CHAPTER 2

Random Variables and Probability Distributions

Random Variables

Suppose that to each point of a sample space we assign a number. We then have a *function* defined on the sample space. This function is called a *random variable* (or *stochastic variable*) or more precisely a *random function* (*stochastic function*). It is usually denoted by a capital letter such as X or Y. In general, a random variable has some specified physical, geometrical, or other significance.

EXAMPLE 2.1 Suppose that a coin is tossed twice so that the sample space is $S = \{HH, HT, TH, TT\}$. Let X represent the number of heads that can come up. With each sample point we can associate a number for X as shown in Table 2-1. Thus, for example, in the case of HH (i.e., 2 heads), $X = 2$ while for TH (1 head), $X = 1$. It follows that X is a random variable.

Table 2-1

Sample Point	HH	HT	TH	TT
X	2	1	1	0

It should be noted that many other random variables could also be defined on this sample space, for example, the square of the number of heads or the number of heads minus the number of tails.

A random variable that takes on a finite or countably infinite number of values (see page 4) is called a *discrete random variable* while one which takes on a noncountably infinite number of values is called a *nondiscrete random variable*.

Discrete Probability Distributions

Let X be a discrete random variable, and suppose that the possible values that it can assume are given by x_1, x_2, x_3, . . . , arranged in some order. Suppose also that these values are assumed with probabilities given by

$$P(X = x_k) = f(x_k) \qquad k = 1, 2, \ldots \tag{1}$$

It is convenient to introduce the *probability function*, also referred to as *probability distribution*, given by

$$P(X = x) = f(x) \tag{2}$$

For $x = x_k$, this reduces to (1) while for other values of x, $f(x) = 0$.

In general, $f(x)$ is a probability function if

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

where the sum in 2 is taken over all possible values of x.

EXAMPLE 2.2 Find the probability function corresponding to the random variable X of Example 2.1. Assuming that the coin is fair, we have

$$P(HH) = \frac{1}{4} \qquad P(HT) = \frac{1}{4} \qquad P(TH) = \frac{1}{4} \qquad P(TT) = \frac{1}{4}$$

Then

$$P(X = 0) = P(TT) = \frac{1}{4}$$

$$P(X = 1) = P(HT \cup TH) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(X = 2) = P(HH) = \frac{1}{4}$$

The probability function is thus given by Table 2-2.

Table 2-2

x	0	1	2
$f(x)$	1/4	1/2	1/4

Distribution Functions for Random Variables

The *cumulative distribution function*, or briefly the *distribution function*, for a random variable X is defined by

$$F(x) = P(X \leq x) \tag{3}$$

where x is any real number, i.e., $-\infty < x < \infty$.

The distribution function $F(x)$ has the following properties:

1. $F(x)$ is nondecreasing [i.e., $F(x) \leq F(y)$ if $x \leq y$].
2. $\lim_{x\to-\infty} F(x) = 0$; $\lim_{x\to\infty} F(x) = 1$.
3. $F(x)$ is continuous from the right [i.e., $\lim_{h\to 0^+} F(x + h) = F(x)$ for all x].

Distribution Functions for Discrete Random Variables

The distribution function for a discrete random variable X can be obtained from its probability function by noting that, for all x in $(-\infty, \infty)$,

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u) \tag{4}$$

where the sum is taken over all values u taken on by X for which $u \leq x$.

If X takes on only a finite number of values $x_1, x_2, \ldots, x_n$, then the distribution function is given by

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leq x < x_3 \\ \vdots & \vdots \\ f(x_1) + \cdots + f(x_n) & x_n \leq x < \infty \end{cases} \tag{5}$$

EXAMPLE 2.3 (a) Find the distribution function for the random variable X of Example 2.2. (b) Obtain its graph.

(a) The distribution function is

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ \frac{1}{4} & 0 \leq x < 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & 2 \leq x < \infty \end{cases}$$

(b) The graph of $F(x)$ is shown in Fig. 2-1.

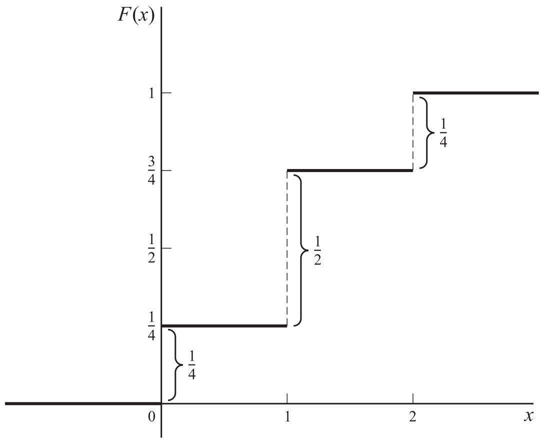

Fig. 2-1

The following things about the above distribution function, which are true in general, should be noted.

1. The magnitudes of the jumps at 0, 1, 2 are $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ which are precisely the probabilities in Table 2-2. This fact enables one to obtain the probability function from the distribution function.
2. Because of the appearance of the graph of Fig. 2-1, it is often called a *staircase function* or *step function*. The value of the function at an integer is obtained from the higher step; thus the value at 1 is $\frac{3}{4}$ and not $\frac{1}{4}$. This is expressed mathematically by stating that the distribution function is *continuous from the right* at 0, 1, 2.
3. As we proceed from left to right (i.e. going *upstairs*), the distribution function either remains the same or increases, taking on values from 0 to 1. Because of this, it is said to be a *monotonically increasing function*.

It is clear from the above remarks and the properties of distribution functions that the probability function of a discrete random variable can be obtained from the distribution function by noting that

$$f(x) = F(x) - \lim_{u \to x^-} F(u). \tag{6}$$

Continuous Random Variables

A nondiscrete random variable X is said to be *absolutely continuous*, or simply *continuous*, if its distribution function may be represented as

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du \qquad (-\infty < x < \infty) \tag{7}$$

where the function $f(x)$ has the properties

1. $f(x) \geq 0$
2. $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$

It follows from the above that if X is a continuous random variable, then the probability that X takes on any one particular value is zero, whereas the *interval probability* that X lies *between two different values*, say, a and b, is given by

$$P(a < X < b) = \int_{a}^{b} f(x)\,dx \tag{8}$$

EXAMPLE 2.4 If an individual is selected at random from a large group of adult males, the probability that his height X is precisely 68 inches (i.e., 68.000 . . . inches) would be zero. However, there is a probability greater than zero than X is between 67.000 . . . inches and 68.500 . . . inches, for example.

A function $f(x)$ that satisfies the above requirements is called a *probability function* or *probability distribution* for a continuous random variable, but it is more often called a *probability density function* or simply *density function*. Any function $f(x)$ satisfying Properties 1 and 2 above will automatically be a density function, and required probabilities can then be obtained from (8).

EXAMPLE 2.5 (a) Find the constant c such that the function

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

is a density function, and (b) compute $P(1 < X < 2)$.

(a) Since $f(x)$ satisfies Property 1 if $c \geq 0$, it must satisfy Property 2 in order to be a density function. Now

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^3 cx^2\,dx = \left.\frac{cx^3}{3}\right|_0^3 = 9c$$

and since this must equal 1, we have $c = 1/9$.

(b) $$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2\,dx = \left.\frac{x^3}{27}\right|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

In case $f(x)$ is continuous, which we shall assume unless otherwise stated, the probability that X is equal to any particular value is zero. In such case we can replace either or both of the signs $<$ in (8) by $\leq$. Thus, in Example 2.5,

$$P(1 \leq X \leq 2) = P(1 \leq X < 2) = P(1 < X \leq 2) = P(1 < X < 2) = \frac{7}{27}$$

EXAMPLE 2.6 (a) Find the distribution function for the random variable of Example 2.5. (b) Use the result of (a) to find $P(1 < x \leq 2)$.

(a) We have

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

If $x < 0$, then $F(x) = 0$. If $0 \leq x < 3$, then

$$F(x) = \int_0^x f(u)\,du = \int_0^x \frac{1}{9}u^2\,du = \frac{x^3}{27}$$

If $x \geq 3$, then

$$F(x) = \int_0^3 f(u)\,du + \int_3^x f(u)\,du = \int_0^3 \frac{1}{9}u^2\,du + \int_3^x 0\,du = 1$$

Thus the required distribution function is

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3/27 & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Note that $F(x)$ increases monotonically from 0 to 1 as is required for a distribution function. It should also be noted that $F(x)$ in this case is continuous.

(b) We have

$$P(1 < X \leq 2) = P(X \leq 2) - P(X \leq 1)$$
$$= F(2) - F(1)$$
$$= \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27}$$

as in Example 2.5.

The probability that X is between x and $x + \Delta x$ is given by

$$P(x \leq X \leq x + \Delta x) = \int_x^{x+\Delta x} f(u)\,du \tag{9}$$

so that if Δx is small, we have approximately

$$P(x \leq X \leq x + \Delta x) = f(x)\Delta x \tag{10}$$

We also see from (7) on differentiating both sides that

$$\frac{dF(x)}{dx} = f(x) \tag{11}$$

at all points where $f(x)$ is continuous; i.e., the derivative of the distribution function is the density function.

It should be pointed out that random variables exist that are neither discrete nor continuous. It can be shown that the random variable X with the following distribution function is an example.

$$F(x) = \begin{cases} 0 & x < 1 \\ \dfrac{x}{2} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

In order to obtain (11), we used the basic property

$$\frac{d}{dx}\int_a^x f(u)\,du = f(x) \tag{12}$$

which is one version of the Fundamental Theorem of Calculus.

Graphical Interpretations

If $f(x)$ is the density function for a random variable X, then we can represent $y = f(x)$ graphically by a curve as in Fig. 2-2. Since $f(x) \geq 0$, the curve cannot fall below the x axis. The entire area bounded by the curve and the x axis must be 1 because of Property 2 on page 36. Geometrically the probability that X is between a and b, i.e., $P(a < X < b)$, is then represented by the area shown shaded, in Fig. 2-2.

The distribution function $F(x) = P(X \leq x)$ is a monotonically increasing function which increases from 0 to 1 and is represented by a curve as in Fig. 2-3.

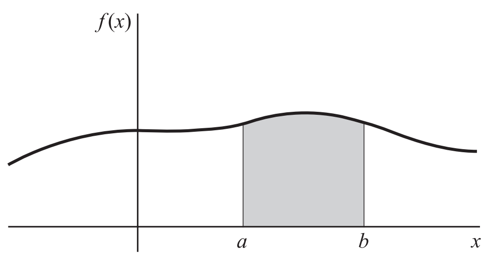

Fig. 2-2

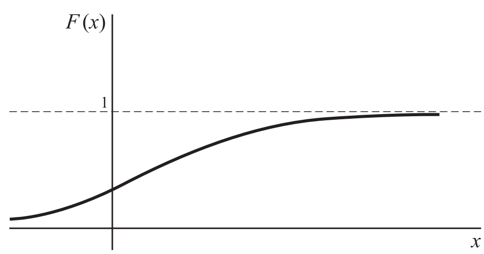

Fig. 2-3

Joint Distributions

The above ideas are easily generalized to two or more random variables. We consider the typical case of two random variables that are either both discrete or both continuous. In cases where one variable is discrete and the other continuous, appropriate modifications are easily made. Generalizations to more than two variables can also be made.

1. DISCRETE CASE. If X and Y are two discrete random variables, we define the *joint probability function* of X and Y by

$$P(X = x, Y = y) = f(x, y) \tag{13}$$

where 1. $f(x, y) \geq 0$

2. $\sum_x \sum_y f(x, y) = 1$

i.e., the sum over all values of x and y is 1.

Suppose that X can assume any one of m values $x_1, x_2, \ldots, x_m$ and Y can assume any one of n values $y_1, y_2, \ldots, y_n$. Then the probability of the event that $X = x_j$ and $Y = y_k$ is given by

$$P(X = x_j, Y = y_k) = f(x_j, y_k) \tag{14}$$

A joint probability function for X and Y can be represented by a *joint probability table* as in Table 2-3. The probability that $X = x_j$ is obtained by adding all entries in the row corresponding to x_i and is given by

$$P(X = x_j) = f_1(x_j) = \sum_{k=1}^{n} f(x_j, y_k) \tag{15}$$

Table 2-3

$X \backslash Y$	y_1	y_2	$\cdots$	y_n	Totals ↓
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	$\cdots$	$f(x_1, y_n)$	$f_1(x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	$\cdots$	$f(x_2, y_n)$	$f_1(x_2)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	$\cdots$	$f(x_m, y_n)$	$f_1(x_m)$
Totals →	$f_2(y_1)$	$f_2(y_2)$	$\cdots$	$f_2(y_n)$	1 ← Grand Total

For $j = 1, 2, \ldots, m$, these are indicated by the entry totals in the extreme right-hand column or margin of Table 2-3. Similarly the probability that $Y = y_k$ is obtained by adding all entries in the column corresponding to y_k and is given by

$$P(Y = y_k) = f_2(y_k) = \sum_{j=1}^{m} f(x_j, y_k) \tag{16}$$

For $k = 1, 2, \ldots, n$, these are indicated by the entry totals in the bottom row or margin of Table 2-3.

Because the probabilities (15) and (16) are obtained from the margins of the table, we often refer to $f_1(x_j)$ and $f_2(y_k)$ [or simply $f_1(x)$ and $f_2(y)$] as the *marginal probability functions* of X and Y, respectively.

It should also be noted that

$$\sum_{j=1}^{m} f_1(x_j) = 1 \quad \sum_{k=1}^{n} f_2(y_k) = 1 \tag{17}$$

which can be written

$$\sum_{j=1}^{m}\sum_{k=1}^{n} f(x_j, y_k) = 1 \tag{18}$$

This is simply the statement that the total probability of all entries is 1. The *grand total* of 1 is indicated in the lower right-hand corner of the table.

The *joint distribution function* of X and Y is defined by

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x}\sum_{v \leq y} f(u, v) \tag{19}$$

In Table 2-3, $F(x, y)$ is the sum of all entries for which $x_j \leq x$ and $y_k \leq y$.

2. CONTINUOUS CASE. The case where both variables are continuous is obtained easily by analogy with the discrete case on replacing sums by integrals. Thus the *joint probability function* for the random variables X and Y (or, as it is more commonly called, the *joint density function* of X and Y) is defined by

1. $f(x, y) \geq 0$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$

Graphically $z = f(x, y)$ represents a surface, called the *probability surface*, as indicated in Fig. 2-4. The total volume bounded by this surface and the xy plane is equal to 1 in accordance with Property 2 above. The probability that X lies between a and b while Y lies between c and d is given graphically by the shaded volume of Fig. 2-4 and mathematically by

$$P(a < X < b, c < Y < d) = \int_{x=a}^{b}\int_{y=c}^{d} f(x, y)\,dx\,dy \tag{20}$$

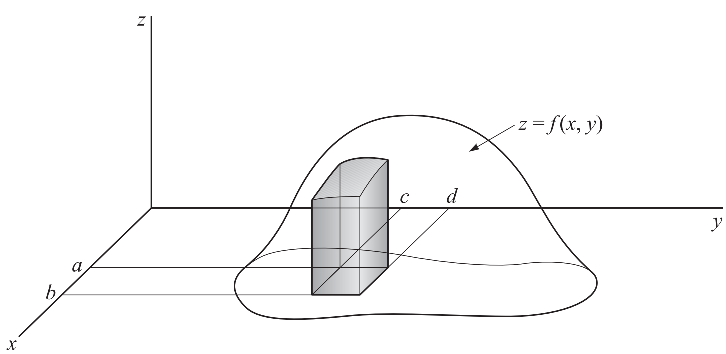

Fig. 2-4

More generally, if A represents any event, there will be a region $\mathcal{R}_A$ of the xy plane that corresponds to it. In such case we can find the probability of A by performing the integration over $\mathcal{R}_A$, i.e.,

$$P(A) = \iint_{\mathcal{R}_A} f(x, y)\,dx\,dy \tag{21}$$

The *joint distribution function* of X and Y in this case is defined by

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{y} f(u, v)\,du\,dv \tag{22}$$

It follows in analogy with (11), page 38, that

$$\frac{\partial^2 F}{\partial x\,\partial y} = f(x, y) \tag{23}$$

i.e., the density function is obtained by differentiating the distribution function with respect to x and y.

From (22) we obtain

$$P(X \le x) = F_1(x) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{\infty} f(u, v)\,du\,dv \tag{24}$$

$$P(Y \le y) = F_2(y) = \int_{u=-\infty}^{\infty} \int_{v=-\infty}^{y} f(u, v)\,du\,dv \tag{25}$$

We call (24) and (25) the *marginal distribution functions*, or simply the *distribution functions*, of X and Y, respectively. The derivatives of (24) and (25) with respect to x and y are then called the *marginal density functions,* or simply the *density functions*, of X and Y and are given by

$$f_1(x) = \int_{v=-\infty}^{\infty} f(x, v)\,dv \qquad f_2(y) = \int_{u=-\infty}^{\infty} f(u, y)\,du \tag{26}$$

Independent Random Variables

Suppose that X and Y are discrete random variables. If the events $X = x$ and $Y = y$ are independent events for all x and y, then we say that X and Y are *independent random variables*. In such case,

$$P(X = x, Y = y) = P(X = x)P(Y = y) \tag{27}$$

or equivalently

$$f(x, y) = f_1(x)f_2(y) \tag{28}$$

Conversely, if for all x and y the joint probability function $f(x, y)$ can be expressed as the product of a function of x alone and a function of y alone (which are then the marginal probability functions of X and Y), X and Y are independent. If, however, $f(x, y)$ cannot be so expressed, then X and Y are *dependent*.

If X and Y are continuous random variables, we say that they are *independent random variables* if the events $X \le x$ and $Y \le y$ are independent events for all x and y. In such case we can write

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y) \tag{29}$$

or equivalently

$$F(x, y) = F_1(x)F_2(y) \tag{30}$$

where $F_1(z)$ and $F_2(y)$ are the (marginal) distribution functions of X and Y, respectively. Conversely, X and Y are independent random variables if for all x and y, their joint distribution function $F(x, y)$ can be expressed as a product of a function of x alone and a function of y alone (which are the marginal distributions of X and Y, respectively). If, however, $F(x, y)$ cannot be so expressed, then X and Y are dependent.

For continuous independent random variables, it is also true that the joint density function $f(x, y)$ is the product of a function of x alone, $f_1(x)$, and a function of y alone, $f_2(y)$, and these are the (marginal) density functions of X and Y, respectively.

Change of Variables

Given the probability distributions of one or more random variables, we are often interested in finding distributions of other random variables that depend on them in some specified manner. Procedures for obtaining these distributions are presented in the following theorems for the case of discrete and continuous variables.

1. DISCRETE VARIABLES

Theorem 2-1 Let X be a discrete random variable whose probability function is $f(x)$. Suppose that a discrete random variable U is defined in terms of X by $U = \phi(X)$, where to each value of X there corresponds one and only one value of U and conversely, so that $X = \psi(U)$. Then the probability function for U is given by

$$g(u) = f[\psi(u)] \tag{31}$$

Theorem 2-2 Let X and Y be discrete random variables having joint probability function $f(x, y)$. Suppose that two discrete random variables U and V are defined in terms of X and Y by $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$, where to each pair of values of X and Y there corresponds one and only one pair of values of U and V and conversely, so that $X = \psi_1(U, V)$, $Y = \psi_2(U, V)$. Then the joint probability function of U and V is given by

$$g(u, v) = f[\psi_1(u, v), \psi_2(u, v)] \tag{32}$$

2. CONTINUOUS VARIABLES

Theorem 2-3 Let X be a continuous random variable with probability density $f(x)$. Let us define $U = \phi(X)$ where $X = \psi(U)$ as in Theorem 2-1. Then the probability density of U is given by $g(u)$ where

$$g(u)|du| = f(x)|dx| \tag{33}$$

or

$$g(u) = f(x)\left|\frac{dx}{du}\right| = f[\psi(u)]\,|\psi'(u)| \tag{34}$$

Theorem 2-4 Let X and Y be continuous random variables having joint density function $f(x, y)$. Let us define $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$ where $X = \psi_1(U, V)$, $Y = \psi_2(U, V)$ as in Theorem 2-2. Then the joint density function of U and V is given by $g(u, v)$ where

$$g(u, v)|du\,dv| = f(x, y)|dx\,dy| \tag{35}$$

or

$$g(u, v) = f(x, y)\left|\frac{\partial(x, y)}{\partial(u, v)}\right| = f[\psi_1(u, v), \psi_2(u, v)]\,|J| \tag{36}$$

In (36) the *Jacobian determinant*, or briefly *Jacobian*, is given by

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\[2ex] \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} \tag{37}$$

Probability Distributions of Functions of Random Variables

Theorems 2-2 and 2-4 specifically involve joint probability functions of two random variables. In practice one often needs to find the probability distribution of some specified function of several random variables. Either of the following theorems is often useful for this purpose.

Theorem 2-5 Let X and Y be continuous random variables and let $U = \phi_1(X, Y)$, $V = X$ (the second choice is arbitrary). Then the density function for U is the marginal density obtained from the joint density of U and V as found in Theorem 2-4. A similar result holds for probability functions of discrete variables.

Theorem 2-6 Let $f(x, y)$ be the joint density function of X and Y. Then the density function $g(u)$ of the random variable $U = \phi_1(X, Y)$ is found by differentiating with respect to u the distribution function given by

$$G(u) = P[\phi_1(X, Y) \leq u] = \iint_{\mathcal{R}} f(x, y)\, dx\, dy \tag{38}$$

Where $\mathcal{R}$ is the region for which $\phi_1(x, y) \leq u$.

Convolutions

As a particular consequence of the above theorems, we can show (see Problem 2.23) that the density function of the sum of two continuous random variables X and Y, i.e., of $U = X + Y$, having joint density function $f(x, y)$ is given by

$$g(u) = \int_{-\infty}^{\infty} f(x, u - x)\, dx \tag{39}$$

In the special case where X and Y are independent, $f(x, y) = f_1(x)f_2(y)$, and (39) reduces to

$$g(u) = \int_{-\infty}^{\infty} f_1(x) f_2(u - x)\, dx \tag{40}$$

which is called the *convolution* of f_1 and f_2, abbreviated, $f_1 * f_2$.

The following are some important properties of the convolution:

1. $f_1 * f_2 = f_2 * f_1$
2. $f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3$
3. $f_1 * (f_2 + f_3) = f_1 * f_2 + f_1 * f_3$

These results show that f_1, f_2, f_3 obey the *commutative*, *associative*, and *distributive laws* of algebra with respect to the operation of convolution.

Conditional Distributions

We already know that if $P(A) > 0$,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{41}$$

If X and Y are discrete random variables and we have the events (A: $X = x$), (B: $Y = y$), then (41) becomes

$$P(Y = y \mid X = x) = \frac{f(x, y)}{f_1(x)} \tag{42}$$

where $f(x, y) = P(X = x, Y = y)$ is the joint probability function and $f_1(x)$ is the marginal probability function for X. We define

$$f(y \mid x) \equiv \frac{f(x, y)}{f_1(x)} \tag{43}$$

and call it the *conditional probability function of Y given X*. Similarly, the conditional probability function of X given Y is

$$f(x \mid y) \equiv \frac{f(x, y)}{f_2(y)} \tag{44}$$

We shall sometimes denote $f(x \mid y)$ and $f(y \mid x)$ by $f_1(x \mid y)$ and $f_2(y \mid x)$, respectively.

These ideas are easily extended to the case where X, Y are continuous random variables. For example, the *conditional density function of Y given X* is

$$f(y \mid x) \equiv \frac{f(x, y)}{f_1(x)} \tag{45}$$

where $f(x, y)$ is the joint density function of X and Y, and $f_1(x)$ is the marginal density function of X. Using (45) we can, for example, find that the probability of Y being between c and d given that $x < X < x + dx$ is

$$P(c < Y < d \mid x < X < x + dx) = \int_c^d f(y \mid x)\,dy \tag{46}$$

Generalizations of these results are also available.

Applications to Geometric Probability

Various problems in probability arise from geometric considerations or have geometric interpretations. For example, suppose that we have a target in the form of a plane region of area K and a portion of it with area K_1, as in Fig. 2-5. Then it is reasonable to suppose that the probability of hitting the region of area K_1 is proportional to K_1. We thus define

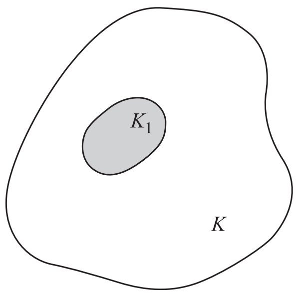

Fig. 2-5

$$P(\text{hitting region of area } K_1) = \frac{K_1}{K} \tag{47}$$

where it is assumed that the probability of hitting the target is 1. Other assumptions can of course be made. For example, there could be less probability of hitting outer areas. The type of assumption used defines the probability distribution function.

SOLVED PROBLEMS

Discrete random variables and probability distributions

2.1. Suppose that a pair of fair dice are to be tossed, and let the random variable X denote the sum of the points. Obtain the probability distribution for X.

The sample points for tosses of a pair of dice are given in Fig. 1-9, page 14. The random variable X is the sum of the coordinates for each point. Thus for (3, 2) we have $X = 5$. Using the fact that all 36 sample points are equally probable, so that each sample point has probability $1/36$, we obtain Table 2-4. For example, corresponding to $X = 5$, we have the sample points (1, 4), (2, 3), (3, 2), (4, 1), so that the associated probability is $4/36$.

Table 2-4

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

2.2. Find the probability distribution of boys and girls in families with 3 children, assuming equal probabilities for boys and girls.

Problem 1.37 treated the case of n mutually independent trials, where each trial had just two possible outcomes, A and A', with respective probabilities p and $q = 1 - p$. It was found that the probability of getting exactly x A's in the n trials is ${}_nC_x p^x q^{n-x}$. This result applies to the present problem, under the assumption that successive births (the "trials") are independent as far as the sex of the child is concerned. Thus, with A being the event "a boy," $n = 3$, and $p = q = \frac{1}{2}$, we have

$$P(\text{exactly } x \text{ boys}) = P(X = x) = {}_3C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x} = {}_3C_x\left(\frac{1}{2}\right)^3$$

where the random variable X represents the number of boys in the family. (Note that X is defined on the sample space of 3 trials.) The probability function for X,

$$f(x) = {}_3C_x\left(\frac{1}{2}\right)^3$$

is displayed in Table 2-5.

Table 2-5

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

Discrete distribution functions

2.3. (a) Find the distribution function $F(x)$ for the random variable X of Problem 2.1, and (b) graph this distribution function.

(a) We have $F(x) = P(X \le x) = \sum_{u \le x} f(u)$. Then from the results of Problem 2.1, we find

$$F(x) = \begin{cases} 0 & -\infty < x < 2 \\ 1/36 & 2 \le x < 3 \\ 3/36 & 3 \le x < 4 \\ 6/36 & 4 \le x < 5 \\ \vdots & \vdots \\ 35/36 & 11 \le x < 12 \\ 1 & 12 \le x < \infty \end{cases}$$

(b) See Fig. 2-6.

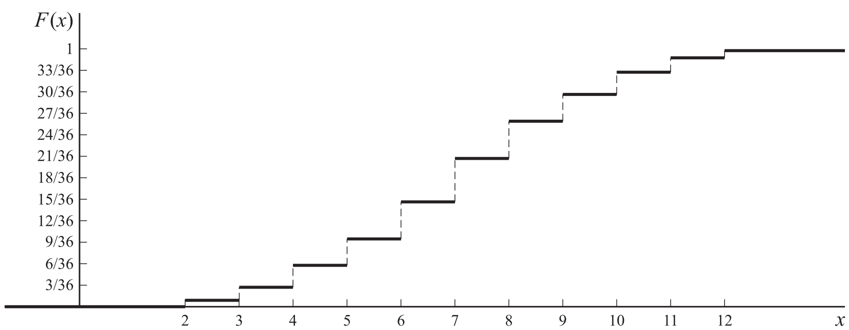

Fig. 2-6

2.4. (a) Find the distribution function $F(x)$ for the random variable X of Problem 2.2, and (b) graph this distribution function.

(a) Using Table 2-5 from Problem 2.2, we obtain

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ 1/8 & 0 \le x < 1 \\ 1/2 & 1 \le x < 2 \\ 7/8 & 2 \le x < 3 \\ 1 & 3 \le x < \infty \end{cases}$$

(b) The graph of the distribution function of (a) is shown in Fig. 2-7.

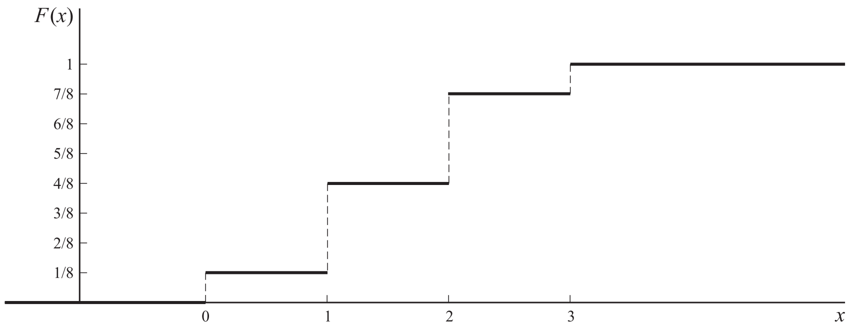

Fig. 2-7

Continuous random variables and probability distributions

2.5. A random variable X has the density function $f(x) = c/(x^2 + 1)$, where $-\infty < x < \infty$. (*a*) Find the value of the constant c. (b) Find the probability that X^2 lies between $1/3$ and 1.

(a) We must have $\int_{-\infty}^{\infty} f(x)\,dx = 1$, i.e.,

$$\int_{-\infty}^{\infty} \frac{c\,dx}{x^2 + 1} = c \tan^{-1} x \Big|_{-\infty}^{\infty} = c\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = 1$$

so that $c = 1/\pi$.

(b) If $\frac{1}{3} \le X^2 \le 1$, then either $\frac{\sqrt{3}}{3} \le X \le 1$ or $-1 \le X \le -\frac{\sqrt{3}}{3}$. Thus the required probability is

$$\frac{1}{\pi}\int_{-1}^{-\sqrt{3}/3} \frac{dx}{x^2 + 1} + \frac{1}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2 + 1} = \frac{2}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2 + 1}$$

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

$$= \frac{2}{\pi}\left(\frac{\pi}{4} - \frac{\pi}{6}\right) = \frac{1}{6}$$

2.6. Find the distribution function corresponding to the density function of Problem 2.5.

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \frac{1}{\pi}\int_{-\infty}^{x} \frac{du}{u^2 + 1} = \frac{1}{\pi}\left[\tan^{-1} u \Big|_{-\infty}^{x}\right]$$

$$= \frac{1}{\pi}[\tan^{-1} x - \tan^{-1}(-\infty)] = \frac{1}{\pi}\left[\tan^{-1} x + \frac{\pi}{2}\right]$$

$$= \frac{1}{2} + \frac{1}{\pi}\tan^{-1} x$$

2.7. The distribution function for a random variable X is

$$F(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the density function, (b) the probability that $X > 2$, and (c) the probability that $-3 < X \leq 4$.

(a)
$$f(x) = \frac{d}{dx}F(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x < 0 \end{cases}$$

(b)
$$P(X > 2) = \int_2^\infty 2e^{-2u}\,du = -e^{-2u}\Big|_2^\infty = e^{-4}$$

Another method

By definition, $P(X \leq 2) = F(2) = 1 - e^{-4}$. Hence,

$$P(X > 2) = 1 - (1 - e^{-4}) = e^{-4}$$

(c)
$$\begin{aligned} P(-3 < X \leq 4) &= \int_{-3}^{4} f(u)\,du = \int_{-3}^{0} 0\,du + \int_0^4 2e^{-2u}\,du \\ &= -e^{-2u}\Big|_0^4 = 1 - e^{-8} \end{aligned}$$

Another method

$$\begin{aligned} P(-3 < X \leq 4) &= P(X \leq 4) - P(X \leq -3) \\ &= F(4) - F(-3) \\ &= (1 - e^{-8}) - (0) = 1 - e^{-8} \end{aligned}$$

Joint distributions and independent variables

2.8. The joint probability function of two discrete random variables X and Y is given by $f(x, y) = c(2x + y)$, where x and y can assume all integers such that $0 \leq x \leq 2$, $0 \leq y \leq 3$, and $f(x, y) = 0$ otherwise.

(a) Find the value of the constant c. (c) Find $P(X \geq 1, Y \leq 2)$.
(b) Find $P(X = 2, Y = 1)$.

(a) The sample points (x, y) for which probabilities are different from zero are indicated in Fig. 2-8. The probabilities associated with these points, given by $c(2x + y)$, are shown in Table 2-6. Since the grand total, $42c$, must equal 1, we have $c = 1/42$.

Table 2-6

X \ Y	0	1	2	3	Totals ↓
0	0	c	$2c$	$3c$	$6c$
1	$2c$	$3c$	$4c$	$5c$	$14c$
2	$4c$	$5c$	$6c$	$7c$	$22c$
Totals →	$6c$	$9c$	$12c$	$15c$	$42c$

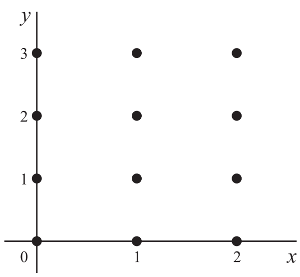

Fig. 2-8

(b) From Table 2-6 we see that

$$P(X = 2, Y = 1) = 5c + \frac{5}{42}$$

(c) From Table 2-6 we see that

$$\begin{aligned} P(X \geq 1, Y \leq 2) &= \sum_{x \geq 1} \sum_{y \leq 2} f(x, y) \\ &= (2c + 3c + 4c)(4c + 5c + 6c) \\ &= 24c = \frac{24}{42} = \frac{4}{7} \end{aligned}$$

as indicated by the entries shown shaded in the table.

2.9. Find the marginal probability functions (a) of X and (b) of Y for the random variables of Problem 2.8.

(a) The marginal probability function for X is given by $P(X = x) = f_1(x)$ and can be obtained from the margin totals in the right-hand column of Table 2-6. From these we see that

$$P(X = x) = f_1(x) = \begin{cases} 6c = 1/7 & x = 0 \\ 14c = 1/3 & x = 1 \\ 22c = 11/21 & x = 2 \end{cases}$$

Check: $\frac{1}{7} + \frac{1}{3} + \frac{11}{21} = 1$

(b) The marginal probability function for Y is given by $P(Y = y) = f_2(y)$ and can be obtained from the margin totals in the last row of Table 2-6. From these we see that

$$P(Y = y) = f_2(y) = \begin{cases} 6c = 1/7 & y = 0 \\ 9c = 3/14 & y = 1 \\ 12c = 2/7 & y = 2 \\ 15c = 5/14 & y = 3 \end{cases}$$

Check: $\frac{1}{7} + \frac{3}{14} + \frac{2}{7} + \frac{5}{14} = 1$

2.10. Show that the random variables X and Y of Problem 2.8 are dependent.

If the random variables X and Y are independent, then we must have, for all x and y,

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

But, as seen from Problems 2.8(b) and 2.9,

$$P(X = 2, Y = 1) = \frac{5}{42} \qquad P(X = 2) = \frac{11}{21} \qquad P(Y = 1) = \frac{3}{14}$$

so that

$$P(X = 2, Y = 1) \neq P(X = 2)P(Y = 1)$$

The result also follows from the fact that the joint probability function $(2x + y)/42$ cannot be expressed as a function of x alone times a function of y alone.

2.11. The joint density function of two continuous random variables X and Y is

$$f(x, y) = \begin{cases} cxy & 0 < x < 4, 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the value of the constant c.
(b) Find $P(1 < X < 2, 2 < Y < 3)$.
(c) Find $P(X \geq 3, Y \leq 2)$.

(a) We must have the total probability equal to 1, i.e.,

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

Using the definition of $f(x, y)$, the integral has the value

$$\int_{x=0}^{4}\int_{y=1}^{5} cxy\,dx\,dy = c\int_{x=0}^{4}\left[\int_{y=1}^{5} xy\,dy\right]dx$$

$$= c\int_{z=0}^{4} \frac{xy^2}{2}\bigg|_{y=1}^{5} dx = c\int_{x=0}^{4}\left(\frac{25x}{2} - \frac{x}{2}\right)dx$$

$$= c\int_{x=0}^{4} 12x\,dx = c(6x^2)\bigg|_{x=0}^{4} = 96c$$

Then $96c = 1$ and $c = 1/96$.

(b) Using the value of c found in (a), we have

$$P(1 < X < 2, 2 < Y < 3) = \int_{x=1}^{2}\int_{y=2}^{3} \frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=1}^{2}\left[\int_{y=2}^{3} xy\,dy\right]dx = \frac{1}{96}\int_{x=1}^{2} \frac{xy^2}{2}\bigg|_{y=2}^{3} dx$$

$$= \frac{1}{96}\int_{x=1}^{2} \frac{5x}{2}\,dx = \frac{5}{192}\left(\frac{x^2}{2}\right)\bigg|_{1}^{2} = \frac{5}{128}$$

(c)

$$P(X \geq 3, Y \leq 2) = \int_{x=3}^{4}\int_{y=1}^{2} \frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=3}^{4}\left[\int_{y=1}^{2} xy\,dy\right]dx = \frac{1}{96}\int_{x=3}^{4} \frac{xy^2}{2}\bigg|_{y=1}^{2} dx$$

$$= \frac{1}{96}\int_{x=3}^{4} \frac{3x}{2}\,dx = \frac{7}{128}$$

2.12. Find the marginal distribution functions (a) of X and (b) of Y for Problem 2.11.

(a) The marginal distribution function for X if $0 \leq x < 4$ is

$$F_1(x) = P(X \leq x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv$$

$$= \int_{u=0}^{x}\int_{v=1}^{5} \frac{uv}{96}\,du\,dv$$

$$= \frac{1}{96}\int_{u=0}^{x}\left[\int_{v=1}^{5} uv\,dv\right]du = \frac{x^2}{16}$$

For $x \geq 4$, $F_1(x) = 1$; for $x < 0$, $F_1(x) = 0$. Thus

$$F_1(x) = \begin{cases} 0 & x < 0 \\ x^{2/16} & 0 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

As $F_1(x)$ is continuous at $x = 0$ and $x = 4$, we could replace $<$ by $\leq$ in the above expression.

(b) The marginal distribution function for Y if $1 \le y < 5$ is

$$F_2(y) = P(Y \le y) = \int_{u=-\infty}^{\infty} \int_{v=1}^{y} f(u, v)\,du\,dv$$

$$= \int_{u=0}^{4} \int_{v=1}^{y} \frac{uv}{96}\,du\,dv = \frac{y^2 - 1}{24}$$

For $y \ge 5$, $F_2(y) = 1$. For $y < 1$, $F_2(y) = 0$. Thus

$$F_2(y) = \begin{cases} 0 & y < 1 \\ (y^2 - 1)/24 & 1 \le y < 5 \\ 1 & y \ge 5 \end{cases}$$

As $F_2(y)$ is continuous at $y = 1$ and $y = 5$, we could replace $<$ by $\le$ in the above expression.

2.13. Find the joint distribution function for the random variables X, Y of Problem 2.11.

From Problem 2.11 it is seen that the joint density function for X and Y can be written as the product of a function of x alone and a function of y alone. In fact, $f(x, y) = f_1(x)f_2(y)$, where

$$f_1(x) = \begin{cases} c_1 x & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases} \qquad f_2(y) = \begin{cases} c_2 y & 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

and $c_1 c_2 = c = 1/96$. It follows that X and Y are independent, so that their joint distribution function is given by $F(x, y) = F_1(x)F_2(y)$. The marginal distributions $F_1(x)$ and $F_2(y)$ were determined in Problem 2.12, and Fig. 2-9 shows the resulting piecewise definition of $F(x, y)$.

2.14. In Problem 2.11 find $P(X + Y < 3)$.

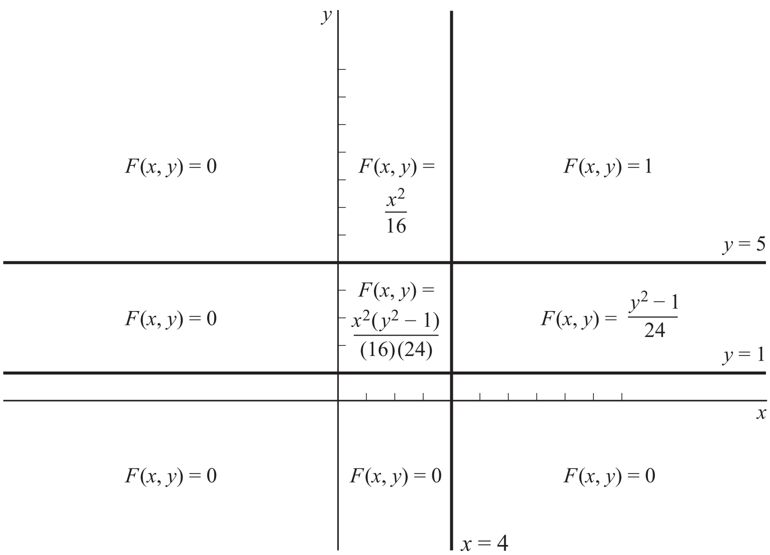

Fig. 2-9

In Fig. 2-10 we have indicated the square region $0 < x < 4$, $1 < y < 5$ within which the joint density function of X and Y is different from zero. The required probability is given by

$$P(X + Y < 3) = \iint_{\mathcal{R}} f(x, y)\,dx\,dy$$

where $\mathcal{R}$ is the part of the square over which $x + y < 3$, shown shaded in Fig. 2-10. Since $f(x, y) = xy/96$ over $\mathcal{R}$, this probability is given by

$$\int_{x=0}^{2}\int_{y=1}^{3-x}\frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=0}^{2}\left[\int_{y=1}^{3-x} xy\,dy\right]dx$$

$$= \frac{1}{96}\int_{x=0}^{2}\frac{xy^2}{2}\bigg|_{y=1}^{3-x} dx = \frac{1}{192}\int_{x=0}^{2}[x(3-x)^2 - x] = \frac{1}{48}$$

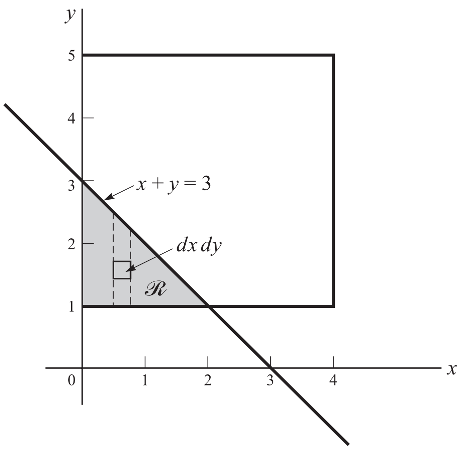

Fig. 2-10

Change of variables

2.15. Prove Theorem 2-1, page 42.

The probability function for U is given by

$$g(u) = P(U = u) = P[\phi(X) = u] = P[X = \psi(u)] = f[\psi(u)]$$

In a similar manner Theorem 2-2, page 42, can be proved.

2.16. Prove Theorem 2-3, page 42.

Consider first the case where $u = \phi(x)$ or $x = \psi(u)$ is an increasing function, i.e., u increases as x increases (Fig. 2-11). There, as is clear from the figure, we have

(1) $$P(u_1 < U < u_2) = P(x_1 < X < x_2)$$

or

(2) $$\int_{u_1}^{u_2} g(u)\,du = \int_{x_1}^{x_2} f(x)\,dx$$

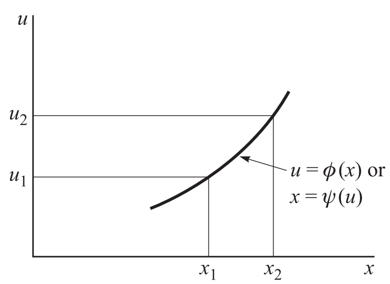

Fig. 2-11

Letting $x = \psi(u)$ in the integral on the right, (2) can be written

$$\int_{u_1}^{u_2} g(u)\,du = \int_{u_1}^{u_2} f[\psi(u)]\psi'(u)\,du$$

This can hold for all u_1 and u_2 only if the integrands are identical, i.e.,

$$g(u) = f[\psi(u)]\psi'(u)$$

This is a special case of (34), page 42, where $\psi'(u) > 0$ (i.e., the slope is positive). For the case where $\psi'(u) \leq 0$, i.e., u is a decreasing function of x, we can also show that (34) holds (see Problem 2.67). The theorem can also be proved if $\psi'(u) \geq 0$ or $\psi'(u) < 0$.

2.17. Prove Theorem 2-4, page 42.

We suppose first that as x and y increase, u and v also increase. As in Problem 2.16 we can then show that

$$P(u_1 < U < u_2, v_1 < V < v_2) = P(x_1 < X < x_2, y_1 < Y < y_2)$$

or
$$\int_{v_1}^{u_2}\int_{v_1}^{v_2} g(u, v)\,du\,dv = \int_{x_1}^{x_2}\int_{y_1}^{y_2} f(x, y)\,dx\,dy$$

Letting $x = \psi_1(u, v)$, $y = \psi_2(u, v)$ in the integral on the right, we have, by a theorem of advanced calculus,

$$\int_{v_1}^{u_2}\int_{v_1}^{v_2} g(u, v)\,du\,dv = \int_{u_1}^{u_2}\int_{v_1}^{v_2} f[\psi_1(u, v), \psi_2(u, v)]J\,du\,dv$$

where
$$J = \frac{\partial(x, y)}{\partial(u, v)}$$

is the *Jacobian*. Thus

$$g(u, v) = f[\psi_1(u, v), \psi_2(u, v)]J$$

which is (36), page 42, in the case where $J > 0$. Similarly, we can prove (36) for the case where $J < 0$.

2.18. The probability function of a random variable X is

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

Find the probability function for the random variable $U = X^4 + 1$.

Since $U = X^4 + 1$, the relationship between the values u and x of the random variables U and X is given by $u = x^4 + 1$ or $x = \sqrt[4]{u - 1}$, where $u = 2, 17, 82, \ldots$ and the real positive root is taken. Then the required probability function for U is given by

$$g(u) = \begin{cases} 2^{-\sqrt[4]{u-1}} & u = 2, 17, 82, \ldots \\ 0 & \text{otherwise} \end{cases}$$

using Theorem 2-1, page 42, or Problem 2.15.

2.19. The probability function of a random variable X is given by

$$f(x) = \begin{cases} x^2/81 & -3 < x < 6 \\ 0 & \text{otherwise} \end{cases}$$

Find the probability density for the random variable $U = \frac{1}{3}(12 - X)$.

We have $u = \frac{1}{3}(12 - x)$ or $x = 12 - 3u$. Thus to each value of x there is one and only one value of u and conversely. The values of u corresponding to $x = -3$ and $x = 6$ are $u = 5$ and $u = 2$, respectively. Since $\psi'(u) = dx/du = -3$, it follows by Theorem 2-3, page 42, or Problem 2.16 that the density function for U *is*

$$g(u) = \begin{cases} (12 - 3u)^2/27 & 2 < u < 5 \\ 0 & \text{otherwise} \end{cases}$$

Check:
$$\int_2^5 \frac{(12 - 3u)^2}{27}\, du = -\frac{(12 - 3u)^3}{243}\bigg|_2^5 = 1$$

2.20. Find the probability density of the random variable $U = X^2$ where X is the random variable of Problem 2.19.

We have $u = x^2$ or $x = \pm\sqrt{u}$. Thus to each value of x there corresponds one and only one value of u, but to each value of $u \neq 0$ there correspond *two* values of x. The values of x for which $-3 < x < 6$ correspond to values of u for which $0 \leq u < 36$ as shown in Fig. 2-12.

As seen in this figure, the interval $-3 < x \leq 3$ corresponds to $0 \leq u \leq 9$ while $3 < x < 6$ corresponds to $9 < u < 36$. In this case we cannot use Theorem 2-3 directly but can proceed as follows. The distribution function for U is

$$G(u) = P(U \leq u)$$

Now if $0 \leq u \leq 9$, we have

$$G(u) = P(U \leq u) = P(X^2 \leq u) = P(-\sqrt{u} \leq X \leq \sqrt{u})$$

$$= \int_{-\sqrt{u}}^{\sqrt{u}} f(x)\, dx$$

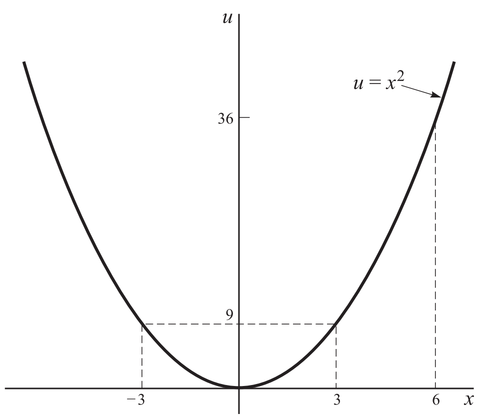

Fig. 2-12

But if $9 < u < 36$, we have

$$G(u) = P(U \leq u) = P(-3 < X < \sqrt{u}) = \int_{-3}^{\sqrt{u}} f(x)\, dx$$

Since the density function $g(u)$ is the derivative of $G(u)$, we have, using (12),

$$g(u) = \begin{cases} \dfrac{f(\sqrt{u}) + f(-\sqrt{u})}{2\sqrt{u}} & 0 \le u \le 9 \\ \dfrac{f(\sqrt{u})}{2\sqrt{u}} & 9 < u < 36 \\ 0 & \text{otherwise} \end{cases}$$

Using the given definition of $f(x)$, this becomes

$$g(u) = \begin{cases} \sqrt{u}/81 & 0 \le u \le 9 \\ \sqrt{u}/162 & 9 < u < 36 \\ 0 & \text{otherwise} \end{cases}$$

Check:

$$\int_0^9 \frac{\sqrt{u}}{81}\,du + \int_9^{36} \frac{\sqrt{u}}{162}\,du = \frac{2u^{3/2}}{243}\bigg|_0^9 + \frac{u^{3/2}}{243}\bigg|_9^{36} = 1$$

2.21. If the random variables X and Y have joint density function

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(see Problem 2.11), find the density function of $U = X + 2Y$.

Method 1
Let $u = x + 2y$, $v = x$, the second relation being chosen arbitrarily. Then simultaneous solution yields $x = v$, $y = \frac{1}{2}(u - v)$. Thus the region $0 < x < 4$, $1 < y < 5$ corresponds to the region $0 < v < 4$, $2 < u - v < 10$ shown shaded in Fig. 2-13.

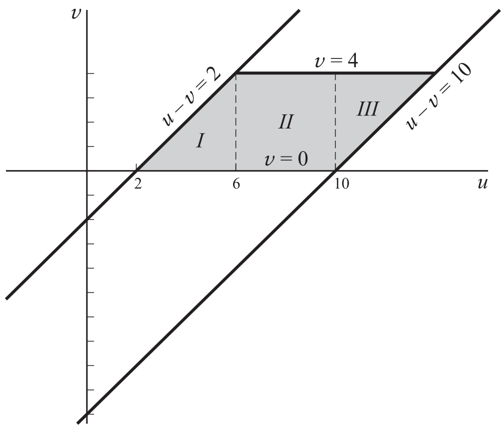

Fig. 2-13

The Jacobian is given by

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

Then by Theorem 2-4 the joint density function of U and V is

$$g(u, v) = \begin{cases} v(u - v)/384 & 2 < u - v < 10, 0 < v < 4 \\ 0 & \text{otherwise} \end{cases}$$

The marginal density function of U is given by

$$g_1(u) = \begin{cases} \displaystyle\int_{v=0}^{u-2} \frac{v(u - v)}{384}\,dv & 2 < u < 6 \\ \displaystyle\int_{v=0}^{4} \frac{v(u - v)}{384}\,dv & 6 < u < 10 \\ \displaystyle\int_{v=u-10}^{4} \frac{v(u - v)}{384}\,dv & 10 < u < 14 \\ 0 & \text{otherwise} \end{cases}$$

as seen by referring to the shaded regions *I*, *II*, *III* of Fig. 2-13. Carrying out the integrations, we find

$$g_1(u) = \begin{cases} (u - 2)^2(u + 4)/2304 & 2 < u < 6 \\ (3u - 8)/144 & 6 < u < 10 \\ (348u - u^3 - 2128)/2304 & 10 < u < 14 \\ 0 & \text{otherwise} \end{cases}$$

A check can be achieved by showing that the integral of $g_1(u)$ is equal to 1.

Method 2

The distribution function of the random variable $X + 2Y$ is given by

$$P(X + 2Y \leq u) = \iint\limits_{x+2y \leq u} f(x, y)\,dx\,dy = \iint\limits_{\substack{x+2y \leq u \\ 0<x<4 \\ 1<y<5}} \frac{xy}{96}\,dx\,dy$$

For $2 < u < 6$, we see by referring to Fig. 2-14, that the last integral equals

$$\int_{x=0}^{u-2}\int_{y=1}^{(u-x)/2} \frac{xy}{96}\,dx\,dy = \int_{x=0}^{u-2}\left[\frac{x(u - x)^2}{768} - \frac{x}{192}\right]dx$$

The derivative of this with respect to u is found to be $(u - 2)^2(u + 4)/2304$. In a similar manner we can obtain the result of Method 1 for $6 < u < 10$, etc.

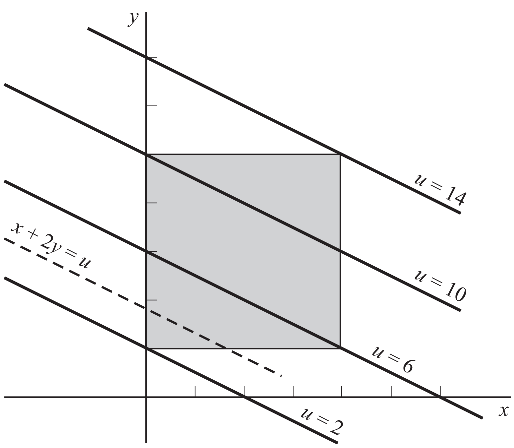

Fig. 2-14

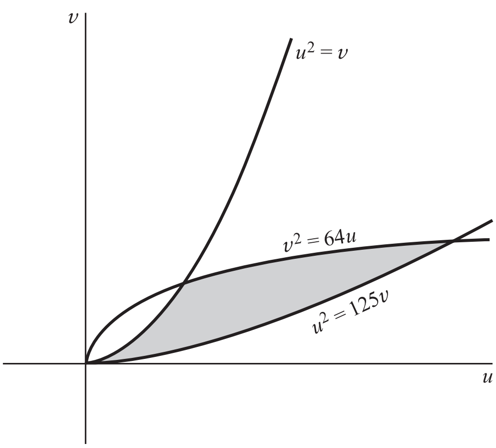

Fig. 2-15

2.22. If the random variables X and Y have joint density function

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4, 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(see Problem 2.11), find the joint density function of $U = XY^2$, $V = X^2Y$.

Consider $u = xy^2$, $v = x^2y$. Dividing these equations, we obtain $y/x = u/v$ so that $y = ux/v$. This leads to the simultaneous solution $x = v^{2/3}u^{-1/3}$, $y = u^{2/3}v^{-1/3}$. The image of $0 < x < 4$, $1 < y < 5$ in the uv-plane is given by

$$0 < v^{2/3}u^{-1/3} < 4 \qquad 1 < u^{2/3}v^{-1/3} < 5$$

which are equivalent to

$$v^2 < 64u \qquad v < u^2 < 125v$$

This region is shown shaded in Fig. 2-15.

The Jacobian is given by

$$J = \begin{vmatrix} -\frac{1}{3}v^{2/3}u^{-4/3} & \frac{2}{3}v^{-1/3}u^{-1/3} \\ \frac{2}{3}u^{-1/3}v^{-1/3} & -\frac{1}{3}u^{2/3}v^{-4/3} \end{vmatrix} = -\frac{1}{3}u^{-2/3}v^{-2/3}$$

Thus the joint density function of U and V is, by Theorem 2-4,

$$g(u, v) = \begin{cases} \dfrac{(v^{2/3}u^{-1/3})(u^{2/3}v^{-1/3})}{96}\left(\frac{1}{3}u^{-2/3}v^{-2/3}\right) & v^2 < 64u, v < u^2 < 125v \\ 0 & \text{otherwise} \end{cases}$$

or

$$g(u, v) = \begin{cases} u^{-1/3}v^{-1/3}/288 & v^2 < 64u, \quad v < u^2 < 125v \\ 0 & \text{otherwise} \end{cases}$$

Convolutions

2.23. Let X and Y be random variables having joint density function $f(x, y)$. Prove that the density function of $U = X + Y$ is

$$g(u) = \int_{-\infty}^{\infty} f(v, u - v)dv$$

Method 1

Let $U = X + Y$, $V = X$, where we have arbitrarily added the second equation. Corresponding to these we have $u = x + y$, $v = x$ or $x = v$, $y = u - v$. The Jacobian of the transformation is given by

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

Thus by Theorem 2-4, page 42, the joint density function of U and V is

$$g(u, v) = f(v, u - v)$$

It follows from (26), page 41, that the marginal density function of U is

$$g(u) = \int_{-\infty}^{\infty} f(v, u - v)dv$$

Method 2

The distribution function of $U = X + Y$ is equal to the double integral of $f(x, y)$ taken over the region defined by $x + y \leq u$, i.e.,

$$G(u) = \iint_{x+y \leq u} f(x, y)\, dx\, dy$$

Since the region is below the line $x + y = u$, as indicated by the shading in Fig. 2-16, we see that

$$G(u) = \int_{x=-\infty}^{\infty} \left[\int_{y=-\infty}^{u-x} f(x, y)\, dy \right] dx$$

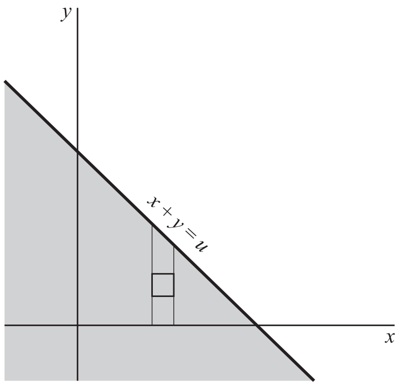

Fig. 2-16

The density function of U is the derivative of $G(u)$ with respect to u and is given by

$$g(u) = \int_{-\infty}^{\infty} f(x, u - x)\, dx$$

using (12) first on the x integral and then on the y integral.

2.24. Work Problem 2.23 if X and Y are independent random variables having density functions $f_1(x)$, $f_2(y)$, respectively.

In this case the joint density function is $f(x, y) = f_1(x) f_2(y)$, so that by Problem 2.23 the density function of $U = X + Y$ is

$$g(u) = \int_{-\infty}^{\infty} f_1(v) f_2(u - v) dv = f_1 * f_2$$

which is the *convolution* of f_1 and f_2.

2.25. If X and Y are independent random variables having density functions

$$f_1(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad f_2(y) = \begin{cases} 3e^{-3y} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

find the density function of their sum, $U = X + Y$.

By Problem 2.24 the required density function is the convolution of f_1 and f_2 and is given by

$$g(u) = f_1 * f_2 = \int_{-\infty}^{\infty} f_1(v) f_2(u - v)\, dv$$

In the integrand f_1 vanishes when $v < 0$ and f_2 vanishes when $v > u$. Hence

$$g(u) = \int_0^u (2e^{-2v})(3e^{-3(u-v)})\, dv$$

$$= 6e^{-3u} \int_0^u e^v\, dv = 6e^{-3u}(e^u - 1) = 6(e^{-2u} - e^{3u})$$

if $u \geq 0$ and $g(u) = 0$ if $u < 0$.

Check: $$\int_{-\infty}^{\infty} g(u)\,du = 6\int_{0}^{\infty}(e^{-2u} - e^{-3u})\,du = 6\left(\frac{1}{2} - \frac{1}{3}\right) = 1$$

2.26. Prove that $f_1 * f_2 = f_2 * f_1$ (Property 1, page 43).

We have

$$f_1 * f_2 = \int_{v=-\infty}^{\infty} f_1(v) f_2(u - v)\,dv$$

Letting $w = u - v$ so that $v = u - w$, $dv = -dw$, we obtain

$$f_1 * f_2 = \int_{w=\infty}^{-\infty} f_1(u - w) f_2(w)(-dw) = \int_{w=-\infty}^{\infty} f_2(w) f_1(u - w)\,dw = f_2 * f_1$$

Conditional distributions

2.27. Find (a) $f(y \mid 2)$, (b) $P(Y = 1 \mid X = 2)$ for the distribution of Problem 2.8.

(a) Using the results in Problems 2.8 and 2.9, we have

$$f(y \mid x) = \frac{f(x, y)}{f_1(x)} = \frac{(2x + y)/42}{f_1(x)}$$

so that with $x = 2$

$$f(y \mid 2) = \frac{(4 + y)/42}{11/21} = \frac{4 + y}{22}$$

(b) $$P(Y = 1 \mid X = 2) = f(1 \mid 2) = \frac{5}{22}$$

2.28. If X and Y have the joint density function

$$f(x, y) = \begin{cases} \frac{3}{4} + xy & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

find (a) $f(y \mid x)$, (b) $P(Y > \frac{1}{2} \mid \frac{1}{2} < X < \frac{1}{2} + dx)$.

(a) For $0 < x < 1$,

$$f_1(x) = \int_0^1 \left(\frac{3}{4} + xy\right) dy = \frac{3}{4} + \frac{x}{2}$$

and $$f(y \mid x) = \frac{f(x, y)}{f_1(x)} = \begin{cases} \dfrac{3 + 4xy}{3 + 2x} & 0 < y < 1 \\ 0 & \text{other } y \end{cases}$$

For other values of x, $f(y \mid x)$ is not defined.

(b) $$P(Y > \tfrac{1}{2} \mid \tfrac{1}{2} < X < \tfrac{1}{2} + dx) = \int_{1/2}^{\infty} f(y \mid \tfrac{1}{2})\,dy = \int_{1/2}^{1} \frac{3 + 2y}{4}\,dy = \frac{9}{16}$$

2.29. The joint density function of the random variables X and Y is given by

$$f(x, y) = \begin{cases} 8xy & 0 \leq x \leq 1,\ 0 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the marginal density of X, (b) the marginal density of Y, (c) the conditional density of X, (d) the conditional density of Y.

The region over which $f(x, y)$ is different from zero is shown shaded in Fig. 2-17.

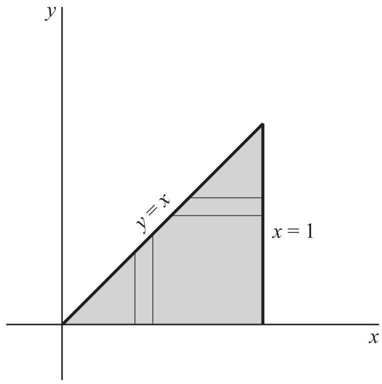

Fig. 2-17

(a) To obtain the marginal density of X, we fix x and integrate with respect to y from 0 to x as indicated by the vertical strip in Fig. 2-17. The result is

$$f_1(x) = \int_{y=0}^{x} 8xy\,dy = 4x^3$$

for $0 < x < 1$. For all other values of x, $f_1(x) = 0$.

(b) Similarly, the marginal density of Y is obtained by fixing y and integrating with respect to x from $x = y$ to $x = 1$, as indicated by the horizontal strip in Fig. 2-17. The result is, for $0 < y < 1$,

$$f_2(y) = \int_{x=y}^{1} 8xy\,dx = 4y(1 - y^2)$$

For all other values of y, $f_2(y) = 0$.

(c) The conditional density function of X is, for $0 < y < 1$,

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)} = \begin{cases} 2x/(1 - y^2) & y \le x \le 1 \\ 0 & \text{other } x \end{cases}$$

The conditional density function is not defined when $f_2(y) = 0$.

(d) The conditional density function of Y is, for $0 < x < 1$,

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)} = \begin{cases} 2y/x^2 & 0 \le y \le x \\ 0 & \text{other } y \end{cases}$$

The conditional density function is not defined when $f_1(x) = 0$.

Check: $$\int_0^1 f_1(x)\,dx = \int_0^1 4x^3\,dx = 1, \quad \int_0^1 f_2(y)\,dy = \int_0^1 4y(1 - y^2)\,dy = 1$$

$$\int_y^1 f_1(x \mid y)\,dx = \int_y^1 \frac{2x}{1 - y^2}\,dx = 1$$

$$\int_0^x f_2(y \mid x)\,dy = \int_0^x \frac{2y}{x^2}\,dy = 1$$

2.30. Determine whether the random variables of Problem 2.29 are independent.

In the shaded region of Fig. 2-17, $f(x, y) = 8xy$, $f_1(x) = 4x^3$, $f_2(y) = 4y(1 - y^2)$. Hence $f(x, y) \ne f_1(x)f_2(y)$, and thus X and Y are dependent.

It should be noted that it does not follow from $f(x, y) = 8xy$ that $f(x, y)$ can be expressed as a function of x alone times a function of y alone. This is because the restriction $0 \le y \le x$ occurs. If this were replaced by some restriction on y not depending on x (as in Problem 2.21), such a conclusion would be valid.

Applications to geometric probability

2.31. A person playing darts finds that the probability of the dart striking between r and $r + dr$ is

$$P(r \leq R \leq r + dr) = c\left[1 - \left(\frac{r}{a}\right)^2\right]dr$$

Here, R is the distance of the hit from the center of the target, c is a constant, and a is the radius of the target (see Fig. 2-18). Find the probability of hitting the bull's-eye, which is assumed to have radius b. Assume that the target is always hit.

The density function is given by

$$f(r) = c\left[1 - \left(\frac{r}{a}\right)^2\right]$$

Since the target is always hit, we have

$$c\int_0^a \left[1 - \left(\frac{r}{a}\right)^2\right]dr = 1$$

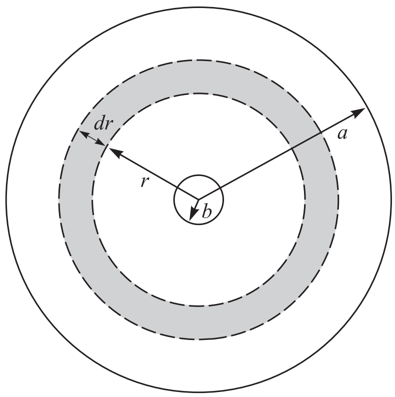

Fig. 2-18

from which $c = 3/2a$. Then the probability of hitting the bull's-eye is

$$\int_0^b f(r)\,dr = \frac{3}{2a}\int_0^b \left[1 - \left(\frac{r}{a}\right)^2\right]dr = \frac{b(3a^2 - b^2)}{2a^3}$$

2.32. Two points are selected at random in the interval $0 \leq x \leq 1$. Determine the probability that the sum of their squares is less than 1.

Let X and Y denote the random variables associated with the given points. Since equal intervals are assumed to have equal probabilities, the density functions of X and Y are given, respectively, by

(1) $$f_1(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \qquad f_2(y) = \begin{cases} 1 & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Then since X and Y are independent, the joint density function is given by

(2) $$f(x, y) = f_1(x)f_2(y) = \begin{cases} 1 & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

It follows that the required probability is given by

(3) $$P(X^2 + Y^2 \leq 1) = \iint_{\mathscr{R}} dx\,dy$$

where $\mathscr{R}$ is the region defined by $x^2 + y^2 \leq 1, x \geq 0, y \geq 0$, which is a quarter of a circle of radius 1 (Fig. 2-19). Now since (3) represents the area of $\mathscr{R}$, we see that the required probability is $\pi/4$.

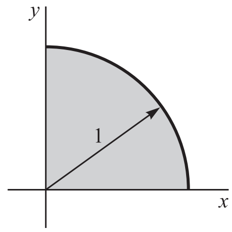

Fig. 2-19

Miscellaneous problems

2.33. Suppose that the random variables X and Y have a joint density function given by

$$f(x, y) = \begin{cases} c(2x + y) & 2 < x < 6, 0 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the constant c, (b) the marginal distribution functions for X and Y, (c) the marginal density functions for X and Y, (d) $P(3 < X < 4, Y > 2)$, (e) $P(X > 3)$, (f) $P(X + Y > 4)$, (g) the joint distribution function, (h) whether X and Y are independent.

(a) The total probability is given by

$$\int_{x=2}^{6}\int_{y=0}^{5} c(2x + y)\,dx\,dy = \int_{x=2}^{6} c\left(2xy + \frac{y^2}{2}\right)\Big|_0^5 dx$$

$$= \int_{x=2}^{6} c\left(10x + \frac{25}{2}\right)dx = 210c$$

For this to equal 1, we must have $c = 1/210$.

(b) The marginal distribution function for X is

$$F_1(x) = P(X \le x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv$$

$$= \begin{cases} \displaystyle\int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} 0\,du\,dv = 0 & x < 2 \\ \displaystyle\int_{u=2}^{x}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = \frac{2x^2 + 5x - 18}{84} & 2 \le x < 6 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = 1 & x \ge 6 \end{cases}$$

The marginal distribution function for Y is

$$F_2(y) = P(Y \le y) = \int_{u=-\infty}^{\infty}\int_{v=-\infty}^{y} f(u, v)\,du\,dv$$

$$= \begin{cases} \displaystyle\int_{u=-\infty}^{\infty}\int_{v=-8}^{y} 0\,du\,dv = 0 & y < 0 \\ \displaystyle\int_{u=0}^{6}\int_{v=0}^{y} \frac{2u + v}{210}\,du\,dv = \frac{y^2 + 16y}{105} & 0 \le y < 5 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = 1 & y \ge 5 \end{cases}$$

(c) The marginal density function for X is, from part (b),

$$f_1(x) = \frac{d}{dx}F_1(x) = \begin{cases} (4x + 5)/84 & 2 < x < 6 \\ 0 & \text{otherwise} \end{cases}$$

The marginal density function for Y is, from part (b),

$$f_2(y) = \frac{d}{dy}F_2(y) = \begin{cases} (2y + 16)/105 & 0 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(d)
$$P(3 < X < 4, Y > 2) = \frac{1}{210}\int_{x=3}^{4}\int_{y=2}^{5}(2x + y)\,dx\,dy = \frac{3}{20}$$

(e)
$$P(X > 3) = \frac{1}{210}\int_{x=3}^{6}\int_{y=0}^{5}(2x + y)\,dx\,dy = \frac{23}{28}$$

(f)
$$P(X + Y > 4) = \iint_{\mathcal{R}} f(x, y)\,dx\,dy$$

where $\mathcal{R}$ is the shaded region of Fig. 2-20. Although this can be found, it is easier to use the fact that

$$P(X + Y > 4) = 1 - P(X + Y \le 4) = 1 - \iint_{\mathcal{R}} f(x, y)\,dx\,dy$$

where $\mathcal{R}'$ is the cross-hatched region of Fig. 2-20. We have

$$P(X + Y \le 4) = \frac{1}{210}\int_{x=2}^{4}\int_{y=0}^{4-x}(2x + y)\,dx\,dy = \frac{2}{35}$$

Thus $P(X + Y > 4) = 33/35$.

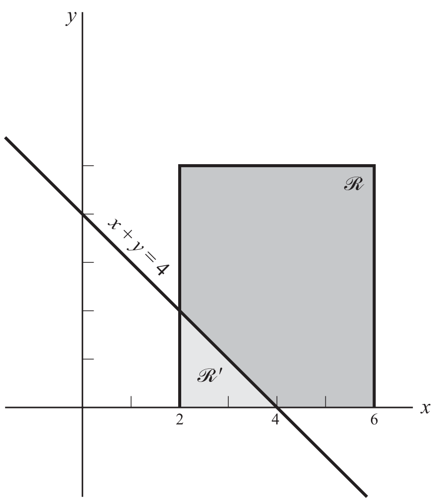

Fig. 2-20

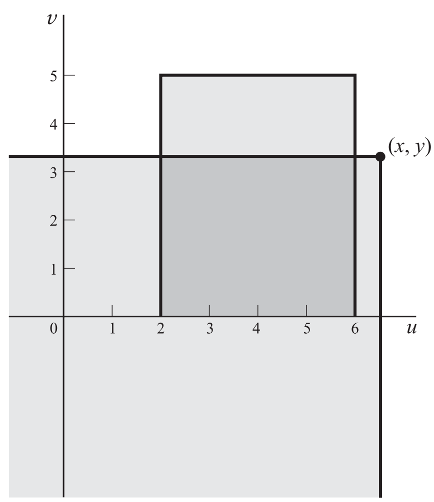

Fig. 2-21

(g) The joint distribution function is

$$F(x, y) = P(X \le x, Y \le y) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{y} f(u, v)\,du\,dv$$

In the uv plane (Fig. 2-21) the region of integration is the intersection of the quarter plane $u \le x$, $v \le y$ and the rectangle $2 < u < 6$, $0 < v < 5$ [over which $f(u, v)$ is nonzero]. For (x, y) located as in the figure, we have

$$F(x, y) = \int_{u=2}^{6}\int_{v=0}^{y}\frac{2u + v}{210}\,du\,dv = \frac{16y + y^2}{105}$$

When (x, y) lies inside the rectangle, we obtain another expression, etc. The complete results are shown in Fig. 2-22.

(h) The random variables are dependent since

$$f(x, y) \neq f_1(x)f_2(y)$$

or equivalently, $F(x, y) \neq F_1(x)F_2(y)$.

2.34. Let X have the density function

$$f(x) = \begin{cases} 6x(1 - x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find a function $Y = h(X)$ which has the density function

$$g(y) = \begin{cases} 12y^3(1 - y^2) & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

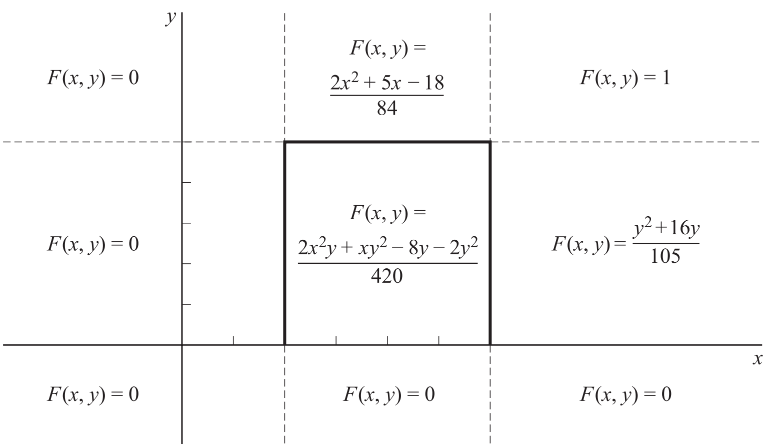

Fig. 2-22

We assume that the unknown function h is such that the intervals $X \leq x$ and $Y \leq y + h(x)$ correspond in a one-one, continuous fashion. Then $P(X \leq x) = P(Y \leq y)$, i.e., the distribution functions of X and Y must be equal. Thus, for $0 < x, y < 1$,

$$\int_0^x 6u(1 - u)\,du = \int_0^y 12v^3(1 - v^2)\,dv$$

or

$$3x^2 - 2x^3 = 3y^4 - 2y^6$$

By inspection, $x = y^2$ or $y = h(x) = +\sqrt{x}$ is a solution, and this solution has the desired properties. Thus $Y = +\sqrt{X}$.

2.35. Find the density function of $U = XY$ if the joint density function of X and Y is $f(x, y)$.

Method 1

Let $U = XY$ and $V = X$, corresponding to which $u = xy$, $v = x$ or $x = v$, $y = u/v$. Then the Jacobian is given by

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ v^{-1} & -uv^{-2} \end{vmatrix} = -v^{-1}$$

Thus the joint density function of U and V is

$$g(u, v) = \frac{1}{|v|} f\left(v, \frac{u}{v}\right)$$

from which the marginal density function of U is obtained as

$$g(u) = \int_{-\infty}^{\infty} g(u, v)\,dv = \int_{-\infty}^{\infty} \frac{1}{|v|} f\left(v, \frac{u}{v}\right) dv$$

Method 2

The distribution function of U is

$$G(u) = \iint\limits_{xy \le u} f(x, y)\,dx\,dy$$

For $u \ge 0$, the region of integration is shown shaded in Fig. 2-23. We see that

$$G(u) = \int_{-\infty}^{0} \left[\int_{u/x}^{\infty} f(x, y)\,dy\right] dx + \int_{0}^{\infty} \left[\int_{-\infty}^{u/x} f(x, y)\,dy\right] dx$$

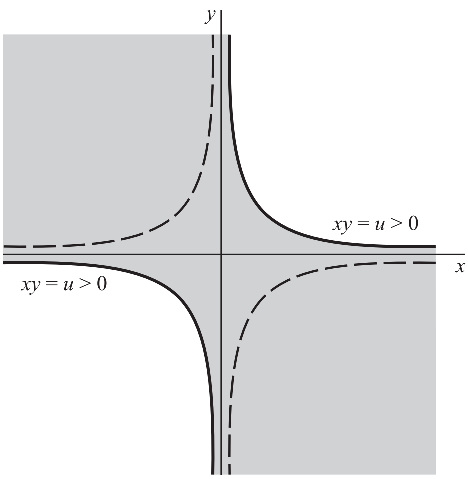

Fig. 2-23

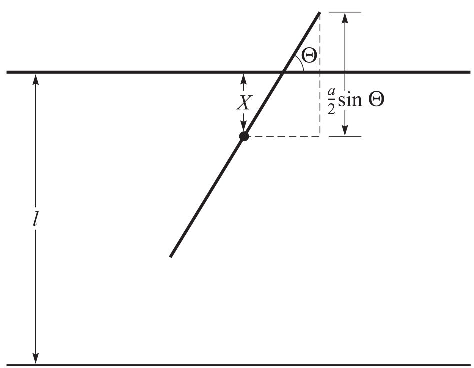

Fig. 2-24

Differentiating with respect to u, we obtain

$$g(u) = \int_{-\infty}^{0} \left(\frac{-1}{x}\right) f\left(x, \frac{u}{x}\right) dx + \int_{0}^{\infty} \frac{1}{x} f\left(x, \frac{u}{x}\right) dx = \int_{-\infty}^{\infty} \frac{1}{|x|} f\left(x, \frac{u}{x}\right) dx$$

The same result is obtained for $u < 0$, when the region of integration is bounded by the dashed hyperbola in Fig. 2-24.

2.36. A floor has parallel lines on it at equal distances l from each other. A needle of length $a < l$ is dropped at random onto the floor. Find the probability that the needle will intersect a line. (This problem is known as *Buffon's needle problem.*)

Let X be a random variable that gives the distance of the midpoint of the needle to the nearest line (Fig. 2-24). Let Θ be a random variable that gives the acute angle between the needle (or its extension) and the line. We denote by x and θ any particular values of X and Θ. It is seen that X can take on any value between 0 and $l/2$, so that $0 \le x \le l/2$. Also Θ can take on any value between 0 and $\pi/2$. It follows that

$$P(x < X \le x + dx) = \frac{2}{l}\,dx \qquad P(\theta \le \Theta + d\theta) = \frac{2}{\pi}\,d\theta$$

i.e., the density functions of X and Θ are given by $f_1(x) = 2/l$, $f_2(\theta) = 2/\pi$. As a check, we note that

$$\int_0^{l/2} \frac{2}{l}\,dx = 1 \qquad \int_0^{\pi/2} \frac{2}{\pi}\,d\theta = 1$$

Since X and Θ are independent the joint density function is

$$f(x, \theta) = \frac{2}{l} \cdot \frac{2}{\pi} = \frac{4}{l\pi}$$

From Fig. 2-24 it is seen that the needle actually hits a line when $X \le (a/2)\sin\Theta$. The probability of this event is given by

$$\frac{4}{l\pi}\int_{\theta=0}^{\pi/2}\int_{x=0}^{(a/2)\sin\theta} dx\,d\theta = \frac{2a}{l\pi}$$

When the above expression is equated to the frequency of hits observed in actual experiments, accurate values of π are obtained. This indicates that the probability model described above is appropriate.

2.37. Two people agree to meet between 2:00 P.M. and 3:00 P.M., with the understanding that each will wait no longer than 15 minutes for the other. What is the probability that they will meet?

Let X and Y be random variables representing the times of arrival, measured in fractions of an hour after 2:00 P.M., of the two people. Assuming that equal intervals of time have equal probabilities of arrival, the density functions of X and Y are given respectively by

$$f_1(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_2(y) = \begin{cases} 1 & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Then, since X and Y are independent, the joint density function is

(1) $$f(x, y) = f_1(x)f_2(y) = \begin{cases} 1 & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Since 15 minutes $= \frac{1}{4}$ hour, the required probability is

(2) $$P\left(|X - Y| \le \frac{1}{4}\right) = \iint_{\mathscr{R}} dx\,dy$$

where 5 is the region shown shaded in Fig. 2-25. The right side of (2) is the area of this region, which is equal to $1 - (\frac{3}{4})(\frac{3}{4}) = \frac{7}{16}$, since the square has area 1, while the two corner triangles have areas $\frac{1}{2}(\frac{3}{4})(\frac{3}{4})$ each. Thus the required probability is $7/16$.

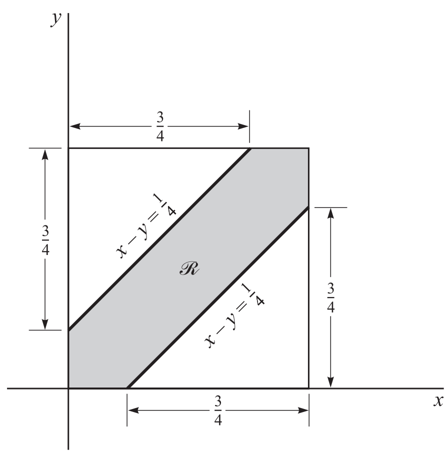

Fig. 2-25

SUPPLEMENTARY PROBLEMS

Discrete random variables and probability distributions

2.38. A coin is tossed three times. If X is a random variable giving the number of heads that arise, construct a table showing the probability distribution of X.

2.39. An urn holds 5 white and 3 black marbles. If 2 marbles are to be drawn at random without replacement and X denotes the number of white marbles, find the probability distribution for X.

2.40. Work Problem 2.39 if the marbles are to be drawn with replacement.

2.41. Let Z be a random variable giving the number of heads minus the number of tails in 2 tosses of a fair coin. Find the probability distribution of Z. Compare with the results of Examples 2.1 and 2.2.

2.42. Let X be a random variable giving the number of aces in a random draw of 4 cards from an ordinary deck of 52 cards. Construct a table showing the probability distribution of X.

Discrete distribution functions

2.43. The probability function of a random variable X is shown in Table 2-7. Construct a table giving the distribution function of X.

Table 2-7

x	1	2	3
$f(x)$	1/2	1/3	1/6

Table 2-8

x	1	2	3	4
$F(x)$	1/8	3/8	3/4	1

2.44. Obtain the distribution function for (a) Problem 2.38, (b) Problem 2.39, (c) Problem 2.40.

2.45. Obtain the distribution function for (a) Problem 2.41, (b) Problem 2.42.

2.46. Table 2-8 shows the distribution function of a random variable X. Determine (a) the probability function, (b) $P(1 \le X \le 3)$, (c) $P(X \ge 2)$, (d) $P(X < 3)$, (e) $P(X > 1.4)$.

Continuous random variables and probability distributions

2.47. A random variable X has density function

$$f(x) = \begin{cases} ce^{-3x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

Find (a) the constant c, (b) $P(1 < X < 2)$, (c) $P(X \ge 3)$, (d) $P(X < 1)$.

2.48. Find the distribution function for the random variable of Problem 2.47. Graph the density and distribution functions, describing the relationship between them.

2.49. A random variable X has density function

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

Find (a) the constant c, (b) $P(X > 2)$, (c) $P(1/2 < X < 3/2)$.

2.50. Find the distribution function for the random variable X of Problem 2.49.

2.51. The distribution function of a random variable X is given by

$$F(x) = \begin{cases} cx^3 & 0 \le x < 3 \\ 1 & x \ge 3 \\ 0 & x < 0 \end{cases}$$

If $P(X = 3) = 0$, find (a) the constant c, (b) the density function, (c) $P(X > 1)$, (d) $P(1 < X < 2)$.

2.52. Can the function

$$F(x) = \begin{cases} c(1 - x^2) & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

be a distribution function? Explain.

2.53. Let X be a random variable having density function

$$f(x) = \begin{cases} cx & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the value of the constant c, (b) $P(\frac{1}{2} < X < \frac{3}{2})$, (c) $P(X > 1)$, (d) the distribution function.

Joint distributions and independent variables

2.54. The joint probability function of two discrete random variables X and Y is given by $f(x, y) = cxy$ for $x = 1, 2, 3$ and $y = 1, 2, 3$, and equals zero otherwise. Find (a) the constant c, (b) $P(X = 2, Y = 3)$, (c) $P(1 \le X \le 2, Y \le 2)$, (d) $P(X \ge 2)$, (e) $P(Y < 2)$, (f) $P(X = 1)$, (g) $P(Y = 3)$.

2.55. Find the marginal probability functions of (a) X and (b) Y for the random variables of Problem 2.54. (c) Determine whether X and Y are independent.

2.56. Let X and Y be continuous random variables having joint density function

$$f(x, y) = \begin{cases} c(x^2 + y^2) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Determine (a) the constant c, (b) $P(X < \frac{1}{2}, Y > \frac{1}{2})$, (c) $P(\frac{1}{4} < X < \frac{3}{4})$, (d) $P(Y < \frac{1}{2})$, (e) whether X and Y are independent.

2.57. Find the marginal distribution functions (a) of X and (b) of Y for the density function of Problem 2.56.

Conditional distributions and density functions

2.58. Find the conditional probability function (a) of X given Y, (b) of Y given X, for the distribution of Problem 2.54.

2.59. Let

$$f(x, y) = \begin{cases} x + y & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the conditional density function of (a) X given Y, (b) Y given X.

2.60. Find the conditional density of (a) X given Y, (b) Y given X, for the distribution of Problem 2.56.

2.61. Let

$$f(x, y) = \begin{cases} e^{-(x+y)} & x \ge 0, y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

be the joint density function of X and Y. Find the conditional density function of (a) X given Y, (b) Y given X.

Change of variables

2.62. Let X have density function

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Find the density function of $Y = X^2$.

2.63. (a) If the density function of X is $f(x)$ find the density function of X^3. (b) Illustrate the result in part (a) by choosing

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

and check the answer.

2.64. If X has density function $f(x) = 2(\pi)^{-1/2}e^{-x^2/2}$, $-\infty < x < \infty$, find the density function of $Y = X^2$.

2.65. Verify that the integral of $g_1(u)$ in Method 1 of Problem 2.21 is equal to 1.

2.66. If the density of X is $f(x) = 1/\pi(x^2 + 1)$, $-\infty < x < \infty$, find the density of $Y = \tan^{-1} X$.

2.67. Complete the work needed to find $g_1(u)$ in Method 2 of Problem 2.21 and check your answer.

2.68. Let the density of X be

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

Find the density of (a) $3X - 2$, (b) $X^3 + 1$.

2.69. Check by direct integration the joint density function found in Problem 2.22.

2.70. Let X and Y have joint density function

$$f(x, y) = \begin{cases} e^{-(x+y)} & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

If $U = X/Y$, $V = X + Y$, find the joint density function of U and V.

2.71. Use Problem 2.22 to find the density function of (a) $U = XY^2$, (b) $V = X^2Y$.

2.72. Let X and Y be random variables having joint density function $f(x, y) = (2\pi)^{-1} e^{-(x^2+y^2)}$, $-\infty < x < \infty$, $-\infty < y < \infty$. If R and Θ are new random variables such that $X = R\cos\Theta$, $Y = R\sin\Theta$, show that the density function of R is

$$g(r) = \begin{cases} re^{-r^2/2} & r \geq 0 \\ 0 & r < 0 \end{cases}$$

2.73. Let

$$f(x, y) = \begin{cases} 1 & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

be the joint density function of X and Y. Find the density function of $Z = XY$.

Convolutions

2.74. Let X and Y be identically distributed independent random variables with density function

$$f(t) = \begin{cases} 1 & 0 \le t \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of $X + Y$ and check your answer.

2.75. Let X and Y be identically distributed independent random variables with density function

$$f(t) = \begin{cases} e^{-t} & t \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of $X + Y$ and check your answer.

2.76. Work Problem 2.21 by first making the transformation $2Y = Z$ and then using convolutions to find the density function of $U = X + Z$.

2.77. If the independent random variables X_1 and X_2 are identically distributed with density function

$$f(t) = \begin{cases} te^{-t} & t \ge 0 \\ 0 & t < 0 \end{cases}$$

find the density function of $X_1 + X_2$.

Applications to geometric probability

2.78. Two points are to be chosen at random on a line segment whose length is $a > 0$. Find the probability that the three line segments thus formed will be the sides of a triangle.

2.79. It is known that a bus will arrive at random at a certain location sometime between 3:00 P.M. and 3:30 P.M. A man decides that he will go at random to this location between these two times and will wait at most 5 minutes for the bus. If he misses it, he will take the subway. What is the probability that he will take the subway?

2.80. Two line segments, AB and CD, have lengths 8 and 6 units, respectively. Two points P and Q are to be chosen at random on AB and CD, respectively. Show that the probability that the area of a triangle will have height AP and that the base CQ will be greater than 12 square units is equal to $(1 - \ln 2)/2$.

Miscellaneous problems

2.81. Suppose that $f(x) = c/3^x$, $x = 1, 2, \ldots$, is the probability function for a random variable X. (a) Determine c. (b) Find the distribution function. (c) Graph the probability function and the distribution function. (d) Find $P(2 \le X < 5)$. (e) Find $P(X \ge 3)$.

2.82. Suppose that

$$f(x) = \begin{cases} cxe^{-2x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

is the density function for a random variable X. (a) Determine c. (b) Find the distribution function. (c) Graph the density function and the distribution function. (d) Find $P(X \ge 1)$. (e) Find $P(2 \le X < 3)$.

2.83. The probability function of a random variable X is given by

$$f(x) = \begin{cases} 2p & x = 1 \\ p & x = 2 \\ 4p & x = 3 \\ 0 & \text{otherwise} \end{cases}$$

where p is a constant. Find (a) $P(0 \leq X < 3)$, (b) $P(X > 1)$.

2.84. (a) Prove that for a suitable constant c,

$$F(x) = \begin{cases} 0 & x \leq 0 \\ c(1 - e^{-x})^2 & x > 0 \end{cases}$$

is the distribution function for a random variable X, and find this c. (b) Determine $P(1 < X < 2)$.

2.85. A random variable X has density function

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

Find the density function of the random variable $Y = X^2$ and check your answer.

2.86. Two independent random variables, X and Y, have respective density functions

$$f(x) = \begin{cases} c_1 e^{-2x} & x > 0 \\ 0 & x \leq 0 \end{cases} \qquad g(y) = \begin{cases} c_2 y e^{-3y} & y > 0 \\ 0 & y \leq 0 \end{cases}$$

Find (a) c_1 and c_2, (b) $P(X + Y > 1)$, (c) $P(1 < X < 2, Y \geq 1)$, (d) $P(1 < X < 2)$, (e) $P(Y \geq 1)$.

2.87. In Problem 2.86 what is the relationship between the answers to (c), (d), and (e)? Justify your answer.

2.88. Let X and Y be random variables having joint density function

$$f(x, y) = \begin{cases} c(2x + y) & 0 < x < 1, 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) the constant c, (b) $P(X > \frac{1}{2}, Y < \frac{3}{2})$, (c) the (marginal) density function of X, (d) the (marginal) density function of Y.

2.89. In Problem 2.88 is $P(X > \frac{1}{2}, Y < \frac{3}{2}) = P(X > \frac{1}{2})P(Y < \frac{3}{2})$? Why?

2.90. In Problem 2.86 find the density function (a) of X^2, (b) of $X + Y$.

2.91. Let X and Y have joint density function

$$f(x, y) = \begin{cases} 1/y & 0 < x < y, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Determine whether X and Y are independent, (b) Find $P(X > \frac{1}{2})$. (c) Find $P(X < \frac{1}{2}, Y > \frac{1}{3})$. (d) Find $P(X + Y > \frac{1}{2})$.

2.92. Generalize (a) Problem 2.74 and (b) Problem 2.75 to three or more variables.

2.93. Let X and Y be identically distributed independent random variables having density function $f(u) = (2\pi)^{-1/2}e^{-u^2/2}$, $-\infty < u < \infty$. Find the density function of $Z = X^2 + Y^2$.

2.94. The joint probability function for the random variables X and Y is given in Table 2-9. (a) Find the marginal probability functions of X and Y. (b) Find $P(1 \leq X < 3, Y \geq 1)$. (c) Determine whether X and Y are independent.

Table 2-9

$X \backslash Y$	0	1	2
0	1/18	1/9	1/6
1	1/9	1/18	1/9
2	1/6	1/6	1/18

2.95. Suppose that the joint probability function of random variables X and Y is given by

$$f(x, y) = \begin{cases} cxy & 0 \leq x \leq 2, 0 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$

(a) Determine whether X and Y are independent. (b) Find $P(\frac{1}{2} < X < 1)$. (c) Find $P(Y \geq 1)$. (d) Find $P(\frac{1}{2} < X < 1, Y \geq 1)$.

2.96. Let X and Y be independent random variables each having density function

$$f(u) = \frac{\lambda^u e^{-\lambda}}{u} \qquad u = 0, 1, 2, \ldots$$

where $\lambda > 0$. Prove that the density function of $X + Y$ is

$$g(u) = \frac{(2\lambda)^u e^{-2\lambda}}{u!} \qquad u = 0, 1, 2, \ldots$$

2.97. A stick of length L is to be broken into two parts. What is the probability that one part will have a length of more than double the other? State clearly what assumptions would you have made. Discuss whether you believe these assumptions are realistic and how you might improve them if they are not.

2.98. A floor is made up of squares of side l. A needle of length $a < l$ is to be tossed onto the floor. Prove that the probability of the needle intersecting at least one side is equal to $a(4l - a)/\pi l^2$.

2.99. For a needle of given length, what should be the side of a square in Problem 2.98 so that the probability of intersection is a maximum? Explain your answer.

2.100. Let

$$f(x, y, z) = \begin{cases} 24xy^2z^3 & 0 < x < 1, 0 < y < 1, 0 < z < 1 \\ 0 & \text{otherwise} \end{cases}$$

be the joint density function of three random variables X, Y, and Z. Find (a) $P(X > \frac{1}{2}, Y < \frac{1}{2}, Z > \frac{1}{2})$, (b) $P(Z < X + Y)$.

2.101. A cylindrical stream of particles, of radius a, is directed toward a hemispherical target ABC with center at O as indicated in Fig. 2-26. Assume that the distribution of particles is given by

$$f(r) = \begin{cases} 1/a & 0 < r < a \\ 0 & \text{otherwise} \end{cases}$$

where r is the distance from the axis OB. Show that the distribution of particles along the target is given by

$$g(\theta) = \begin{cases} \cos\theta & 0 < \theta < \pi/2 \\ 0 & \text{otherwise} \end{cases}$$

where θ is the angle that line OP (from O to any point P on the target) makes with the axis.

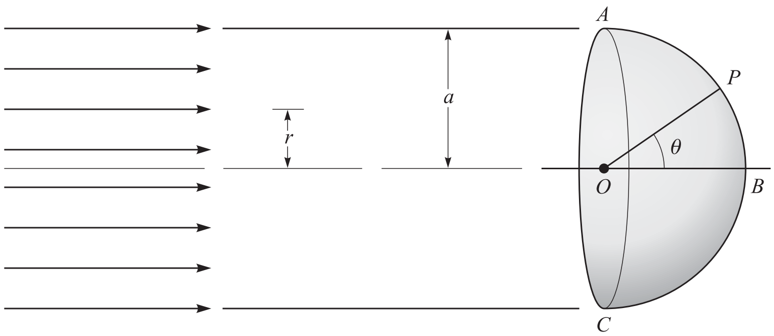

Fig. 2-26

2.102. In Problem 2.101 find the probability that a particle will hit the target between $\theta = 0$ and $\theta = \pi/4$.

2.103. Suppose that random variables X, Y, and Z have joint density function

$$f(x, y, z) = \begin{cases} 1 - \cos \pi x \cos \pi y \cos \pi z & 0 < x < 1, 0 < y < 1, 0 < z < 1 \\ 0 & \text{otherwise} \end{cases}$$

Show that although any two of these random variables are independent, i.e., their marginal density function factors, all three are not independent.

ANSWERS TO SUPPLEMENTARY PROBLEMS

2.38.

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

2.39.

x	0	1	2
$f(x)$	3/28	15/28	5/14

2.40.

x	0	1	2
$f(x)$	9/64	15/32	25/64

2.42.

x	0	1	2	3	4
$f(x)$	$\frac{194{,}580}{270{,}725}$	$\frac{69{,}184}{270{,}725}$	$\frac{6768}{270{,}725}$	$\frac{192}{270{,}725}$	$\frac{1}{270{,}725}$

2.43.

x	0	1	2	3
$f(x)$	1/8	1/2	7/8	1

2.46. (a)

x	1	2	3	4
$f(x)$	1/8	1/4	3/8	1/4

(b) 3/4 (c) 7/8 (d) 3/8 (e) 7/8

2.47. (a) 3 (b) $e^{-3} - e^{-6}$ (c) e^{-9} (d) $1 - e^{-3}$

2.48. $F(x) = \begin{cases} 1 - e^{-3x} & x \geq 0 \\ 0 & x \leq 0 \end{cases}$

2.49. (a) 6/29 (b) 15/29 (c) 19/116

2.50. $F(x) = \begin{cases} 0 & x \leq 1 \\ (2x^3 - 2)/29 & 1 \leq x \leq 2 \\ (3x^2 + 2)/29 & 2 \leq x \leq 3 \\ 1 & x \geq 3 \end{cases}$

2.51. (a) 1/27 (b) $f(x) = \begin{cases} x^2/9 & 0 \leq x < 3 \\ 0 & \text{otherwise} \end{cases}$ (c) 26/27 (d) 7/27

2.53. (a) 1/2 (b) 1/2 (c) 3/4 (d) $F(x) = \begin{cases} 0 & x \leq 0 \\ x^2/4 & 0 \leq x \leq 2 \\ 1 & x \geq 2 \end{cases}$

2.54. (a) 1/36 (b) 1/6 (c) 1/4 (d) 5/6 (e) 1/6 (f) 1/6 (g) 1/2

2.55. (a) $f_1(x) = \begin{cases} x/6 & x = 1, 2, 3 \\ 0 & \text{other } x \end{cases}$ (b) $f_2(y) = \begin{cases} y/6 & y = 1, 2, 3 \\ 0 & \text{other } y \end{cases}$

2.56. (a) 3/2 (b) 1/4 (c) 29/64 (d) 5/16

2.57. (a) $F_1(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{2}(x^3 + x) & 0 \leq x \leq 1 \\ 1 & x \geq 1 \end{cases}$ (b) $F_2(y) = \begin{cases} 0 & y \leq 0 \\ \frac{1}{2}(y^3 + y) & 0 \leq y \leq 1 \\ 1 & y \geq 1 \end{cases}$

2.58. (a) $f(x \mid y) = f_1(x)$ for $y = 1, 2, 3$ (see Problem 2.55)

(b) $f(y \mid x) = f_2(y)$ for $x = 1, 2, 3$ (see Problem 2.55)

2.59. (a) $f(x \mid y) = \begin{cases} (x + y)/(y + \frac{1}{2}) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{other } x, 0 \leq y \leq 1 \end{cases}$

(b) $f(y \mid x) = \begin{cases} (x + y)/(x + \frac{1}{2}) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & 0 \leq x \leq 1, \text{other } y \end{cases}$

2.60. (a) $f(x \mid y) = \begin{cases} (x^2 + y^2)/(y^2 + \frac{1}{3}) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{other } x, 0 \leq y \leq 1 \end{cases}$

(b) $f(y \mid x) = \begin{cases} (x^2 + y^2)/(x^2 + \frac{1}{3}) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & 0 \leq x \leq 1, \text{other } y \end{cases}$

2.61. (a) $f(x \mid y) = \begin{cases} e^{-x} & x \geq 0, y \geq 0 \\ 0 & x < 0, y \geq 0 \end{cases}$ (b) $f(y \mid x) = \begin{cases} e^{-y} & x \geq 0, y \geq 0 \\ 0 & x \geq 0, y < 0 \end{cases}$

2.62. $e^{-\sqrt{y}}/2\sqrt{y}$ for $y > 0$; 0 otherwise

2.64. $(2\pi)^{-1/2} y^{-1/2} e^{-y/2}$ for $y > 0$; 0 otherwise

2.66. $1/\pi$ for $-\pi/2 < y < \pi/2$; 0 otherwise

2.68. (a) $g(y) = \begin{cases} \frac{1}{6} & -5 < y < 1 \\ 0 & \text{otherwise} \end{cases}$ (b) $g(y) = \begin{cases} \frac{1}{6}(1 - y)^{-2/3} & 0 < y < 1 \\ \frac{1}{6}(y - 1)^{-2/3} & 1 < y < 2 \\ 0 & \text{otherwise} \end{cases}$

2.70. $ve^{-v}/(1 + u)^2$ for $u \geq 0$, $v \geq 0$; 0 otherwise

2.73. $g(z) = \begin{cases} -\ln z & 0 < z < 1 \\ 0 & \text{otherwise} \end{cases}$ **2.77.** $g(x) = \begin{cases} x^3 e^{-x}/6 & x \geq 0 \\ 0 & x < 0 \end{cases}$

2.74. $g(u) = \begin{cases} u & 0 \leq u \leq 1 \\ 2 - u & 1 \leq u \leq 2 \\ 0 & \text{otherwise} \end{cases}$ **2.78.** 1/4

2.75. $g(u) = \begin{cases} ue^{-u} & u \geq 0 \\ 0 & u < 0 \end{cases}$ **2.79.** 61/72

2.81. (a) 2 (b) $F(x) = \begin{cases} 0 & x < 1 \\ 1 - 3^{-y} & y \leq x < y + 1;\ y = 1, 2, 3, \ldots \end{cases}$ (d) 26/81 (e) 1/9

2.82. (a) 4 (b) $F(x) = \begin{cases} 1 - e^{-2x}(2x + 1) & x \geq 0 \\ 0 & x < 0 \end{cases}$ (d) $3e^{-2}$ (e) $5e^{-4} - 7e^{-6}$

2.83. (a) 3/7 (b) 5/7 **2.84.** (a) $c = 1$ (b) $e^{-4} - 3e^{-2} + 2e^{-1}$

2.86. (a) $c_1 = 2, c_2 = 9$ (b) $9e^{-2} - 14e^{-3}$ (c) $4e^{-5} - 4e^{-7}$ (d) $e^{-2} - e^{-4}$ (e) $4e^{-3}$

2.88. (a) 1/4 (b) 27/64 (c) $f_1(x) = \begin{cases} x + \frac{1}{2} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$ (d) $f_2(y) = \begin{cases} \frac{1}{4}(y + 1) & 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$

2.90. (a) $\begin{cases} e^{-2y}/\sqrt{y} & y > 0 \\ 0 & \text{otherwise} \end{cases}$ (b) $\begin{cases} 18e^{-2u} & u > 0 \\ 0 & \text{otherwise} \end{cases}$

2.91. (b) $\frac{1}{2}(1 - \ln 2)$ (c) $\frac{1}{6} + \frac{1}{2}\ln 2$ (d) $\frac{1}{2}\ln 2$ **2.95.** (b) 15/256 (c) 9/16 (d) 0

2.93. $g(z) = \begin{cases} \frac{1}{2}e^{-z/2} & z \geq 0 \\ 0 & z < 0 \end{cases}$ **2.100.** (a) 45/512 (b) 1/14

2.94. (b) 7/18 **2.102.** $\sqrt{2}/2$

CHAPTER 3

Mathematical Expectation

Definition of Mathematical Expectation

A very important concept in probability and statistics is that of the *mathematical expectation*, *expected value*, or briefly the *expectation*, of a random variable. For a discrete random variable X having the possible values $x_1, \ldots, x_n$, the expectation of X is defined as

$$E(X) = x_1 P(X = x_1) + \cdots + x_n P(X = x_n) = \sum_{j=1}^{n} x_j P(X = x_j) \tag{1}$$

or equivalently, if $P(X = x_j) = f(x_j)$,

$$E(X) = x_1 f(x_1) + \cdots + x_n f(x_n) = \sum_{j=1}^{n} x_j f(x_j) = \sum x f(x) \tag{2}$$

where the last summation is taken over all appropriate values of x. As a special case of (2), where the probabilities are all equal, we have

$$E(X) = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{3}$$

which is called the *arithmetic mean*, or simply the *mean*, of $x_1, x_2, \ldots, x_n$.

If X takes on an infinite number of values $x_1, x_2, \ldots$, then $E(X) = \sum_{j=1}^{\infty} x_j f(x_j)$ provided that the infinite series converges absolutely.

For a continuous random variable X having density function $f(x)$, the expectation of X is defined as

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx \tag{4}$$

provided that the integral converges absolutely.

The expectation of X is very often called the *mean* of X and is denoted by μ_X, or simply μ, when the particular random variable is understood.

The mean, or expectation, of X gives a single value that acts as a representative or average of the values of X, and for this reason it is often called a *measure of central tendency.* Other measures are considered on page 83.

EXAMPLE 3.1 Suppose that a game is to be played with a single die assumed fair. In this game a player wins \$20 if a 2 turns up, \$40 if a 4 turns up; loses \$30 if a 6 turns up; while the player neither wins nor loses if any other face turns up. Find the expected sum of money to be won.

Let X be the random variable giving the amount of money won on any toss. The possible amounts won when the die turns up $1, 2, \ldots, 6$ are $x_1, x_2, \ldots, x_6$, respectively, while the probabilities of these are $f(x_1), f(x_2), \ldots, f(x_6)$. The probability function for X is displayed in Table 3-1. Therefore, the expected value or expectation is

$$E(X) = (0)\left(\frac{1}{6}\right) + (20)\left(\frac{1}{6}\right) + (0)\left(\frac{1}{6}\right) + (40)\left(\frac{1}{6}\right) + (0)\left(\frac{1}{6}\right) + (-30)\left(\frac{1}{6}\right) = 5$$

Table 3-1

x_j	0	+20	0	+40	0	−30
$f(x_j)$	1/6	1/6	1/6	1/6	1/6	1/6

It follows that the player can expect to win \$5. In a fair game, therefore, the player should be expected to pay \$5 in order to play the game.

EXAMPLE 3.2 The density function of a random variable X is given by

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

The expected value of X is then

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_0^2 x\left(\frac{1}{2}x\right)dx = \int_0^2 \frac{x^2}{2}\,dx = \left.\frac{x^3}{6}\right|_0^2 = \frac{4}{3}$$

Functions of Random Variables

Let X be a discrete random variable with probability function $f(x)$. Then $Y = g(X)$ is also a discrete random variable, and the probability function of Y is

$$h(y) = P(Y = y) = \sum_{\{x|g(x)=y\}} P(X = x) = \sum_{\{x|g(x)=y\}} f(x)$$

If X takes on the values $x_1, x_2, \ldots, x_n$, and Y the values $y_1, y_2, \ldots, y_m$ $(m \le n)$, then $y_1h(y_1) + y_2h(y_2) + \cdots + y_mh(y_m) = g(x_1)f(x_1) + g(x_2)f(x_2) + \cdots + g(x_n)f(x_n)$. Therefore,

$$\begin{aligned} E[g(X)] &= g(x_1)f(x_1) + g(x_2)f(x_2) + \cdots + g(x_n)f(x_n) \\ &= \sum_{j=1}^{n} g(x_j)f(x_j) = \sum g(x)f(x) \end{aligned} \tag{5}$$

Similarly, if X is a continuous random variable having probability density $f(x)$, then it can be shown that

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx \tag{6}$$

Note that (5) and (6) do not involve, respectively, the probability function and the probability density function of $Y = g(X)$.

Generalizations are easily made to functions of two or more random variables. For example, if X and Y are two continuous random variables having joint density function $f(x, y)$, then the expectation of $g(X, Y)$ is given by

$$E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f(x, y)\,dx\,dy \tag{7}$$

EXAMPLE 3.3 If X is the random variable of Example 3.2,

$$E(3X^2 - 2X) = \int_{-\infty}^{\infty} (3x^2 - 2x)f(x)\,dx = \int_0^2 (3x^2 - 2x)\left(\frac{1}{2}x\right)dx = \frac{10}{3}$$

Some Theorems on Expectation

Theorem 3-1 If c is any constant, then

$$E(cX) = cE(X) \tag{8}$$

Theorem 3-2 If X and Y are any random variables, then

$$E(X + Y) = E(X) + E(Y) \tag{9}$$

Theorem 3-3 If X and Y are independent random variables, then

$$E(XY) = E(X)E(Y) \tag{10}$$

Generalizations of these theorems are easily made.

The Variance and Standard Deviation

We have already noted on page 75 that the expectation of a random variable X is often called the *mean* and is denoted by μ. Another quantity of great importance in probability and statistics is called the *variance* and is defined by

$$\text{Var}(X) = E[(X - \mu)^2] \tag{11}$$

The variance is a nonnegative number. The positive square root of the variance is called the *standard deviation* and is given by

$$\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{E[(X - \mu)^2]} \tag{12}$$

Where no confusion can result, the standard deviation is often denoted by σ instead of σ_X, and the variance in such case is σ^2.

If X is a discrete random variable taking the values $x_1, x_2, \ldots, x_n$ and having probability function $f(x)$, then the variance is given by

$$\sigma_X^2 = E[(X - \mu)^2] = \sum_{j=1}^{n}(x_j - \mu)^2 f(x_j) = \sum (x - \mu)^2 f(x) \tag{13}$$

In the special case of (13) where the probabilities are all equal, we have

$$\sigma^2 = [(x_1 - \mu)^2 + (x_2 - \mu)^2 + \cdots + (x_n - \mu)^2]/n \tag{14}$$

which is the variance for a set of n numbers $x_1, \ldots, x_n$.

If X takes on an infinite number of values $x_1, x_2, \ldots$, then $\sigma_X^2 = \sum_{j=1}^{\infty}(x_j - \mu)^2 f(x_j)$, provided that the series converges.

If X is a continuous random variable having density function $f(x)$, then the variance is given by

$$\sigma_X^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty}(x - \mu)^2 f(x)\,dx \tag{15}$$

provided that the integral converges.

The variance (or the standard deviation) is a measure of the *dispersion*, or *scatter*, of the values of the random variable about the mean μ. If the values tend to be concentrated near the mean, the variance is small; while if the values tend to be distributed far from the mean, the variance is large. The situation is indicated graphically in Fig. 3-1 for the case of two continuous distributions having the same mean μ.

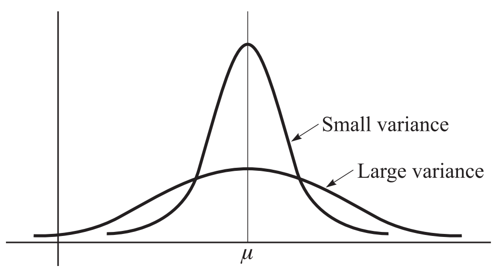

Fig. 3-1

EXAMPLE 3.4 Find the variance and standard deviation of the random variable of Example 3.2. As found in Example 3.2, the mean is $\mu = E(X) = 4/3$. Then the variance is given by

$$\sigma^2 = E\left[\left(X - \frac{4}{3}\right)^2\right] = \int_{-\infty}^{\infty}\left(x - \frac{4}{3}\right)^2 f(x)\,dx = \int_0^2\left(x - \frac{4}{3}\right)^2\left(\frac{1}{2}x\right)dx = \frac{2}{9}$$

and so the standard deviation is $\sigma = \sqrt{\frac{2}{9}} = \frac{\sqrt{2}}{3}$

Note that if X has certain *dimensions* or *units*, such as *centimeters* (cm), then the variance of X has units cm^2 while the standard deviation has the same unit as X, i.e., cm. It is for this reason that the standard deviation is often used.

Some Theorems on Variance

Theorem 3-4
$$\sigma^2 = E[(X - \mu)^2] = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2 \tag{16}$$

where $\mu = E(X)$.

Theorem 3-5 If c is any constant,

$$\operatorname{Var}(cX) = c^2 \operatorname{Var}(X) \tag{17}$$

Theorem 3-6 The quantity $E[(X - a)^2]$ is a minimum when $a = \mu = E(X)$.

Theorem 3-7 If X and Y are independent random variables,

$$\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) \qquad \text{or} \qquad \sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y \tag{18}$$

$$\operatorname{Var}(X - Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) \qquad \text{or} \qquad \sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y \tag{19}$$

Generalizations of Theorem 3-7 to more than two independent variables are easily made. In words, the variance of a sum of independent variables equals the sum of their variances.

Standardized Random Variables

Let X be a random variable with mean μ and standard deviation $\sigma\,(\sigma > 0)$. Then we can define an associated *standardized random variable* given by

$$X^* = \frac{X - \mu}{\sigma} \tag{20}$$

An important property of X^* is that it has a mean of zero and a variance of 1, which accounts for the name *standardized,* i.e.,

$$E(X^*) = 0, \qquad \operatorname{Var}(X^*) = 1 \tag{21}$$

The values of a standardized variable are sometimes called *standard scores,* and X is then said to be expressed in *standard units* (i.e., σ is taken as the unit in measuring $X - \mu$).

Standardized variables are useful for comparing different distributions.

Moments

The *rth moment of a random variable X about the mean* μ, also called the *rth central moment*, is defined as

$$\mu_r = E[(X - \mu)^r] \tag{22}$$

where $r = 0, 1, 2, \ldots$. It follows that $\mu_0 = 1$, $\mu_1 = 0$, and $\mu_2 = \sigma^2$, i.e., the second central moment or second moment about the mean is the variance. We have, assuming absolute convergence,

$$\mu_r = \sum (x - \mu)^r f(x) \qquad \text{(discrete variable)} \tag{23}$$

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x)\, dx \qquad \text{(continuous variable)} \tag{24}$$

The *rth moment of X about the origin*, also called the *rth raw moment*, is defined as

$$\mu'_r = E(X^r) \tag{25}$$

where $r = 0, 1, 2, \ldots$, and in this case there are formulas analogous to (23) and (24) in which $\mu = 0$.

The relationship between these moments is given by

$$\mu_r = \mu'_r - \binom{r}{1}\mu'_{r-1}\mu + \cdots + (-1)^j \binom{r}{j}\mu'_{r-j}\mu^j + \cdots + (-1)^r \mu'_0 \mu^r \tag{26}$$

As special cases we have, using $\mu'_1 = \mu$ and $\mu'_0 = 1$,

$$\begin{aligned} \mu_2 &= \mu'_2 - \mu^2 \\ \mu_3 &= \mu'_3 - 3\mu'_2\mu + 2\mu^3 \\ \mu_4 &= \mu'_4 - 4\mu'_3\mu + 6\mu'_2\mu^2 - 3\mu^4 \end{aligned} \tag{27}$$

Moment Generating Functions

The *moment generating function* of X is defined by

$$M_X(t) = E(e^{tX}) \tag{28}$$

that is, assuming convergence,

$$M_X(t) = \sum e^{tx} f(x) \qquad \text{(discrete variable)} \tag{29}$$

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx \qquad \text{(continuous variable)} \tag{30}$$

We can show that the Taylor series expansion is [Problem 3.15(a)]

$$M_X(t) = 1 + \mu t + \mu'_2 \frac{t^2}{2!} + \cdots + \mu'_r \frac{t^r}{r!} + \cdots \tag{31}$$

Since the coefficients in this expansion enable us to find the moments, the reason for the name *moment generating function* is apparent. From the expansion we can show that [Problem 3.15(b)]

$$\mu'_r = \left.\frac{d^r}{dt^r} M_X(t)\right|_{t=0} \tag{32}$$

i.e., μ'_r is the rth derivative of $M_X(t)$ evaluated at $t = 0$. Where no confusion can result, we often write $M(t)$ instead of $M_X(t)$.

Some Theorems on Moment Generating Functions

Theorem 3-8 If $M_X(t)$ is the moment generating function of the random variable X and a and b ($b \neq 0$) are constants, then the moment generating function of $(X + a)/b$ is

$$M_{(X+a)/b}(t) = e^{at/b} M_X\left(\frac{t}{b}\right) \tag{33}$$

Theorem 3-9 If X and Y are independent random variables having moment generating functions $M_X(t)$ and $M_Y(t)$, respectively, then

$$M_{X+Y}(t) = M_X(t)\,M_Y(t) \tag{34}$$

Generalizations of Theorem 3-9 to more than two independent random variables are easily made. In words, the moment generating function of a sum of independent random variables is equal to the product of their moment generating functions.

Theorem 3-10 **(Uniqueness Theorem)** Suppose that X and Y are random variables having moment generating functions $M_X(t)$ and $M_Y(t)$, respectively. Then X and Y have the same probability distribution if and only if $M_X(t) = M_Y(t)$ identically.

Characteristic Functions

If we let $t = i\omega$, where i is the imaginary unit, in the moment generating function we obtain an important function called the *characteristic function.* We denote this by

$$\phi_X(\omega) = M_X(i\omega) = E(e^{i\omega X}) \tag{35}$$

It follows that

$$\phi_X(\omega) = \sum e^{i\omega x} f(x) \qquad \text{(discrete variable)} \tag{36}$$

$$\phi_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\,dx \qquad \text{(continuous variable)} \tag{37}$$

Since $|e^{i\omega x}| = 1$, the series and the integral always converge absolutely.

The corresponding results (31) and (32) become

$$\phi_X(\omega) = 1 + i\mu\omega - \mu_2'\frac{\omega^2}{2!} + \cdots + i^r\mu_r'\frac{\omega^r}{r!} + \cdots \tag{38}$$

where

$$\mu_r' = (-1)^r i^r \frac{d^r}{d\omega^r}\phi_X(\omega)\bigg|_{\omega=0} \tag{39}$$

When no confusion can result, we often write $\phi(\omega)$ instead of $\phi_X(\omega)$.

Theorems for characteristic functions corresponding to Theorems 3-8, 3-9, and 3-10 are as follows.

Theorem 3-11 If $\phi_X(\omega)$ is the characteristic function of the random variable X and a and b ($b \neq 0$) are constants, then the characteristic function of $(X + a)/b$ is

$$\phi_{(X+a)/b}(\omega) = e^{ai\omega/b}\phi_X\left(\frac{\omega}{b}\right) \tag{40}$$

Theorem 3-12 If X and Y are independent random variables having characteristic functions $\phi_X(\omega)$ and $\phi_Y(\omega)$, respectively, then

$$\phi_{X+Y}(\omega) = \phi_X(\omega)\,\phi_Y(\omega) \tag{41}$$

More generally, the characteristic function of a sum of independent random variables is equal to the product of their characteristic functions.

Theorem 3-13 **(Uniqueness Theorem)** Suppose that X and Y are random variables having characteristic functions $\phi_X(\omega)$ and $\phi_Y(\omega)$, respectively. Then X and Y have the same probability distribution if and only if $\phi_X(\omega) = \phi_Y(\omega)$ identically.

An important reason for introducing the characteristic function is that (37) represents the *Fourier transform* of the density function $f(x)$. From the theory of Fourier transforms, we can easily determine the density function from the characteristic function. In fact,

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega x}\phi_X(\omega)\,d\omega \tag{42}$$

which is often called an *inversion formula*, or *inverse Fourier transform.* In a similar manner we can show in the discrete case that the probability function $f(x)$ can be obtained from (36) by use of *Fourier series*, which is the analog of the Fourier integral for the discrete case. See Problem 3.39.

Another reason for using the characteristic function is that it always exists whereas the moment generating function may not exist.

Variance for Joint Distributions. Covariance

The results given above for one variable can be extended to two or more variables. For example, if X and Y are two continuous random variables having joint density function $f(x, y)$, the means, or expectations, of X and Y are

$$\mu_X = E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x, y)\,dx\,dy, \qquad \mu_Y = E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x, y)\,dx\,dy \tag{43}$$

and the variances are

$$\begin{aligned} \sigma_X^2 &= E[(X - \mu_X)^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)^2 f(x, y)\,dx\,dy \\ \sigma_Y^2 &= E[(Y - \mu_Y)^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y - \mu_Y)^2 f(x, y)\,dx\,dy \end{aligned} \tag{44}$$

Note that the marginal density functions of X and Y are not directly involved in (43) and (44).

Another quantity that arises in the case of two variables X and Y is the *covariance* defined by

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \tag{45}$$

In terms of the joint density function $f(x, y)$, we have

$$\sigma_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y)f(x, y)\,dx\,dy \tag{46}$$

Similar remarks can be made for two discrete random variables. In such cases (43) and (46) are replaced by

$$\mu_X = \sum_x\sum_y xf(x, y) \qquad \mu_Y = \sum_x\sum_y yf(x, y) \tag{47}$$

$$\sigma_{XY} = \sum_x\sum_y (x - \mu_X)(y - \mu_Y)f(x, y) \tag{48}$$

where the sums are taken over all the discrete values of X and Y.

The following are some important theorems on covariance.

Theorem 3-14

$$\sigma_{XY} = E(XY) - E(X)E(Y) = E(XY) - \mu_X\mu_Y \tag{49}$$

Theorem 3-15 If X and Y are independent random variables, then

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = 0 \tag{50}$$

Theorem 3-16

$$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X, Y) \tag{51}$$

or

$$\sigma_{X\pm Y}^2 = \sigma_X^2 + \sigma_Y^2 \pm 2\sigma_{XY} \tag{52}$$

Theorem 3-17

$$|\sigma_{XY}| \le \sigma_X\sigma_Y \tag{53}$$

The converse of Theorem 3-15 is not necessarily true. If X and Y are independent, Theorem 3-16 reduces to Theorem 3-7.

Correlation Coefficient

If X and Y are independent, then $\mathrm{Cov}(X, Y) = \sigma_{XY} = 0$. On the other hand, if X and Y are completely dependent, for example, when $X = Y$, then $\mathrm{Cov}(X, Y) = \sigma_{XY} = \sigma_X\sigma_Y$. From this we are led to a *measure of the dependence* of the variables X and Y given by

$$\rho = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \tag{54}$$

We call ρ the *correlation coefficient*, or *coefficient of correlation.* From Theorem 3-17 we see that $-1 \leq \rho \leq 1$. In the case where $\rho = 0$ (i.e., the covariance is zero), we call the variables X and Y *uncorrelated.* In such cases, however, the variables may or may not be independent. Further discussion of correlation cases will be given in Chapter 8.

Conditional Expectation, Variance, and Moments

If X and Y have joint density function $f(x, y)$, then as we have seen in Chapter 2, the conditional density function of Y given X is $f(y \mid x) = f(x, y)/f_1(x)$ where $f_1(x)$ is the marginal density function of X. We can define the *conditional expectation*, or *conditional mean*, of Y given X by

$$E(Y \mid X = x) = \int_{-\infty}^{\infty} y f(y \mid x)\, dy \tag{55}$$

where "$X = x$" is to be interpreted as $x < X \leq x + dx$ in the continuous case. Theorems 3-1 and 3-2 also hold for conditional expectation.

We note the following properties:

1. $E(Y \mid X = x) = E(Y)$ when X and Y are independent.
2. $E(Y) = \displaystyle\int_{-\infty}^{\infty} E(Y \mid X = x) f_1(x)\, dx.$

It is often convenient to calculate expectations by use of Property 2, rather than directly.

EXAMPLE 3.5 The average travel time to a distant city is c hours by car or b hours by bus. A woman cannot decide whether to drive or take the bus, so she tosses a coin. What is her expected travel time?

Here we are dealing with the joint distribution of the outcome of the toss, X, and the travel time, Y, where $Y = Y_{\text{car}}$ if $X = 0$ and $Y = Y_{\text{bus}}$ if $X = 1$. Presumably, both Y_{car} and Y_{bus} are independent of X, so that by Property 1 above

$$E(Y \mid X = 0) = E(Y_{\text{car}} \mid X = 0) = E(Y_{\text{car}}) = c$$

and

$$E(Y \mid X = 1) = E(Y_{\text{bus}} \mid X = 1) = E(Y_{\text{bus}}) = b$$

Then Property 2 (with the integral replaced by a sum) gives, for a fair coin,

$$E(Y) = E(Y \mid X = 0)P(X = 0) + E(Y \mid X = 1)P(X = 1) = \frac{c + b}{2}$$

In a similar manner we can define the *conditional variance* of Y given X as

$$E[(Y - \mu_2)^2 \mid X = x] = \int_{-\infty}^{\infty} (y - \mu_2)^2 f(y \mid x)\, dy \tag{56}$$

where $\mu_2 = E(Y \mid X = x)$. Also we can define the rth *conditional moment* of Y about any value a given X as

$$E[(Y - a)^r \mid X = x] = \int_{-\infty}^{\infty} (y - a)^r f(y \mid x)\, dy \tag{57}$$

The usual theorems for variance and moments extend to conditional variance and moments.

Chebyshev's Inequality

An important theorem in probability and statistics that reveals a general property of discrete or continuous random variables having finite mean and variance is known under the name of *Chebyshev's inequality.*

Theorem 3-18 **(Chebyshev's Inequality)** Suppose that X is a random variable (discrete or continuous) having mean μ and variance σ^2, which are finite. Then if ϵ is any positive number,

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \tag{58}$$

or, with $\epsilon = k\sigma$,

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \tag{59}$$

EXAMPLE 3.6 Letting $k = 2$ in Chebyshev's inequality (59), we see that

$$P(|X - \mu| \geq 2\sigma) \leq 0.25 \qquad \text{or} \qquad P(|X - \mu| < 2\sigma) \geq 0.75$$

In words, the probability of X differing from its mean by more than 2 standard deviations is less than or equal to 0.25; equivalently, the probability that X will lie within 2 standard deviations of its mean is greater than or equal to 0.75. This is quite remarkable in view of the fact that we have not even specified the probability distribution of X.

Law of Large Numbers

The following theorem, called the *law of large numbers*, is an interesting consequence of Chebyshev's inequality.

Theorem 3-19 **(Law of Large Numbers):** Let $X_1, X_2, \ldots, X_n$ be mutually independent random variables (discrete or continuous), each having finite mean μ and variance σ^2. Then if $S_n = X_1 + X_2 + \cdots + X_n (n = 1, 2, \ldots)$,

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0 \tag{60}$$

Since S_n/n is the arithmetic mean of $X_1, \ldots, X_n$, this theorem states that the probability of the arithmetic mean S_n/n differing from its expected value μ by more than ϵ approaches zero as $n \to \infty$. A stronger result, which we might expect to be true, is that $\lim\limits_{n\to\infty} S_n/n = \mu$, but this is actually false. However, we can prove that $\lim\limits_{n\to\infty} S_n/n = \mu$ *with probability one.* This result is often called the *strong law of large numbers*, and, by contrast, that of Theorem 3-19 is called the *weak law of large numbers.* When the "law of large numbers" is referred to without qualification, the weak law is implied.

Other Measures of Central Tendency

As we have already seen, the mean, or expectation, of a random variable X provides a measure of central tendency for the values of a distribution. Although the mean is used most, two other measures of central tendency are also employed. These are the *mode* and the *median.*

1. **MODE.** The *mode* of a discrete random variable is that value which occurs most often or, in other words, has the greatest probability of occurring. Sometimes we have two, three, or more values that have relatively large probabilities of occurrence. In such cases, we say that the distribution is *bimodal*, *trimodal*, or *multimodal*, respectively. The mode of a continuous random variable X is the value (or values) of X where the probability density function has a relative maximum.
2. **MEDIAN.** The *median* is that value x for which $P(X < x) \leq \frac{1}{2}$ and $P(X > x) \leq \frac{1}{2}$. In the case of a continuous distribution we have $P(X < x) = \frac{1}{2} = P(X > x)$, and the median separates the density curve into two parts having equal areas of $1/2$ each. In the case of a discrete distribution a unique median may not exist (see Problem 3.34).

Percentiles

It is often convenient to subdivide the area under a density curve by use of ordinates so that the area to the left of the ordinate is some percentage of the total unit area. The values corresponding to such areas are called *percentile values*, or briefly *percentiles*. Thus, for example, the area to the left of the ordinate at x_α in Fig. 3-2 is α. For instance, the area to the left of $x_{0.10}$ would be 0.10, or 10%, and $x_{0.10}$ would be called the *10th percentile* (also called the *first decile*). The median would be the *50th percentile* (or *fifth decile*).

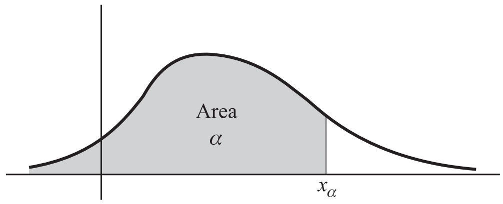

Fig. 3-2

Other Measures of Dispersion

Just as there are various measures of central tendency besides the mean, there are various measures of dispersion or scatter of a random variable besides the variance or standard deviation. Some of the most common are the following.

1. **SEMI-INTERQUARTILE RANGE.** If $x_{0.25}$ and $x_{0.75}$ represent the 25th and 75th percentile values, the difference $x_{0.75} - x_{0.25}$ is called the *interquartile range* and $\frac{1}{2}(x_{0.75} - x_{0.25})$ is the *semi-interquartile range.*
2. **MEAN DEVIATION.** The *mean deviation* (M.D.) of a random variable X is defined as the expectation of $|X - \mu|$, i.e., assuming convergence,

$$\text{M.D.}(X) = E[|X - \mu|] = \sum |x - \mu| f(x) \qquad \text{(discrete variable)} \tag{61}$$

$$\text{M.D.}(X) = E[|X - \mu|] = \int_{-\infty}^{\infty} |x - \mu| f(x)\,dx \qquad \text{(continuous variable)} \tag{62}$$

Skewness and Kurtosis

1. **SKEWNESS.** Often a distribution is not symmetric about any value but instead has one of its tails longer than the other. If the longer tail occurs to the right, as in Fig. 3-3, the distribution is said to be *skewed to the right*, while if the longer tail occurs to the left, as in Fig. 3-4, it is said to be *skewed to the left*. Measures describing this asymmetry are called *coefficients of skewness*, or briefly *skewness*. One such measure is given by

$$\alpha_3 = \frac{E[(X - \mu)^3]}{\sigma^3} = \frac{\mu_3}{\sigma^3} \tag{63}$$

The measure σ_3 will be positive or negative according to whether the distribution is skewed to the right or left, respectively. For a symmetric distribution, $\sigma_3 = 0$.

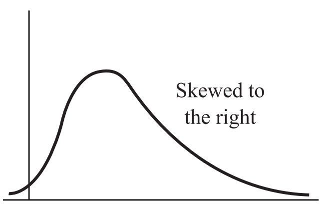

Fig. 3-3

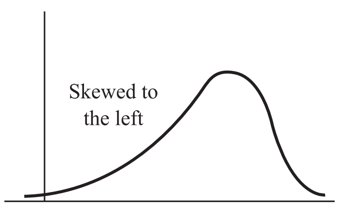

Fig. 3-4

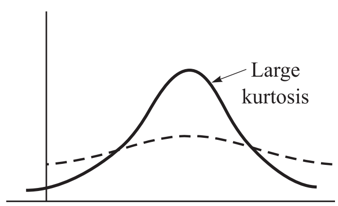

Fig. 3-5

2. **KURTOSIS.** In some cases a distribution may have its values concentrated near the mean so that the distribution has a large peak as indicated by the solid curve of Fig. 3-5. In other cases the distribution may be

relatively flat as in the dashed curve of Fig. 3-5. Measures of the degree of peakedness of a distribution are called *coefficients of kurtosis*, or briefly *kurtosis*. A measure often used is given by

$$\alpha_4 = \frac{E[(X - \mu)^4]}{\sigma^4} = \frac{\mu_4}{\sigma^4} \tag{64}$$

This is usually compared with the normal curve (see Chapter 4), which has a coefficient of kurtosis equal to 3. See also Problem 3.41.

SOLVED PROBLEMS

Expectation of random variables

3.1. In a lottery there are 200 prizes of \$5, 20 prizes of \$25, and 5 prizes of \$100. Assuming that 10,000 tickets are to be issued and sold, what is a fair price to pay for a ticket?

Let X be a random variable denoting the amount of money to be won on a ticket. The various values of X together with their probabilities are shown in Table 3-2. For example, the probability of getting one of the 20 tickets giving a \$25 prize is $20/10{,}000 = 0.002$. The expectation of X in dollars is thus

$$E(X) = (5)(0.02) + (25)(0.002) + (100)(0.0005) + (0)(0.9775) = 0.2$$

or 20 cents. Thus the fair price to pay for a ticket is 20 cents. However, since a lottery is usually designed to raise money, the price per ticket would be higher.

Table 3-2

x (dollars)	5	25	100	0
$P(X = x)$	0.02	0.002	0.0005	0.9775

3.2. Find the expectation of the sum of points in tossing a pair of fair dice.

Let X and Y be the points showing on the two dice. We have

$$E(X) = E(Y) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + \cdots + 6\left(\frac{1}{6}\right) = \frac{7}{2}$$

Then, by Theorem 3-2,

$$E(X + Y) = E(X) + E(Y) = 7$$

3.3. Find the expectation of a discrete random variable X whose probability function is given by

$$f(x) = \left(\frac{1}{2}\right)^x \qquad (x = 1, 2, 3, \ldots)$$

We have

$$E(X) = \sum_{x=1}^{\infty} x\left(\frac{1}{2}\right)^x = \frac{1}{2} + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{8}\right) + \cdots$$

To find this sum, let $\quad S = \frac{1}{2} + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{8}\right) + 4\left(\frac{1}{16}\right) + \cdots$

Then $\quad \frac{1}{2}S = \frac{1}{4} + 2\left(\frac{1}{8}\right) + 3\left(\frac{1}{16}\right) + \cdots$

Subtracting, $\quad \frac{1}{2}S = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 1$

Therefore, $S = 2$.

3.4. A continuous random variable X has probability density given by

$$f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

Find (a) $E(X)$, (b) $E(X^2)$.

(a)
$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_{0}^{\infty} x(2e^{-2x})\,dx = 2\int_{0}^{\infty} xe^{-2x}\,dx$$
$$= 2\left[(x)\left(\frac{e^{-2x}}{-2}\right) - (1)\left(\frac{e^{-2x}}{4}\right)\right]\Bigg|_0^{\infty} = \frac{1}{2}$$

(b)
$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = 2\int_{0}^{\infty} x^2 e^{-2x}\,dx$$
$$= 2\left[(x^2)\left(\frac{e^{-2x}}{-2}\right) - (2x)\left(\frac{e^{-2x}}{4}\right) + (2)\left(\frac{e^{-2x}}{-8}\right)\right]\Bigg|_0^{\infty} = \frac{1}{2}$$

3.5. The joint density function of two random variables X and Y is given by

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $E(X)$, (b) $E(Y)$, (c) $E(XY)$, (d) $E(2X + 3Y)$.

(a)
$$E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} x\left(\frac{xy}{96}\right)dx\,dy = \frac{8}{3}$$

(b)
$$E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} y\left(\frac{xy}{96}\right)dx\,dy = \frac{31}{9}$$

(c)
$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (xy)f(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} (xy)\left(\frac{xy}{96}\right)dx\,dy = \frac{248}{27}$$

(d)
$$E(2X + 3Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (2x + 3y)f(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} (2x + 3y)\left(\frac{xy}{96}\right)dx\,dy = \frac{47}{3}$$

Another method

(c) Since X and Y are independent, we have, using parts (a) and (b),

$$E(XY) = E(X)E(Y) = \left(\frac{8}{3}\right)\left(\frac{31}{9}\right) = \frac{248}{27}$$

(d) By Theorems 3-1 and 3-2, pages 76–77, together with (a) and (b),

$$E(2X + 3Y) = 2E(X) + 3E(Y) = 2\left(\frac{8}{3}\right) + 3\left(\frac{31}{9}\right) = \frac{47}{3}$$

3.6. Prove Theorem 3-2, page 77.

Let $f(x, y)$ be the joint probability function of X and Y, assumed discrete. Then

$$E(X + Y) = \sum_x \sum_y (x + y)f(x, y)$$
$$= \sum_x \sum_y xf(x, y) + \sum_x \sum_y yf(x, y)$$
$$= E(X) + E(Y)$$

If either variable is continuous, the proof goes through as before, with the appropriate summations replaced by integrations. Note that the theorem is true whether or not X and Y are independent.

3.7. Prove Theorem 3-3, page 77.

Let $f(x, y)$ be the joint probability function of X and Y, assumed discrete. If the variables X and Y are independent, we have $f(x, y) = f_1(x)f_2(y)$. Therefore,

$$\begin{aligned} E(XY) &= \sum_x \sum_y xyf(x, y) = \sum_x \sum_y xyf_1(x)f_2(y) \\ &= \sum_x \left[xf_1(x) \sum_y yf_2(y) \right] \\ &= \sum_x [(xf_1(x)E(y)] \\ &= E(X)E(Y) \end{aligned}$$

If either variable is continuous, the proof goes through as before, with the appropriate summations replaced by integrations. Note that the validity of this theorem hinges on whether $f(x, y)$ can be expressed as a function of x multiplied by a function of y, for all x and y, i.e., on whether X and Y are independent. For dependent variables it is not true in general.

Variance and standard deviation

3.8. Find (a) the variance, (b) the standard deviation of the sum obtained in tossing a pair of fair dice.

(a) Referring to Problem 3.2, we have $E(X) = E(Y) = 1/2$. Moreover,

$$E(X^2) = E(Y^2) = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + \cdots + 6^2\left(\frac{1}{6}\right) = \frac{91}{6}$$

Then, by Theorem 3-4,

$$\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

and, since X and Y are independent, Theorem 3-7 gives

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = \frac{35}{6}$$

(b)
$$\sigma_{X+Y} = \sqrt{\mathrm{Var}(X + Y)} = \sqrt{\frac{35}{6}}$$

3.9. Find (a) the variance, (b) the standard deviation for the random variable of Problem 3.4.

(a) As in Problem 3.4, the mean of X is $\mu = E(X) = \frac{1}{2}$. Then the variance is

$$\mathrm{Var}(X) = E[(X - \mu)^2] = E\left[\left(X - \frac{1}{2}\right)^2\right] = \int_{-\infty}^{\infty}\left(x - \frac{1}{2}\right)^2 f(x)\,dx$$

$$= \int_0^{\infty}\left(x - \frac{1}{2}\right)^2 (2e^{-2x})\,dx = \frac{1}{4}$$

Another method
By Theorem 3-4,

$$\mathrm{Var}(X) = E[(X - \mu)^2] = E(X^2) - [E(X)]^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

(b)
$$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{\frac{1}{4}} = \frac{1}{2}$$

3.10. Prove Theorem 3-4, page 78.

We have

$$\begin{aligned} E[(X-\mu)^2] &= E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

3.11. Prove Theorem 3-6, page 78.

$$\begin{aligned} E[(X-a)^2] &= E[\{(X-\mu) + (\mu - a)\}^2] \\ &= E[(X-\mu)^2 + 2(X-\mu)(\mu-a) + (\mu-a)^2] \\ &= E[(X-\mu)^2] + 2(\mu-a)E(X-\mu) + (\mu-a)^2 \\ &= E[(X-\mu)^2] + (\mu-a)^2 \end{aligned}$$

since $E(X-\mu) = E(X) - \mu = 0$. From this we see that the minimum value of $E[(X-a)^2]$ occurs when $(\mu - a)^2 = 0$, i.e., when $a = \mu$.

3.12. If $X^* = (X-\mu)/\sigma$ is a standardized random variable, prove that (a) $E(X^*) = 0$, (b) $\text{Var}(X^*) = 1$.

(a)
$$E(X^*) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}[E(X-\mu)] = \frac{1}{\sigma}[E(X) - \mu] = 0$$

since $E(X) = \mu$.

(b)
$$\text{Var}(X^*) = \text{Var}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}E[(X-\mu)^2] = 1$$

using Theorem 3-5, page 78, and the fact that $E[(X-\mu)^2] = \sigma^2$.

3.13. Prove Theorem 3-7, page 78.

$$\begin{aligned} \text{Var}(X+Y) &= E[\{(X+Y) - (\mu_X + \mu_Y)\}^2] \\ &= E[\{(X-\mu_X) + (Y-\mu_Y)\}^2] \\ &= E[(X-\mu_X)^2 + 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2] \\ &= E[(X-\mu_X)^2] + 2E[(X-\mu_X)(Y-\mu_Y)] + E[(Y-\mu_Y)^2] \\ &= \text{Var}(X) + \text{Var}(Y) \end{aligned}$$

using the fact that

$$E[(X-\mu_X)(Y-\mu_Y)] = E(X-\mu_X)E(Y-\mu_Y) = 0$$

since X and Y, and therefore $X - \mu_X$ and $Y - \mu_Y$, are independent. The proof of (*19*), page 78, follows on replacing Y by $-Y$ and using Theorem 3-5.

Moments and moment generating functions

3.14. Prove the result (26), page 79.

$$\begin{aligned} \mu_r &= E[(X-\mu)^r] \\ &= E\left[X^r - \binom{r}{1}X^{r-1}\mu + \cdots + (-1)^j\binom{r}{j}X^{r-j}\mu^j \right. \\ &\qquad \left. + \cdots + (-1)^{r-1}\binom{r}{r-1}X\mu^{r-1} + (-1)^r\mu^r\right] \end{aligned}$$

$$= E(X^r) - \binom{r}{1}E(X^{r-1})\mu + \cdots + (-1)^j\binom{r}{j}E(X^{r-j})\mu^j$$

$$+ \cdots + (-1)^{r-1}\binom{r}{r-1}E(X)\mu^{r-1} + (-1)^r\mu^r$$

$$= \mu'_r - \binom{r}{1}\mu'_{r-1}\mu + \cdots + (-1)^j\binom{r}{j}\mu'_{r-j}\mu^j$$

$$+ \cdots + (-1)^{r-1}r\mu^r + (-1)^{-r}\mu^r$$

where the last two terms can be combined to give $(-1)^{r-1}(r-1)\mu^r$.

3.15. Prove (a) result (31), (b) result (32), page 79.

(a) Using the power series expansion for e^u (3., Appendix A), we have

$$M_X(t) = E(e^{tX}) = E\left(1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdots\right)$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \cdots$$

$$= 1 + \mu t + \mu'_2\frac{t^2}{2!} + \mu'_3\frac{t^3}{3!} + \cdots$$

(b) This follows immediately from the fact known from calculus that if the Taylor series of $f(t)$ about $t = a$ is

$$f(t) = \sum_{n=0}^{\infty} c_n(t-a)^n$$

then

$$c_n = \frac{1}{n!}\frac{d^n}{dt^n}f(t)\bigg|_{t=a}$$

3.16. Prove Theorem 3-9, page 80.

Since X and Y are independent, any function of X and any function of Y are independent. Hence,

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E(e^{tX}e^{tY}) = E(e^{tX})E(e^{tY}) = M_X(t)M_Y(t)$$

3.17. The random variable X can assume the values 1 and -1 with probability $\frac{1}{2}$ each. Find (a) the moment generating function, (b) the first four moments about the origin.

(a)
$$E(e^{tX}) = e^{t(1)}\left(\frac{1}{2}\right) + e^{t(-1)}\left(\frac{1}{2}\right) = \frac{1}{2}(e^t + e^{-t})$$

(b) We have
$$e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots$$

$$e^{-t} = 1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \frac{t^4}{4!} - \cdots$$

Then (1)
$$\frac{1}{2}(e^t + e^{-t}) = 1 + \frac{t^2}{2!} + \frac{t^4}{4!} + \cdots$$

But (2)
$$M_X(t) = 1 + \mu t + \mu'_2\frac{t^2}{2!} + \mu'_3\frac{t^3}{3!} + \mu'_4\frac{t^4}{4!} + \cdots$$

Then, comparing (1) and (2), we have

$$\mu = 0, \qquad \mu'_2 = 1, \qquad \mu'_3 = 0, \qquad \mu'_4 = 1, \ldots$$

The odd moments are all zero, and the even moments are all one.

3.18. A random variable X has density function given by

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the moment generating function, (b) the first four moments about the origin.

(a)
$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$
$$= \int_0^{\infty} e^{tx}(2e^{-2x})\,dx = 2\int_0^{\infty} e^{(t-2)x}\,dx$$
$$= \frac{2e^{(t-2)x}}{t-2}\bigg|_0^{\infty} = \frac{2}{2-t}, \quad \text{assuming } t < 2$$

(b) If $|t| < 2$ we have

$$\frac{2}{2-t} = \frac{1}{1-t/2} = 1 + \frac{t}{2} + \frac{t^2}{4} + \frac{t^3}{8} + \frac{t^4}{16} + \cdots$$

But
$$M(t) = 1 + \mu t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \mu_4' \frac{t^4}{4!} + \cdots$$

Therefore, on comparing terms, $\mu = \frac{1}{2}, \mu_2' = \frac{1}{2}, \mu_3' = \frac{3}{4}, \mu_4' = \frac{3}{2}$.

3.19. Find the first four moments (a) about the origin, (b) about the mean, for a random variable X having density function

$$f(x) = \begin{cases} 4x(9-x^2)/81 & 0 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

(a)
$$\mu_1' = E(X) = \frac{4}{81}\int_0^3 x^2(9-x^2)\,dx = \frac{8}{5} = \mu$$
$$\mu_2' = E(X^2) = \frac{4}{81}\int_0^3 x^3(9-x^2)\,dx = 3$$
$$\mu_3' = E(X^3) = \frac{4}{81}\int_0^3 x^4(9-x^2)\,dx = \frac{216}{35}$$
$$\mu_4' = E(X^4) = \frac{4}{81}\int_0^3 x^5(9-x^2)\,dx = \frac{27}{2}$$

(b) Using the result (27), page 79, we have

$$\mu_1 = 0$$
$$\mu_2 = 3 - \left(\frac{8}{5}\right)^2 = \frac{11}{25} = \sigma^2$$
$$\mu_3 = \frac{216}{35} - 3(3)\left(\frac{8}{5}\right) + 2\left(\frac{8}{5}\right)^3 = -\frac{32}{875}$$
$$\mu_4 = \frac{27}{2} - 4\left(\frac{216}{35}\right)\left(\frac{8}{5}\right) + 6(3)\left(\frac{8}{5}\right)^2 - 3\left(\frac{8}{5}\right)^4 = \frac{3693}{8750}$$

Characteristic functions

3.20. Find the characteristic function of the random variable X of Problem 3.17.

The characteristic function is given by

$$E(e^{i\omega X}) = e^{i\omega(1)}\left(\frac{1}{2}\right) + e^{i\omega(-1)}\left(\frac{1}{2}\right) = \frac{1}{2}(e^{i\omega} + e^{-i\omega}) = \cos\omega$$

using Euler's formulas,

$$e^{i\theta} = \cos\theta + i\sin\theta \qquad e^{-i\theta} = \cos\theta - i\sin\theta$$

with $\theta = \omega$. The result can also be obtained from Problem 3.17(a) on putting $t = i\omega$.

3.21. Find the characteristic function of the random variable X having density function given by

$$f(x) = \begin{cases} 1/2a & |x| < a \\ 0 & \text{otherwise} \end{cases}$$

The characteristic function is given by

$$E(e^{i\omega X}) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\,dx = \frac{1}{2a}\int_{-a}^{a} e^{i\omega x}\,dx$$
$$= \frac{1}{2a}\frac{e^{i\omega x}}{i\omega}\bigg|_{-a}^{a} = \frac{e^{ia\omega} - e^{-ia\omega}}{2ia\omega} = \frac{\sin a\omega}{a\omega}$$

using Euler's formulas (see Problem 3.20) with $\theta = a\omega$.

3.22. Find the characteristic function of the random variable X having density function $f(x) = ce^{-a|x|}$, $-\infty < x < \infty$, where $a > 0$, and c is a suitable constant.

Since $f(x)$ is a density function, we must have

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

so that

$$c\int_{-\infty}^{\infty} e^{-a|x|}\,dx = c\left[\int_{-\infty}^{0} e^{-a(-x)}\,dx + \int_{0}^{\infty} e^{-a(x)}\,dx\right]$$
$$= c\frac{e^{ax}}{a}\bigg|_{-\infty}^{0} + c\frac{e^{-ax}}{-a}\bigg|_{0}^{\infty} = \frac{2c}{a} = 1$$

Then $c = a/2$. The characteristic function is therefore given by

$$E(e^{i\omega X}) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\,dx$$
$$= \frac{a}{2}\left[\int_{-\infty}^{0} e^{i\omega x}e^{-a(-x)}\,dx + \int_{0}^{\infty} e^{i\omega x}e^{-a(x)}\,dx\right]$$
$$= \frac{a}{2}\left[\int_{-\infty}^{0} e^{(a+i\omega)x}\,dx + \int_{0}^{\infty} e^{-(a-i\omega)x}\,dx\right]$$
$$= \frac{a}{2}\frac{e^{(a+i\omega)x}}{a + i\omega}\bigg|_{-\infty}^{0} + a\frac{e^{-(a-i\omega)x}}{-(a - i\omega)}\bigg|_{0}^{\infty}$$
$$= \frac{a}{2(a + i\omega)} + \frac{a}{2(a - i\omega)} = \frac{a^2}{a^2 + \omega^2}$$

Covariance and correlation coefficient

3.23. Prove Theorem 3-14, page 81.

By definition the covariance of X and Y is

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$
$$= E[XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y]$$
$$= E(XY) - \mu_X E(Y) - \mu_Y E(X) + E(\mu_X\mu_Y)$$
$$= E(XY) - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y$$
$$= E(XY) - \mu_X\mu_Y$$
$$= E(XY) - E(X)E(Y)$$

3.24. Prove Theorem 3-15, page 81.

If X and Y are independent, then $E(XY) = E(X)E(Y)$. Therefore, by Problem 3.23,

$$\sigma_{XY} = \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$$

3.25. Find (a) $E(X)$, (b) $E(Y)$, (c) $E(XY)$, (d) $E(X^2)$, (e) $E(Y^2)$, (f) Var (X), (g) Var (Y), (h) Cov (X, Y), (i) ρ, if the random variables X and Y are defined as in Problem 2.8, pages 47–48.

(a)
$$E(X) = \sum_x \sum_y x f(x, y) = \sum_x x\left[\sum_y f(x, y)\right]$$
$$= (0)(6c) + (1)(14c) + (2)(22c) = 58c = \frac{58}{42} = \frac{29}{21}$$

(b)
$$E(Y) = \sum_x \sum_y y f(x, y) = \sum_y y\left[\sum_x f(x, y)\right]$$
$$= (0)(6c) + (1)(9c) + (2)(12c) + (3)(15c) = 78c = \frac{78}{42} = \frac{13}{7}$$

(c)
$$E(XY) = \sum_x \sum_y xy f(x, y)$$
$$= (0)(0)(0) + (0)(1)(c) + (0)(2)(2c) + (0)(3)(3c)$$
$$+ (1)(0)(2c) + (1)(1)(3c) + (1)(2)(4c) + (1)(3)(5c)$$
$$+ (2)(0)(4c) + (2)(1)(5c) + (2)(2)(6c) + (2)(3)(7c)$$
$$= 102c = \frac{102}{42} = \frac{17}{7}$$

(d)
$$E(X^2) = \sum_x \sum_y x^2 f(x, y) = \sum_x x^2\left[\sum_y f(x, y)\right]$$
$$= (0)^2(6c) + (1)^2(14c) + (2)^2(22c) = 102c = \frac{102}{42} = \frac{17}{7}$$

(e)
$$E(Y^2) = \sum_x \sum_y y^2 f(x, y) = \sum_y y^2\left[\sum_x f(x, y)\right]$$
$$= (0)^2(6c) + (1)^2(9c) + (2)^2(12c) + (3)^2(15c) = 192c = \frac{192}{42} = \frac{32}{7}$$

(f)
$$\sigma_X^2 = \text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{17}{7} - \left(\frac{29}{21}\right)^2 = \frac{230}{441}$$

(g)
$$\sigma_Y^2 = \text{Var}(Y) = E(Y^2) - [E(Y)]^2 = \frac{32}{7} - \left(\frac{13}{7}\right)^2 = \frac{55}{49}$$

(h)
$$\sigma_{XY} = \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{17}{7} - \left(\frac{29}{21}\right)\left(\frac{13}{7}\right) = -\frac{20}{147}$$

(i)
$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-20/147}{\sqrt{230/441}\sqrt{55/49}} = \frac{-20}{\sqrt{230}\sqrt{55}} = -0.2103 \text{ approx.}$$

3.26. Work Problem 3.25 if the random variables X and Y are defined as in Problem 2.33, pages 61–63.

Using $c = 1/210$, we have:

(a)
$$E(X) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5} (x)(2x + y)\,dx\,dy = \frac{268}{63}$$

(b)
$$E(Y) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5} (y)(2x + y)\,dx\,dy = \frac{170}{63}$$

(c)
$$E(XY) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5} (xy)(2x + y)\,dx\,dy = \frac{80}{7}$$

(d) $$E(X^2) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5}(x^2)(2x + y)\,dx\,dy = \frac{1220}{63}$$

(e) $$E(Y^2) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5}(y^2)(2x + y)\,dx\,dy = \frac{1175}{126}$$

(f) $$\sigma_X^2 = \text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{1220}{63} - \left(\frac{268}{63}\right)^2 = \frac{5036}{3969}$$

(g) $$\sigma_Y^2 = \text{Var}(Y) = E(Y^2) - [E(Y)]^2 = \frac{1175}{126} - \left(\frac{170}{63}\right)^2 = \frac{16{,}225}{7938}$$

(h) $$\sigma_{XY} = \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{80}{7} - \left(\frac{268}{63}\right)\left(\frac{170}{63}\right) = -\frac{200}{3969}$$

(i) $$\rho = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{-200/3969}{\sqrt{5036/3969}\sqrt{16{,}225/7938}} = \frac{-200}{\sqrt{2518}\sqrt{16{,}225}} = -0.03129 \text{ approx.}$$

Conditional expectation, variance, and moments

3.27. Find the conditional expectation of Y given $X = 2$ in Problem 2.8, pages 47–48.

As in Problem 2.27, page 58, the conditional probability function of Y given $X = 2$ is

$$f(y\,|\,2) = \frac{4 + y}{22}$$

Then the conditional expectation of Y given $X = 2$ is

$$E(Y\,|\,X = 2) = \sum_y y\left(\frac{4 + y}{22}\right)$$

where the sum is taken over all y corresponding to $X = 2$. This is given by

$$E(Y\,|\,X = 2) = (0)\left(\frac{4}{22}\right) + 1\left(\frac{5}{22}\right) + 2\left(\frac{6}{22}\right) + 3\left(\frac{7}{22}\right) = \frac{19}{11}$$

3.28. Find the conditional expectation of (a) Y given X, (b) X given Y in Problem 2.29, pages 58–59.

(a) $$E(Y\,|\,X = x)\int_{-\infty}^{\infty} yf_2(y\,|\,x)\,dy = \int_0^x y\left(\frac{2y}{x^2}\right)dy = \frac{2x}{3}$$

(b) $$E(X\,|\,Y = y) = \int_{-\infty}^{\infty} xf_1(x\,|\,y)\,dx = \int_y^1 x\left(\frac{2x}{1 - y^2}\right)dx$$

$$= \frac{2(1 - y^3)}{3(1 - y^2)} = \frac{2(1 + y + y^2)}{3(1 + y)}$$

3.29. Find the conditional variance of Y given X for Problem 2.29, pages 58–59.

The required variance (second moment about the mean) is given by

$$E[(Y - \mu_2)^2\,|\,X = x] = \int_{-\infty}^{\infty}(y - \mu_2)^2 f_2(y\,|\,x)\,dy = \int_0^x\left(y - \frac{2x}{3}\right)^2\left(\frac{2y}{x^2}\right)dy = \frac{x^2}{18}$$

where we have used the fact that $\mu_2 = E(Y\,|\,X = x) = 2x/3$ from Problem 3.28(a).

Chebyshev's inequality

3.30. Prove Chebyshev's inequality.

We shall present the proof for continuous random variables. A proof for discrete variables is similar if integrals are replaced by sums. If $f(x)$ is the density function of X, then

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty}(x - \mu)^2 f(x)\,dx$$

Since the integrand is nonnegative, the value of the integral can only decrease when the range of integration is diminished. Therefore,

$$\sigma^2 \geq \int_{|x-\mu| \geq \epsilon} (x - \mu)^2 f(x)\,dx \geq \int_{|x-\mu| \geq \epsilon} \epsilon^2 f(x)\,dx = \epsilon^2 \int_{|x-\mu| \geq \epsilon} f(x)\,dx$$

But the last integral is equal to $P(|X - \mu| \geq \epsilon)$. Hence,

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

3.31. For the random variable of Problem 3.18, (a) find $P(|X - \mu| > 1)$. (b) Use Chebyshev's inequality to obtain an upper bound on $P(|X - \mu| > 1)$ and compare with the result in (a).

(a) From Problem 3.18, $\mu = 1/2$. Then

$$P(|X - \mu| < 1) = P\left(\left|X - \frac{1}{2}\right| < 1\right) = P\left(-\frac{1}{2} < X < \frac{3}{2}\right)$$

$$= \int_0^{3/2} 2e^{-2x}\,dx = 1 - e^{-3}$$

Therefore $$P\left(\left|X - \frac{1}{2}\right| \geq 1\right) = 1 - (1 - e^{-3}) = e^{-3} = 0.04979$$

(b) From Problem 3.18, $\sigma^2 = \mu_2' - \mu^2 = 1/4$. Chebyshev's inequality with $\epsilon = 1$ then gives

$$P(|X - \mu| \geq 1) \leq \sigma^2 = 0.25$$

Comparing with (a), we see that the bound furnished by Chebyshev's inequality is here quite crude. In practice, Chebyshev's inequality is used to provide estimates when it is inconvenient or impossible to obtain exact values.

Law of large numbers

3.32. Prove the law of large numbers stated in Theorem 3-19, page 83.

We have $$E(X_1) = E(X_2) = \cdots = E(X_n) = \mu$$

$$\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = \cdots = \mathrm{Var}(X_n) = \sigma^2$$

Then $$E\left(\frac{S_n}{n}\right) = E\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{1}{n}[E(X_1) + \cdots + E(X_n)] = \frac{1}{n}(n\mu) = \mu$$

$$\mathrm{Var}(S_n) = \mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = n\sigma^2$$

so that $$\mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\mathrm{Var}(S_n) = \frac{\sigma^2}{n}$$

where we have used Theorem 3-5 and an extension of Theorem 3-7.

Therefore, by Chebyshev's inequality with $X = S_n/n$, we have

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$$

Taking the limit as $n \to \infty$, this becomes, as required,

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0$$

Other measures of central tendency

3.33. The density function of a continuous random variable X is

$$f(x) = \begin{cases} 4x(9 - x^2)/81 & 0 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the mode. (b) Find the median. (c) Compare mode, median, and mean.

(a) The mode is obtained by finding where the density $f(x)$ has a relative maximum. The relative maxima of $f(x)$ occur where the derivative is zero, i.e.,

$$\frac{d}{dx}\left[\frac{4x(9 - x^2)}{81}\right] = \frac{36 - 12x^2}{81} = 0$$

Then $x = \sqrt{3} = 1.73$ approx., which is the required mode. Note that this does give the maximum since the second derivative, $-24x/81$, is negative for $x = \sqrt{3}$.

(b) The median is that value a for which $P(X \le a) = 1/2$. Now, for $0 < a < 3$,

$$P(X \le a) = \frac{4}{81}\int_0^a x(9 - x^2)\,dx = \frac{4}{81}\left(\frac{9a^2}{2} - \frac{a^4}{4}\right)$$

Setting this equal to $1/2$, we find that

$$2a^4 - 36a^2 + 81 = 0$$

from which

$$a^2 = \frac{36 \pm \sqrt{(36)^2 - 4(2)(81)}}{2(2)} = \frac{36 \pm \sqrt{648}}{4} = 9 \pm \frac{9}{2}\sqrt{2}$$

Therefore, the required median, which must lie between 0 and 3, is given by

$$a^2 = 9 - \frac{9}{2}\sqrt{2}$$

from which $a = 1.62$ approx.

(c) $$E(X) = \frac{4}{81}\int_0^3 x^2(9 - x^2)\,dx = \frac{4}{81}\left(3x^3 - \frac{x^5}{5}\right)\Bigg|_0^3 = 1.60$$

which is practically equal to the median. The mode, median, and mean are shown in Fig. 3-6.

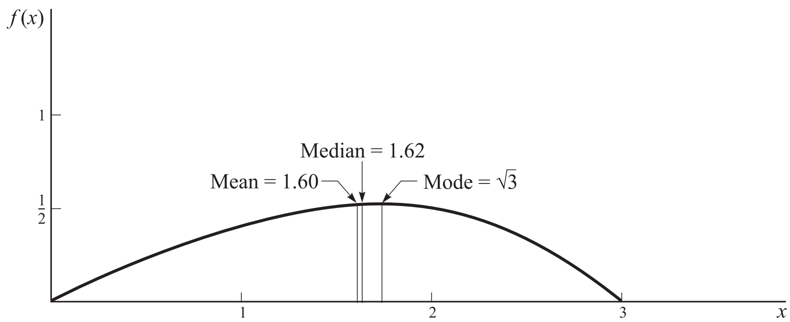

Fig. 3-6

3.34. A discrete random variable has probability function $f(x) = 1/2^x$ where $x = 1, 2, \ldots$. Find (a) the mode, (b) the median, and (c) compare them with the mean.

(a) The mode is the value x having largest associated probability. In this case it is $x = 1$, for which the probability is $1/2$.

(b) If x is any value between 1 and 2, $P(X < x) = \frac{1}{2}$ and $P(X > x) = \frac{1}{2}$. Therefore, *any number* between 1 and 2 could represent the median. For convenience, we choose the midpoint of the interval, i.e., $3/2$.

(c) As found in Problem 3.3, $\mu = 2$. Therefore, the ordering of the three measures is just the reverse of that in Problem 3.33.

Percentiles

3.35. Determine the (a) 10th, (b) 25th, (c) 75th percentile values for the distribution of Problem 3.33.

From Problem 3.33(b) we have

$$P(X \leq a) = \frac{4}{81}\left(\frac{9a^2}{2} - \frac{a^4}{4}\right) = \frac{18a^2 - a^4}{81}$$

(a) The 10th percentile is the value of a for which $P(X \leq a) = 0.10$, i.e., the solution of $(18a^2 - a^4)/81 = 0.10$. Using the method of Problem 3.33, we find $a = 0.68$ approx.

(b) The 25th percentile is the value of a such that $(18a^2 - a^4)/81 = 0.25$, and we find $a = 1.098$ approx.

(c) The 75th percentile is the value of a such that $(18a^2 - a^4)/81 = 0.75$, and we find $a = 2.121$ approx.

Other measures of dispersion

3.36. Determine, (a) the semi-interquartile range, (b) the mean deviation for the distribution of Problem 3.33.

(a) By Problem 3.35 the 25th and 75th percentile values are 1.098 and 2.121, respectively. Therefore,

$$\text{Semi-interquartile range} = \frac{2.121 - 1.098}{2} = 0.51 \text{ approx.}$$

(b) From Problem 3.33 the mean is $\mu = 1.60 = 8/5$. Then

$$\begin{aligned}
\text{Mean deviation} = \text{M.D.} = E(|X - \mu|) &= \int_{-\infty}^{\infty} |x - \mu| f(x)\,dx \\
&= \int_0^3 \left| x - \frac{8}{5} \right| \left[\frac{4x}{81}(9 - x^2)\right] dx \\
&= \int_0^{8/5} \left(\frac{8}{5} - x\right)\left[\frac{4x}{81}(9 - x^2)\right] dx + \int_{8/5}^3 \left(x - \frac{8}{5}\right)\left[\frac{4x}{81}(9 - x^2)\right] dx \\
&= 0.555 \text{ approx.}
\end{aligned}$$

Skewness and kurtosis

3.37. Find the coefficient of (a) skewness, (b) kurtosis for the distribution of Problem 3.19.

From Problem 3.19(b) we have

$$\sigma^2 = \frac{11}{25} \qquad \mu_3 = -\frac{32}{875} \qquad \mu_4 = \frac{3693}{8750}$$

(a) Coefficient of skewness $= \alpha_3 = \dfrac{\mu_3}{\sigma^3} = -0.1253$

(b) Coefficient of kurtosis $= \alpha_4 = \dfrac{\mu_4}{\sigma^4} = 2.172$

It follows that there is a moderate skewness to the left, as is indicated in Fig. 3-6. Also the distribution is somewhat less peaked than the normal distribution, which has a kurtosis of 3.

Miscellaneous problems

3.38. If $M(t)$ is the moment generating function for a random variable X, prove that the mean is $\mu = M'(0)$ and the variance is $\sigma^2 = M''(0) - [M'(0)]^2$.

From (32), page 79, we have on letting $r = 1$ and $r = 2$,

$$\mu_1' = M'(0) \qquad \mu_2' = M''(0)$$

Then from (27)

$$\mu = M'(0) \qquad \mu_2 = \sigma^2 = M''(0) - [M'(0)]^2$$

3.39. Let X be a random variable that takes on the values $x_k = k$ with probabilities p_k where $k = \pm 1, \dots, \pm n$. (a) Find the characteristic function $\phi(\omega)$ of X, (b) obtain p_k in terms of $\phi(\omega)$.

(a) The characteristic function is

$$\phi(\omega) = E(e^{i\omega X}) = \sum_{k=-n}^{n} e^{i\omega x_k} p_k = \sum_{k=-n}^{n} p_k e^{ik\omega}$$

(b) Multiply both sides of the expression in (a) by $e^{-ij\omega}$ and integrate with respect to ω from 0 to 2π. Then

$$\int_{\omega=0}^{2\pi} e^{-ij\omega}\phi(\omega)\,d\omega = \sum_{k=-n}^{n} p_k \int_{\omega=0}^{2\pi} e^{i(k-j)\omega}\,d\omega = 2\pi p_j$$

since

$$\int_{\omega=0}^{2\pi} e^{i(k-j)\omega}\,d\omega = \begin{cases} \left.\dfrac{e^{i(k-j)\omega}}{i(k-j)}\right|_0^{2\pi} = 0 & k \neq j \\ 2\pi & k = j \end{cases}$$

Therefore,

$$p_j = \frac{1}{2\pi}\int_{\omega=0}^{2\pi} e^{-ij\omega}\phi(\omega)\,d\omega$$

or, replacing j by k,

$$p_k = \frac{1}{2\pi}\int_{\omega=0}^{2\pi} e^{-ik\omega}\phi(\omega)\,d\omega$$

We often call $\sum_{k=-n}^{n} p_k e^{ik\omega}$ (where n can theoretically be infinite) the *Fourier series* of $\phi(\omega)$ and p_k the *Fourier coefficients.* For a continuous random variable, the Fourier series is replaced by the Fourier integral (see page 81).

3.40. Use Problem 3.39 to obtain the probability distribution of a random variable X whose characteristic function is $\phi(\omega) = \cos\omega$.

From Problem 3.39

$$\begin{aligned} p_k &= \frac{1}{2\pi}\int_{\omega=0}^{2\pi} e^{-ik\omega}\cos\omega\,d\omega \\ &= \frac{1}{2\pi}\int_{\omega=0}^{2\pi} e^{-ik\omega}\left[\frac{e^{i\omega}+e^{-i\omega}}{2}\right]d\omega \\ &= \frac{1}{4\pi}\int_{\omega=0}^{2\pi} e^{i(1-k)\omega}\,d\omega + \frac{1}{4\pi}\int_{\omega=0}^{2\pi} e^{-i(1+k)\omega}\,d\omega \end{aligned}$$

If $k = 1$, we find $p_1 = \frac{1}{2}$; if $k = -1$, we find $p_{-1} = \frac{1}{2}$. For all other values of k, we have $p_k = 0$. Therefore, the random variable is given by

$$X = \begin{cases} 1 & \text{probability } 1/2 \\ -1 & \text{probability } 1/2 \end{cases}$$

As a check, see Problem 3.20.

3.41. Find the coefficient of (a) skewness, (b) kurtosis of the distribution defined by the *normal curve*, having density

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

(a) The distribution has the appearance of Fig. 3-7. By symmetry, $\mu_1' = \mu = 0$ and $\mu_3' = 0$. Therefore the coefficient of skewness is zero.

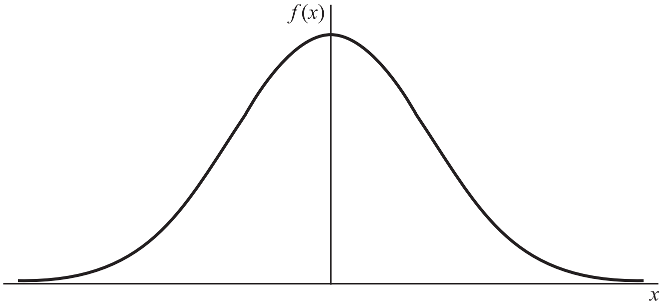

Fig. 3-7

(b) We have

$$\mu_2' = E(X^2) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2 e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} x^2 e^{-x^2/2}\,dx$$
$$= \frac{2}{\sqrt{\pi}}\int_0^{\infty} v^{1/2}e^{-v}\,dv$$
$$= \frac{2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) = \frac{2}{\sqrt{\pi}}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = 1$$

where we have made the transformation $x^2/2 = v$ and used properties of the gamma function given in (2) and (5) of Appendix A. Similarly we obtain

$$\mu_4' = E(X^4) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^4 e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} x^4 e^{-x^2/2}\,dx$$
$$= \frac{4}{\sqrt{\pi}}\int_0^{\infty} v^{3/2}e^{-v}\,dv$$
$$= \frac{4}{\sqrt{\pi}}\Gamma\left(\frac{5}{2}\right) = \frac{4}{\sqrt{\pi}}\cdot\frac{3}{2}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = 3$$

Now

$$\sigma^2 = E[(X-\mu)^2] = E(X)^2 = \mu_2' = 1$$
$$\mu_4 = E[(X-\mu)^4] = E(X^4) = \mu_4' = 3$$

Thus the coefficient of kurtosis is

$$\frac{\mu_4}{\sigma^4} = 3$$

3.42. Prove that $-1 \le \rho \le 1$ (see page 82).

For any real constant c, we have

$$E[\{Y - \mu_Y - c(X-\mu)\}^2] \ge 0$$

Now the left side can be written

$$E[(Y-\mu_Y)^2] + c^2E[(X-\mu_X)^2] - 2cE[(X-\mu_X)(Y-\mu_Y)] = \sigma_Y^2 + c^2\sigma_X^2 - 2c\sigma_{XY}$$
$$= \sigma_Y^2 + \sigma_X^2\left(c^2 - \frac{2c\sigma_{XY}}{\sigma_X^2}\right)$$
$$= \sigma_Y^2 + \sigma_X^2\left(c^2 - \frac{\sigma_{XY}}{\sigma_X^2}\right)^2 - \frac{\sigma_{XY}^2}{\sigma_X^2}$$
$$= \frac{\sigma_X^2\sigma_Y^2 - \sigma_{XY}^2}{\sigma_X^2} + \sigma_X^2\left(c - \frac{\sigma_{XY}}{\sigma_X^2}\right)^2$$

In order for this last quantity to be greater than or equal to zero for every value of c, we must have

$$\sigma_X^2\sigma_Y^2 - \sigma_{XY}^2 \geq 0 \quad \text{or} \quad \frac{\sigma_{XY}^2}{\sigma_X^2\sigma_Y^2} \leq 1$$

which is equivalent to $\rho^2 \leq 1$ or $-1 \leq \rho \leq 1$.

SUPPLEMENTARY PROBLEMS

Expectation of random variables

3.43. A random variable X is defined by $X = \begin{cases} -2 & \text{prob. } 1/3 \\ 3 & \text{prob. } 1/2. \\ 1 & \text{prob. } 1/6 \end{cases}$ Find (a) $E(X)$, (b) $E(2X + 5)$, (c) $E(X^2)$.

3.44. Let X be a random variable defined by the density function $f(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$.

Find (a) $E(X)$, (b) $E(3X - 2)$, (c) $E(X^2)$.

3.45. The density function of a random variable X is $f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$.

Find (a) $E(X)$, (b) $E(X^2)$, (c) $E[(X - 1)^2]$.

3.46. What is the expected number of points that will come up in 3 successive tosses of a fair die? Does your answer seem reasonable? Explain.

3.47. A random variable X has the density function $f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$. Find $E(e^{2X/3})$.

3.48. Let X and Y be independent random variables each having density function

$$f(u) = \begin{cases} 2e^{-2u} & u \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $E(X + Y)$, (b) $E(X^2 + Y^2)$, (c) $E(XY)$.

3.49. Does (a) $E(X + Y) = E(X) + E(Y)$, (b) $E(XY) = E(X)E(Y)$, in Problem 3.48? Explain.

3.50. Let X and Y be random variables having joint density function

$$f(x, y) = \begin{cases} \frac{3}{5}x(x + y) & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $E(X)$, (b) $E(Y)$, (c) $E(X + Y)$, (d) $E(XY)$.

3.51. Does (a) $E(X + Y) = E(X) + E(Y)$, (b) $E(XY) = E(X)E(Y)$, in Problem 3.50? Explain.

3.52. Let X and Y be random variables having joint density

$$f(x, y) = \begin{cases} 4xy & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $E(X)$, (b) $E(Y)$, (c) $E(X + Y)$, (d) $E(XY)$.

3.53. Does (a) $E(X + Y) = E(X) + E(Y)$, (b) $E(XY) = E(X)E(Y)$, in Problem 3.52? Explain.

3.54. Let $f(x, y) = \begin{cases} \frac{1}{4}(2x + y) & 0 \le x \le 1, 0 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$. Find (a) $E(X)$, (b) $E(Y)$, (c) $E(X^2)$, (d) $E(Y^2)$, (e) $E(X + Y)$, (f) $E(XY)$.

3.55. Let X and Y be independent random variables such that

$$X = \begin{cases} 1 & \text{prob. } 1/3 \\ 0 & \text{prob. } 2/3 \end{cases} \qquad Y = \begin{cases} 2 & \text{prob. } 3/4 \\ -3 & \text{prob. } 1/4 \end{cases}$$

Find (a) $E(3X + 2Y)$, (b) $E(2X^2 - Y^2)$, (c) $E(XY)$, (d) $E(X^2Y)$.

3.56. Let $X_1, X_2, \ldots, X_n$ be n random variables which are identically distributed such that

$$X_k = \begin{cases} 1 & \text{prob. } 1/2 \\ 2 & \text{prob. } 1/3 \\ -1 & \text{prob. } 1/6 \end{cases}$$

Find (a) $E(X_1 + X_2 + \cdots + X_n)$, (b) $E(X_1^2 + X_2^2 + \cdots + X_n^2)$.

Variance and standard deviation

3.57. Find (a) the variance, (b) the standard deviation of the number of points that will come up on a single toss of a fair die.

3.58. Let X be a random variable having density function

$$f(x) = \begin{cases} 1/4 & -2 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $\text{Var}(X)$, (b) σ_X.

3.59. Let X be a random variable having density function

$$f(x) = \begin{cases} e^{-x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $\text{Var}(X)$, (b) σ_X.

3.60. Find the variance and standard deviation for the random variable X of (a) Problem 3.43, (b) Problem 3.44.

3.61. A random variable X has $E(X) = 2$, $E(X^2) = 8$. Find (a) $\text{Var}(X)$, (b) σ_X.

3.62. If a random variable X is such that $E[(X - 1)^2] = 10$, $E[(X - 2)^2] = 6$ find (a) $E(X)$, (b) $\text{Var}(X)$, (c) σ_X.

Moments and moment generating functions

3.63. Find (a) the moment generating function of the random variable

$$X = \begin{cases} 1/2 & \text{prob. } 1/2 \\ -1/2 & \text{prob. } 1/2 \end{cases}$$

and (b) the first four moments about the origin.

3.64. (a) Find the moment generating function of a random variable X having density function

$$f(x) = \begin{cases} x/2 & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

(b) Use the generating function of (a) to find the first four moments about the origin.

3.65. Find the first four moments about the mean in (a) Problem 3.43, (b) Problem 3.44.

3.66. (a) Find the moment generating function of a random variable having density function

$$f(x) = \begin{cases} e^{-x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

and (b) determine the first four moments about the origin.

3.67. In Problem 3.66 find the first four moments about the mean.

3.68. Let X have density function $f(x) = \begin{cases} 1/(b-a) & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$. Find the kth moment about (a) the origin, (b) the mean.

3.69. If $M(t)$ is the moment generating function of the random variable X, prove that the 3rd and 4th moments about the mean are given by

$$\mu_3 = M'''(0) - 3M''(0)M'(0) + 2[M'(0)]^3$$
$$\mu_4 = M^{(\text{iv})}(0) - 4M'''(0)M'(0) + 6M''(0)[M'(0)]^2 - 3[M'(0)]^4$$

Characteristic functions

3.70. Find the characteristic function of the random variable $X = \begin{cases} a & \text{prob. } p \\ b & \text{prob. } q = 1 - p \end{cases}$.

3.71. Find the characteristic function of a random variable X that has density function

$$f(x) = \begin{cases} 1/2a & |x| \le a \\ 0 & \text{otherwise} \end{cases}$$

3.72. Find the characteristic function of a random variable with density function

$$f(x) = \begin{cases} x/2 & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

3.73. Let $X_k = \begin{cases} 1 & \text{prob. } 1/2 \\ -1 & \text{prob. } 1/2 \end{cases}$ be independent random variables ($k = 1, 2, \ldots, n$). Prove that the characteristic function of the random variable

$$\frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}}$$

is $[\cos(\omega/\sqrt{n})]^n$.

3.74. Prove that as $n \to \infty$ the characteristic function of Problem 3.73 approaches $e^{-\omega^2/2}$. (*Hint*: Take the logarithm of the characteristic function and use L'Hospital's rule.)

Covariance and correlation coefficient

3.75. Let X and Y be random variables having joint density function

$$f(x, y) = \begin{cases} x + y & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $\text{Var}(X)$, (b) $\text{Var}(Y)$, (c) σ_X, (d) σ_Y, (e) σ_{XY}, (f) ρ.

3.76. Work Problem 3.75 if the joint density function is $f(x, y) = \begin{cases} e^{-(x+y)} & x \ge 0,\ y \ge 0 \\ 0 & \text{otherwise} \end{cases}$.

3.77. Find (a) $\text{Var}(X)$, (b) $\text{Var}(Y)$, (c) σ_X, (d) σ_Y, (e) σ_{XY}, (f) ρ, for the random variables of Problem 2.56.

3.78. Work Problem 3.77 for the random variables of Problem 2.94.

3.79. Find (a) the covariance, (b) the correlation coefficient of two random variables X and Y if $E(X) = 2$, $E(Y) = 3$, $E(XY) = 10$, $E(X^2) = 9$, $E(Y^2) = 16$.

3.80. The correlation coefficient of two random variables X and Y is $-\frac{1}{4}$ while their variances are 3 and 5. Find the covariance.

Conditional expectation, variance, and moments

3.81. Let X and Y have joint density function

$$f(x, y) = \begin{cases} x + y & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the conditional expectation of (a) Y given X, (b) X given Y.

3.82. Work Problem 3.81 if $f(x, y) = \begin{cases} 2e^{-(x+2y)} & x \ge 0,\ y \ge 0 \\ 0 & \text{otherwise} \end{cases}$

3.83. Let X and Y have the joint probability function given in Table 2-9, page 71. Find the conditional expectation of (a) Y given X, (b) X given Y.

3.84. Find the conditional variance of (a) Y given X, (b) X given Y for the distribution of Problem 3.81.

3.85. Work Problem 3.84 for the distribution of Problem 3.82.

3.86. Work Problem 3.84 for the distribution of Problem 2.94.

Chebyshev's inequality

3.87. A random variable X has mean 3 and variance 2. Use Chebyshev's inequality to obtain an upper bound for (a) $P(|X - 3| \ge 2)$, (b) $P(|X - 3| \ge 1)$.

3.88. Prove Chebyshev's inequality for a discrete variable X. *(Hint*: See Problem 3.30.)

3.89. A random variable X has the density function $f(x) = \frac{1}{2}e^{-|x|}$, $-\infty < x < \infty$. (a) Find $P(|X - \mu| > 2)$. (b) Use Chebyshev's inequality to obtain an upper bound on $P(|X - \mu| > 2)$ and compare with the result in (a).

Law of large numbers

3.90. Show that the (weak) law of large numbers can be stated as

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) = 1$$

and interpret.

3.91. Let X_k ($k = 1, \ldots, n$) be n independent random variables such that

$$X_k = \begin{cases} 1 & \text{prob. } p \\ 0 & \text{prob. } q = 1 - p \end{cases}$$

(a) If we interpret X_k to be the number of heads on the kth toss of a coin, what interpretation can be given to $S_n = X_1 + \cdots + X_n$?

(b) Show that the law of large numbers in this case reduces to

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) = 0$$

and interpret this result.

Other measures of central tendency

3.92. Find (a) the mode, (b) the median of a random variable X having density function

$$f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and (c) compare with the mean.

3.93. Work Problem 3.100 if the density function is

$$f(x) = \begin{cases} 4x(1 - x^2) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

3.94. Find (a) the median, (b) the mode for a random variable X defined by

$$X = \begin{cases} 2 & \text{prob. } 1/3 \\ -1 & \text{prob. } 2/3 \end{cases}$$

and (c) compare with the mean.

3.95. Find (a) the median, (b) the mode of the set of numbers 1, 3, 2, 1, 5, 6, 3, 3, and (c) compare with the mean.

Percentiles

3.96. Find the (a) 25th, (b) 75th percentile values for the random variable having density function

$$f(x) = \begin{cases} 2(1 - x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

3.97. Find the (a) 10th, (b) 25th, (c) 75th, (d) 90th percentile values for the random variable having density function

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

where c is an appropriate constant.

Other measures of dispersion

3.98. Find (a) the semi-interquartile range, (b) the mean deviation for the random variable of Problem 3.96.

3.99. Work Problem 3.98 for the random variable of Problem 3.97.

3.100. Find the mean deviation of the random variable X in each of the following cases.

(a) $f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$ (b) $f(x) = \dfrac{1}{\pi(1 + x^2)}, \quad -\infty < x < \infty.$

3.101. Obtain the probability that the random variable X differs from its mean by more than the semi-interquartile range in the case of (a) Problem 3.96, (b) Problem 3.100(a).

Skewness and kurtosis

3.102. Find the coefficient of (a) skewness, (b) kurtosis for the distribution of Problem 3.100(a).

3.103. If

$$f(x) = \begin{cases} c\left(1 - \dfrac{|x|}{a}\right) & |x| \leq a \\ 0 & |x| > a \end{cases}$$

where c is an appropriate constant, is the density function of X, find the coefficient of (a) skewness, (b) kurtosis.

3.104. Find the coefficient of (a) skewness, (b) kurtosis, for the distribution with density function

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

Miscellaneous problems

3.105. Let X be a random variable that can take on the values 2, 1, and 3 with respective probabilities $1/3$, $1/6$, and $1/2$. Find (a) the mean, (b) the variance, (c) the moment generating function, (d) the characteristic function, (e) the third moment about the mean.

3.106. Work Problem 3.105 if X has density function

$$f(x) = \begin{cases} c(1 - x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where c is an appropriate constant.

3.107. Three dice, assumed fair, are tossed successively. Find (a) the mean, (b) the variance of the sum.

3.108. Let X be a random variable having density function

$$f(x) = \begin{cases} cx & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

where c is an appropriate constant. Find (a) the mean, (b) the variance, (c) the moment generating function, (d) the characteristic function, (e) the coefficient of skewness, (f) the coefficient of kurtosis.

3.109. Let X and Y have joint density function

$$f(x, y) = \begin{cases} cxy & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find (a) $E(X^2 + Y^2)$, (b) $E(\sqrt{X^2 + Y^2})$.

3.110. Work Problem 3.109 if X and Y are independent identically distributed random variables having density function $f(u) = (2\pi)^{-1/2}e^{-u^2/2}$, $-\infty < u < \infty$.

3.111. Let X be a random variable having density function

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

and let $Y = X^2$. Find (a) $E(X)$, (b) $E(Y)$, (c) $E(XY)$.

ANSWERS TO SUPPLEMENTARY PROBLEMS

3.43. (a) 1 (b) 7 (c) 6 **3.44.** (a) $3/4$ (b) $1/4$ (c) $3/5$

3.45. (a) 1 (b) 2 (c) 1 **3.46.** 10.5 **3.47.** 3

3.48. (a) 1 (b) 1 (c) $1/4$

3.50. (a) $7/10$ (b) $6/5$ (c) $19/10$ (d) $5/6$

3.52. (a) $2/3$ (b) $2/3$ (c) $4/3$ (d) $4/9$

3.54. (a) $7/12$ (b) $7/6$ (c) $5/12$ (d) $5/3$ (e) $7/4$ (f) $2/3$

3.55. (a) $5/2$ (b) $-55/12$ (c) $1/4$ (d) $1/4$

3.56. (a) n (b) $2n$ **3.57.** (a) $35/12$ (b) $\sqrt{35/12}$

3.58. (a) $4/3$ (b) $\sqrt{4/3}$ **3.59.** (a) 1 (b) 1

3.60. (a) $\text{Var}(X) = 5, \sigma_X = \sqrt{5}$ (b) $\text{Var}(X) = 3/80, \sigma_X = \sqrt{15}/20$

3.61. (a) 4 (b) 2 **3.62.** (a) $7/2$ (b) $15/4$ (c) $\sqrt{15}/2$

3.63. (a) $\frac{1}{2}(e^{t/2} + e^{-t/2}) = \cosh(t/2)$ (b) $\mu = 0, \mu_2' = 1, \mu_3' = 0, \mu_4' = 1$

3.64. (a) $(1 + 2te^{2t} - e^{2t})/2t^2$ (b) $\mu = 4/3, \mu_2' = 2, \mu_3' = 16/5, \mu_4' = 16/3$

3.65. (a) $\mu_1 = 0, \mu_2 = 5, \mu_3 = -5, \mu_4 = 35$ (b) $\mu_1 = 0, \mu_2 = 3/80, \mu_3 = -121/160, \mu_4 = 2307/8960$

3.66. (a) $1/(1 - t), |t| < 1$ (b) $\mu = 1, \mu_2' = 2, \mu_3' = 6, \mu_4' = 24$

3.67. $\mu_1 = 0, \mu_2 = 1, \mu_3 = 2, \mu_4 = 33$

3.68. (a) $(b^{k+1} - a^{k+1})/(k + 1)(b - a)$ (b) $[1 + (-1)^k](b - a)^k/2^{k+1}(k + 1)$

3.70. $pe^{i\omega a} + qe^{i\omega b}$ **3.71.** $(\sin a\omega)/a\omega$ **3.72.** $(e^{2i\omega} - 2i\omega e^{2i\omega} - 1)/2\omega^2$

3.75. (a) $11/144$ (b) $11/144$ (c) $\sqrt{11}/12$ (d) $\sqrt{11}/12$ (e) $-1/144$ (f) $-1/11$

3.76. (a) 1 (b) 1 (c) 1 (d) 1 (e) 0 (f) 0

3.77. (a) $73/960$ (b) $73/960$ (c) $\sqrt{73/960}$ (d) $\sqrt{73/960}$ (e) $-1/64$ (f) $-15/73$

3.78. (a) $233/324$ (b) $233/324$ (c) $\sqrt{233}/18$ (d) $\sqrt{233}/18$ (e) $-91/324$ (f) $-91/233$

3.79. (a) 4 (b) $4/\sqrt{35}$ **3.80.** $-\sqrt{15}/4$

3.81. (a) $(3x + 2)/(6x + 3)$ for $0 \le x \le 1$ (b) $(3y + 2)/(6y + 3)$ for $0 \le y \le 1$

3.82. (a) $1/2$ for $x \ge 0$ (b) 1 for $y \ge 0$

3.83. (a)

X	0	1	2
$E(Y \mid X)$	$4/3$	1	$5/7$

(b)

Y	0	1	2
$E(X \mid Y)$	$4/3$	$7/6$	$1/2$

3.84. (a) $\dfrac{6x^2 + 6x + 1}{18(2x + 1)^2}$ for $0 \le x \le 1$ (b) $\dfrac{6y^2 + 6y + 1}{18(2y + 1)^2}$ for $0 \le y \le 1$

3.85. (a) $1/9$ (b) 1

3.86. (a)

X	0	1	2
$\mathrm{Var}(Y \mid X)$	$5/9$	$4/5$	$24/49$

(b)

Y	0	1	2
$\mathrm{Var}(X \mid Y)$	$5/9$	$29/36$	$7/12$

3.87. (a) $1/2$ (b) 2 (useless) **3.89.** (a) e^{-2} (b) 0.5

3.92. (a) $+0$ (b) $\ln 2$ (c) 1 **3.93.** (a) $1/\sqrt{3}$ (b) $\sqrt{1 - (1/\sqrt{2})}$ (c) $8/15$

3.94. (a) does not exist (b) -1 (c) 0 **3.95.** (a) 3 (b) 3 (c) 3

3.96. (a) $1 - \frac{1}{2}\sqrt{3}$ (b) $1/2$

3.97. (a) $\sqrt{1 - (3/\sqrt{10})}$ (b) $\sqrt{1 - (\sqrt{3}/2)}$ (c) $\sqrt{1/2}$ (d) $\sqrt{1 - (1/\sqrt{10})}$

3.98. (a) 1 (b) $(\sqrt{3} - 1)/4$ (c) $16/81$

3.99. (a) 1 (b) 0.17 (c) 0.051 **3.100.** (a) $1 - 2e^{-1}$ (b) does not exist

3.101. (a) $(5 - 2\sqrt{3})/3$ (b) $(3 - 2e^{-1}\sqrt{3})/3$

3.102. (a) 2 (b) 9 **3.103.** (a) 0 (b) $24/5a$ **3.104.** (a) 2 (b) 9

3.105. (a) 7/3 (b) 5/9 (c) $(e^t + 2e^{2t} + 3e^{3t})/6$ (d) $(e^{i\omega} + 2e^{2i\omega} + 3e^{3i\omega})/6$ (e) $-7/27$

3.106. (a) 1/3 (b) 1/18 (c) $2(e^t - 1 - t)/t^2$ (d) $-2(e^{i\omega} - 1 - i\omega)/\omega^2$ (*e*) 1/135

3.107. (a) 21/2 (b) 35/4

3.108. (a) 4/3 (b) 2/9 (c) $(1 + 2te^{2t} - e^{2t})/2t^2$ (d) $-(1 + 2i\omega e^{2i\omega} - e^{2i\omega})/2\omega^2$
(e) $-2\sqrt{18}/15$ (f) 12/5

3.109. (a) 1 (b) $8(2\sqrt{2} - 1)/15$

3.110. (a) 2 (b) $\sqrt{2\pi}/2$

3.111. (a) 0 (b) 1/3 (c) 0