5 Joint Probability Distributions

CHAPTER OUTLINE

LEARNING OBJECTIVES

After careful study of this chapter you should be able to do the following:

1. Use joint probability mass functions and joint probability density functions to calculate probabilities
2. Calculate marginal and conditional probability distributions from joint probability distributions
3. Use the multinomial distribution to determine probabilities
4. Interpret and calculate covariances and correlations between random variables
5. Understand properties of a bivariate normal distribution and be able to draw contour plots for the probability density function
6. Calculate means and variance for linear combinations of random variables and calculate probabilities for linear combinations of normally distributed random variables
7. Determine the distribution of a general function of a random variable

5-1 Two Discrete Random Variables

Example 5-1

In the development of a new receiver for the transmission of digital information, each received bit is rated as *acceptable, suspect,* or *unacceptable,* depending on the quality of the received signal, with probabilities 0.9, 0.08, and 0.02, respectively. Assume that the ratings of each bit are independent.

In the first four bits transmitted, let

$$X \text{ denote the number of acceptable bits}$$

$$Y \text{ denote the number of suspect bits}$$

Then, the distribution of X is binomial with $n = 4$ and $p = 0.9$, and the distribution of Y is binomial with $n = 4$ and $p = 0.08$. However, because only four bits are being rated, the possible values of X and Y are restricted to the points shown in the graph in Fig. 5-1. Although the possible values of X are 0, 1, 2, 3, or 4, if $y = 3$, $x = 0$ or 1. By specifying the probability of each of the points in Fig. 5-1, we specify the joint probability distribution of X and Y. Similarly to an individual random variable, we define the range of the random variables (X, Y) to be the set of points (x, y) in two-dimensional space for which the probability that $X = x$ and $Y = y$ is positive.

5-1 Two Discrete Random Variables

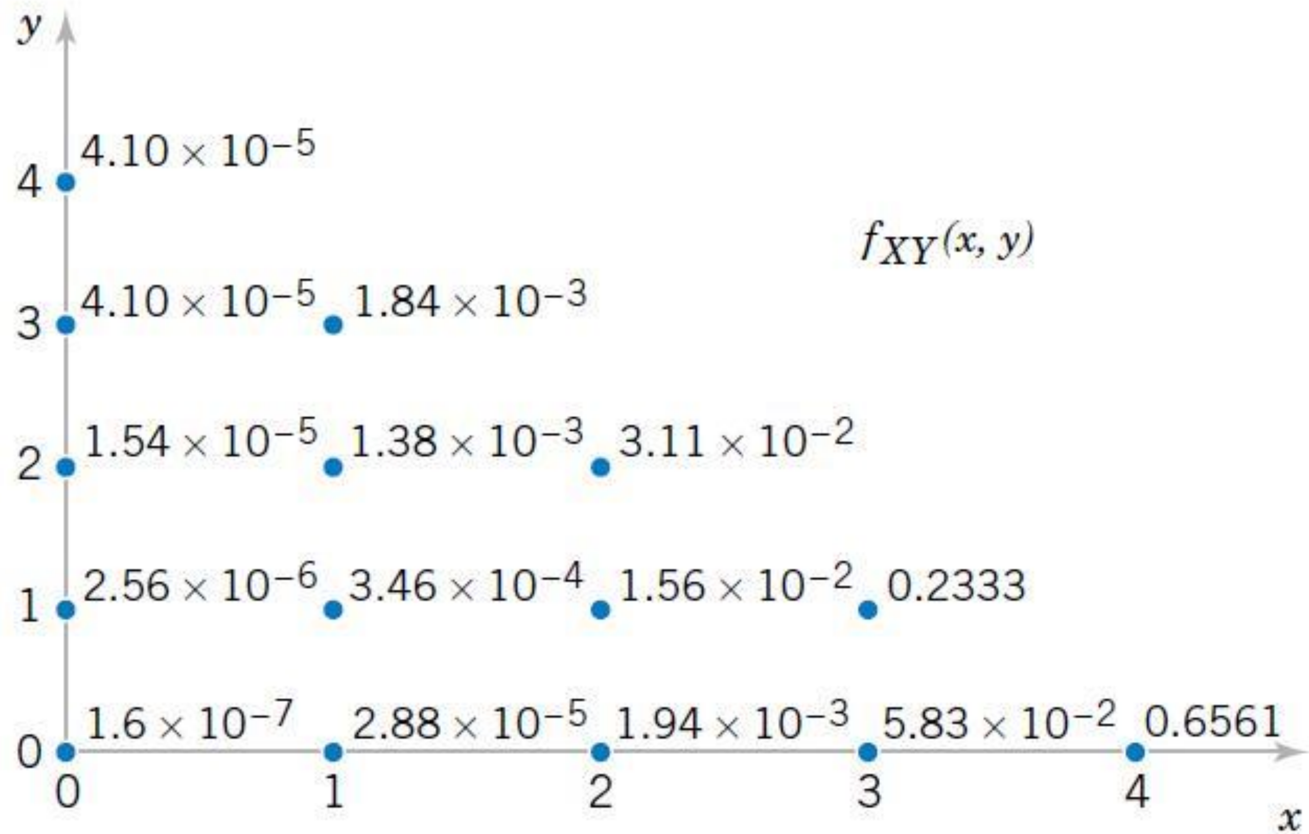


Figure 5-1 Joint probability distribution of X and Y in Example 5-1.

5-1 Two Discrete Random Variables

5-1.1 Joint Probability Distributions

The **joint probability mass function** of the discrete random variables X and Y, denoted as $f_{XY}(x, y)$, satisfies

(1) $f_{XY}(x, y) \geq 0$

(2) $\sum_x \sum_y f_{XY}(x, y) = 1$

(3) $f_{XY}(x, y) = P(X = x, Y = y)$ (5-1)

5-1 Two Discrete Random Variables

5-1.2 Marginal Probability Distributions

• The individual probability distribution of a random variable is referred to as its **marginal probability distribution**.

• In general, the marginal probability distribution of X can be determined from the joint probability distribution of X and other random variables. For example, to determine $P(X = x)$, we sum $P(X = x, Y = y)$ over all points in the range of (X, Y) for which $X = x$. Subscripts on the probability mass functions distinguish between the random variables.

5-1 Two Discrete Random Variables

Example 5-2

The joint probability distribution of X and Y in Fig. 5-1 can be used to find the marginal probability distribution of X. For example,

$$\begin{aligned} P(X = 3) &= P(X = 3, Y = 0) + P(X = 3, Y = 1) \\ &= 0.0583 + 0.2333 = 0.292 \end{aligned}$$

As expected, this probability matches the result obtained from the binomial probability distribution for X; that is, $P(X = 3) = \binom{4}{3}0.9^3 0.1^1 = 0.292$. The marginal probability distribution for X is found by summing the probabilities in each column, whereas the marginal probability distribution for Y is found by summing the probabilities in each row. The results are shown in Fig. 5-2.

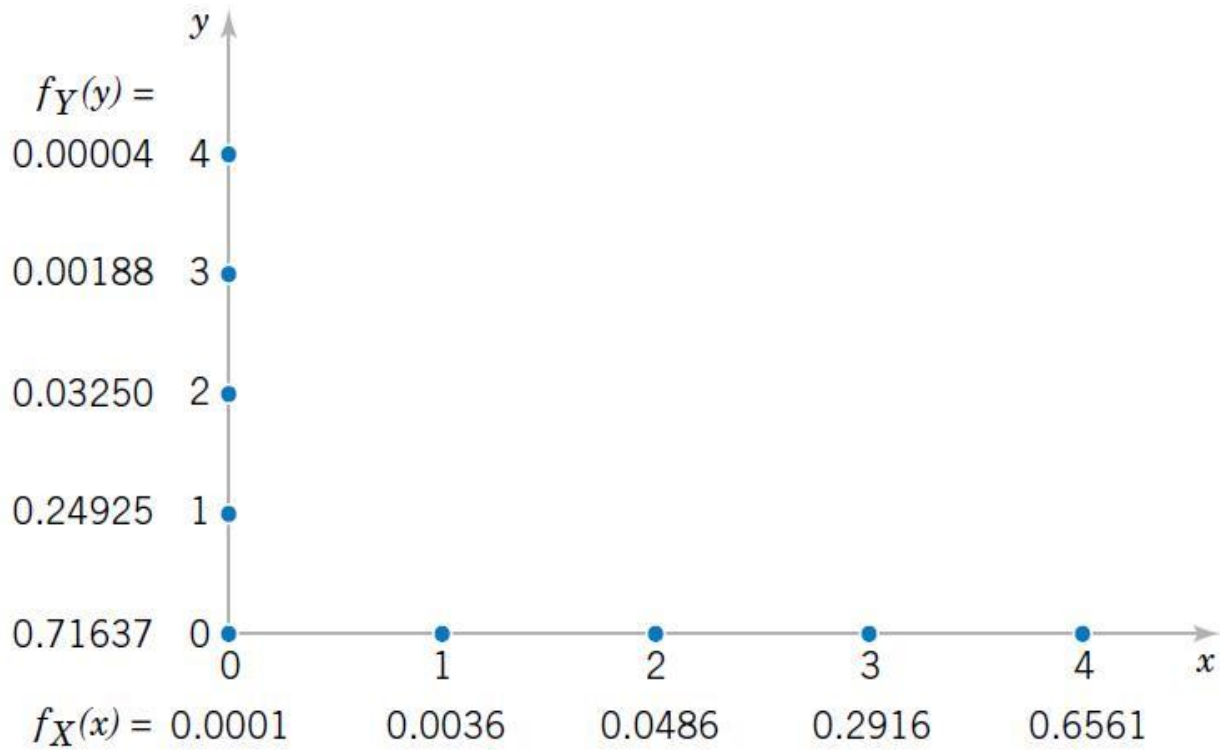


Figure 5-2 Marginal probability distributions of X and Y from Fig. 5-1.

5-1 Two Discrete Random Variables

Definition: Marginal Probability Mass Functions

If X and Y are discrete random variables with joint probability mass function $f_{XY}(x, y)$, then the **marginal probability mass functions** of X and Y are

$$f_X(x) = P(X = x) = \sum_y f_{XY}(x, y) \quad \text{and} \quad f_Y(y) = P(Y = y) = \sum_x f_{XY}(x, y) \tag{5-2}$$

where the first sum is over all points in the range of (X, Y) for which $X = x$ and the second sum is over all points in the range of (X, Y) for which $Y = y$

If the marginal probability distribution of X has the probability mass $f_X(x)$ function then

$$E(X) = \mu_X = \sum_x x f_X(x)$$

$$V(X) = \sigma_X^2 = \sum_x (x - \mu_x)^2 f_X(x)$$

In Example 5-1, $E(X)$ can be found as

$$
\begin{aligned}
E(X) = {} & 0[f_{XY}(0,0) + f_{XY}(0,1) + f_{XY}(0,2) + f_{XY}(0,3) + f_{XY}(0,4)] \\
& + 1[f_{XY}(1,0) + f_{XY}(1,1) + f_{XY}(1,2) + f_{XY}(1,3)] \\
& + 2[f_{XY}(2,0) + f_{XY}(2,1) + f_{XY}(2,2)] \\
& + 3[f_{XY}(3,0) + f_{XY}(3,1)] \\
& + 4[f_{XY}(4,0)] \\
= {} & 0[0.0001] + 1[0.0036] + 2[0.0486] + 3[0.02916] + 4[0.6561] = 3.6
\end{aligned}
$$

Alternatively, because the marginal probability distribution of X is binomial,

$$E(X) = np = 4(0.9) = 3.6$$

The calculation using the joint probability distribution can be used to determine $E(X)$ even in cases in which the marginal probability distribution of X is not known. As practice, you can use the joint probability distribution to verify that $E(Y) = 0.32$ in Example 5-1.

Also,

$$V(X) = np(1 - p) = 4(0.9)(1 - 0.9) = 0.36$$

Verify that the same result can be obtained from the joint probability distribution of X and Y.

5-1 Two Discrete Random Variables

5-1.3 Conditional Probability Distributions

Given discrete random variables X and Y with joint probability mass function $f_{XY}(x, y)$ the **conditional probability mass function** of Y given $X = x$ is

$$f_{Y|x}(y) = f_{XY}(x, y)/f_X(x) \qquad \text{for } f_X(x) > 0 \qquad (5\text{-}3)$$

Given discrete random variables X and Y with joint probability mass function $f_{XY}(x, y)$ the **conditional probability mass function** of Y given $X = x$ is

$$f_{Y|x}(y) = f_{XY}(x, y)/f_X(x) \qquad \text{for } f_X(x) > 0 \tag{5-3}$$

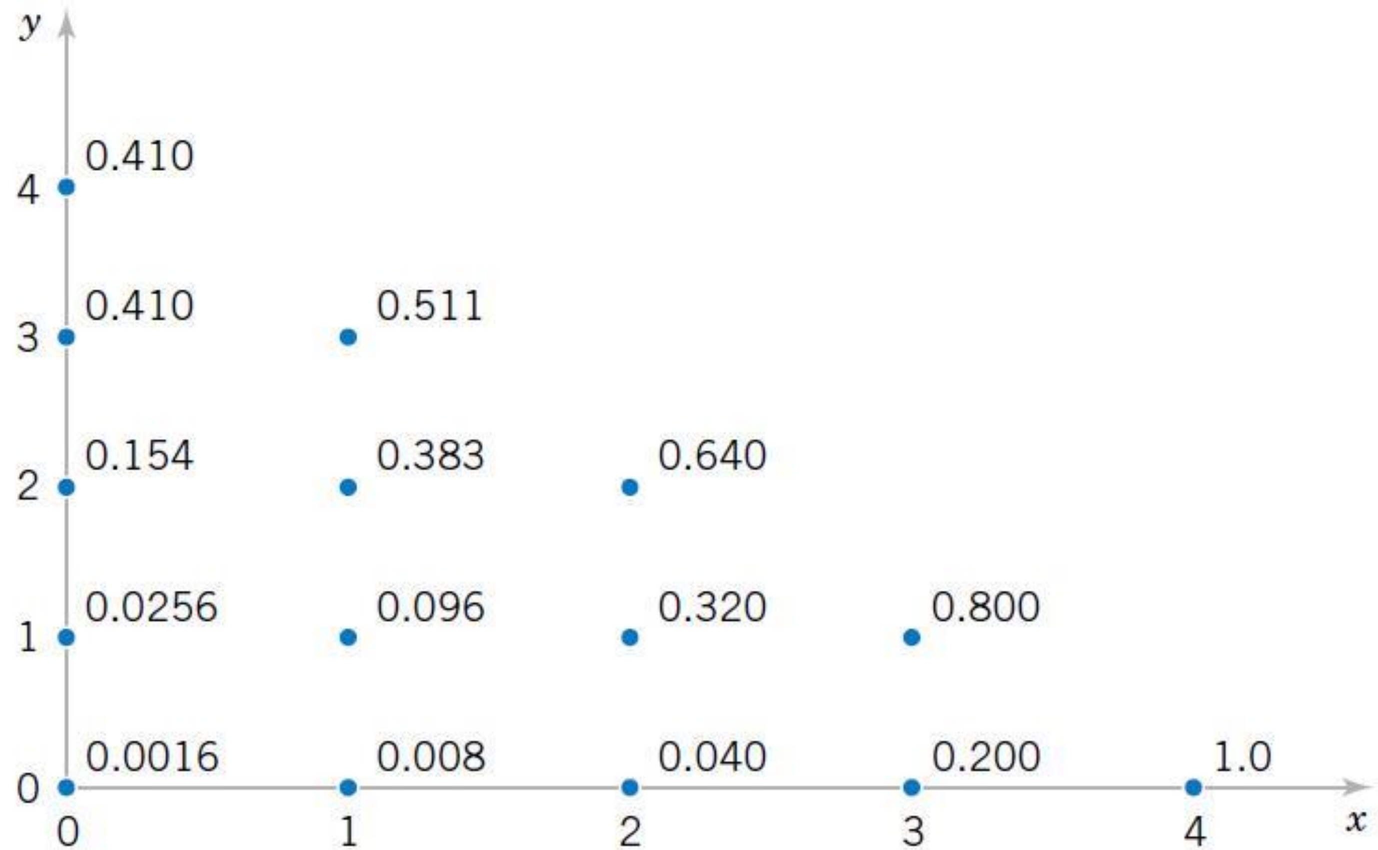


Figure 5-3 Conditional probability distributions of Y given $X = x$, $f_{Y|x}(y)$ in Example 5-6.

5-1 Two Discrete Random Variables

5-1.3 Conditional Probability Distributions

Because a conditional probability mass function $f_{Y|x}(y)$ is a probability mass function for all y in R_x, the following properties are satisfied:

(1) $f_{Y|x}(y) \geq 0$

(2) $\sum_{y} f_{Y|x}(y) = 1$

(3) $P(Y = y | X = x) = f_{Y|x}(y)$ (5-4)

5-1 Two Discrete Random Variables

Definition: Conditional Mean and Variance

The **conditional mean** of Y given $X = x$, denoted as $E(Y|x)$ or $\mu_{Y|x}$, is

$$E(Y|x) = \sum_y y f_{Y|x}(y) \tag{5-5}$$

and the **conditional variance** of Y given $X = x$, denoted as $V(Y|x)$ or $\sigma^2_{Y|x}$, is

$$V(Y|x) = \sum_y (y - \mu_{Y|x})^2 f_{Y|x}(y) = \sum_y y^2 f_{Y|x}(y) - \mu^2_{Y|x}$$

5-1 Two Discrete Random Variables

For the random variables in Example 5-1, the conditional mean of Y given $X = 2$ is obtained from the conditional distribution in Fig. 5-3:

$$E(Y|2) = \mu_{Y|2} = 0(0.040) + 1(0.320) + 2(0.640) = 1.6$$

The conditional mean is interpreted as the expected number of acceptable bits given that two of the four bits transmitted are suspect. The conditional variance of Y given $X = 2$ is

$$V(Y|2) = (0 - \mu_{Y|2})^2(0.040) + (1 - \mu_{Y|2})^2(0.320) + (2 - \mu_{Y|2})^2(0.640) = 0.32$$

5-1 Two Discrete Random Variables

5-1.4 Independence

Example 5-6

In a plastic molding operation, each part is classified as to whether it conforms to color and length specifications. Define the random variable X and Y as

$$X = \begin{cases} 1 & \text{if the part conforms to color specifications} \\ 0 & \text{otherwise} \end{cases}$$

$$Y = \begin{cases} 1 & \text{if the part conforms to length specifications} \\ 0 & \text{otherwise} \end{cases}$$

Assume the joint probability distribution of X and Y is defined by $f_{XY}(x, y)$ in Fig. 5-4(a). The marginal probability distributions of X and Y are also shown in Fig. 5-4(a). Note that $f_{XY}(x, y) = f_X(x)\, f_Y(y)$. The conditional probability mass function $f_{Y|x}(y)$ is shown in Fig. 5-4(b). Notice that for any x, $f_{Y|x}(y) = f_Y(y)$. That is, knowledge of whether or not the part meets color specifications does not change the probability that it meets length specifications.

Example 5-8

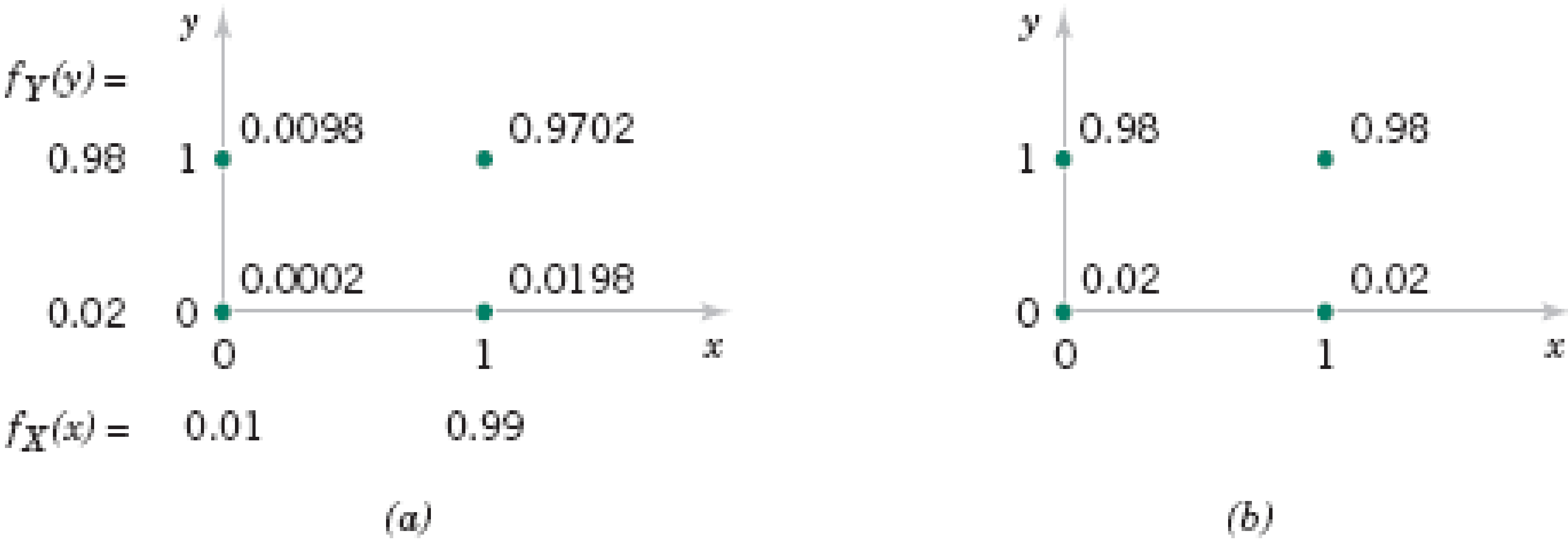


Figure 5-4 (a)Joint and marginal probability distributions of X and Y in Example 5-8. (b) Conditional probability distribution of Y given $X = x$ in Example 5-8.

5-1 Two Discrete Random Variables

5-1.4 Independence

For discrete random variables X and Y, if any one of the following properties is true, the others are also true, and X and Y are **independent.**

(1) $f_{XY}(x, y) = f_X(x) f_Y(y)$ for all x and y

(2) $f_{Y|x}(y) = f_Y(y)$ for all x and y with $f_X(x) > 0$

(3) $f_{X|y}(x) = f_X(x)$ for all x and y with $f_Y(y) > 0$

(4) $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ for any sets A and B in the range of X and Y, respectively. (5-6)

5-1 Two Discrete Random Variables

5-1.5 Multiple Discrete Random Variables

Definition: Joint Probability Mass Function

The **joint probability mass function** of $X_1, X_2, \ldots, X_p$ is

$$f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p) = P(X_1 = x_1, X_2 = x_2, \ldots, X_p = x_p) \qquad (5\text{-}7)$$

for all points $(x_1, x_2, \ldots, x_p)$ in the range of $X_1, X_2, \ldots, X_p$.

A marginal probability distribution is a simple extension of the result for two random variables.

5-1 Two Discrete Random Variables

5-1.5 Multiple Discrete Random Variables

Definition: Marginal Probability Mass Function

If $X_1, X_2, X_3, \ldots, X_p$ are discrete random variables with joint probability mass function $f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)$, the **marginal probability mass function** of any X_i is

$$f_{X_i}(x_i) = P(X_i = x_i) = \sum f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p) \tag{5-8}$$

where the sum is over the points in the range of $(X_1, X_2, \ldots, X_p)$ for which $X_i = x_i$.

5-1 Two Discrete Random Variables

Example 5-8

Points that have positive probability in the joint probability distribution of three random variables X_1, X_2, X_3 are shown in Fig. 5-5. The range is the nonnegative integers with $x_1 + x_2 + x_3 = 3$. The marginal probability distribution of X_2 is found as follows.

$$P(X_2 = 0) = f_{X_1X_2X_3}(3, 0, 0) + f_{X_1X_2X_3}(0, 0, 3) + f_{X_1X_2X_3}(1, 0, 2) + f_{X_1X_2X_3}(2, 0, 1)$$
$$P(X_2 = 1) = f_{X_1X_2X_3}(2, 1, 0) + f_{X_1X_2X_3}(0, 1, 2) + f_{X_1X_2X_3}(1, 1, 1)$$
$$P(X_2 = 2) = f_{X_1X_2X_3}(1, 2, 0) + f_{X_1X_2X_3}(0, 2, 1)$$
$$P(X_2 = 3) = f_{X_1X_2X_3}(0, 3, 0)$$

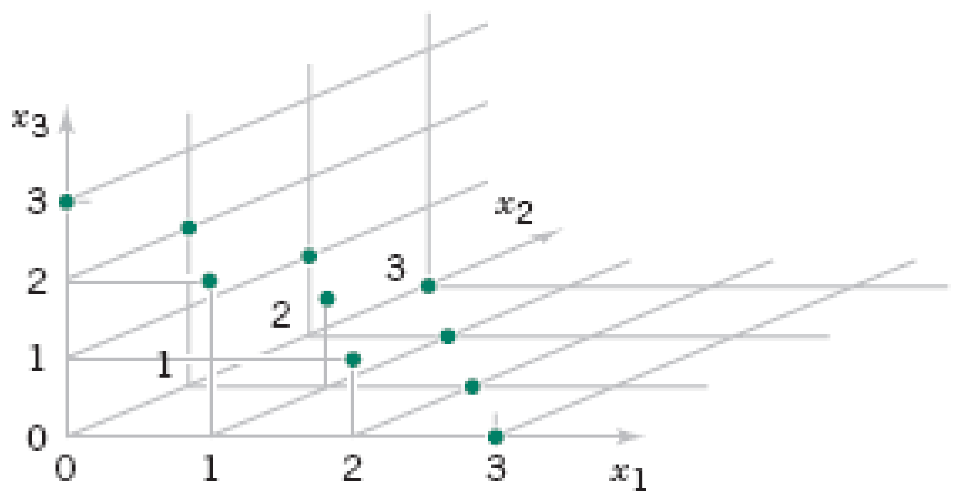


Figure 5-5 Joint probability distribution of X_1, X_2, and X_3.

5-1 Two Discrete Random Variables

5-1.5 Multiple Discrete Random Variables

Mean and Variance from Joint Probability

$$E(X_i) = \sum x_i f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)$$

and

$$V(X_i) = \sum (x_i - \mu_{X_i})^2 f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p) \qquad (5\text{-}9)$$

where the sum is over all points in the range of $X_1, X_2, \ldots, X_p$.

5-1 Two Discrete Random Variables

5-1.5 Multiple Discrete Random Variables

Distribution of a Subset of Random Variables

If $X_1, X_2, X_3, \ldots, X_p$ are discrete random variables with joint probability mass function $f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)$, the **joint probability mass function** of $X_1, X_2, \ldots, X_k$, $k < p$, is

$$f_{X_1X_2\ldots X_k}(x_1, x_2, \ldots, x_k) = P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)$$
$$= \sum_{R_{x_1x_2\ldots x_k}} P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) \qquad (5\text{-}10)$$

where the sum is over all points in the range of $X_1, X_2, \ldots, X_p$ for which $X_1 = x_1$, $X_2 = x_2, \ldots, X_k = x_k$.

5-1 Two Discrete Random Variables

5-1.5 Multiple Discrete Random Variables

Conditional Probability Distributions

Discrete variables $X_1, X_2, \ldots, X_p$ are **independent** if and only if

$$f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p) = f_{X_1}(x_1)\, f_{X_2}(x_2) \ldots f_{X_p}(x_p) \qquad (5\text{-}11)$$

for *all* $x_1, x_2, \ldots, x_p$.

5-1 Two Discrete Random Variables

5-1.6 Multinomial Probability Distribution

Suppose a random experiment consists of a series of n trials. Assume that

(1) The result of each trial is classified into one of k classes.

(2) The probability of a trial generating a result in class 1, class 2, … , class k is constant over the trials and equal to $p_1, p_2, \ldots, p_k$, respectively.

(3) The trials are independent.

The random variables $X_1, X_2, \ldots, X_k$ that denote the number of trials that result in class 1, class 2, … , class k, respectively, have a **multinomial distribution** and the joint probability mass function is

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1!x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k} \qquad (5\text{-}12)$$

for $x_1 + x_2 + \cdots + x_k = n$ and $p_1 + p_2 + \cdots + p_k = 1$.

5-1 Two Discrete Random Variables

5-1.6 Multinomial Probability Distribution

Each trial in a multinomial random experiment can be regarded as either generating or not generating a result in class i, for each $i = 1, 2, \ldots, k$. Because the random variable X_i is the number of trials that result in class i, X_i has a binomial distribution.

If $X_1, X_2, \ldots, X_k$ have a multinomial distribution, the marginal probability distribution of X_i is binomial with

$$E(X_i) = np_i \quad \text{and} \quad V(X_i) = np_i(1 - p_i) \tag{5-13}$$

Example

Show that the following function satisfies the properties of a joint probability mass function.

x	y	$f_{XY}(x, y)$
1	1	1/4
1.5	2	1/8
1.5	3	1/4
2.5	4	1/4
3	5	1/8

Example. cont

- It is clear that $f_{XY}(x,y) \geq 0$ for all ***R*,** which represent the set of all possible values of (X,Y).

- Then, $\sum_{R} f_{XY}(x,y) = \frac{1}{4} + \frac{1}{8} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} = 1$

Now, determine the following:

(a) $P(X < 2.5, Y < 3)$

By looking to the table:

$$P(X < 2.5, Y < 3) = f_{XY}(1.5,2) + f_{XY}(1,1)$$

$$= \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$$

Example. cont

(b) $P(X < 2.5)$

By looking to the table:

$$P(X < 2.5) = f_{XY}(1.5,2) + f_{XY}(1.5,3) + f_{XY}(1,1)$$

$$= \frac{1}{8} + \frac{1}{4} + \frac{1}{4} = \frac{5}{8}$$

(c) $P(Y < 3)$

By looking to the table:

$$P(Y < 3) = f_{XY}(1.5,2) + f_{XY}(1,1)$$

$$= \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$$

Example. cont

(d) $P(X > 1.8, Y > 4.7)$

By looking to the table:

$$P(X > 1.8, Y > 4.7) = f_{XY}(3,5) = \frac{1}{8}$$

(e) Find $E(X), E(Y), V(X)$ *and* $V(Y)$

remember:

$$E(X) = \mu_X = \sum_x x f_X(x)$$

$$V(X) = \sigma_X^2 = \sum_x (x - \mu_x)^2 f_X(x)$$

Example. cont

$$E(X) = \mu_X = \sum_x x f_X(x)$$

$$= \left(1 \times \frac{1}{4}\right) + \left(1.5 \times \left[\frac{1}{8} + \frac{1}{4}\right]\right) + \left(2.5 \times \frac{1}{4}\right) + \left(3 \times \frac{1}{8}\right) = 1.8125$$

$$E(Y) = \mu_Y = \sum_y y f_Y(y)$$

$$= \left(1 \times \frac{1}{4}\right) + \left(2 \times \frac{1}{8}\right) + \left(3 \times \frac{1}{4}\right) + \left(4 \times \frac{1}{4}\right) + \left(5 \times \frac{1}{8}\right) = 2.875$$

Example. cont

$$V(X) = \sigma_X^2 = \sum_x (x - \mu_x)^2 f_X(x) = E(X^2) - [E(X)]^2$$

$$= \left[\left(1^2 \times \frac{1}{4}\right) + \left(1.5^2 \times \left(\frac{1}{8} + \frac{1}{4}\right)\right) + \left(2.5^2 \times \frac{1}{4}\right) + \left(3^2 \times \frac{1}{8}\right)\right] - 1.8125^2 = 0.4961$$

$$V(Y) = \sigma_Y^2 = \sum_Y (y - \mu_Y)^2 f_Y(y) = E(Y^2) - [E(Y)]^2$$

$$= \left[\left(1^2 \times \frac{1}{4}\right) + \left(2^2 \times \frac{1}{8}\right) + \left(3^2 \times \frac{1}{4}\right) + \left(4^2 \times \frac{1}{4}\right) + \left(5^2 \times \frac{1}{8}\right)\right] - 2.875^2 = 1.8594$$

Example. cont

Marginal probability distribution of the random variable X.

x	$f_X(x)$
1	1/4
1.5	1/8+1/4= 3/8
2.5	1/4
3	1/8

Example. cont

(f) Conditional probability distribution of Y given that X=1.5.

$$f_{Y|1.5}(y) = \frac{f_{XY}(1.5, y)}{f_X(1.5)}, where \quad f_X(1.5) = \frac{3}{8}, then$$

y	$f_{Y\|1.5}(y)$
2	(1/8) / (3/8) = (1/3)
3	(1/4) / (3/8) = (2/3)

(h) Conditional probability distribution of X given that Y=2.

$$f_{X|2}(x) = \frac{f_{XY}(x,2)}{f_Y(2)}, where \quad f_Y(2) = \frac{1}{8}, then$$

x	$f_{X\|2}(x)$
1.5	(1/8) / (1/8) = 1

Example. cont

(i) Find $E(Y \mid X = 1.5)$

$$E(Y \mid X = 1.5) = \left(2 \times \frac{1}{3}\right) + \left(3 \times \frac{2}{3}\right)$$

(j) Are *X* and ***Y*** independent?

$$Since \quad f_{Y|1.5}(y) \neq f_Y(y)$$

Then, ***X*** and ***Y*** are not independent.

Example

Determine the value of c that makes the function

$$f_{XY}(x, y) = c(x + y)$$

A joint probability mass function over the nine points with

$$x = 1,2,3 \quad and \quad y = 1,2,3$$

$$f_{XY}(x,y)=c(x+y)$$

Solution: From the joint probability mass function properties.

$$\sum_{x,y} f_{XY}(x,y)=1$$

$$then$$

$$\sum_{x,y} f_{XY}(x,y)= f_{XY}(1,1)+ f_{XY}(1,2)+ f_{XY}(1,3)$$

$$+ f_{XY}(2,1)+ f_{XY}(2,2)+ f_{XY}(2,3)$$

$$+ f_{XY}(3,1)+ f_{XY}(3,2)+ f_{XY}(3,3)$$

$$=c(2+3+4+3+4+5+4+5+6)=1$$

$$=36c=1,$$

$$then,\ c=\frac{1}{36}$$

Example cont.

Now, determine the followings:

$(a) \quad P(X=1, Y<4)$

solution,

$$P(X=1, Y<4) = f_{XY}(1,1) + f_{XY}(1,2) + f_{XY}(1,3)$$

$$= \frac{1}{36}(2+3+4) = \frac{1}{4}$$

$(b) \quad P(X=1)$

solution,

the same result as in part (a),

becasue all possible values of Y (1,2,3) *are included in part*(a)

when $X=1$

Example cont.

(c) determine, the marginal probability distribution of the random variable ***X***.

$$f_X(x) = f_{XY}(x,1) + f_{XY}(x,2) + f_{XY}(x,3)$$

x	**marginal probability distribution**
1	1/4
2	1/3
3	5/12

Go to covariance and correlation…..

Correlation determines the degree of similarity between two signals. If the signals are identical, then the correlation coefficient is 1; if they are totally different, the correlation coefficient is 0, and if they are identical except that the phase is shifted by (i.e. mirrored), then the correlation coefficient is -1.

When two independent signals are compared, the procedure is known as *cross-correlation*, and when the same signal is compared to phase shifted copies of itself, the procedure is known as*autocorrelation*.

5-2 Two Continuous Random Variables

5-2.1 Joint Probability Distribution

Definition

A **joint probability density function** for the continuous random variables X and Y, denoted as $f_{XY}(x, y)$, satisfies the following properties:

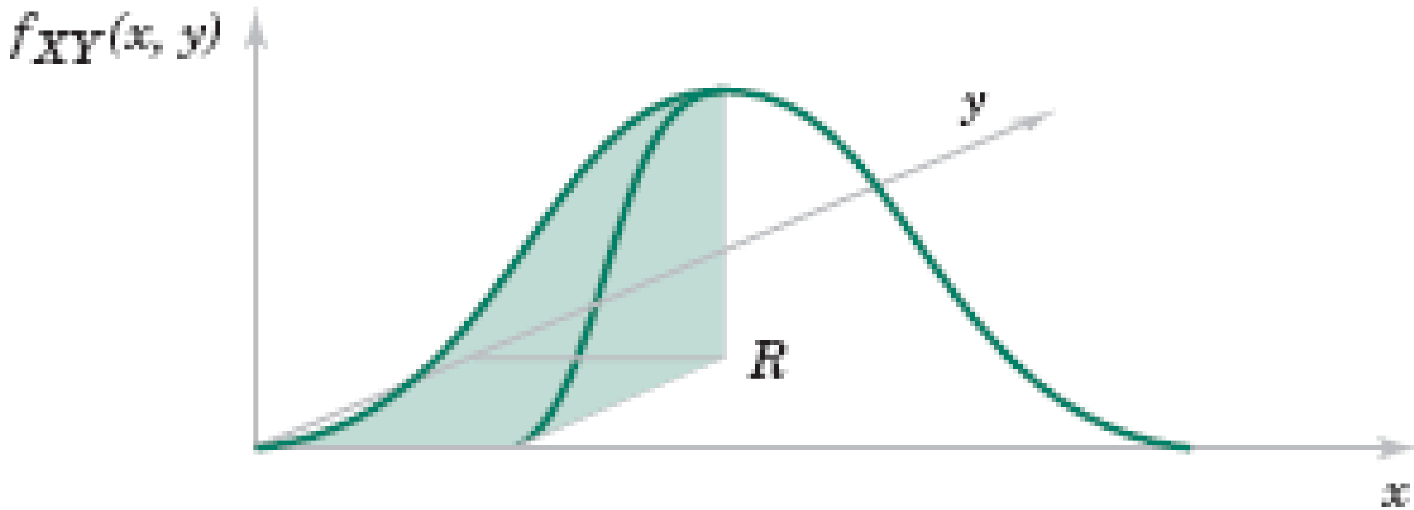
(1) $f_{XY}(x, y) \geq 0$ for all x, y

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

(3) For any region R of two-dimensional space

$$P((X, Y) \in R) = \iint_R f_{XY}(x, y)\, dx\, dy \qquad (5\text{-}14)$$

5-2 Two Continuous Random Variables



Probability that (X, Y) is in the region R is determined by the volume of $f_{XY}(x, y)$ over the region R.

Figure 5-6 Joint probability density function for random variables *X* and *Y*.

5-2 Two Continuous Random Variables

Example 5-12

Let the random variable X denote the time until a computer server connects to your machine (in milliseconds), and let Y denote the time until the server authorizes you as a valid user (in milliseconds). Each of these random variables measures the wait from a common starting time and $X < Y$. Assume that the joint probability density function for X and Y is

$$f_{XY}(x, y) = 6 \times 10^{-6} \exp(-0.001x - 0.002y) \quad \text{for } x < y$$

Reasonable assumptions can be used to develop such a distribution, but for now, our focus is only on the joint probability density function.

5-2 Two Continuous Random Variables

Example 5-12

The region with nonzero probability is shaded in Fig. 5-8. The property that this joint probability density function integrates to 1 can be verified by the integral of $f_{XY}(x, y)$ over this region as follows:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dy\,dx = \int_{0}^{\infty}\left(\int_{x}^{\infty} 6\times 10^{-6} e^{-0.001x-0.002y}\,dy\right)dx$$

$$= 6\times 10^{-6}\int_{0}^{\infty}\left(\int_{x}^{\infty} e^{-0.002y}\,dy\right)e^{-0.001x}\,dx$$

$$= 6\times 10^{-6}\int_{0}^{\infty}\left(\frac{e^{-0.002x}}{0.002}\right)e^{-0.001x}\,dx$$

$$= 0.003\left(\int_{0}^{\infty} e^{-0.003x}\,dx\right) = 0.003\left(\frac{1}{0.003}\right) = 1$$

5-2 Two Continuous Random Variables

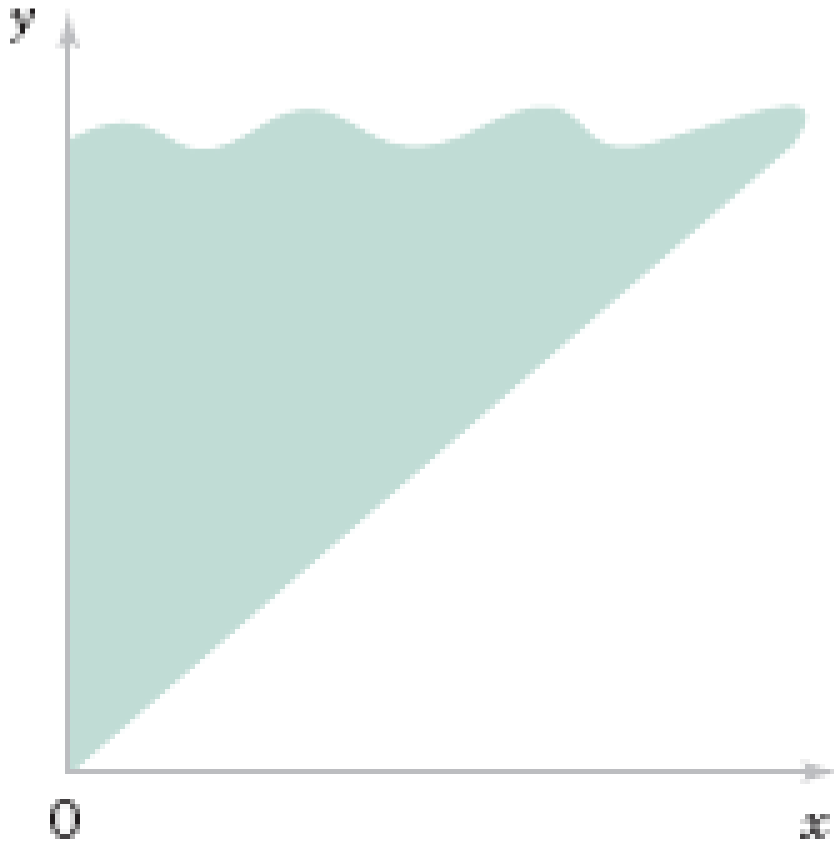


Figure 5-8 The joint probability density function of *X* and *Y* is nonzero over the shaded region.

5-2 Two Continuous Random Variables

Example 5-12

The probability that $X < 1000$ and $Y < 2000$ is determined as the integral over the darkly shaded region in Fig. 5-9.

$$
\begin{aligned}
P(X \le 1000, Y \le 2000) &= \int_0^{1000}\int_x^{2000} f_{XY}(x, y)\, dy\, dx \\
&= 6 \times 10^{-6} \int_0^{1000} \left(\int_x^{2000} e^{-0.002y}\, dy \right) e^{-0.001x}\, dx \\
&= 6 \times 10^{-6} \int_0^{1000} \left(\frac{e^{-0.002x} - e^{-4}}{0.002} \right) e^{-0.001x}\, dx \\
&= 0.003 \int_0^{1000} e^{-0.003x} - e^{-4} e^{-0.001x}\, dx \\
&= 0.003 \left[\left(\frac{1 - e^{-3}}{0.003} \right) - e^{-4} \left(\frac{1 - e^{-1}}{0.001} \right) \right] \\
&= 0.003\,(316.738 - 11.578) = 0.915
\end{aligned}
$$

5-2 Two Continuous Random Variables

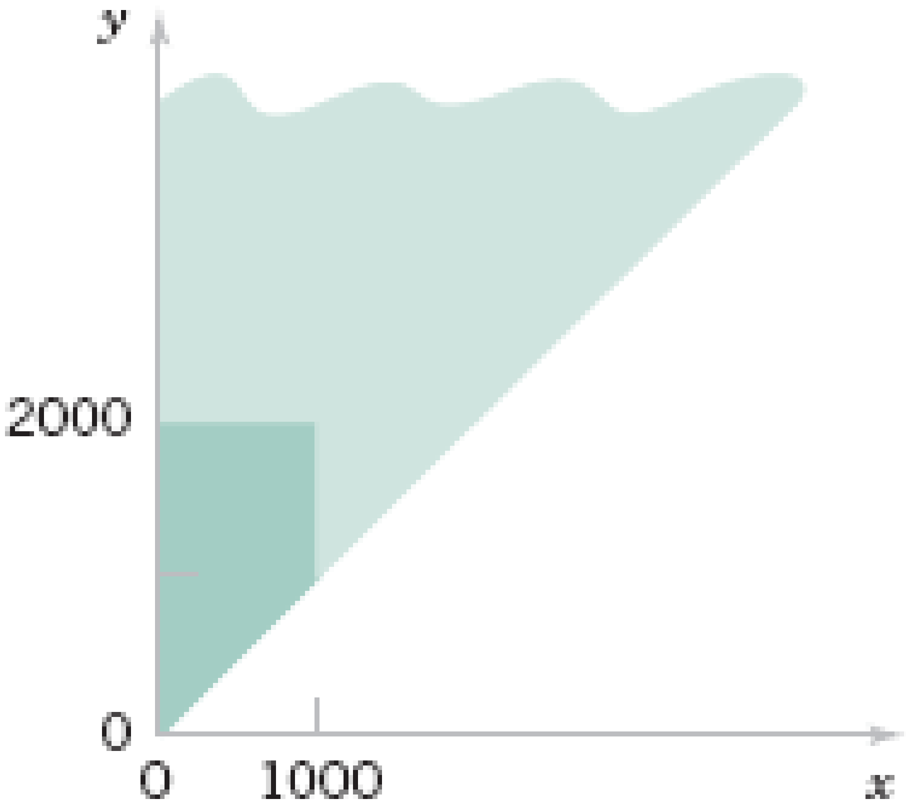


Figure 5-9 Region of integration for the probability that $X < 1000$ and $Y < 2000$ is darkly shaded.

5-2 Two Continuous Random Variables

5-2.2 Marginal Probability Distributions

Definition

If the joint probability density function of continuous random variables X and Y is $f_{XY}(x, y)$, the **marginal probability density functions** of X and Y are

$$f_X(x) = \int_y f_{XY}(x, y)\, dy \quad \text{and} \quad f_Y(y) = \int_x f_{XY}(x, y)\, dx \qquad (5\text{-}15)$$

where the first integral is over all points in the range of (X, Y) for which $X = x$ and the second integral is over all points in the range of (X, Y) for which $Y = y$

5-2 Two Continuous Random Variables

Example 5-13

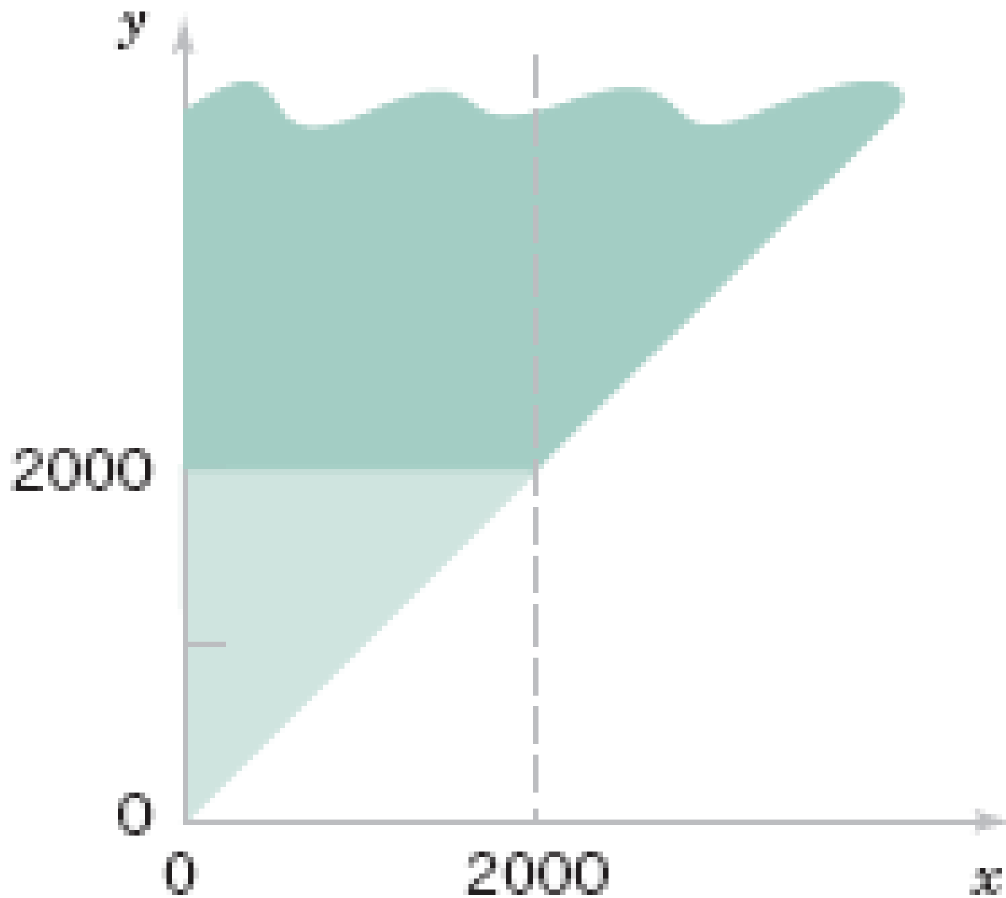
For the random variables that denote times in Example 5-12, calculate the probability that Y exceeds 2000 milliseconds.

This probability is determined as the integral of $f_{XY}(x, y)$ over the darkly shaded region in Fig. 5-10. The region is partitioned into two parts and different limits of integration are determined for each part.

$$P(Y > 2000) = \int_{0}^{2000} \left(\int_{2000}^{\infty} 6 \times 10^{-6} e^{-0.001x - 0.002y}\, dy \right) dx + \int_{2000}^{\infty} \left(\int_{x}^{\infty} 6 \times 10^{-6} e^{-0.001x - 0.002y}\, dy \right) dx$$

5-2 Two Continuous Random Variables

Figure 5-10 Region of integration for the probability that *Y* < 2000 is darkly shaded and it is partitioned into two regions with *x* < 2000 and and *x* > 2000.



5-2 Two Continuous Random Variables

Example 5-13

The first integral is

$$6 \times 10^{-6} \int_0^{2000} \left(\frac{e^{-0.002y}}{-0.002} \bigg|_{2000}^{\infty} \right) e^{-0.001x}\, dx = \frac{6 \times 10^{-6}}{0.002} e^{-4} \int_0^{2000} e^{-0.001x}\, dx$$

$$= \frac{6 \times 10^{-6}}{0.002} e^{-4} \left(\frac{1 - e^{-2}}{0.001} \right) = 0.0475$$

The second integral is

$$6 \times 10^{-6} \int_{2000}^{\infty} \left(\frac{e^{-0.002y}}{-0.002} \bigg|_{x}^{\infty} \right) e^{-0.001x}\, dx = \frac{6 \times 10^{-6}}{0.002} \int_{2000}^{\infty} e^{-0.003x}\, dx$$

$$= \frac{6 \times 10^{-6}}{0.002} \left(\frac{e^{-6}}{0.003} \right) = 0.0025$$

5-2 Two Continuous Random Variables

Example 5-13

Therefore,

$$P(Y > 2000) = 0.0475 + 0.0025 = 0.05.$$

Alternatively, the probability can be calculated from the marginal probability distribution of Y as follows. For $y > 0$

$$f_Y(y) = \int_0^y 6 \times 10^{-6} e^{-0.001x-0.002y}\, dx = 6 \times 10^{-6} e^{-0.002y} \int_0^y e^{-0.001x}\, dx$$

$$= 6 \times 10^{-6} e^{-0.002y} \left(\frac{e^{-0.001x}}{-0.001} \Bigg|_0^y \right) = 6 \times 10^{-6} e^{-0.002y} \left(\frac{1 - e^{-0.001y}}{0.001} \right)$$

$$= 6 \times 10^{-3} e^{-0.002y} (1 - e^{-0.001y}) \quad \text{for } y > 0$$

5-2 Two Continuous Random Variables

Example 5-13

We have obtained the marginal probability density function of Y. Now,

$$P(Y > 2000) = 6 \times 10^{-3} \int_{2000}^{\infty} e^{-0.002y}(1 - e^{-0.001y})\, dy$$

$$= 6 \times 10^{-3} \left[\left(\frac{e^{-0.002y}}{-0.002} \Bigg|_{2000}^{\infty} \right) - \left(\frac{e^{-0.003y}}{-0.003} \Bigg|_{2000}^{\infty} \right) \right]$$

$$= 6 \times 10^{-3} \left[\frac{e^{-4}}{0.002} - \frac{e^{-6}}{0.003} \right] = 0.05$$

5-2 Two Continuous Random Variables

5-2.3 Conditional Probability Distributions

Definition

Given continuous random variables X and Y with joint probability density function $f_{XY}(x, y)$, the **conditional probability density function** of Y given $X = x$ is

$$f_{Y|x}(y) = \frac{f_{XY}(x, y)}{f_X(x)} \quad \text{for} \quad f_X(x) > 0 \tag{5-16}$$

5-2 Two Continuous Random Variables

5-2.3 Conditional Probability Distributions

Because the conditional probability density function $f_{Y|x}(y)$ is a probability density function for all y in R_x, the following properties are satisfied:

(1) $f_{Y|x}(y) \geq 0$

(2) $\int_{R_x} f_{Y|x}(y)\,dy = 1$

(3) $P(Y \in B | X = x) = \int_B f_{Y|x}(y)\,dy$ for any set B in the range of Y

(5-17)

5-2 Two Continuous Random Variables

Example 5-14

For the random variables that denote times in Example 5-12, determine the conditional probability density function for Y given that $X = x$.

First the marginal density function of x is determined. For $x > 0$

$$f_X(x) = \int_x^\infty 6 \times 10^{-6} e^{-0.001x - 0.002y}\, dy = 6 \times 10^{-6} e^{-0.001x} \left(\frac{e^{-0.002y}}{-0.002} \Bigg|_x^\infty \right)$$

$$= 6 \times 10^{-6} e^{-0.001x} \left(\frac{e^{-0.002x}}{0.002} \right) = 0.003 e^{-0.003x} \quad \text{for} \quad x > 0$$

This is an exponential distribution with $\lambda = 0.003$. Now, for $0 < x$ and $x < y$ the conditional probability density function is

$$f_{Y|x}(y) = f_{XY}(x, y)/f_x(x) = \frac{6 \times 10^{-6} e^{-0.001x - 0.002y}}{0.003 e^{-0.003x}}$$

$$= 0.002 e^{0.002x - 0.002y} \quad \text{for } 0 < x \quad \text{and} \quad x < y$$

5-2 Two Continuous Random Variables

Example 5-14

The conditional probability density function of Y, given that $x = 1500$, is nonzero on the solid line in Fig. 5-11.

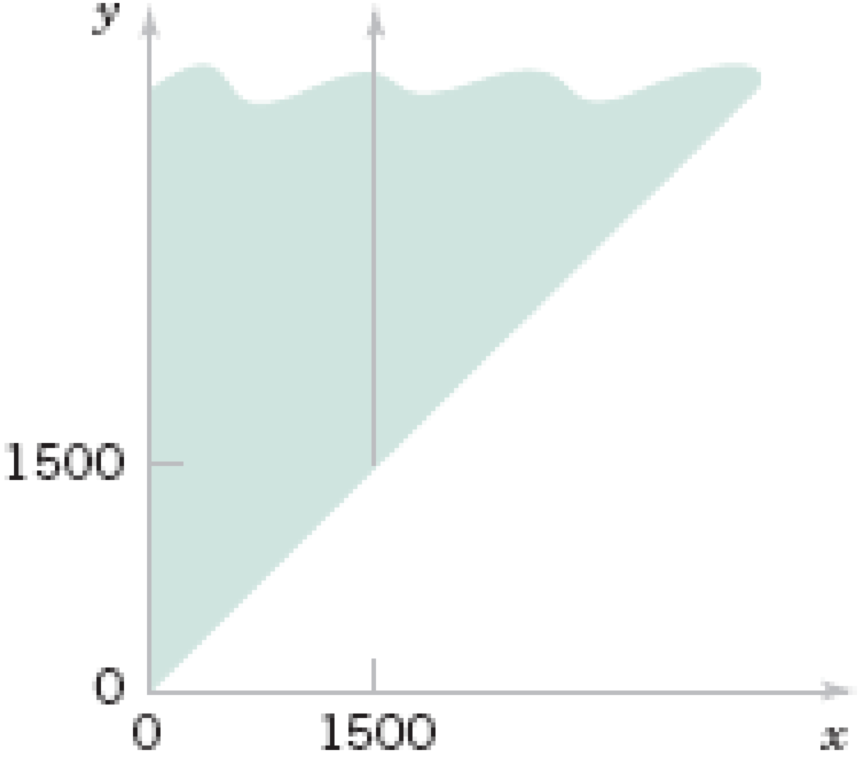


Figure 5-11 The conditional probability density function for Y, given that x = 1500, is nonzero over the solid line.

5-2 Two Continuous Random Variables

Definition: Conditional Mean and Variance

The **conditional mean** of Y given $X = x$, denoted as $E(Y|x)$ or $\mu_{Y|x}$, is

$$E(Y|x) = \int y f_{Y|x}(y)\, dy$$

and the **conditional variance** of Y given $X = x$, denoted as $V(Y|x)$ or $\sigma^2_{Y|x}$, is

$$V(Y|x) = \int_{-\infty} (y - \mu_{Y|x})^2 f_{Y|x}(y)\, dy = \int y^2 f_{Y|x}(y)\, dy - \mu^2_{Y|x} \qquad (5\text{-}18)$$

5-2 Two Continuous Random Variables

5-2.4 Independence

Definition

For continuous random variables X and Y, if any one of the following properties is true, the others are also true, and X and Y are said to be **independent.**

(1) $f_{XY}(x, y) = f_X(x) f_Y(y)$ for all x and y

(2) $f_{Y|x}(y) = f_Y(y)$ for all x and y with $f_X(x) > 0$

(3) $f_{X|y}(x) = f_X(x)$ for all x and y with $f_Y(y) > 0$

(4) $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ for any sets A and B in the range of X and Y, respectively. (5-19)

5-2 Two Continuous Random Variables

Example 5-16

For the joint distribution of times in Example 5-12, the

- Marginal distribution of Y was determined in Example 5-13.
- Conditional distribution of Y given $X = x$ was determined in Example 5-14.

Because the marginal and conditional probability densities are not the same for all values of x, property (2) of Equation 5-18 implies that the random variables are not independent. The fact that these variables are not independent can be determined quickly by noticing that the range of (X, Y), shown in Fig. 5-8, is not rectangular. Consequently, knowledge of X changes the interval of values for Y that receives nonzero probability.

5-2 Two Continuous Random Variables

Example 5-18

Let the random variables X and Y denote the lengths of two dimensions of a machined part, respectively. Assume that X and Y are independent random variables, and further assume that the distribution of X is normal with mean 10.5 millimeters and variance 0.0025 (millimeter)^2 and that the distribution of Y is normal with mean 3.2 millimeters and variance 0.0036 (millimeter)^2. Determine the probability that $10.4 < X < 10.6$ and $3.15 < Y < 3.25$.

Because X and Y are independent,

$$P(10.4 < X < 10.6, 3.15 < Y < 3.25) = P(10.4 < X < 10.6)P(3.15 < Y < 3.25)$$

$$= P\left(\frac{10.4 - 10.5}{0.05} < Z < \frac{10.6 - 10.5}{0.05}\right) P\left(\frac{3.15 - 3.2}{0.06} < Z < \frac{3.25 - 3.2}{0.06}\right)$$

$$= P(-2 < Z < 2)P(-0.833 < Z < 0.833) = 0.566$$

where Z denotes a standard normal random variable.

5-2 Two Continuous Random Variables

Example 5-20

In an electronic assembly, let the random variables X_1, X_2, X_3, X_4 denote the lifetimes of four components in hours. Suppose that the joint probability density function of these variables is

$$f_{X_1X_2X_3X_4}(x_1, x_2, x_3, x_4) = 9 \times 10^{-2} e^{-0.001x_1 - 0.002x_2 - 0.0015x_3 - 0.003x_4}$$
$$\text{for } x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0$$

What is the probability that the device operates for more than 1000 hours without any failures?

The requested probability is $P(X_1 > 1000, X_2 > 1000, X_3 > 1000, X_4 > 1000)$, which equals the multiple integral of $f_{X_1X_2X_3X_4}(x_1, x_2, x_3, x_4)$ over the region $x_1 > 1000$, $x_2 > 1000$, $x_3 > 1000$, $x_4 > 1000$. The joint probability density function can be written as a product of exponential functions, and each integral is the simple integral of an exponential function. Therefore,

$$P(X_1 > 1000, X_2 > 1000, X_3 > 1000, X_4 > 1000) = e^{-1-2-1.5-3} = 0.00055$$

5-2 Two Continuous Random Variables

Definition: Marginal Probability Density Function

If the joint probability density function of continuous random variables $X_1, X_2, \ldots, X_p$ is $f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)$, the **marginal probability density function** of X_i is

$$f_{X_i}(x_i) = \int\int_{R_{x_i}} \cdots \int f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)\, dx_1\, dx_2 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_p \qquad (5\text{-}21)$$

where the integral is over all points in the range of $X_1, X_2, \ldots, X_p$ for which $X_i = x_i$.

5-2 Two Continuous Random Variables

Mean and Variance from Joint Distribution

$$E(X_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)\, dx_1\, dx_2 \ldots dx_p$$

and (5-22)

$$V(X_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_i - \mu_{X_i})^2 f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)\, dx_1\, dx_2 \ldots dx_p$$

5-2 Two Continuous Random Variables

Distribution of a Subset of Random Variables

If the joint probability density function of continuous random variables $X_1, X_2, \ldots, X_p$ is $f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)$, the **probability density function** of $X_1, X_2, \ldots, X_k$, $k < p$, is

$$f_{X_1 X_2 \ldots X_k}(x_1, x_2, \ldots, x_k)$$
$$= \int \underset{R_{x_1 x_2 \ldots x_k}}{\int} \cdots \int f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)\, dx_{k+1}\, dx_{k+2} \ldots dx_p \qquad (5\text{-}23)$$

where the integral is over all points in the range of $X_1, X_2, \ldots, X_k$ for which $X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k$.

5-2 Two Continuous Random Variables

Conditional Probability Distribution

Definition

Continuous random variables $X_1, X_2, \ldots, X_p$ are **independent** if and only if

$$f_{X_1 X_2 \ldots X_p}(x_1, x_2 \ldots, x_p) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_p}(x_p) \quad \text{for } \textit{all } x_1, x_2, \ldots, x_p \qquad (5\text{-}24)$$

5-2 Two Continuous Random Variables

Example 5-23

Suppose X_1, X_2, and X_3 represent the thickness in micrometers of a substrate, an active layer, and a coating layer of a chemical product. Assume that X_1, X_2, and X_3 are independent and normally distributed with $\mu_1 = 10000$, $\mu_2 = 1000$, $\mu_3 = 80$, $\sigma_1 = 250$, $\sigma_2 = 20$, and $\sigma_3 = 4$, respectively.

The specifications for the thickness of the substrate, active layer, and coating layer are $9200 < x_1 < 10800$, $950 < x_2 < 1050$, and $75 < x_3 < 85$, respectively. What proportion of chemical products meets all thickness specifications? Which one of the three thicknesses has the least probability of meeting specifications?

The requested probability is $P(9200 < X_1 < 10800, 950 < X_2 < 1050, 75 < X_3 < 85$. Because the random variables are independent,

$$\begin{aligned} &P(9200 < X_1 < 10800, 950 < X_2 < 1050, 75 < X_3 < 85) \\ &= P(9200 < X_1 < 10800)P(950 < X_2 < 1050)P(75 < X_3 < 85) \end{aligned}$$

5-2 Two Continuous Random Variables

Example 5-23

After standardizing, the above equals

$$P(-3.2 < Z < 3.2)P(-2.5 < Z < 2.5)P(-1.25 < Z < 1.25)$$

where Z is a standard normal random variable. From the table of the standard normal distribution, the above equals

$$(0.99862)(0.98758)(0.78870) = 0.7778$$

The thickness of the coating layer has the least probability of meeting specifications. Consequently, a priority should be to reduce variability in this part of the process.

5-3 Covariance and Correlation

Definition: Expected Value of a Function of Two Random Variables

$$E[h(X,Y)] = \begin{cases} \sum\sum_{R} h(x,y) f_{XY}(x,y) & X, Y \text{ discrete} \\ \iint_{R} h(x,y) f_{XY}(x,y)\, dx\, dy & X, Y \text{ continuous} \end{cases} \tag{5-25}$$

5-3 Covariance and Correlation

Example 5-24

For the joint probability distribution of the two random variables in Fig. 5-12, calculate $E[(X - \mu_X)(Y - \mu_Y)]$.

The result is obtained by multiplying $x - \mu_X$ times $y - \mu_Y$, times $f_{XY}(x, y)$ for each point in the range of (X, Y). First, μ_X and μ_Y are determined from Equation 5-3 as

$$\mu_X = 1 \times 0.3 + 3 \times 0.7 = 2.4$$

and

$$\mu_Y = 1 \times 0.3 + 2 \times 0.4 + 3 \times 0.3 = 2.0$$

Therefore,

$$\begin{aligned} E[(X - \mu_X)(Y - \mu_Y)] &= (1 - 2.4)(1 - 2.0) \times 0.1 \\ &\quad + (1 - 2.4)(2 - 2.0) \times 0.2 + (3 - 2.4)(1 - 2.0) \times 0.2 \\ &\quad + (3 - 2.4)(2 - 2.0) \times 0.2 + (3 - 2.4)(3 - 2.0) \times 0.3 = 0.2 \end{aligned}$$

5-3 Covariance and Correlation

Example 5-24

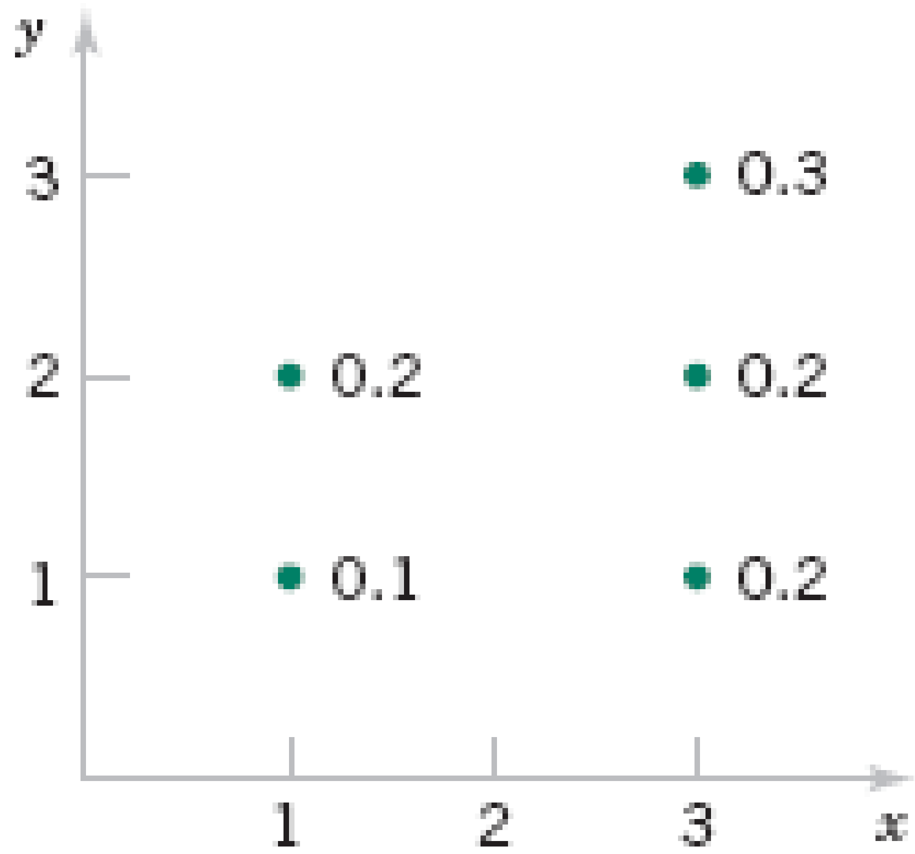


Figure 5-12 Joint distribution of *X* and *Y* for Example 5-24.

5-3 Covariance and Correlation

Definition

The **covariance** between the random variables X and Y, denoted as $\text{cov}(X, Y)$ or σ_{XY}, is

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X\mu_Y \tag{5-26}$$

Covariance is a measure of **linear relationship** between the random variables. If the relationship between the random variables is nonlinear, the covariance might not be sensitive to the relationship. This is illustrated in Fig. 5-13(d). The only points with nonzero probability are the points on the circle. There is an identifiable relationship between the variables. Still, the covariance is zero.

5-3 Covariance and Correlation

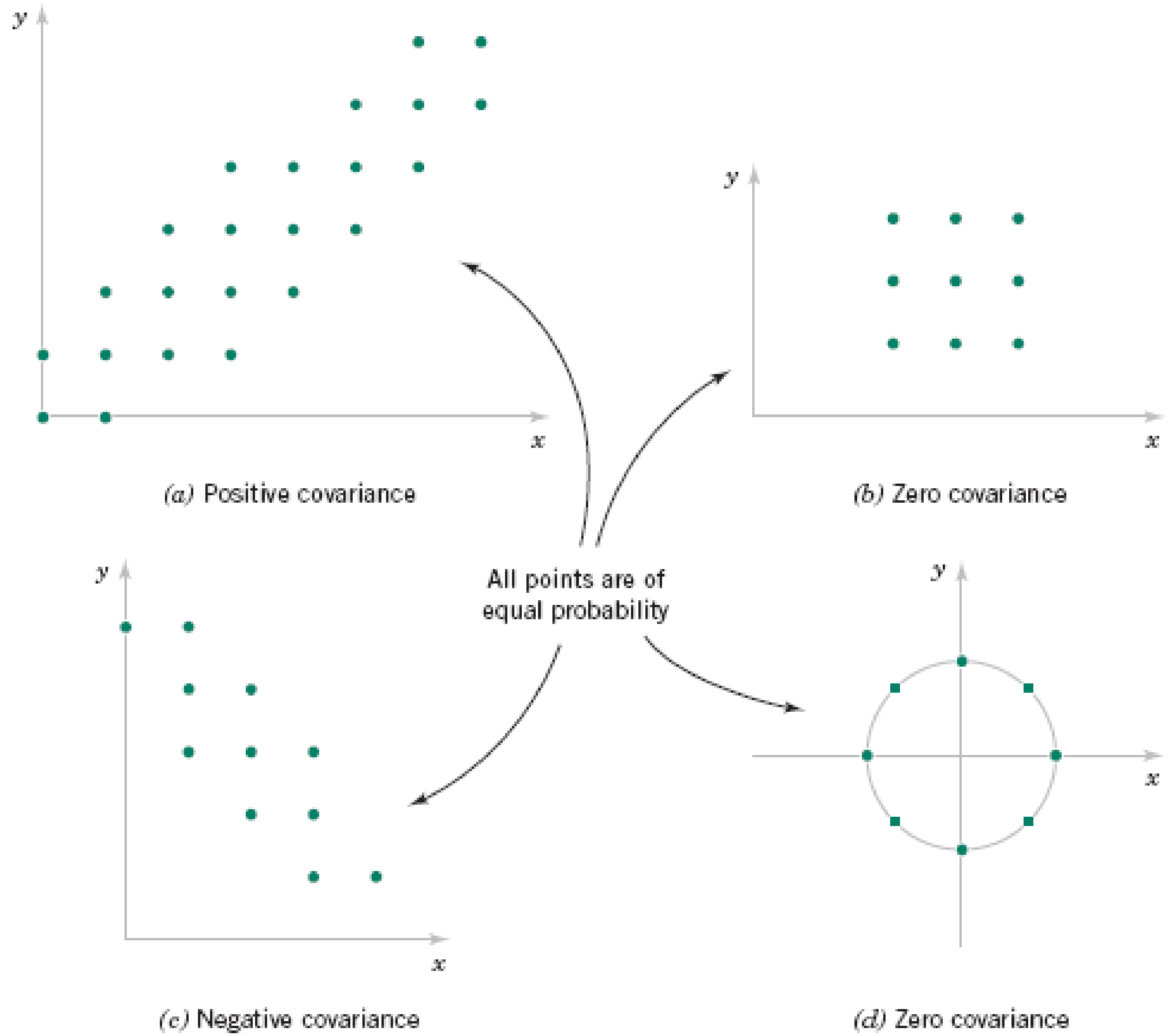


Figure 5-13 Joint probability distributions and the sign of covariance between *X* and *Y*.

5-3 Covariance and Correlation

Definition

The **correlation** between random variables X and Y, denoted as ρ_{XY}, is

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{V(X)\,V(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{5-27}$$

For any two random variables X and Y

$$-1 \le \rho_{XY} \le +1 \tag{5-28}$$

5-3 Covariance and Correlation

Example 5-26

For the discrete random variables X and Y with the joint distribution shown in Fig. 5-14, determine σ_{XY} and ρ_{XY}.

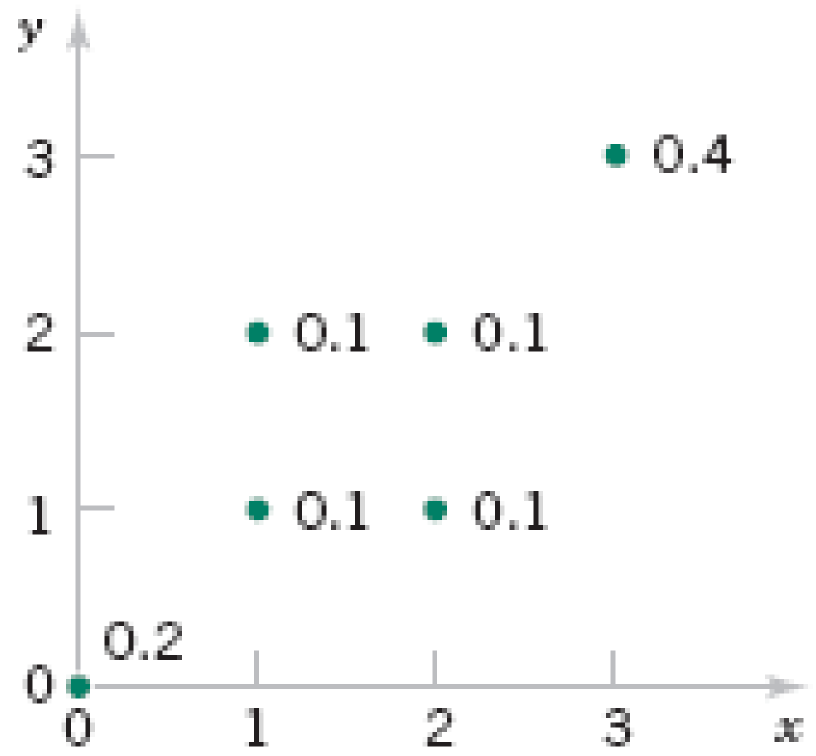


Figure 5-14 Joint distribution for Example 5-26.

5-3 Covariance and Correlation

Example 5-26 (continued)

The calculations for $E(XY)$, $E(X)$, and $V(X)$ are as follows.

$$\begin{aligned}
E(XY) &= 0 \times 0 \times 0.2 + 1 \times 1 \times 0.1 + 1 \times 2 \times 0.1 + 2 \times 1 \times 0.1 \\
&\quad + 2 \times 2 \times 0.1 + 3 \times 3 \times 0.4 = 4.5 \\
E(X) &= 0 \times 0.2 + 1 \times 0.2 + 2 \times 0.2 + 3 \times 0.4 = 1.8 \\
V(X) &= (0 - 1.8)^2 \times 0.2 + (1 - 1.8)^2 \times 0.2 + (2 - 1.8)^2 \times 0.2 \\
&\quad + (3 - 1.8)^2 \times 0.4 = 1.36
\end{aligned}$$

Because the marginal probability distribution of Y is the same as for X, $E(Y) = 1.8$ and $V(Y) = 1.36$. Consequently,

$$\sigma_{XY} = E(XY) - E(X)E(Y) = 4.5 - (1.8)(1.8) = 1.26$$

Furthermore,

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{1.26}{(\sqrt{1.36})(\sqrt{1.36})} = 0.926$$

5-3 Covariance and Correlation

If X and Y are independent random variables,

$$\sigma_{XY} = \rho_{XY} = 0 \tag{5-29}$$

But if, cov(X,Y)=0, then we can not say they are independent.

Example 5-28

For the two random variables in Fig. 5-16, show that $\sigma_{XY} = 0$.

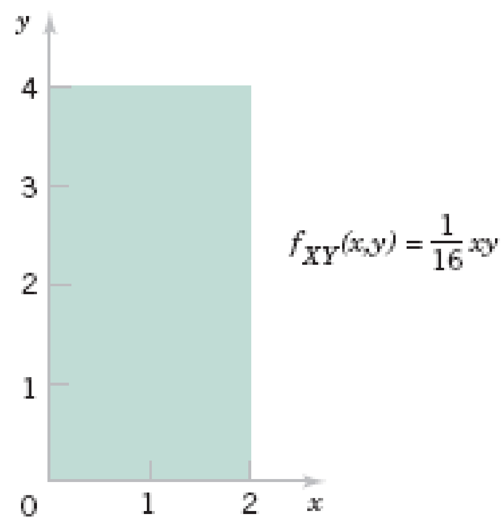


Figure 5-16 Random variables with zero covariance from Example 5-28.

5-3 Covariance and Correlation

Example 5-28 (continued)

The two random variables in this example are continuous random variables. In this case $E(XY)$ is defined as the double integral over the range of (X, Y). That is,

$$E(XY) = \int_0^4 \int_0^2 xy f_{XY}(x, y)\,dx\,dy = \frac{1}{16}\int_0^4 \left[\int_0^2 x^2y^2\,dx\right]dy = \frac{1}{16}\int_0^4 y^2\left[x^3/3\,\Big|_0^2\right]$$

$$= \frac{1}{16}\int_0^4 y^2[8/3]\,dy = \frac{1}{6}\left[y^3/3\,\Big|_0^4\right] = \frac{1}{6}[64/3] = 32/9$$

5-3 Covariance and Correlation

Example 5-28 (continued)

Also,

$$E(X) = \int_0^4 \int_0^2 x f_{XY}(x,y)\,dx\,dy = \frac{1}{16}\int_0^4 \left[\int_0^2 x^2\,dx\right] dy = \frac{1}{16}\int_0^4 \left[x^3/3 \Big|_0^2\right] dy$$

$$= \frac{1}{16}\left[y^2/2 \Big|_0^4\right][8/3] = \frac{1}{6}[16/2] = 4/3$$

$$E(Y) = \int_0^4 \int_0^2 y f_{XY}(x,y)\,dx\,dy = \frac{1}{16}\int_0^4 y^2\left[\int_0^2 x\,dx\right] dy = \frac{1}{16}\int_0^4 y^2\left[x^2/2 \Big|_0^2\right] dy$$

$$= \frac{2}{16}\left[y^3/3 \Big|_0^4\right] = \frac{1}{8}[64/3] = 8/3$$

5-3 Covariance and Correlation

Example 5-28 (continued)

Thus,

$$E(XY) - E(X)E(Y) = 32/9 - (4/3)(8/3) = 0$$

It can be shown that these two random variables are independent. You can check that $f_{XY}(x, y) = f_X(x) f_Y(y)$ for all x and y.

However, if the correlation between two random variables is zero, we *cannot* immediately conclude that the random variables are independent. Figure 5-13(d) provides an example.

5-4 Bivariate Normal Distribution

Definition

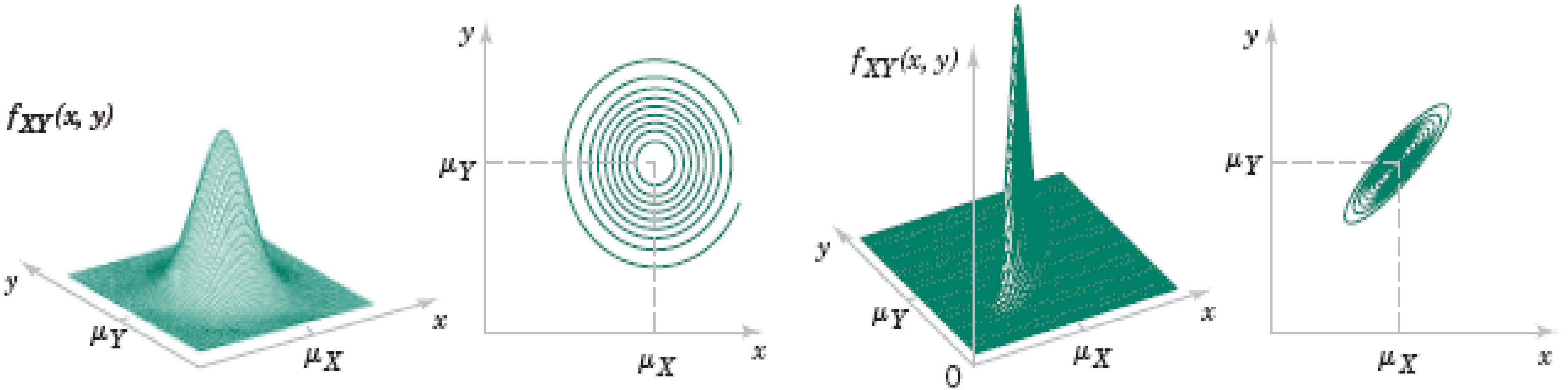
The probability density function of a **bivariate normal distribution** is

$$f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]\right\} \quad (5\text{-}30)$$

for $-\infty < x < \infty$ and $-\infty < y < \infty$, with parameters $\sigma_X > 0$, $\sigma_Y > 0$, $-\infty < \mu_X < \infty$, $-\infty < \mu_Y < \infty$, and $-1 < \rho < 1$.

5-4 Bivariate Normal Distribution

Figure 5-17. Examples of bivariate normal distributions.

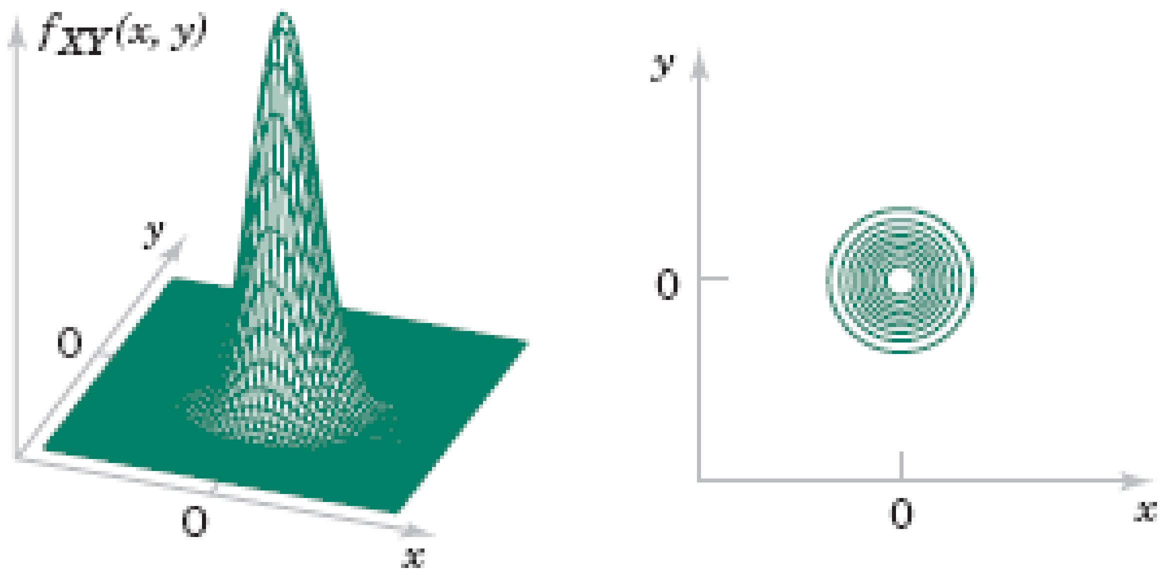


5-4 Bivariate Normal Distribution

Example 5-30

The joint probability density function $f_{XY}(x, y) = \dfrac{1}{\sqrt{2\pi}} e^{-0.5(x^2+y^2)}$ is a special case of a bivariate normal distribution with $\sigma_X = 1$, $\sigma_Y = 1$, $\mu_X = 0$, $\mu_Y = 0$, and $\rho = 0$. This probability density function is illustrated in Fig. 5-18. Notice that the contour plot consists of concentric circles about the origin.

Figure 5-18



5-4 Bivariate Normal Distribution

Marginal Distributions of Bivariate Normal Random Variables

If X and Y have a bivariate normal distribution with joint probability density $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$, the **marginal probability distributions** of X and Y are normal with means μ_X and μ_Y and standard deviations σ_X and σ_Y, respectively. (5-31)

5-4 Bivariate Normal Distribution

Figure 5-19 illustrates that the marginal probability distributions of X and Y are normal. Furthermore, as the notation suggests, ρ represents the correlation between X and Y. The following result is left as an exercise.

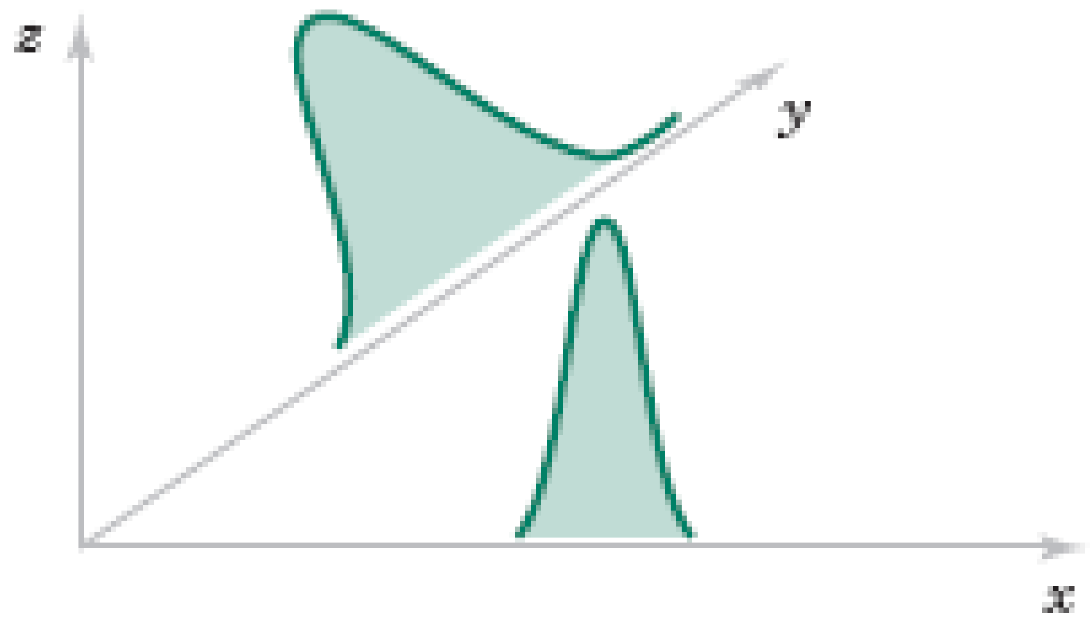


Figure 5-19 Marginal probability density functions of a bivariate normal distributions.

5-4 Bivariate Normal Distribution

If X and Y have a bivariate normal distribution with joint probability density function $f_{XY}(x, y, \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho)$, the correlation between X and Y is ρ. (5-32)

If X and Y have a bivariate normal distribution with $\rho = 0$, X and Y are independent. (5-33)

5-4 Bivariate Normal Distribution

Example 5-31

Suppose that the X and Y dimensions of an injection-molded part have a bivariate normal distribution with $\sigma_X = 0.04$, $\sigma_Y = 0.08$. $\mu_X = 3.00$. $\mu_Y = 7.70$, and $\rho = 0.8$. Then, the probability that a part satisfies both specifications is

$$P(2.95 < X < 3.05, 7.60 < Y < 7.80)$$

This probability can be obtained by integrating $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_X \mu_Y, \rho)$ over the region $2.95 < x < 3.05$ and $7.60 < y < 7.80$, as shown in Fig. 5-7. Unfortunately, there is often no closed-form solution to probabilities involving bivariate normal distributions. In this case, the integration must be done numerically.

5-5 Linear Combinations of Random Variables (Important)

Definition

Given random variables $X_1, X_2, \ldots, X_p$ and constants $c_1, c_2, \ldots, c_p$,

$$Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p \qquad (5\text{-}34)$$

is a **linear combination** of $X_1, X_2, \ldots, X_p$.

Mean of a Linear Combination

If $Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p$,

$$E(Y) = c_1E(X_1) + c_2E(X_2) + \cdots + c_pE(X_p) \qquad (5\text{-}35)$$

5-5 Linear Combinations of Random Variables (Important)

Variance of a Linear Combination

If $X_1, X_2, \ldots, X_p$ are random variables, and $Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p$, then in general

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \cdots + c_p^2V(X_p) + 2\sum_{i<j}\sum c_ic_j \operatorname{cov}(X_i, X_j) \quad (5\text{-}36)$$

If $X_1, X_2, \ldots, X_p$ are **independent**,

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \cdots + c_p^2V(X_p) \quad (5\text{-}37)$$

5-5 Linear Combinations of Random Variables

Example 5-33

An important use of equation 5-37 is in **error propagation** that is presented in the following example.

A semiconductor product consists of three layers. If the variances in thickness of the first, second, and third layers are 25, 40, and 30 nanometers squared, what is the variance of the thickness of the final product.

Let X_1, X_2, X_3, and X be random variables that denote the thickness of the respective layers, and the final product. Then

$$X = X_1 + X_2 + X_3$$

The variance of X is obtained from equaion 5-39

$$V(X) = V(X_1) + V(X_2) + V(X_3) = 25 + 40 + 30 = 95 \text{ nm}^2$$

Consequently, the standard deviation of thickness of the final product is $95^{1/2} = 9.75$ nm and this shows how the variation in each layer is propagated to the final product.

5-5 Linear Combinations of Random Variables

Mean and Variance of an Average

If $\overline{X} = (X_1 + X_2 + \cdots + X_p)/p$ with $E(X_i) = \mu$ for $i = 1, 2, \ldots, p$

$$E(\overline{X}) = \mu \tag{5-38a}$$

if $X_1, X_2, \ldots, X_p$ are also independent with $V(X_i) = \sigma^2$ for $i = 1, 2, \ldots, p$,

$$V(\overline{X}) = \frac{\sigma^2}{p} \tag{5-38b}$$

5-5 Linear Combinations of Random Variables

Reproductive Property of the Normal Distribution

If $X_1, X_2, \ldots, X_p$ are independent, normal random variables with $E(X_i) = \mu_i$ and $V(X_i) = \sigma_i^2$, for $i = 1, 2, \ldots, p$,

$$Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p$$

is a normal random variable with

$$E(Y) = c_1\mu_1 + c_2\mu_2 + \cdots + c_p\mu_p$$

and

$$V(Y) = c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + \cdots + c_p^2\sigma_p^2 \qquad (5\text{-}39)$$

5-5 Linear Combinations of Random Variables

Example 5-34

Let the random variables X_1 and X_2 denote the length and width, respectively, of a manufactured part. Assume that X_1 is normal with $E(X_1) = 2$ centimeters and standard deviation 0.1 centimeter and that X_2 is normal with $E(X_2) = 5$ centimeters and standard deviation 0.2 centimeter. Also, assume that X_1 and X_2 are independent. Determine the probability that the perimeter exceeds 14.5 centimeters.

Then, $Y = 2X_1 + 2X_2$ is a normal random variable that represents the perimeter of the part. We obtain, $E(Y) = 14$ centimeters and the variance of Y is

$$V(Y) = 4 \times 0.1^2 + 4 \times 0.2^2 = 0.2$$

Now,

$$\begin{aligned} P(Y > 14.5) &= P[(Y - \mu_Y)/\sigma_Y > (14.5 - 14)/\sqrt{0.2}] \\ &= P(Z > 1.12) = 0.13 \end{aligned}$$

5-6 General Functions of Random Variables

A Discrete Random Variable

Suppose that X is a **discrete** random variable with probability distribution $f_X(x)$. Let $Y = h(X)$ define a one-to-one transformation between the values of X and Y so that the equation $y = h(x)$ can be solved uniquely for x in terms of y. Let this solution be $x = u(y)$. Then the probability mass function of the random variable Y is

$$f_Y(y) = f_X[u(y)] \tag{5-40}$$

5-6 General Functions of Random Variables

Example 5-36

Let X be a geometric random variable with probability distribution

$$f_X(x) = p(1-p)^{x-1}, \qquad x = 1, 2, \ldots$$

Find the probability distribution of $Y = X^2$.

Since $X \geq 0$, the transformation is one to one; that is, $y = x^2$ and $x = \sqrt{y}$. Therefore, Equation 5-40 indicates that the distribution of the random variable Y is

$$f_Y(y) = f(\sqrt{y}) = p(1-p)^{\sqrt{y}-1}, \qquad y = 1, 4, 9, 16, \ldots$$

5-6 General Functions of Random Variables

A Continuous Random Variable

Suppose that X is a **continuous** random variable with probability distribution $f_X(x)$. The function $Y = h(X)$ is a one-to-one transformation between the values of Y and X so that the equation $y = h(x)$ can be uniquely solved for x in terms of y. Let this solution be $x = u(y)$. The probability distribution of Y is

$$f_Y(y) = f_X[u(y)]\,|J| \tag{5-41}$$

where $J = u'(y)$ is called the **Jacobian** of the transformation and the absolute value of J is used.

5-6 General Functions of Random Variables

Example 5-37

Let X be a continuous random variable with probability distribution

$$f_X(x) = \frac{x}{8}, \quad 0 \le x < 4$$

Find the probability distribution of $Y = h(X) = 2X + 4$.

Note that $y = h(x) = 2x + 4$ is an increasing function of x. The inverse solution is $x = u(y) = (y - 4)/2$, and from this we find the Jacobian to be $J = u'(y) = dx/dy = 1/2$. Therefore, from S5-3 the probability distribution of Y is

$$f_Y(y) = \frac{(y-4)/2}{8}\left(\frac{1}{2}\right) = \frac{y-4}{32}, \qquad 4 \le y \le 12$$

IMPORTANT TERMS AND CONCEPTS

Bivariate distribution
Bivariate normal distribution
Conditional mean
Conditional probability density function
Conditional probability mass function
Conditional variance
Contour plots
Correlation
Covariance
Error propagation
General functions of random variables
Independence
Joint probability density function
Joint probability mass function
Linear functions of random variables
Marginal probability distribution
Multinomial distribution
Reproductive property of the normal distribution