

$$\sigma^2 = E\{\mathbf{x}^2\} - E^2\{\mathbf{x}\} = p - p^2 = pq$$ ◀

EXAMPLE 5-24 ▶ A Poisson distributed random variable with parameter λ takes the values $0, 1, \ldots$ with probabilities

$$P\{\mathbf{x} = k\} = e^{-\lambda}\frac{\lambda^k}{k!}$$

We shall show that its mean and variance both equal λ:

$$E\{\mathbf{x}\} = \lambda \qquad E\{\mathbf{x}^2\} = \lambda^2 + \lambda \qquad \sigma^2 = \lambda \tag{5-63}$$

Proof. We differentiate twice the Taylor expansion of e^λ:

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

$$e^\lambda = \sum_{k=0}^{\infty} k \frac{\lambda^{k-1}}{k!} = \frac{1}{\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!}$$

$$e^\lambda = \sum_{k=1}^{\infty} k(k-1) \frac{\lambda^{k-2}}{k!} = \frac{1}{\lambda^2} \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} - \frac{1}{\lambda^2} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!}$$

Hence

$$E\{\mathbf{x}\} = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = \lambda \qquad E\{\mathbf{x}^2\} = e^{-\lambda} \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} = \lambda^2 + \lambda$$

and (5-63) results.

Poisson points. As we have shown in (4-117), the number $\mathbf{n}$ of Poisson points in an interval of length t_0 is a Poisson distributed random variable with parameter $a = \lambda t_0$. From this it follows that

$$E\{\mathbf{n}\} = \lambda t_0 \qquad \sigma_n^2 = \lambda t_0 \tag{5-64}$$

This shows that the density λ of Poisson points equals the expected number of points per unit time. ◀

Notes 1. The variance σ^2 of a random variable $\mathbf{x}$ is a measure of the concentration of $\mathbf{x}$ near its mean η. Its relative frequency interpretation (empirical estimate) is the average of $(x_i - \eta)^2$:

$$\sigma^2 \simeq \frac{1}{n} \sum (x_i - \eta)^2 \tag{5-65}$$

where x_i are the observed values of $\mathbf{x}$. This average can be used as the estimate of σ^2 only if η is known. If it is unknown, we replace it by its estimate $\bar{x}$ and we change n to $n - 1$. This yields the estimate

$$\sigma^2 \simeq \frac{1}{n-1} \sum (x_i - \bar{x})^2 \qquad \bar{x} = \frac{1}{n} \sum x_i \tag{5-66}$$

known as the *sample variance* of $\mathbf{x}$ [see (7-65)]. The reason for changing n to $n - 1$ is explained later.

2. A simpler measure of the concentration of $\mathbf{x}$ near η is the first absolute central moment $M = E\{|\mathbf{x} - \eta|\}$. Its empirical estimate is the average of $|x_i - \eta|$:

$$M \simeq \frac{1}{n} \sum |x_i - \eta|$$

If η is unknown, it is replaced by $\bar{x}$. This estimate avoids the computation of squares.

5-4 MOMENTS

The following quantities are of interest in the study of random variables:

Moments

$$m_n = E\{\mathbf{x}^n\} = \int_{-\infty}^{\infty} x^n f(x)\, dx \tag{5-67}$$

Central moments

$$\mu_n = E\{(\mathbf{x}-\eta)^n\} = \int_{-\infty}^{\infty} (x-\eta)^n f(x)\,dx \tag{5-68}$$

Absolute moments

$$E\{|\mathbf{x}|^n\} \qquad E\{|\mathbf{x}-\eta|^n\} \tag{5-69}$$

Generalized moments

$$E\{(\mathbf{x}-a)^n\} \qquad E\{|\mathbf{x}-a|^n\} \tag{5-70}$$

We note that

$$\mu_n = E\{(\mathbf{x}-\eta)^n\} = E\left\{\sum_{k=0}^{n}\binom{n}{k}\mathbf{x}^k(-\eta)^{n-k}\right\}$$

Hence

$$\mu_n = \sum_{k=0}^{n}\binom{n}{k} m_k(-\eta)^{n-k} \tag{5-71}$$

Similarly,

$$m_n = E\{[(\mathbf{x}-\eta)+\eta]^n\} = E\left\{\sum_{k=0}^{n}\binom{n}{k}(\mathbf{x}-\eta)^k\eta^{n-k}\right\}$$

Hence

$$m_n = \sum_{k=0}^{n}\binom{n}{k}\mu_k\eta^{n-k} \tag{5-72}$$

In particular,

$$\mu_0 = m_0 = 1 \qquad m_1 = \eta \qquad \mu_1 = 0 \qquad \mu_2 = \sigma^2$$

and

$$\mu_3 = m_3 - 3\eta m_2 + 2\eta^3 \qquad m_3 = \mu_3 + 3\eta\sigma^2 + \eta^3$$

Notes 1. If the function $f(x)$ is interpreted as mass density on the x axis, then $E\{\mathbf{x}\}$ equals its center of gravity, $E\{\mathbf{x}^2\}$ equals the moment of inertia with respect to the origin, and σ^2 equals the central moment of inertia. The standard deviation σ is the radius of gyration.

2. The constants η and σ give only a limited characterization of $f(x)$. Knowledge of other moments provides additional information that can be used, for example, to distinguish between two densities with the same η and σ. In fact, if m_n is known for every n, then, under certain conditions, $f(x)$ is determined uniquely [see also (5-105)]. The underlying theory is known in mathematics as the *moment problem.*

3. The moments of a random variable are not arbitrary numbers but must satisfy various inequalities [see (5-92)]. For example [see (5-61)]

$$\sigma^2 = m_2 - m_1^2 \ge 0$$

Similarly, since the quadratic

$$E\{(\mathbf{x}^n - a)^2\} = m_{2n} - 2am_n + a^2$$

is nonnegative for any a, its discriminant cannot be positive. Hence

$$m_{2n} \ge m_n^2$$

Normal random variables. We shall show that if

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$$

then

$$E\{\mathbf{x}^n\} = \begin{cases} 0 & n = 2k+1 \\ 1\cdot 3\cdots(n-1)\sigma^n & n = 2k \end{cases} \tag{5-73}$$

$$E\{|\mathbf{x}|^n\} = \begin{cases} 1\cdot 3\cdots(n-1)\sigma^n & n = 2k \\ 2^k k!\sigma^{2k+1}\sqrt{2/\pi} & n = 2k+1 \end{cases} \tag{5-74}$$

The odd moments of $\mathbf{x}$ are 0 because $f(-x) = f(x)$. To prove the lower part of (5-73), we differentiate k times the identity

$$\int_{-\infty}^{\infty} e^{-\alpha x^2}\,dx = \sqrt{\frac{\pi}{\alpha}}$$

This yields

$$\int_{-\infty}^{\infty} x^{2k} e^{-\alpha x^2}\,dx = \frac{1\cdot 3\cdots(2k-1)}{2^k}\sqrt{\frac{\pi}{\alpha^{2k+1}}}$$

and with $\alpha = 1/2\sigma^2$, (5-73) results.

Since $f(-x) = f(x)$, we have

$$E\{|\mathbf{x}|^{2k+1}\} = 2\int_0^{\infty} x^{2k+1} f(x)\,dx = \frac{2}{\sigma\sqrt{2\pi}}\int_0^{\infty} x^{2k+1} e^{-x^2/2\sigma^2}\,dx$$

With $y = x^2/2\sigma^2$, the above yields

$$\sqrt{\frac{2}{\pi}}\frac{(2\sigma^2)^{k+1}}{2\sigma}\int_0^{\infty} y^k e^{-y}\,dy$$

and (5-74) results because the last integral equals $k!$

We note in particular that

$$E\{\mathbf{x}^4\} = 3\sigma^4 = 3E^2\{\mathbf{x}^2\} \tag{5-75}$$

EXAMPLE 5-25

▶ If $\mathbf{x}$ has a *Rayleigh density*

$$f(x) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} U(x)$$

then

$$E\{\mathbf{x}^n\} = \frac{1}{\sigma^2}\int_0^{\infty} x^{n+1} e^{-x^2/2\sigma^2}\,dx = \frac{1}{2\sigma^2}\int_{-\infty}^{\infty} |x|^{n+1} e^{-x^2/2\sigma^2}\,dx$$

From this and (5-74) it follows that

$$E\{\mathbf{x}^n\} = \begin{cases} 1\cdot 3\cdots n\sigma^n\sqrt{\pi/2} & n = 2k+1 \\ 2^k k!\sigma^{2k} & n = 2k \end{cases} \tag{5-76}$$

In particular,

$$E\{\mathbf{x}\} = \sigma\sqrt{\pi/2} \qquad \operatorname{Var}\{\mathbf{x}\} = (2-\pi/2)\sigma^2 \tag{5-77}$$

◀

EXAMPLE 5-26

▶ If x has a *Maxwell density*

$$f(x) = \frac{\sqrt{2}}{\alpha^3\sqrt{\pi}} x^2 e^{-x^2/2\alpha^2} U(x)$$

then

$$E\{\mathbf{x}^n\} = \frac{1}{\alpha^3\sqrt{2\pi}} \int_{-\infty}^{\infty} |x|^{n+2} e^{-x^2/2\alpha^2}\, dx$$

and (5-74) yields

$$E\{\mathbf{x}^n\} = \begin{cases} 1\cdot 3\cdots(n+1)\alpha^n & n = 2k \\ 2^k k!\alpha^{2k-1}\sqrt{2/\pi} & n = 2k-1 \end{cases} \tag{5-78}$$

In particular,

$$E\{\mathbf{x}\} = 2\alpha\sqrt{2/\pi} \qquad E\{\mathbf{x}^2\} = 3\alpha^2 \tag{5-79}$$

◀

Poisson random variables. The moments of a Poisson distributed random variable are functions of the parameter λ:

$$m_n(\lambda) = E\{\mathbf{x}^n\} = e^{-\lambda} \sum_{k=0}^{\infty} k^n \frac{\lambda^k}{k!} \tag{5-80}$$

$$\mu_n(\lambda) = E\{(\mathbf{x}-\lambda)^n\} = e^{-\lambda} \sum_{k=0}^{\infty} (k-\lambda)^n \frac{\lambda^k}{k!} \tag{5-81}$$

We shall show that they satisfy the recursion equations

$$m_{n+1}(\lambda) = \lambda[m_n(\lambda) + m_n'(\lambda)] \tag{5-82}$$

$$\mu_{n+1}(\lambda) = \lambda[n\mu_{n-1}(\lambda) + \mu_n'(\lambda)] \tag{5-83}$$

Proof. Differentiating (5-80) with respect to λ, we obtain

$$m_n'(\lambda) = -e^{-\lambda} \sum_{k=0}^{\infty} k^n \frac{\lambda^k}{k!} + e^{-\lambda} \sum_{k=0}^{\infty} k^{n+1} \frac{\lambda^{k-1}}{k!} = -m_n(\lambda) + \frac{1}{\lambda} m_{n+1}(\lambda)$$

and (5-82) results. Similarly, from (5-81) it follows that

$$\mu_n'(\lambda) = -e^{-\lambda} \sum_{k=0}^{\infty} (k-\lambda)^n \frac{\lambda^k}{k!} - ne^{-\lambda} \sum_{k=0}^{\infty} (k-\lambda)^{n-1} \frac{\lambda^k}{k!}$$
$$+ e^{-\lambda} \sum_{k=0}^{\infty} (k-\lambda)^n k \frac{\lambda^{k-1}}{k!}$$

Setting $k = (k-\lambda) + \lambda$ in the last sum, we obtain $\mu_n' = -\mu_n - n\mu_{n-1} + (1/\lambda)(\mu_{n+1} + \lambda\mu_n)$ and (5-83) results.

The preceding equations lead to the recursive determination of the moments m_n and μ_n. Starting with the known moments $m_1 = \lambda$, $\mu_1 = 0$, and $\mu_2 = \lambda$ [see (5-63)], we obtain $m_2 = \lambda(\lambda+1)$ and

$$m_3 = \lambda(\lambda^2+\lambda+2\lambda+1) = \lambda^3+3\lambda^2+\lambda \qquad \mu_3 = \lambda(\mu_2'+2\mu_1) = \lambda$$

ESTIMATE OF THE MEAN OF $g(\mathbf{x})$. The mean of the random variable $\mathbf{y} = g(\mathbf{x})$ is given by

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g(x)f(x)\,dx \tag{5-84}$$

Hence, for its determination, knowledge of $f(x)$ is required. However, if $\mathbf{x}$ is concentrated near its mean, then $E\{g(\mathbf{x})\}$ can be expressed in terms of the moments μ_n of $\mathbf{x}$.

Suppose, first, that $f(x)$ is negligible outside an interval $(\eta-\varepsilon, \eta+\varepsilon)$ and in this interval, $g(x) \simeq g(\eta)$. In this case, (5-84) yields

$$E\{g(\mathbf{x})\} \simeq g(\eta)\int_{\eta-\varepsilon}^{\eta+\varepsilon} f(x)\,dx \simeq g(\eta)$$

This estimate can be improved if $g(x)$ is approximated by a polynomial

$$g(x) \simeq g(\eta) + g'(\eta)(x-\eta) + \cdots + g^{(n)}(\eta)\frac{(x-\eta)^n}{n!}$$

Inserting into (5-84), we obtain

$$E\{g(\mathbf{x})\} \simeq g(\eta) + g''(\eta)\frac{\sigma^2}{2} + \cdots + g^{(n)}(\eta)\frac{\mu_n}{n!} \tag{5-85}$$

In particular, if $g(x)$ is approximated by a parabola, then

$$\eta_y = E\{g(\mathbf{x})\} \simeq g(\eta) + g''(\eta)\frac{\sigma^2}{2} \tag{5-86}$$

And if it is approximated by a straight line, then $\eta_y \simeq g(\eta)$. This shows that the slope of $g(x)$ has no effect on η_y; however, as we show next, it affects the variance σ_y^2 of $\mathbf{y}$.

Variance. We maintain that the first-order estimate of σ_y^2 is given by

$$\sigma_y^2 \simeq |g'(\eta)|^2\sigma^2 \tag{5-87}$$

Proof. We apply (5-86) to the function $g^2(x)$. Since its second derivative equals $2(g')^2 + 2gg''$, we conclude that

$$\sigma_y^2 + \eta_y^2 = E\{g^2(\mathbf{x})\} \simeq g^2 + [(g')^2 + gg'']\sigma^2$$

Inserting the approximation (5-86) for η_y into the above and neglecting the σ^4 term, we obtain (5-87).

EXAMPLE 5-27

▶ A voltage $E = 120$ V is connected across a resistor whose resistance is a random variable $\mathbf{r}$ uniform between 900 and 1100 Ω. Using (5-85) and (5-86), we shall estimate the mean and variance of the resulting current

$$\mathbf{i} = \frac{E}{\mathbf{r}}$$

Clearly, $E\{\mathbf{r}\} = \eta = 10^3, \sigma^2 = 100^2/3$. With $g(r) = E/r$, we have

$$g(\eta) = 0.12 \qquad g'(\eta) = -12 \times 10^{-5} \qquad g''(\eta) = 24 \times 10^{-8}$$

Hence

$$E\{\mathbf{i}\} \simeq 0.12 + 0.0004\,A \qquad \sigma_i^2 \simeq 48 \times 10^{-6}\,A^2$$ ◀

A measure of the concentration of a random variable near its mean η is its variance σ^2. In fact, as the following theorem shows, the probability that $\mathbf{x}$ is outside an arbitrary interval $(\eta - \varepsilon, \eta + \varepsilon)$ is negligible if the ratio σ/ε is sufficiently small. This result, known as the *Chebyshev inequality*, is fundamental.

CHEBYSHEV (TCHEBYCHEFF) INEQUALITY

▶ For any $\varepsilon > 0$,

$$P\{|\mathbf{x} - \eta| \ge \varepsilon\} \le \frac{\sigma^2}{\varepsilon^2} \tag{5-88}$$

Proof. The proof is based on the fact that

$$P\{|\mathbf{x} - \eta| \ge \varepsilon\} = \int_{-\infty}^{-\eta-\varepsilon} f(x)\,dx + \int_{\eta+\varepsilon}^{\infty} f(x)\,dx = \int_{|x-\eta|\ge\varepsilon} f(x)\,dx$$

Indeed

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\eta)^2 f(x)\,dx \ge \int_{|x-\eta|\ge\varepsilon} (x-\eta)^2 f(x)\,dx \ge \varepsilon^2 \int_{|x-\eta|\ge\varepsilon} f(x)\,dx$$

and (5-88) results because the last integral equals $P\{|\mathbf{x} - \eta| \ge \varepsilon\}$. ◀

Notes 1. From (5-88) it follows that, if $\sigma = 0$, then the probability that $\mathbf{x}$ is outside the interval $(\eta - \varepsilon, \eta + \varepsilon)$ equals 0 for any ε; hence $\mathbf{x} = \eta$ with probability 1. Similarly, if

$$E\{\mathbf{x}^2\} = \eta^2 + \sigma^2 = 0 \qquad \text{then} \quad \eta = 0 \quad \sigma = 0$$

hence $\mathbf{x} = 0$ with probability 1.

2. For specific densities, the bound in (5-88) is too high. Suppose, for example, that $\mathbf{x}$ is normal. In this case, $P\{|\mathbf{x} - \eta| \ge 3\sigma\} = 2 - 2G(3) = 0.0027$. Inequality (5-88), however, yields $P\{|\mathbf{x} - \eta| \ge 3\sigma\} \le 1/9$.

The significance of Chebyshev's inequality is the fact that it holds for *any* $f(x)$ and can, therefore be used even if $f(x)$ is not known.

3. The bound in (5-88) can be reduced if various assumptions are made about $f(x)$ [see *Chernoff bound* (Prob. 5-35)].

MARKOV INEQUALITY

▶ If $f(x) = 0$ for $x < 0$, then, for any $\alpha > 0$,

$$P\{\mathbf{x} \ge \alpha\} \le \frac{\eta}{\alpha} \tag{5-89}$$

Proof.

$$E\{\mathbf{x}\} = \int_0^{\infty} x f(x)\,dx \ge \int_{\alpha}^{\infty} x f(x)\,dx \ge \alpha \int_{\alpha}^{\infty} f(x)\,dx$$

and (5-89) results because the last integral equals $P\{\mathbf{x} \ge \alpha\}$. ◀

BIENAYMÉ INEQUALITY

▶ Suppose that $\mathbf{x}$ is an arbitrary random variable and a and n are two arbitrary numbers. Clearly, the random variable $|\mathbf{x}-a|^n$ takes only positive values. Applying (5-89), with $\alpha=\varepsilon^n$, we conclude that

$$P\{|\mathbf{x}-a|^n \ge \varepsilon^n\} \le \frac{E\{|\mathbf{x}-a|^n\}}{\varepsilon^n} \tag{5-90}$$

Hence

$$P\{|\mathbf{x}-a| \ge \varepsilon\} \le \frac{E\{|\mathbf{x}-a|^n\}}{\varepsilon^n} \tag{5-91}$$

This result is known as the *inequality of Bienaymé*. Chebyshev's inequality is a special case obtained with $a=\eta$ and $n=2$. ◀

LYAPUNOV INEQUALITY

▶ Let $\beta_k = E\{|\mathbf{x}|^k\} < \infty$ represent the absolute moments of the random variable $\mathbf{x}$. Then for any k

$$\beta_{k-1}^{1/(k-1)} \le \beta_k^{1/k} \qquad k \ge 1 \tag{5-92}$$

Proof. Consider the random variable

$$\mathbf{y} = a|\mathbf{x}|^{(k-1)/2} + |\mathbf{x}|^{(k+1)/2}$$

Then

$$E\{\mathbf{y}^2\} = a^2\beta_{k-1} + 2a\beta_k + \beta_{k+1} \ge 0$$

implying that the discriminant of the preceding quadratic must be nonpositive. Thus

$$\beta_k^2 \le \beta_{k-1}\beta_{k+1} \quad \text{or} \quad \beta_k^{2k} \le \beta_{k-1}^k\beta_{k+1}^k$$

This gives

$$\beta_1^2 \le \beta_0\beta_2, \qquad \beta_2^4 \le \beta_1^2\beta_3^2, \ldots, \beta_{n-1}^{2(n-1)} \le \beta_{n-2}^{n-1}\beta_n^{n-1}$$

where $\beta_0 = 1$. Multiplying successively we get

$$\beta_1^2 \le \beta_2, \qquad \beta_2^3 \le \beta_3^2, \qquad \beta_3^4 \le \beta_4^3, \ldots, \beta_{k-1}^k \le \beta_k^{k-1}, \quad \text{or} \quad \beta_{k-1}^{1/(k-1)} \le \beta_k^{1/k}$$

Thus, we also obtain

$$\beta_1 \le \beta_2^{1/2} \le \beta_3^{1/3} \le \cdots \le \beta_n^{1/n} \tag{5-93}$$

◀

5-5 CHARACTERISTIC FUNCTIONS

The *characteristic function* of a random variable is by definition the integral

$$\Phi_x(\omega) = \int_{-\infty}^{\infty} f(x)e^{j\omega x}\,dx \tag{5-94}$$

This function is maximum at the origin because $f(x) \ge 0$:

$$|\Phi_x(\omega)| \le \Phi_x(0) = 1 \tag{5-95}$$

If $j\omega$ is changed to s, the resulting integral

$$\Phi(s) = \int_{-\infty}^{\infty} f(x)e^{sx}\,dx \qquad \Phi(j\omega) = \Phi_{\mathbf{x}}(\omega) \tag{5-96}$$

is the *moment (generating) function* of $\mathbf{x}$.

The function

$$\Psi(\omega) = \ln \Phi_{\mathbf{x}}(\omega) = \Psi(j\omega) \tag{5-97}$$

is the *second characteristic function* of $\mathbf{x}$.

Clearly [see (5-58)]

$$\Phi_{\mathbf{x}}(\omega) = E\{e^{j\omega\mathbf{x}}\} \qquad \Phi(s) = E\{e^{s\mathbf{x}}\} \tag{5-98}$$

This leads to the fact that

$$\text{If } \mathbf{y} = a\mathbf{x} + b \quad \text{then} \quad \Phi_{\mathbf{y}}(\omega) = e^{jb\omega}\Phi_{\mathbf{x}}(a\omega) \tag{5-99}$$

because

$$E\{e^{j\omega\mathbf{y}}\} = E\left\{e^{j\omega(a\mathbf{x}+b)}\right\} = e^{jb\omega}E\{e^{ja\omega\mathbf{x}}\}$$

EXAMPLE 5-28

▶ We shall show that the characteristic function of an $N(\eta, \sigma)$ random variable $\mathbf{x}$ equals (see Table 5-2)

$$\Phi_x(\omega) = \exp\left\{j\eta\omega - \tfrac{1}{2}\sigma^2\omega^2\right\} \tag{5-100}$$

Proof. The random variable $\mathbf{z} = (\mathbf{x} - \eta)/\sigma$ is $N(0, 1)$ and its moment function equals

$$\Phi_z(s) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{sz}e^{-z^2/2}\,dz$$

with

$$sz - \frac{z^2}{2} = -\frac{1}{2}(z-s)^2 + \frac{s^2}{2}$$

we conclude that

$$\Phi_z(s) = e^{s^2/2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-(z-s)^2/2}\,dz = e^{s^2/2} \tag{5-101}$$

And since $\mathbf{x} = \sigma\mathbf{z} + \eta$, (5-100) follows from (5-99) and (5-101) with $s = j\omega$. ◀

Inversion formula As we see from (5-94), $\Phi_x(\omega)$ is the Fourier transform of $f(x)$. Hence the properties of characteristic functions are essentially the same as the properties of Fourier transforms. We note, in particular, that $f(x)$ can be expressed in terms of $\Phi(\omega)$

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_x(\omega)e^{-j\omega x}\,d\omega \tag{5-102}$$

Moment theorem. Differentiating (5-96) n times, we obtain

$$\Phi^{(n)}(s) = E\{\mathbf{x}^n e^{s\mathbf{x}}\}$$

Hence

$$\Phi^{(n)}(0) = E\{\mathbf{x}^n\} = m_n \tag{5-103}$$

Thus the derivatives of $\Phi(s)$ at the origin equal the moments of $\mathbf{x}$. This justifies the name "moment function" given to $\Phi(s)$.

In particular,

$$\Phi'(0) = m_1 = \eta \qquad \Phi''(0) = m_2 = \eta^2 + \sigma^2 \tag{5-104}$$

Note Expanding $\Phi(s)$ into a series near the origin and using (5-103), we obtain

$$\Phi(s) = \sum_{n=0}^{\infty} \frac{m_n}{n!} s^n \tag{5-105}$$

This is valid only if all moments are finite and the series converges absolutely near $s = 0$. Since $f(x)$ can be determined in terms of $\Phi(s)$, (5-105) shows that, under the stated conditions, the density of a random variable is uniquely determined if all its moments are known.

EXAMPLE 5-29

▶ We shall determine the moment function and the moments of a random variable $\mathbf{x}$ with *gamma distribution:* (see also Table 5-2)

$$f(x) = \gamma x^{b-1} e^{-cx} U(x) \qquad \gamma = \frac{c^{b+1}}{\Gamma(b+1)}$$

From (4-35) it follows that

$$\Phi(s) = \gamma \int_0^\infty x^{b-1} e^{-(c-s)x}\, dx = \frac{\gamma \Gamma(b)}{(c-s)^b} = \frac{c^b}{(c-s)^b} \tag{5-106}$$

Differentiating with respect to s and setting $s = 0$, we obtain

$$\Phi^{(n)}(0) = \frac{b(b+1)\cdots(b+n-1)}{c^n} = E\{\mathbf{x}^n\}$$

With $n = 1$ and $n = 2$, this yields

$$E\{\mathbf{x}\} = \frac{b}{c} \qquad E\{\mathbf{x}^2\} = \frac{b(b+1)}{c^2} \qquad \sigma^2 = \frac{b}{c^2} \tag{5-107}$$

The ***exponential density*** is a special case obtained with $b = 1, c = \lambda$:

$$f(x) = \lambda e^{-\lambda x} U(x) \quad \Phi(s) = \frac{\lambda}{\lambda - s} \quad E\{\mathbf{x}\} = \frac{1}{\lambda} \quad \sigma^2 = \frac{1}{\lambda^2} \tag{5-108}$$

Chi square: Setting $b = m/2$ and $c = 1/2$ in (5-106), we obtain the moment function of the chi-square density $\chi^2(m)$:

$$\Phi(s) = \frac{1}{\sqrt{(1-2s)^m}} \qquad E\{\mathbf{x}\} = m \qquad \sigma^2 = 2m \tag{5-109}$$

◀

Cumulants. The cumulants λ_n of random variable $\mathbf{x}$ are by definition the derivatives

$$\frac{d^n \Psi(0)}{ds^n} = \lambda_n \tag{5-110}$$

of its second moment function $\Psi(s)$. Clearly [see (5-97)] $\Psi(0) = \lambda_0 = 0$; hence

$$\Psi(s) = \lambda_1 s + \frac{1}{2}\lambda_2 s^2 + \cdots + \frac{1}{n!}\lambda_n s^n + \cdots$$

We maintain that

$$\lambda_1 = \eta \qquad \lambda_2 = \sigma^2 \tag{5-111}$$

Proof. Since $\Phi = e^{\Psi}$, we conclude that

$$\Phi' = \Psi' e^{\Psi} \qquad \Phi' = [\Psi'' + (\Psi')^2]e^{\Psi}$$

With $s = 0$, this yields

$$\Phi'(0) = \Psi'(0) = m_1 \qquad \Phi''(0) = \Psi''(0) + [\Psi'(0)]^2 = m_2$$

and (5-111) results.

Discrete Type

Suppose that $\mathbf{x}$ is a discrete-type random variable taking the values x_i with probability p_i. In this case, (5-94) yields

$$\Phi_x(\omega) = \sum_i p_i e^{j\omega x_i} \tag{5-112}$$

Thus $\Phi_x(\omega)$ is a sum of exponentials. The moment function of $\mathbf{x}$ can be defined as in (5-96). However, if $\mathbf{x}$ takes only integer values, then a definition in terms of z transforms is preferable.

MOMENT GENERATING FUNCTIONS. If $\mathbf{x}$ is a lattice type random variable taking integer values, then its moment generating function is by definition the sum

$$\Gamma(z) = E\{z^{\mathbf{x}}\} = \sum_{n=-\infty}^{+\infty} P\{\mathbf{x} = n\}z^n = \sum_{n=-\infty}^{\infty} p_n z^n \tag{5-113}$$

Thus $\Gamma(1/z)$ is the ordinary z transform of the sequence $p_n = P\{\mathbf{x} = n\}$. With $\Phi_x(\omega)$ as in (5-112), this yields

$$\Phi_x(\omega) = \Gamma(e^{j\omega}) = \sum_{n=-\infty}^{\infty} p_n e^{jn\omega}$$

Thus $\Phi_x(\omega)$ is the *discrete Fourier transform* (DFT) of the sequence $\{p_n\}$, and

$$\Psi(s) = \ln \Gamma(e^s) \tag{5-114}$$

Moment theorem. Differentiating (5-113) k times, we obtain

$$\Gamma^{(k)}(z) = E\{\mathbf{x}(\mathbf{x}-1)\cdots(\mathbf{x}-k+1)z^{\mathbf{x}-k}\}$$

With $z = 1$, this yields

$$\Gamma^{(k)}(1) = E\{\mathbf{x}(\mathbf{x}-1)\cdots(\mathbf{x}-k+1)\} \tag{5-115}$$

We note, in particular, that $\Gamma(1) = 1$ and

$$\Gamma'(1) = E\{\mathbf{x}\} \qquad \Gamma''(1) = E\{\mathbf{x}^2\} - E\{\mathbf{x}\} \tag{5-116}$$

EXAMPLE 5-30

▶ (*a*) If $\mathbf{x}$ takes the values 0 and 1 with $P\{\mathbf{x} = 1\} = p$ and $P\{\mathbf{x} = 0\} = q$, then

$$\Gamma(z) = pz + q$$

$$\Gamma'(1) = E\{\mathbf{x}\} = p \qquad \Gamma''(1) = E\{\mathbf{x}^2\} - E\{\mathbf{x}\} = 0$$

(*b*) If $\mathbf{x}$ has the binomial distribution $B(m, p)$ given by

$$p_n = P\{\mathbf{x} = n\} = \binom{m}{n} p^n q^{m-n} \qquad 0 \le n \le m$$

then

$$\Gamma(z) = \sum_{n=0}^{m} \binom{m}{n} p^n q^{m-n} z^n = (pz + q)^m \tag{5-117}$$

and

$$\Gamma'(1) = mp \qquad \Gamma''(1) = m(m-1)p^2$$

Hence

$$E\{\mathbf{x}\} = mp \qquad \sigma^2 = mpq \tag{5-118}$$

◀

EXAMPLE 5-31

▶ If $\mathbf{x}$ is Poisson distributed with parameter λ,

$$P\{\mathbf{x} = n\} = e^{-\lambda}\frac{\lambda^n}{n!} \qquad n = 0, 1, \ldots$$

then

$$\Gamma(z) = e^{-\lambda} \sum_{n=0}^{\infty} \lambda^n \frac{z^n}{n!} = e^{\lambda(z-1)} \tag{5-119}$$

In this case [see (5-114)]

$$\Psi(s) = \lambda(e^s - 1) \qquad \Psi'(0) = \lambda \qquad \Psi''(0) = \lambda$$

and (5-111) yields $E\{\mathbf{x}\} = \lambda$, $\sigma^2 = \lambda$ in agreement with (5-63). ◀

We can use the characteristic function method to establish the DeMoivre–Laplace theorem in (4-90).

THEOREM 5-2

DEMOIVRE–LAPLACE THEOREM

▶ Let $\mathbf{x} \sim B(n, p)$. Then from (5-117), we obtain the characteristic function of the binomial random variable to be

$$\Phi_x(\omega) = (pe^{j\omega} + q)^n$$

and define

$$\mathbf{y} = \frac{\mathbf{x} - np}{\sqrt{npq}} \tag{5-120}$$

This gives

$$
\begin{aligned}
\Phi_y(\omega) &= E\{e^{jy\omega}\} = e^{-np\omega/\sqrt{npq}}\Phi_x\left(\frac{\omega}{\sqrt{npq}}\right) \\
&= e^{-np\omega/\sqrt{npq}}(pe^{j\omega/\sqrt{npq}}+q)^n \\
&= (pe^{jq\omega/\sqrt{npq}}+qe^{-jp\omega/\sqrt{npq}})^n \\
&= \left\{p\left(1+\frac{jq\omega}{\sqrt{npq}}-\frac{q^2\omega^2}{2npq}+\sum_{k=3}^{\infty}\frac{1}{k!}\left(\frac{jq\omega}{\sqrt{npq}}\right)^k\right)\right. \\
&\qquad \left. +q\left(1-\frac{jp\omega}{\sqrt{npq}}-\frac{p^2\omega^2}{2npq}+\sum_{k=3}^{\infty}\frac{1}{k!}\left(\frac{-jp\omega}{\sqrt{npq}}\right)^k\right)\right\}^n \\
&= \left(1-\frac{\omega^2}{2n}\{1+\phi(n)\}\right)^n \to e^{-\omega^2/2}, \quad \text{as} \quad n\to\infty
\end{aligned}
\tag{5-121}
$$

since

$$
\phi(n) \triangleq 2\sum_{k=3}^{\infty}\frac{1}{k!}\left(\frac{j\omega}{\sqrt{n}}\right)^{k-2}\frac{pq^k+q(-p)^k}{(\sqrt{pq})^k} \to 0, \quad \text{as} \quad n\to\infty
$$

On comparing (5-121) with (5-100), we conclude that as $n \to \infty$, the random variable $\mathbf{y}$ tends to the standard normal distribution, or from (5-120), $\mathbf{x}$ tends to $N(np, npq)$. ◀

In Examples 5-32 and 5-33 we shall exhibit the usefulness of the moment generating function in solving problems. The next example is of historical interest, as it was first proposed and solved by DeMoivre.

EXAMPLE 5-32

▶ An event A occurs in a series of independent trials with constant probability p. If A occurs at least r times in succession, we refer to it as a run of length r. Find the probability of obtaining a run of length r for A in n trials.

SOLUTION

Let p_n denote the probability of the event X_n that represents a run of length r for A in n trials. A run of length r in $n+1$ trials can happen in only two mutually exclusive ways: either there is a run of length r in the first n trials, or a run of length r is obtained only in the last r trials of the $n+1$ trials and not before that. Thus

$$
X_{n+1} = X_n \cup B_{n+1} \tag{5-122}
$$

where

$$
\begin{aligned}
B_{n+1} &= \{\text{No run of length } r \text{ for } A \text{ in the first } n-r \text{ trials}\} \\
&\qquad \cap \{A \text{ does not occur in the } (n-r+1)\text{th trial}\} \\
&\qquad \cap \{\text{Run of length } r \text{ for } A \text{ in the last } r \text{ trials}\} \\
&= \overline{X}_{n-r} \cap \overline{A} \cap \underbrace{A \cap A \cap \cdots \cap A}_{r}
\end{aligned}
$$

Hence by the independence of these events

$$P\{B_{n+1}\} = (1 - p_{n-r})qp'$$

so that from (5-122)

$$p_{n+1} = P\{X_{n+1}\} = P\{X_n\} + P\{B_{n+1}\} = p_n + (1 - p_{n-r})qp^r \tag{5-123}$$

The equation represents an ordinary difference equation with the obvious initial conditions

$$p_0 = p_1 = \cdots = p_{r-1} = 0 \quad \text{and} \quad p_r = p^r \tag{5-124}$$

From (5-123), although it is possible to obtain $p_{r+1} = p'(1+q), \ldots, p_{r+m} = p^r(1+mq)$ for $m \le r-1$, the expression gets quite complicated for large values of n. The method of moment generating functions in (5-113) can be used to obtain a general expression for p_n. Toward this, let

$$q_n \triangleq 1 - p_n \tag{5-125}$$

so that (5-123) translates into (with n replaced by $n + r$)

$$q_{n+r+1} = q_{n+r} - qp^r q_n \qquad n \ge 0 \tag{5-126}$$

with the new initial conditions

$$q_0 = q_1 = \cdots = q_{r-1} = 1 \qquad q_r = 1 - p^r \tag{5-127}$$

Following (5-113), define the moment generating function

$$\phi(z) = \sum_{n=0}^{\infty} q_n z^n \tag{5-128}$$

and using (5-126) we obtain

$$\begin{aligned}
qp^r\phi(z) &= \left(\sum_{n=0}^{\infty} q_{n+r} z^n - \sum_{n=0}^{\infty} q_{n+r+1} z^n\right) \\
&= \frac{\phi(z) - \sum_{k=0}^{r-1} q_k z^k}{z^r} - \frac{\phi(z) - \sum_{k=0}^{r} q_k z^k}{z^{r+1}} \\
&= \frac{(z-1)\phi(z) - \sum_{k=1}^{r} z^k + \left(\sum_{k=0}^{r-1} z^k + (1-p^r)z^r\right)}{z^{r+1}} \\
&= \frac{(z-1)\phi(z) + 1 - p^r z^r}{z^{r+1}}
\end{aligned} \tag{5-129}$$

where we have made use of the initial conditions in (5-127). From (5-129) we get the desired moment generating function to be

$$\phi(z) = \frac{1 - p^r z^r}{1 - z + qp^r z^{r+1}} \tag{5-130}$$

$\phi(z)$ is a rational function in z, and the coefficient of z^n in its power series expansion gives q_n. More explicitly

$$\begin{aligned}
\phi(z) &= (1 - p^r z^r)[1 - z(1 - qp^r z^r)]^{-1} \\
&= (1 - p^r z^r)[1 + \alpha_{1,r} z + \cdots + \alpha_{n,r} z^n + \cdots]
\end{aligned} \tag{5-131}$$

so that the desired probability equals

$$q_n = \alpha_{n,r} - p^r \alpha_{n-r,r} \tag{5-132}$$

where $\alpha_{n,r}$ is the coefficient of z^n in the expansion of $[1 - z(1 - qp^r z^r)]^{-1}$. But

$$[1 - z(1 - qp^r z^r)]^{-1} = \sum_{m=0}^{\infty} z^m (1 - qp^r z^r)^m = \sum_{m=0}^{\infty}\sum_{k=0}^{m} \binom{m}{k} (-1)^k (qp^r)^k z^{m+kr}$$

Let $m + kr = n$ so that $m = n - kr$, and this expression simplifies to

$$[1 - z(1 - qp^r z^r)]^{-1} = \sum_{n=0}^{\infty} \sum_{k=0}^{\lfloor n/(r+1)\rfloor} \binom{n-kr}{k} (-1)^k (qp^r)^k z^n = \sum_{n=0}^{\infty} \alpha_{n,r} z^n$$

and the upper limit on k corresponds to the condition $n - kr \geq k$ so that $\binom{n-kr}{k}$ is well defined. Thus

$$\alpha_{n,r} = \sum_{k=0}^{\lfloor n/(r+1)\rfloor} \binom{n-kr}{k} (-1)^k (qp^r)^k \tag{5-133}$$

With $\alpha_{n,r}$ so obtained, finally the probability of r runs for A in n trials is given by

$$p_n = 1 - q_n = 1 - \alpha_{n,r} + p^r \alpha_{n-r,r} \tag{5-134}$$

For example, if $n = 25$, $r = 6$, $p = q = 1/2$, we get the probability of six successive heads in 25 trials to be 0.15775.

On a more interesting note, suppose for a regular commuter the morning commute takes 45 minutes under the best of conditions, the probability of which is assumed to be 1/5. Then there is a 67% chance for doing the trip within the best time at least once a week. However there is only about 13% chance of repeating it twice in a row in a week. This shows that especially the day after the "lucky day," one should allow extra travel time. Finally if the conditions for the return trip also are assumed to be the same, for a one week period the probability of doing two consecutive trips within the best time is 0.2733. ◀

The following problem has many varients and its solution goes back to Montmort (1708). It has been further generalized by Laplace and many others.

TABLE 5-1
Probability p_n in (5-134)

	$n = 5$		$n = 10$	
r	$p = 1/5$	$p = 1/3$	$p = 1/5$	$p = 1/3$
1	0.6723	0.8683	0.8926	0.9827
2	0.1347	0.3251	0.2733	0.5773
3	0.0208	0.0864	0.0523	0.2026
4	0.0029	0.0206	0.0093	0.0615
5	0.0003	0.0041	0.0016	0.0178
6	—	—	0.0003	0.0050

EXAMPLE 5-33

THE PAIRING PROBLEM

▶ A person writes n letters and addresses n envelopes. Then one letter is randomly placed into each envelope. What is the probability that at least one letter will reach its correct destination? What if $n \to \infty$?

SOLUTION

When a letter is placed into the envelope addressed to the intended person, let us refer to it as a coincidence. Let X_k represent the event that there are exactly k coincidences among the n envelopes. The events $X_0, X_1, \ldots, X_n$ form a partition since they are mutually exclusive and one of these events is bound to happen. Hence by the theorem of total probability

$$p_n(0) + p_n(1) + p_n(2) + \cdots + p_n(n) = 1 \tag{5-135}$$

where

$$p_n(k) \triangleq P\{X_k\} \tag{5-136}$$

To determine $p_n(k)$ let us examine the event X_k. There are $\binom{n}{k}$ number of ways of drawing k letters from a group of n, and to generate k coincidences, each such sequence should go into their intended envelopes with probability

$$\frac{1}{n} \cdot \frac{1}{n-1} \cdots \frac{1}{n-k+1}$$

while the remaining $n-k$ letters present no coincidences at all with probability $p_{n-k}(0)$. By the independence of these events, we get the probability of k coincidences for each sequence of k letters in a group of n to be

$$\frac{1}{n(n-1)\cdots(n-k+1)} p_{n-k}(0)$$

But there are $\binom{n}{k}$ such mutually exclusive sequences, and using (2-20) we get

$$p_n(k) = P\{X_k\} = \binom{n}{k} \frac{1}{n(n-1)\cdots(n-k+1)} p_{n-k}(0) = \frac{p_{n-k}(0)}{k!} \tag{5-137}$$

Since $p_n(n) = 1/n!$, equation (5-137) gives $p_0(0) = 1$. Substituting (5-137) into (5-135) term by term, we get

$$p_n(0) + \frac{p_{n-1}(0)}{1!} + \frac{p_{n-2}(0)}{2!} + \cdots + \frac{p_1(0)}{(n-1)!} + \frac{1}{n!} = 1 \tag{5-138}$$

which gives successively

$$p_1(0) = 0 \qquad p_2(0) = \tfrac{1}{2} \qquad p_3(0) = \tfrac{1}{6}$$

and to obtain an explicit expression for $p_n(0)$, define the moment generating function

$$\phi(z) = \sum_{n=0}^{\infty} p_n(0) z^n \tag{5-139}$$

Then

$$e^z \phi(z) = \left(\sum_{k=0}^{\infty} \frac{z^k}{k!} \right) \left(\sum_{n=0}^{\infty} p_n(0) z^n \right)$$

$$= 1 + z + z^2 + \cdots + z^n + \cdots = \frac{1}{1-z} \tag{5-140}$$

where we have made use of (5-138). Thus

$$\phi(z) = \frac{e^{-z}}{1-z} = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\frac{(-1)^k}{k!}\right)z^n$$

and on comparing with (5-139), we get

$$p_n(0) = \sum_{k=0}^{n}\frac{(-1)^k}{k!} \to \frac{1}{e} = 0.377879 \tag{5-141}$$

and using (5-137)

$$p_n(k) = \frac{1}{k!}\sum_{m=0}^{n-k}\frac{(-1)^m}{m!} \tag{5-142}$$

Thus

$$P\{\text{At least one letter reaches the correct destination}\}$$
$$= 1 - p_n(0) = 1 - \sum_{k=0}^{n}\frac{(-1)^k}{k!} \to 0.63212056 \tag{5-143}$$

Even for moderate n, this probability is close to 0.6321. Thus even if a mail delivery distributes letters in the most causal manner without undertaking any kind of sorting at all, there is still a 63% chance that at least one family will receive some mail addressed to them.

On a more serious note, by the same token, a couple trying to conceive has about 63% chance of succeeding in their efforts under normal conditions. The abundance of living organisms in Nature is a good testimony to the fact that odds are indeed tilted in favor of this process. ◀

Determination of the density of $g(\mathbf{x})$. We show next that characteristic functions can be used to determine the density $f_y(y)$ of the random variable $\mathbf{y} = g(\mathbf{x})$ in terms of the density $f_x(x)$ of $\mathbf{x}$.

From (5-58) it follows that the characteristic function

$$\Phi_y(\omega) = \int_{-\infty}^{\infty} e^{j\omega y} f_y(y)\,dy$$

of the random variable $\mathbf{y} = g(\mathbf{x})$ equals

$$\Phi_y(\omega) = E\{e^{j\omega g(\mathbf{x})}\} = \int_{-\infty}^{\infty} e^{j\omega g(x)} f_x(x)\,dx \tag{5-144}$$

If, therefore, the integral in (5-144) can be written in the form

$$\int_{-\infty}^{\infty} e^{j\omega y} h(y)\,dy$$

it will follow that (uniqueness theorem)

$$f_y(y) = h(y)$$

This method leads to simple results if the transformation $y = g(x)$ is one-to-one.

TABLE 5-2

Random variable	Probability density function $f_x(x)$	Mean	Variance	Characteristic function $\Phi_x(\omega)$
Normal or Gaussian $N(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}$, $-\infty < x < \infty$	μ	σ^2	$e^{j\mu\omega-\sigma^2\omega^2/2}$
Log-normal	$\frac{1}{x\sqrt{2\pi\sigma^2}}e^{-(\ln x-\mu)^2/2\sigma^2}$, $x \geq 0$,			
Exponential $E(\lambda)$	$\lambda e^{-\lambda x}, x \geq 0, \lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$(1 - j\omega/\lambda)^{-1}$
Gamma $G(\alpha, \beta)$	$\frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}}e^{-x/\beta}$, $x \geq 0, \alpha > 0, \beta > 0$	$\alpha\beta$	$\alpha\beta^2$	$(1 - j\omega\beta)^{-\alpha}$
Erlang-k	$\frac{(k\lambda)^k}{(k-1)!}x^{k-1}e^{-k\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{k\lambda^2}$	$(1 - j\omega/k\lambda)^{-k}$
Chi-square $\chi^2(n)$	$\frac{x^{n/2-1}}{2^{n/2}\Gamma(n/2)}e^{-x/2}, x \geq 0$	n	$2n$	$(1 - j2\omega)^{-n/2}$
Weibull	$\alpha x^{\beta-1}e^{-\alpha x^{\beta}/\beta}$, $x \geq 0, \alpha > 0, \beta > 0$	$\left(\frac{\beta}{\alpha}\right)^{1/\beta}\Gamma\left(1+\frac{1}{\beta}\right)$	$\left(\frac{\beta}{\alpha}\right)^{2/\beta}\left[\Gamma\left(1+\frac{2}{\beta}\right) - \left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^2\right]$	—
Rayleigh	$\frac{x}{\sigma^2}e^{-x^2/2\sigma^2}, x \geq 0$	$\sqrt{\frac{\pi}{2}}\sigma$	$(2 - \pi/2)\sigma^2$	$\left(1 + j\sqrt{\frac{\pi}{2}}\sigma\omega\right)e^{-\sigma^2\omega^2/2}$
Uniform $U(a, b)$	$\frac{1}{b-a}, a < x < b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{jb\omega} - e^{-ja\omega}}{j\omega(b-a)}$
Beta $\beta(\alpha, \beta)$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$, $0 < x < 1, \alpha > 0, \beta > 0$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	—
Cauchy	$\frac{\alpha/\pi}{(x-\mu)^2+\alpha^2}$, $-\infty < x < \infty, \alpha > 0$	—	∞	$e^{j\mu\omega}e^{-\alpha\lvert\omega\rvert}$
Rician	$\frac{x}{\sigma^2}e^{-\frac{x^2+a^2}{2\sigma^2}}I_0\left(\frac{ax}{\sigma^2}\right)$, $-\infty < x < \infty, a > 0$	$\sigma\frac{\sqrt{\pi}}{2}[(1+r)I_0(r/2)$ $+rI_1(r/2)]e^{-r/2}$, $r = a^2/2\sigma^2$	—	—
Nakagami	$\frac{2}{\Gamma(m)}\left(\frac{m}{\Omega}\right)^m x^{2m-1}e^{-\frac{m}{\Omega}x^2}$, $x > 0$	$\frac{\Gamma(m+1/2)}{\Gamma(m)}\sqrt{\frac{\Omega}{m}}$	$\Omega\left(1 - \frac{1}{m}\left(\frac{\Gamma(m+1/2)}{\Gamma(m)}\right)^2\right)$	—

TABLE 5-2
(Continued)

Random variable	Probability density function $f_x(x)$	Mean	Variance	Characteristic function $\Phi_x(\omega)$
Students' $t(n)$	$\dfrac{\Gamma((n+1)/2)}{\sqrt{\pi n}\,\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2}$, $-\infty < x < \infty$	0	$\dfrac{n}{n-2}, n > 2$	—
F-distribution	$\dfrac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)}\left(\dfrac{m}{n}\right)^{m/2} x^{m/2-1} \times \left(1+\dfrac{mx}{n}\right)^{-(m+n)/2}, x > 0$	$\dfrac{n}{n-2}, n > 2$	$\dfrac{n^2(2m+2n-4)}{m(n-2)^2(n-4)}, n > 4$	—
Bernoulli	$P(X=1) = p$, $P(X=0) = 1-p = q$	p	$p(1-p)$	$pe^{j\omega}+q$
Binomial $B(n, p)$	$\binom{n}{k}p^k q^{n-k}$, $k = 0, 1, 2, \ldots, n, p+q = 1$	np	npq	$(pe^{j\omega}+q)^n$
Poisson $P(\lambda)$	$e^{-\lambda}\dfrac{\lambda^k}{k!}, k = 0, 1, 2, \ldots, \infty$	λ	λ	$e^{-\lambda(1-e^{j\omega})}$
Hypergeometric	$\dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$, $\max(0, M+n-N) \le k \le \min(M, n)$	$\dfrac{nM}{N}$	$n\dfrac{M}{N}\left(1-\dfrac{M}{N}\right)\left(1-\dfrac{n-1}{N-1}\right)$	—
Geometric	pq^k, $k = 0, 1, 2, \ldots, \infty$	$\dfrac{q}{p}$	$\dfrac{q}{p^2}$	$\dfrac{p}{1-qe^{j\omega}}$
	or pq^{k-1}, $k = 1, 2, \ldots, \infty$, $p+q = 1$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	$\dfrac{p}{e^{-j\omega}-q}$
Pascal or negative binomial $NB(r, p)$	$\binom{r+k-1}{k}p^r q^k$, $k = 0, 1, 2, \ldots, \infty$	$\dfrac{rq}{p}$	$\dfrac{rq}{p^2}$	$\left(\dfrac{p}{1-qe^{-j\omega}}\right)^r$
	or $\binom{k-1}{r-1}p^r q^{k-r}$, $k = r, r+1, \ldots, \infty$, $p+q = 1$	$\dfrac{r}{p}$	$\dfrac{rq}{p^2}$	$\left(\dfrac{p}{e^{-j\omega}-q}\right)^r$
Discrete uniform	$1/N$, $k = 1, 2, \ldots, N$	$\dfrac{N+1}{2}$	$\dfrac{N^2-1}{12}$	$e^{j(N+1)\omega/2}\dfrac{\sin(N\omega/2)}{\sin(\omega/2)}$

EXAMPLE 5-34

▶ Suppose that x is $N(0;\sigma)$ and $y = ax^2$. Inserting into (5-144) and using the evenness of the integrand, we obtain

$$\Phi_y(\omega) = \int_{-\infty}^{\infty} e^{j\omega a x^2} f(x)\,dx = \frac{2}{\sigma\sqrt{2\pi}} \int_0^{\infty} e^{ja\omega x^2} e^{-x^2/2\sigma^2}\,dx$$

As x increases from 0 to ∞, the transformation $y = ax^2$ is one-to-one. Since

$$dy = 2ax\,dx = 2\sqrt{ay}\,dx$$

the last equation yields

$$\Phi_y(\omega) = \frac{2}{\sigma\sqrt{2\pi}} \int_0^{\infty} e^{j\omega y} e^{-y/2a\sigma^2} \frac{dy}{2\sqrt{ay}}$$

Hence

$$f_y(y) = \frac{e^{-y/2a\sigma^2}}{\sigma\sqrt{2\pi a y}} U(y) \tag{5-145}$$

in agreement with (5-7) and (5-22). ◀

EXAMPLE 5-35

▶ We assume finally that x is uniform in the interval $(-\pi/2, \pi/2)$ and $y = \sin x$. In this case

$$\Phi_y(\omega) = \int_{-\infty}^{\infty} e^{j\omega \sin x} f(x)\,dx = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} e^{j\omega \sin x}\,dx$$

As x increases from $-\pi/2$ to $\pi/2$, the function $y = \sin x$ increases from -1 to 1 and

$$dy = \cos x\,dx = \sqrt{1-y^2}\,dx$$

Hence

$$\Phi_y(\omega) = \frac{1}{\pi} \int_{-1}^{1} e^{j\omega y} \frac{dy}{\sqrt{1-y^2}}$$

This leads to the conclusion that

$$f_y(y) = \frac{1}{\pi\sqrt{1-y^2}} \qquad \text{for} \quad |y| < 1$$

and 0 otherwise, in agreement with (5-33). ◀

PROBLEMS

5-1 The random variable x is $N(5, 2)$ and $y = 2x + 4$. Find η_y, σ_y, and $f_y(y)$.

5-2 Find $F_y(y)$ and $f_y(y)$ if $y = -4x + 3$ and $f_x(x) = 2e^{-2x}U(x)$.

5-3 If the random variable x is $N(0, c^2)$ and $g(x)$ is the function in Fig. 5-4, find and sketch the distribution and the density of the random variable $y = g(x)$.

5-4 The random variable x is uniform in the interval $(-2c, 2c)$. Find and sketch $f_y(y)$ and $F_y(y)$ if $y = g(x)$ and $g(x)$ is the function in Fig. 5-3.

5-5 The random variable x is $N(0, b^2)$ and $g(x)$ is the function in Fig. 5-5. Find and sketch $f_y(y)$ and $F_y(y)$.

5-6 The random variable $\mathbf{x}$ is uniform in the interval (0, 1). Find the density of the random variable $\mathbf{y} = -\ln \mathbf{x}$.

5-7 We place at random 200 points in the interval (0, 100). The distance from 0 to the first random point is a random variable $\mathbf{z}$. Find $F_z(z)$ (*a*) exactly and (*b*) using the Poisson approximation.

5-8 If $\mathbf{y} = \sqrt{\mathbf{x}}$, and $\mathbf{x}$ is an exponential random variable, show that $\mathbf{y}$ represents a Rayleigh random variable.

5-9 Express the density $f_y(y)$ of the random variable $\mathbf{y} = g(\mathbf{x})$ in terms of $f_x(x)$ if (*a*) $g(x) = |x|$; (*b*) $g(x) = e^{-x}U(x)$.

5-10 Find $F_y(y)$ and $f_y(y)$ if $F_x(x) = (1 - e^{-2x})U(x)$ and (*a*) $\mathbf{y} = (\mathbf{x} - 1)U(\mathbf{x} - 1)$; (*b*) $\mathbf{y} = \mathbf{x}^2$.

5-11 Show that, if the random variable $\mathbf{x}$ has a Cauchy density with $\alpha = 1$ and $\mathbf{y} = \arctan \mathbf{x}$, then $\mathbf{y}$ is uniform in the interval $(-\pi/2, \pi/2)$.

5-12 The random variable $\mathbf{x}$ is uniform in the interval $(-2\pi, 2\pi)$. Find $f_y(y)$ if (*a*) $\mathbf{y} = \mathbf{x}^3$, (*b*) $\mathbf{y} = \mathbf{x}^4$, and (*c*) $\mathbf{y} = 2\sin(3\mathbf{x} + 40^\circ)$.

5-13 The random variable $\mathbf{x}$ is uniform in the interval $(-1, 1)$. Find $g(x)$ such that if $\mathbf{y} = g(\mathbf{x})$ then $f_y(y) = 2e^{-2y}U(y)$.

5-14 Given that random variable $\mathbf{x}$ is of continuous type, we form the random variable $\mathbf{y} = g(\mathbf{x})$. (*a*) Find $f_y(y)$ if $g(x) = 2F_x(x) + 4$. (*b*) Find $g(x)$ such that $\mathbf{y}$ is uniform in the interval (8, 10).

5-15 A fair coin is tossed 10 times and $\mathbf{x}$ equals the number of heads. (*a*) Find $F_x(x)$. (*b*) Find $F_y(y)$ if $\mathbf{y} = (\mathbf{x} - 3)^2$.

5-16 If $\mathbf{x}$ represents a beta random variable with parameters α and β, show that $1 - \mathbf{x}$ also represents a beta random variable with parameters β and α.

5-17 Let $\mathbf{x}$ represent a chi-square random variable with n degrees of freedom. Then $\mathbf{y} = \mathbf{x}^2$ is known as the chi-distribution with n degrees of freedom. Determine the p.d.f of $\mathbf{y}$.

5-18 Let $\mathbf{x} \sim U(0, 1)$. Show that $\mathbf{y} = -2\log \mathbf{x}$ is $\chi^2(2)$.

5-19 If $\mathbf{x}$ is an exponential random variable with parameter λ, show that $\mathbf{y} = \mathbf{x}^{1/\beta}$ has a Weibull distribution.

5-20 If $\mathbf{t}$ is a random variable of continuous type and $\mathbf{y} = a \sin \omega\mathbf{t}$, show that

$$f_y(y) \xrightarrow[\omega \to \infty]{} \begin{cases} 1/\pi\sqrt{a^2 - y^2} & |y| < a \\ 0 & |y| > a \end{cases}$$

5-21 Show that if $\mathbf{y} = \mathbf{x}^2$, then

$$f_y(y \mid \mathbf{x} \geq 0) = \frac{U(y)}{1 - F_x(0)} \frac{f_x(\sqrt{y})}{2\sqrt{y}}$$

5-22 (*a*) Show that if $\mathbf{y} = a\mathbf{x} + b$, then $\sigma_y = |a|\sigma_x$. (*b*) Find η_y and σ_y if $\mathbf{y} = (\mathbf{x} - \eta_x)/\sigma_x$.

5-23 Show that if $\mathbf{x}$ has a Rayleigh density with parameter α and $\mathbf{y} = b + c\mathbf{x}^2$, then $\sigma_y^2 = 4c^2\sigma^4$.

5-24 If $\mathbf{x}$ is $N(0, 4)$ and $\mathbf{y} = 3\mathbf{x}^2$, find η_y, σ_y, and $f_y(y)$.

5-25 Let $\mathbf{x}$ represent a binomial random variable with parameters n and p. Show that (*a*) $E(\mathbf{x}) = np$; (*b*) $E[\mathbf{x}(\mathbf{x} - 1)] = n(n - 1)p^2$; (*c*) $E[\mathbf{x}(\mathbf{x} - 1)(\mathbf{x} - 2)] = n(n - 1)(n - 2)p^3$; (*d*) Compute $E(\mathbf{x}^2)$ and $E(\mathbf{x}^3)$.

5-26 For a Poisson random variable $\mathbf{x}$ with parameter λ show that (*a*) $P(0 < \mathbf{x} < 2\lambda) > (\lambda - 1)/\lambda$; (*b*) $E[\mathbf{x}(\mathbf{x} - 1)] = \lambda^2$, $E[\mathbf{x}(\mathbf{x} - 1)(\mathbf{x} - 2)] = \lambda^3$.

5-27 Show that if $U = [A_1, \ldots, A_n]$ is a partition of S, then

$$E\{\mathbf{x}\} = E\{\mathbf{x} \mid A_1\}P(A_1) + \cdots + E\{\mathbf{x} \mid A_n\}P(A_n).$$

5-28 Show that if $\mathbf{x} \geq 0$ and $E\{\mathbf{x}\} = \eta$, then $P\{\mathbf{x} \geq \sqrt{\eta}\} \leq \sqrt{\eta}$.

5-29 Using (5-86), find $E\{\mathbf{x}^3\}$ if $\eta_x = 10$ and $\sigma_x = 2$.

5-30 If $\mathbf{x}$ is uniform in the interval (10,12) and $\mathbf{y} = \mathbf{x}^3$. (*a*) find $f_y(y)$; (*b*) find $E\{\mathbf{y}\}$: (i) exactly (ii) using (5-86).

5-31 The random variable $\mathbf{x}$ is $N(100, 9)$. Find approximately the mean of the random variable $\mathbf{y} = 1/\mathbf{x}$ using (5-86).

5-32 (*a*) Show that if m is the median of $\mathbf{x}$, then

$$E\{|\mathbf{x} - a|\} = E\{|\mathbf{x} - m|\} + 2\int_a^m (x - a) f(x)\,dx$$

for any a. (*b*) Find c such that $E\{|\mathbf{x} - c|\}$ is minimum.

5-33 Show that if the random variable $\mathbf{x}$ is $N(\eta; \sigma^2)$, then

$$E\{|\mathbf{x}|\} = \sigma\sqrt{\frac{2}{\pi}}e^{-\eta^2/2\sigma^2} + 2\eta G\left(\frac{\eta}{\sigma}\right) - \eta$$

5-34 Show that if $\mathbf{x}$ and $\mathbf{y}$ are two random variables with densities $f_x(x)$ and $f_y(y)$, respectively, then

$$E\{\log f_x(\mathbf{x})\} \geq E\{\log f_y(\mathbf{x})\}$$

5-35 (*Chernoff bound*) (*a*) Show that for any $\alpha > 0$ and for any real s,

$$P\{e^{s\mathbf{x}} \geq \alpha\} \leq \frac{\Phi(s)}{\alpha} \quad \text{where } \Phi(s) = E\{e^{s\mathbf{x}}\} \tag{i}$$

Hint: Apply (5-89) to the random variable $\mathbf{y} = e^{s\mathbf{x}}$. (*b*) For any A,

$$P\{\mathbf{x} \geq A\} \leq e^{-sA}\Phi(s) \quad s > 0$$

$$P\{\mathbf{x} \leq A\} \leq e^{-sA}\Phi(s) \quad s < 0$$

(*Hint:* Set $\alpha = e^{sA}$ in (i).)

5-36 Show that for any random variable $\mathbf{x}$

$$[E(|\mathbf{x}|^m)]^{1/m} \leq [E(|\mathbf{x}|^n)]^{1/n} \qquad 1 < m < n < \infty$$

5-37 Show that (*a*) if $f(x)$ is a Cauchy density, then $\Phi(\omega) = e^{-\alpha|\omega|}$; (*b*) if $f(x)$ is a Laplace density, then $\Phi(\omega) = \alpha^2/(\alpha^2 + \omega^2)$.

5-38 (*a*) Let $\mathbf{x} \sim G(\alpha, \beta)$. Show that $E\{\mathbf{x}\} = \alpha\beta$, $\mathrm{Var}\{\mathbf{x}\} = \alpha\beta^2$ and $\Phi_X(\omega) = (1 - \beta e^{j\omega})^{-\alpha}$.
(*b*) Let $\mathbf{x} \sim \chi^2(n)$. Show that $E\{\mathbf{x}\} = n$, $\mathrm{Var}\{\mathbf{x}\} = 2n$ and $\Phi_x(\omega) = (1 - 2e^{j\omega})^{-n/2}$.
(*c*) Let $\mathbf{x} \sim B(n, p)$. Show that $E\{\mathbf{x}\} = np$, $\mathrm{Var}\{\mathbf{x}\} = npq$ and $\Phi_x(\omega) = (pe^{j\omega} + q)^n$.
(*d*) Let $\mathbf{x} \sim NB(r, p)$. Show that $\Phi_x(\omega) = p^r(1 - qe^{j\omega})^{-r}$.

5-39 A random variable $\mathbf{x}$ has a *geometric* distribution if

$$P\{\mathbf{x} = k\} = pq^k \quad k = 0, 1, \ldots \quad p + q = 1$$

Find $\Gamma(z)$ and show that $\eta_x = q/p$, $\sigma_x^2 = q/p^2$

5-40 Let $\mathbf{x}$ denote the event *"the number of failures that precede the n^{th} success"* so that $\mathbf{x} + n$ represents the total number of trials needed to generate n successes. In that case, the event $\{\mathbf{x} = k\}$ occurs if and only if the last trial results in a success and among the previous $(\mathbf{x} + n - 1)$ trials there are $n - 1$ successes (or $\mathbf{x}$ failures). This gives an alternate formulation for the *Pascal* (or *negative binomial*) distribution as follows: (see Table 5-2)

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

find $\Gamma(z)$ and show that $\eta_x = nq/p$, $\sigma_x^2 = nq/p^2$.

5-41 Let $\mathbf{x}$ be a negative binomial random variable with parameters r and p. Show that as $p \to 1$ and $r \to \infty$ such that $r(1-p) \to \lambda$, a constant, then

$$P(\mathbf{x} = n + r) \to e^{-\lambda}\frac{\lambda^n}{n!} \qquad n = 0, 1, 2, \ldots$$

5-42 Show that if $E\{\mathbf{x}\} = \eta$, then

$$E\{e^{s\mathbf{x}}\} = e^{s\eta}\sum_{n=0}^{\infty}\mu_n\frac{s^n}{n!} \qquad \mu_n = E\{(\mathbf{x}-\eta)^n\}$$

5-43 Show that if $\Phi_x(\omega_1) = 1$ for some $\omega_1 \neq 0$, then the random variable $\mathbf{x}$ is of lattice type taking the values $x_n = 2\pi n/\omega_1$.

Hint:

$$0 = 1 - \Phi_x(\omega_1) = \int_{-\infty}^{\infty}(1 - e^{j\omega_1 x})f_x(x)\,dx$$

5-44 The random variable $\mathbf{x}$ has zero mean, central moments μ_n, and cumulants λ_n. Show that $\lambda_3 = \mu_3$, $\lambda_4 = \mu_4 - 3\mu_2^2$; if $\mathbf{y}$ is $N(0; \sigma_y^2)$ and $\sigma_y = \sigma_x$, then $E\{\mathbf{x}^4\} = E\{\mathbf{y}^4\} + \lambda_4$.

5-45 The random variable $\mathbf{x}$ takes the values $0, 1, \ldots$ with $P\{\mathbf{x} = k\} = p_k$. Show that if

$$\mathbf{y} = (\mathbf{x}-1)U(\mathbf{x}-1) \qquad \text{then } \Gamma_y(z) = p_0 + z^{-1}[\Gamma_x(z) - p_0]$$

$$\eta_y = \eta_x - 1 + p_0 \qquad E\{\mathbf{y}^2\} = E\{\mathbf{x}^2\} - 2\eta_x + 1 - p_0$$

5-46 Show that, if $\Phi(\omega) = E\{e^{j\omega\mathbf{x}}\}$, then for any a_i,

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\Phi(\omega_i - \omega_j)a_i a_j^* \geq 0$$

Hint:

$$E\left\{\left|\sum_{i=1}^{n} a_j e^{j\omega_i \mathbf{x}}\right|^2\right\} \geq 0$$

5-47 We are given an even convex function $g(x)$ and a random variable $\mathbf{x}$ whose density $f(x)$ is symmetrical as in Fig. P5-47 with a single maximum at $x = \eta$. Show that the mean $E\{g(\mathbf{x}-a)\}$ of the random variable $g(\mathbf{x}-a)$ is minimum if $a = \eta$.

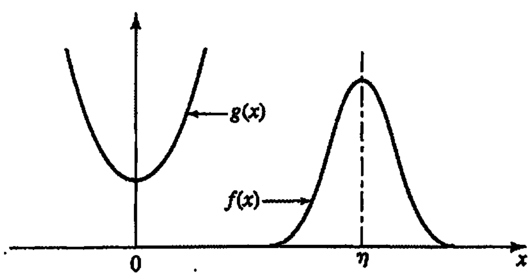


FIGURE P5-47

5-48 The random variable $\mathbf{x}$ is $N(0; \sigma^2)$. (*a*) Using characteristic functions, show that if $g(x)$ is a function such that $g(x)e^{-x^2/2\sigma^2} \to 0$ as $|x| \to \infty$, then (Price's theorem)

$$\frac{dE\{g(\mathbf{x})\}}{dv} = \frac{1}{2}E\left\{\frac{d^2g(\mathbf{x})}{d\mathbf{x}^2}\right\} \qquad v = \sigma^2 \tag{i}$$

(*b*) The moments μ_n of x are functions of v. Using (i), show that

$$\mu_n(v) = \frac{n(n-1)}{2}\int_0^v \mu_{n-2}(\beta)\,d\beta$$

5-49 Show that, if x is an integer-valued random variable with moment function $\Gamma(z)$ as in (5-113), then

$$P\{x = k\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Gamma(e^{j\omega})e^{-jk\omega}\,d\omega$$

5-50 A biased coin is tossed and the first outcome is noted. The tossing is continued until the outcome is the complement of the first outcome, thus completing the first run. Let x denote the length of the first run. Find the p.m.f of x, and show that

$$E\{x\} = \frac{p}{q} + \frac{q}{p}$$

5-51 A box contains N identical items of which $M < N$ are defective ones. A sample of size n is taken from the box, and let x represent the number of defective items in this sample.
(*a*) Find the distribution function of x if the n samples are drawn with replacement.
(*b*) If the n samples are drawn *without* replacement, then show that

$$P\{x = k\} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \qquad \max(0, n+M-N) \le k \le \min(M, N)$$

Find the mean and variance of x. The distribution in (*b*) is known as the *hypergeometric* distribution (see also Problem 3-5). The lottery distribution in (3-39) is an example of this distribution.
(*c*) In (*b*), let $N \to \infty$, $M \to \infty$, such that $M/N \to p$, $0 < p < 1$. Then show that the hypergeometric random variable can be approximated by a Binomial random variable with parameters n and p, provided $n \ll N$.

5-52 A box contains n white and m black marbles. Let x represent the number of draws needed for the rth white marble.
(*a*) If sampling is done with replacement, show that x has a negative binomial distribution with parameters r and $p = n/(m+n)$. (*b*) If sampling is done without replacement, then show that

$$P\{x = k\} = \binom{k-1}{r-1}\frac{\binom{m+n-k}{n-r}}{\binom{m+n}{n}} \qquad k = r, r+1, \ldots, m+n$$

(*c*) For a given k and r, show that the probability distribution in (*b*) tends to a negative binomial distribution as $n + m \to \infty$. Thus, for large population size, sampling with or without replacement is the same.