5-3 Covariance and Correlation

Definition: Expected Value of a Function of Two Random Variables

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

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,

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

5-3 Covariance and Correlation

Example 5-24

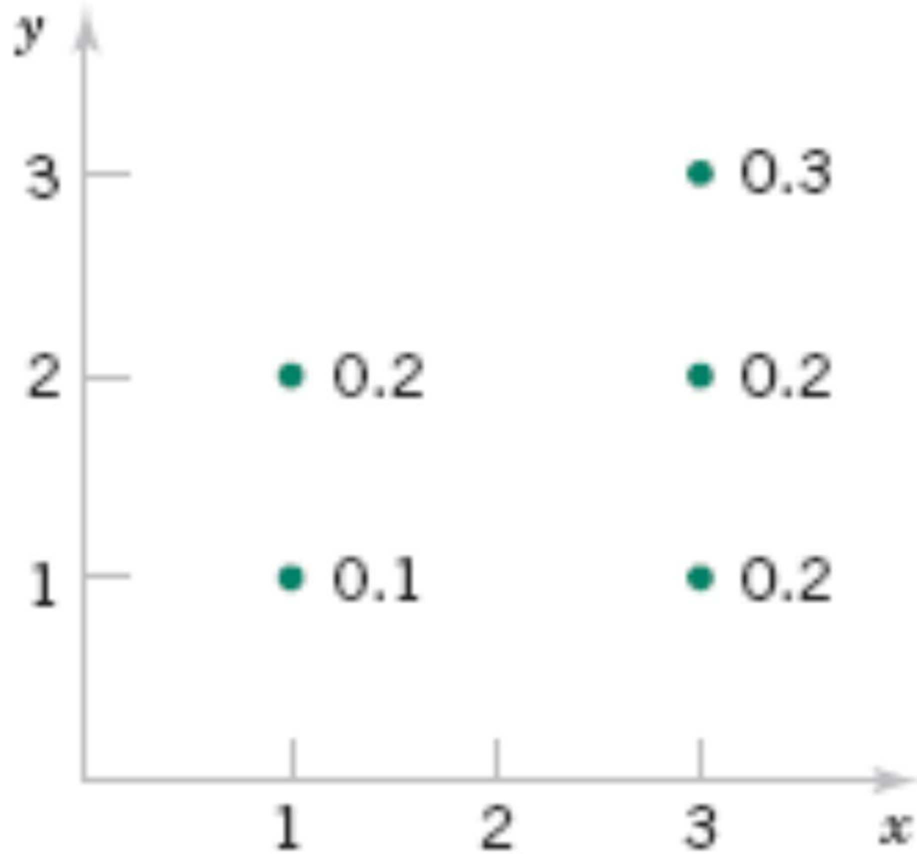


Figure 5-12 Joint distribution of *X* and *Y* for Example 5-24.

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

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

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

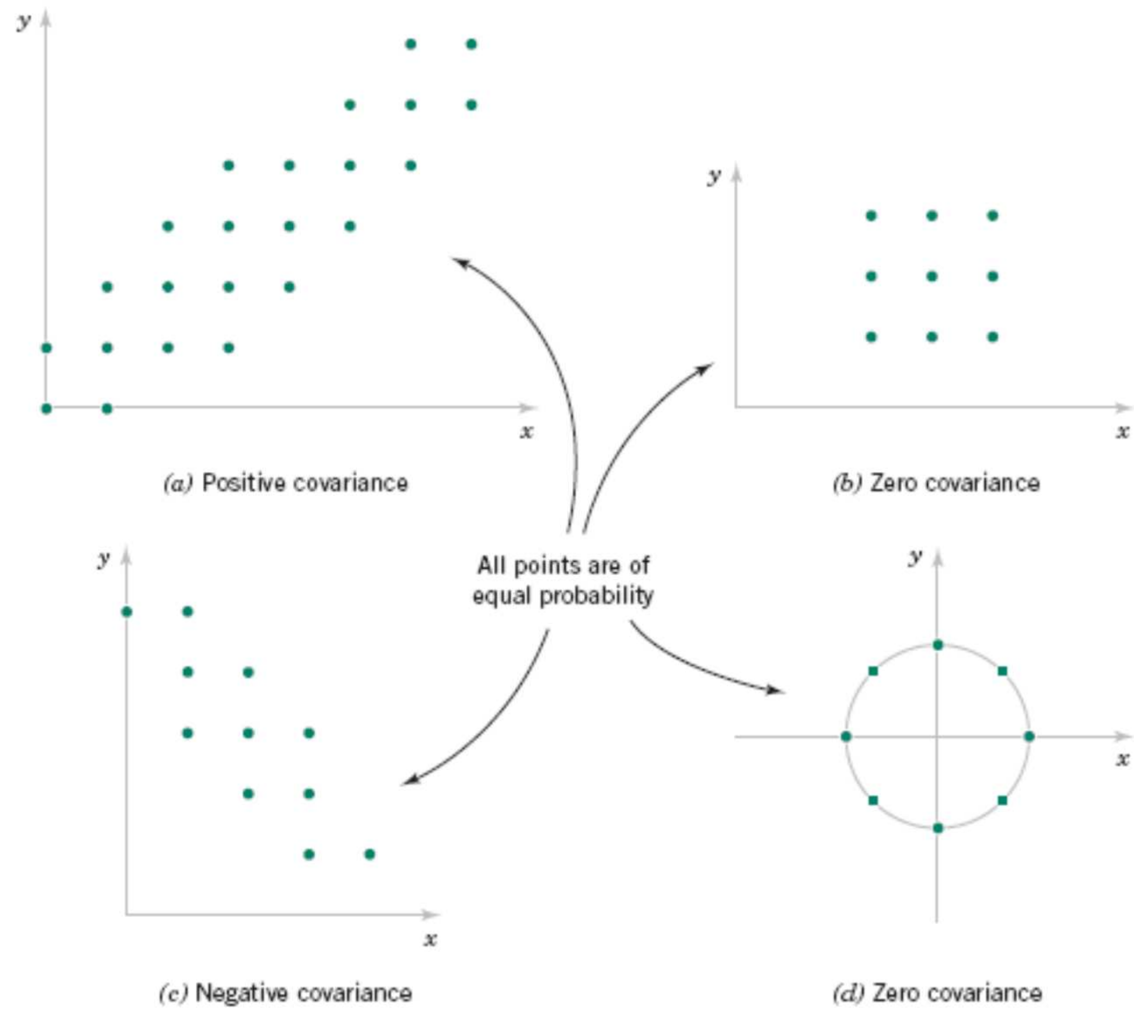


Figure 5-13 Joint probability distributions and the sign of covariance between X and Y.

5-3 Covariance and Correlation

Definition

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

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

For any two random variables X and Y

$$-1 \leq \rho_{XY} \leq +1 \qquad (5\text{-}28)$$

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}.

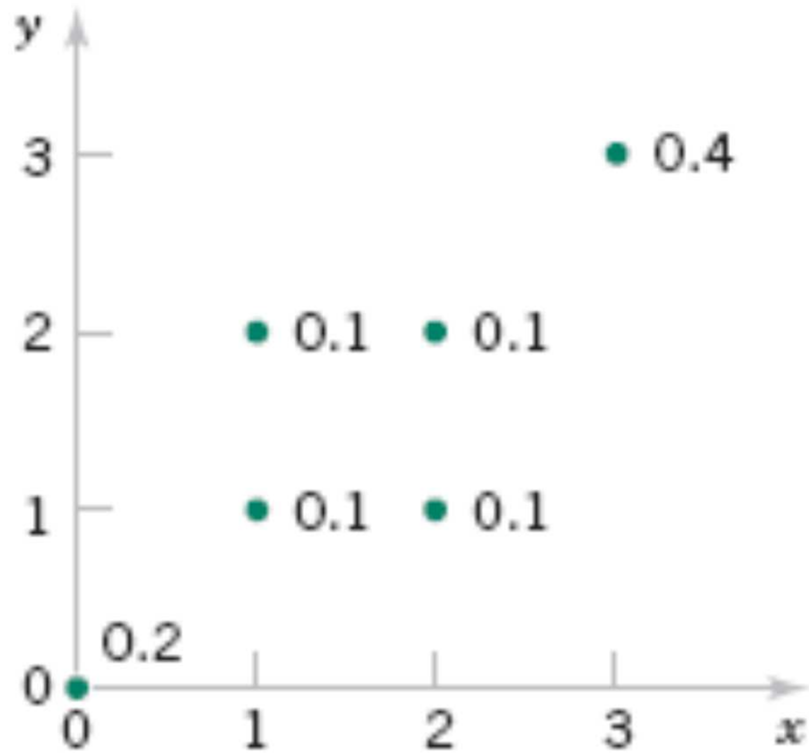


Figure 5-14 Joint distribution for Example 5-26.

5-3 Covariance and Correlation

Example 5-26 (continued)

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

$$E(XY) = 0 \times 0 \times 0.2 + 1 \times 1 \times 0.1 + 1 \times 2 \times 0.1 + 2 \times 1 \times 0.1 + 2 \times 2 \times 0.1 + 3 \times 3 \times 0.4 = 4.5$$

$$E(X) = 0 \times 0.2 + 1 \times 0.2 + 2 \times 0.2 + 3 \times 0.4 = 1.8$$

$$V(X) = (0 - 1.8)^2 \times 0.2 + (1 - 1.8)^2 \times 0.2 + (2 - 1.8)^2 \times 0.2 + (3 - 1.8)^2 \times 0.4 = 1.36$$

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,

$$\sigma_{XY} = \rho_{XY} = 0 \qquad (5\text{-}29)$$

Example 5-28

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

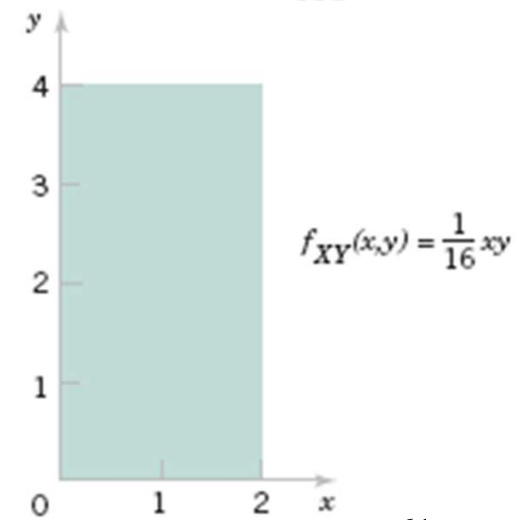


Figure 5-16 Random variables with zero covariance from Example 5-28.

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,

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

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

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

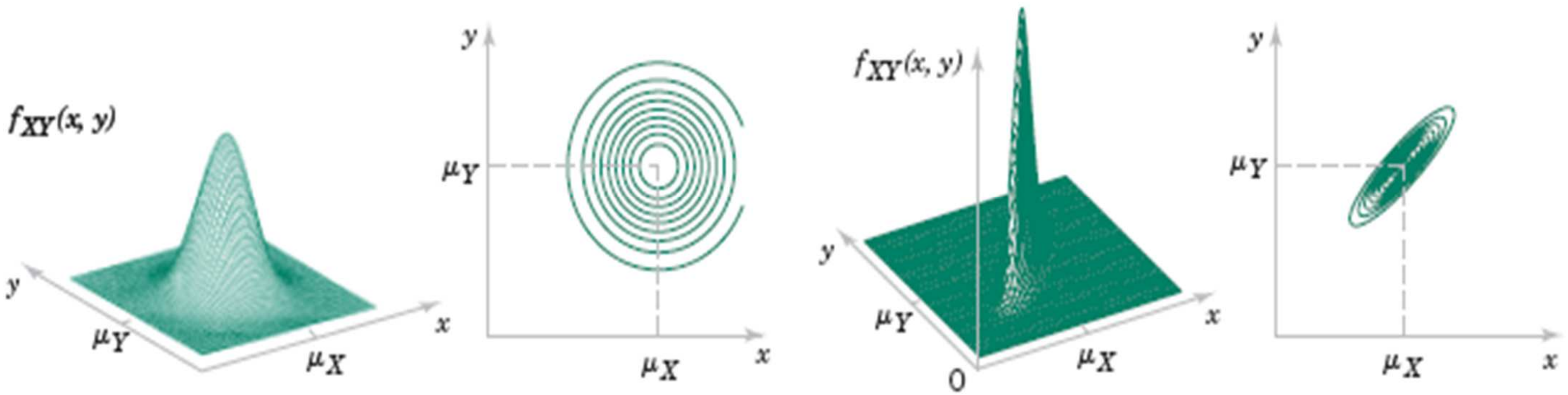
The probability density function of a **bivariate normal distribution** is

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

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 $f_{XY}(x, y) = \frac{1}{\sqrt{2\pi}} e^{-0.5(x^2+y^2)}$ 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.

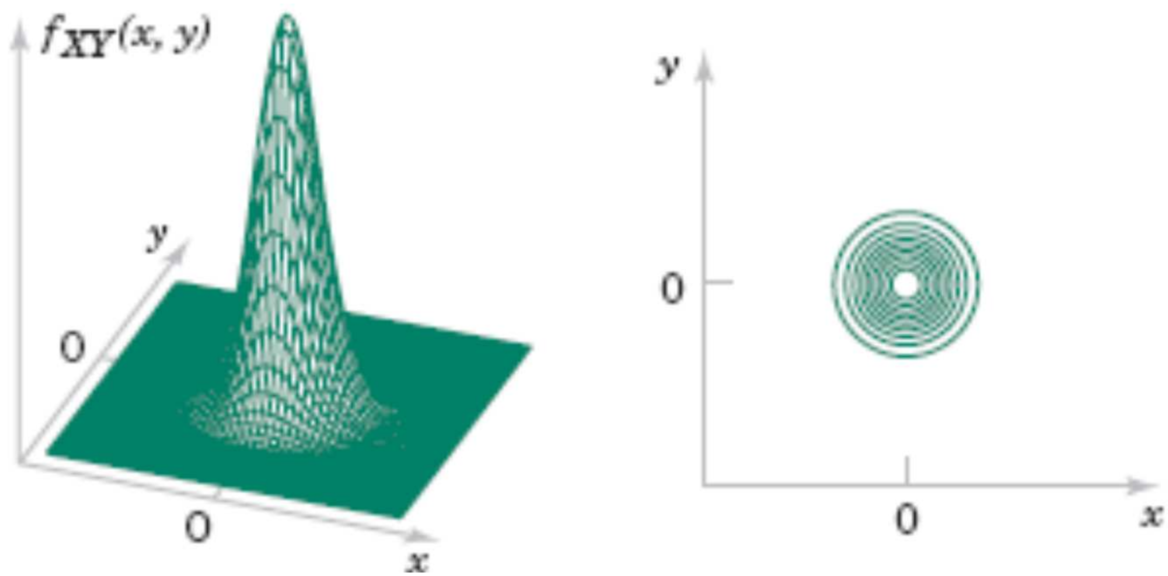


Figure 5-18

5-4 Bivariate Normal Distribution

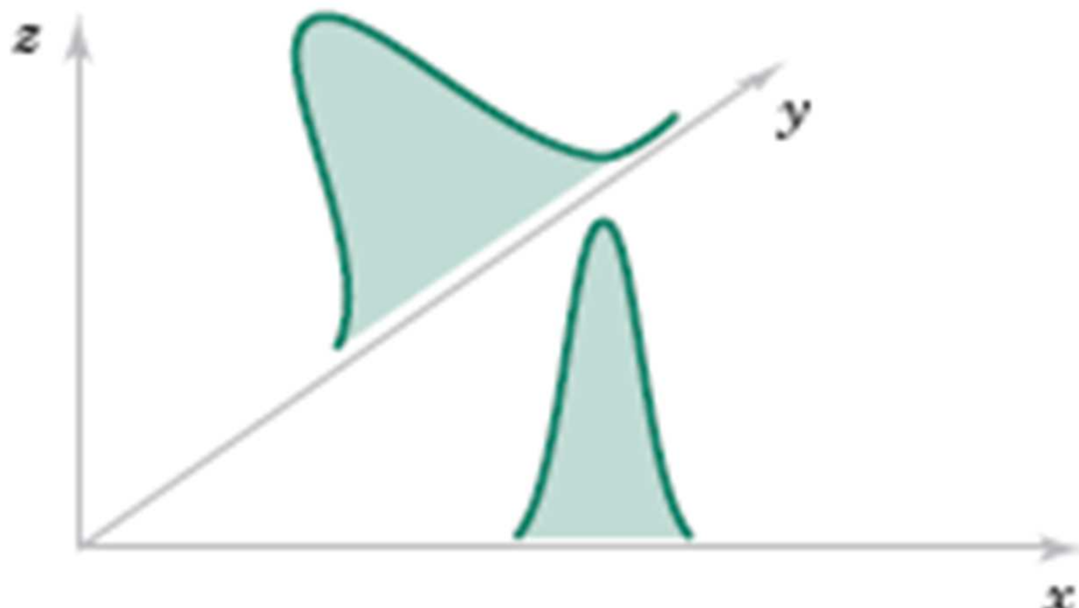
Marginal Distributions of Bivariate Normal Random Variables

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

Definition

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

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

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

Mean of a Linear Combination

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

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

5-5 Linear Combinations of Random Variables

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,

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

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

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

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