in short supply in the United States. Although the number of doses demanded each month is a discrete random variable, the large demands can be approximated with a continuous probability distribution. Suppose that the monthly demands for two of those vaccines, namely measles–mumps–rubella (MMR) and varicella (for chickenpox), are independently, normally distributed with means of 1.1 and 0.55 million doses and standard deviations of 0.3 and 0.1 million doses, respectively. Also suppose that the inventory levels at the beginning of a given month for MMR and varicella vaccines are 1.2 and 0.6 million doses, respectively.

(a) What is the probability that there is no shortage of either vaccine in a month without any vaccine production?

(b) To what should inventory levels be set so that the probability is 90% that there is no shortage of either vaccine in a month without production? Can there be more than one answer? Explain.

5-32. The systolic and diastolic blood pressure values (mm Hg) are the pressures when the heart muscle contracts and relaxes (denoted as Y and X, respectively). Over a collection of individuals, the distribution of diastolic pressure is normal with mean 73 and standard deviation 8. The systolic pressure is conditionally normally distributed with mean $1.6x$ when $X = x$ and standard deviation of 10. Determine the following:

(a) Conditional probability density function $f_{Y|73}(y)$ of Y given $X = 73$

(b) $P(Y < 115 \mid X = 73)$

(c) $E(Y \mid X = 73)$

(d) Recognize the distribution $f_{XY}(x, y)$ and identify the mean and variance of Y and the correlation between X and Y

5-2 Covariance and Correlation

When two or more random variables are defined on a probability space, it is useful to describe how they vary together; that is, it is useful to measure the relationship between the variables. A common measure of the relationship between two random variables is the **covariance**. To define the covariance, we need to describe the expected value of a function of two random variables $h(X, Y)$. The definition simply extends the one for a function of a single random variable.

Expected Value of a Function of Two Random Variables

$$E\left[h(X,Y)\right] = \begin{cases} \sum\sum h(x,y)\, f_{XY}(x,y) & X, Y \text{ discrete} \\ \int\int h(x,y)\, f_{XY}(x,y)\, dx\, dy & X, Y \text{ continuous} \end{cases} \qquad (5\text{-}13)$$

That is, $E\left[h(X,Y)\right]$ can be thought of as the weighted average of $h(x, y)$ for each point in the range of (X, Y). The value of $E\left[h(X,Y)\right]$ represents the average value of $h(X, Y)$ that is expected in a long sequence of repeated trials of the random experiment.

Example 5-19 **Expected Value of a Function of Two Random Variables** For the joint probability distribution of the two random variables in Example 5-1, calculate $E\left[(X-\mu_X)(Y-\mu_Y)\right]$.

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 were determined previously from the marginal distributions for X and Y:

$$\mu_X = 2.35$$

and

$$\mu_Y = 2.49$$

Therefore,

$$\begin{aligned} E[(X-\mu_X)(Y-\mu_Y)] &= (1-2.35)(1-2.49)(0.01) + (2-2.35)(1-2.4)(0.02) + (3-2.35)(1-2.49)(0.25) \\ &+ (1-2.35)(2-2.49)(0.02) + (2-2.35)(2-2.4)(0.03) + (3-2.35)(2-2.49)(0.2) \\ &+ (1-2.35)(3-2.49)(0.02) + (2-2.35)(3-2.4)(0.1) + (3-2.35)(3-2.49)(0.05) \\ &+ (1-2.35)(4-2.49)(0.15) + (2-2.35)(4-2.4)(0.1) + (3-2.35)(4-2.49)(0.05) = -0.5815 \end{aligned}$$

The covariance is defined for both continuous and discrete random variables by the same formula.

Covariance

> The **covariance** between the random variables X and Y, denoted as $\text{cov}(X, Y)$ or σ_{XY}, is
>
> $$\sigma_{XY} = E\left[(X - \mu_X)(Y - \mu_Y)\right] = E(XY) - \mu_X \mu_Y \tag{5-14}$$

If the points in the joint probability distribution of X and Y that receive positive probability tend to fall along a line of positive (or negative) slope, σ_{XY}, is positive (or negative). If the points tend to fall along a line of positive slope, X tends to be greater than μ_X when Y is greater than μ_Y. Therefore, the product of the two terms $x - \mu_X$ and $y - \mu_Y$ tends to be positive. However, if the points tend to fall along a line of negative slope, $x - \mu_X$ tends to be positive when $y - \mu_Y$ is negative, and vice versa. Therefore, the product of $x - \mu_X$ and $y - \mu_Y$ tends to be negative. In this sense, the covariance between X and Y describes the variation between the two random variables. Figure 5-12 assumes all points are equally likely and shows examples of pairs of random variables with positive, negative, and zero covariance.

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. 5-12(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.

The equality of the two expressions for covariance in Equation 5-14 is shown for continuous random variables as follows. By writing the expectations as integrals,

$$E\left[(Y - \mu_Y)(X - \mu_X)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y)\, f_{XY}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[xy - \mu_X y - x\mu_Y + \mu_X \mu_Y\right] f_{XY}(x, y)\, dx\, dy$$

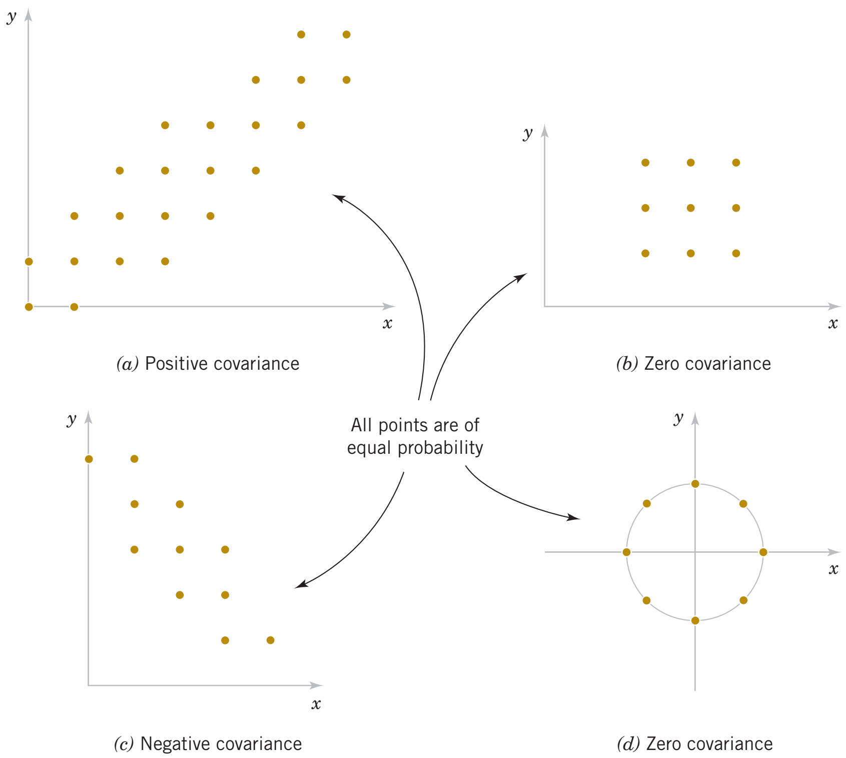

FIGURE 5-12 Joint probability distributions and the sign of covariance between X and Y.

Now

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mu_X y\ f_{XY}(x, y)\,dx\,dy = \mu_X\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x, y)\,dx\,dy\right] = \mu_X\mu_Y$$

Therefore,

$$E\left[(X-\mu_X)(Y-\mu_Y)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x, y)\,dx\,dy - \mu_X\mu_Y - \mu_X\mu_Y + \mu_X\mu_Y$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x, y)\,dx\,dy - \mu_X\mu_Y = E(XY) - \mu_X\mu_Y$$

Example 5-20 In Example 5-1, the random variables X and Y are the number of signal bars and the response time (to the nearest second), respectively. Interpret the covariance between X and Y as positive or negative.

As the signal bars increase, the response time tends to decrease. Therefore, X and Y have a negative covariance. The covariance was calculated to be −0.5815 in Example 5-19.

There is another measure of the relationship between two random variables that is often easier to interpret than the covariance.

Correlation

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

$$\rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \tag{5-15}$$

Because $\sigma_X > 0$ and $\sigma_Y > 0$, if the covariance between X and Y is positive, negative, or zero, the correlation between X and Y is positive, negative, or zero, respectively. The following result can be shown.

For any two random variables X and Y,

$$-1 \le \rho_{XY} \le +1 \tag{5-16}$$

The correlation just scales the covariance by the product of the standard deviation of each variable. Consequently, the correlation is a dimensionless quantity that can be used to compare the linear relationships between pairs of variables in different units.

If the points in the joint probability distribution of X and Y that receive positive probability tend to fall along a line of positive (or negative) slope, ρ_{XY} is near +1 (or −1). If ρ_{XY} equals +1 or −1, it can be shown that the points in the joint probability distribution that receive positive probability fall exactly along a straight line. Two random variables with nonzero correlation are said to be **correlated.** Similar to covariance, the correlation is a measure of the **linear relationship** between random variables.

Example 5-21 **Covariance** For the discrete random variables X and Y with the joint distribution shown in Fig. 5-13, determine σ_{XY} and ρ_{XY}.

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

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

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

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

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,

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

Example 5-22 **Correlation** Suppose that the random variable X has the following distribution: $P(X = 1) = 0.2$, $P(X = 2) = 0.6$, $P(X = 3) = 0.2$. Let $Y = 2X + 5$. That is, $P(Y = 7) = 0.2$, $P(Y = 11) = 0.2$. Determine the correlation between X and Y. Refer to Fig. 5-14.

Because X and Y are linearly related, $\rho = 1$. This can be verified by direct calculations: Try it.

For independent random variables, we do not expect any relationship in their joint probability distribution. The following result is left as an exercise.

If X and Y are independent random variables,

$$\sigma_{XY} = \rho_{XY} = 0 \qquad (5\text{-}17)$$

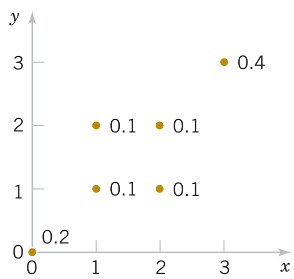

FIGURE 5-13 Joint distribution for Example 5-20.

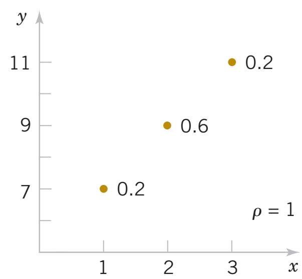

FIGURE 5-14 Joint distribution for Example 5-21.

Example 5-23 **Independence Implies Zero Covariance** For the two random variables in Fig. 5-15, show that $\sigma_{XY} = 0$.

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,

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

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

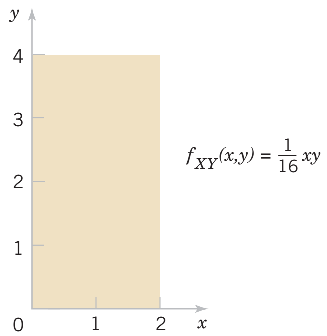

FIGURE 5-15 Random variables with zero covariance from Example 5-22.

Also,

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

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

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

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

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 5-12(d) provides an example.

EXERCISES FOR SECTION 5-2

⊕ Problem available in *WileyPLUS* at instructor's discretion.
⊕ **Go Tutorial** Tutoring problem available in *WileyPLUS* at instructor's discretion.

5-33. ⊕ Determine the covariance and correlation for the following joint probability distribution:

x	1	1	2	4
y	3	4	5	6
$f_{XY}(x,y)$	1/8	1/4	1/2	1/8

5-34. ⊕ Determine the covariance and correlation for the following joint probability distribution:

x	–1	–0.5	0.5	1
y	–2	–1	1	2
$f_{XY}(x,y)$	1/8	1/4	1/2	1/8

5-35. ⊕ Determine the value for c and the covariance and correlation for the joint probability mass function $f_{XY}(x,y) = c(x+y)$ for $x = 1, 2, 3$ and $y = 1, 2, 3$.

5-36. ⊕ Determine the covariance and correlation for the joint proba.bility distribution shown in Fig. 5-10(a) and described in Example 5-10.

5-37. ⊕ Patients are given a drug treatment and then evaluated. Symptoms either improve, degrade, or remain the same with probabilities 0.4, 0.1, 0.5, respectively. Assume that four independent patients are treated and let X and Y

denote the number of patients who improve or degrade. Are X and Y independent? Calculate the covariance and correlation between X and Y.

5-38. For the Transaction Processing Performance Council's benchmark in Exercise 5-10, let X, Y, and Z denote the average number of *selects, updates,* and *inserts* operations required for each type of transaction, respectively. Calculate the following:
(a) Covariance between X and Y
(b) Correlation between X and Y
(c) Covariance between X and Z
(d) Correlation between X and Z

5-39. ⊕ Determine the value for c and the covariance and correlation for the joint probability density function $f_{XY}(x, y) = cxy$ over the range $0 < x < 3$ and $0 < y < x$.

5-40. ⊕ Determine the value for c and the covariance and correlation for the joint probability density function $f_{XY}(x, y) = c$ over the range $0 < x < 5$, $0 < y$, and $x - 1 < y < x + 1$.

5-41. ⊕ Determine the covariance and correlation for the joint probability density function $f_{XY}(x, y) = e^{-x-y}$ over the range $0 < x$ and $0 < y$.

5-42. ⊕ Determine the covariance and correlation for the joint probability density function $f_{XY}(x, y) = 6 \quad 10^{-6} e^{-0.001x - 0.002y}$ over the range $0 < x$ and $x < y$ from Example 5-2.

5-43. The joint probability distribution is

x	–1	0	0	1
y	0	–1	1	0
$f_{XY}(x, y)$	1/4	1/4	1/4	1/4

Show that the correlation between X and Y is zero but X and Y are not independent.

5-44. Determine the covariance and correlation for the CD4 counts in a month and the following month in Exercise 5-30.

5-45. Determine the covariance and correlation for the lengths of the minor and major axes in Exercise 5-29.

5-46. Suppose that X and Y are independent continuous random variables. Show that $\sigma_{XY} = 0$.

5-47. ⊕ Suppose that the correlation between X and Y is ρ. For constants $a, b, c,$ and d, what is the correlation between the random variables $U = aX + b$ and $V = cY + d$?

5-3 Common Joint Distributions

5-3.1 MULTINOMIAL PROBABILITY DISTRIBUTION

The binomial distribution can be generalized to generate a useful joint probability distribution for multiple discrete random variables. The random experiment consists of a series of independent trials. However, the outcome from each trial is categorized into one of k classes. The random variables of interest count the number of outcomes in each class.

Example 5-21 **Digital Channel** We might be interested in a probability such as the following. Of the 20 bits received, what is the probability that 14 are excellent, 3 are good, 2 are fair, and 1 is poor? Assume that the classifications of individual bits are independent events and that the probabilities of E, G, F, and P are 0.6, 0.3, 0.08, and 0.02, respectively. One sequence of 20 bits that produces the specified numbers of bits in each class can be represented as

$$EEEEEEEEEEEEEEGGGFFP$$

Using independence, we find that the probability of this sequence is

$$P(EEEEEEEEEEEEEEGGGFFP) = 0.6^{14} 0.3^{3} 0.08^{2} 0.02^{1}$$
$$= 2.708 \quad 10^{-9}$$

Clearly, all sequences that consist of the same numbers of E's, G's, F's, and P's have the same probability. Consequently, the requested probability can be found by multiplying 2.708×10^{-9} by the number of sequences with 14 E's, 3 G's, 2 F's, and 1 P. The number of sequences is found from Chapter 2 to be

$$\frac{20!}{14!\,3!\,2!\,1!} = 2{,}325{,}600$$

Therefore, the requested probability is

$$P(14 E\text{'s}, 3\ G\text{'s}, 2\ F\text{'s, and } 1\ P) = 2325600(2.708 \quad 10^{-9}) = 0.0063$$

Example 5-24 leads to the following generalization of a binomial experiment and a binomial distribution.

Multinomial Distribution

Suppose that 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

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k} \tag{5-18}$$

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

The multinomial distribution is considered a multivariable extension of the binomial distribution.

Example 5-25 **Digital Channel** In Example 5-24, let the random variables X_1, X_2, X_3, and X_4 denote the number of bits that are E, G, F, and P, respectively, in a transmission of 20 bits. The probability that 12 of the bits received are E, 6 are G, 2 are F, and 0 are P is

$$P(X_1 = 12, X_2 = 6, X_3 = 2, X_4 = 0)$$
$$= \frac{20!}{12!\,6!\,2!\,0!} 0.6^{12} 0.3^{6} 0.08^{2} 0.02^{0} = 0.0358$$

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.

Mean and Variance

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

$$E(X_i) = np_i \quad \text{and} \quad V(X_i) = np_i(1 - p_i) \tag{5-19}$$

Example 5-26 **Marginal Probability Distributions** In Example 5-25, the marginal probability distribution of X_2 is binomial with $n = 20$ and $p = 0.3$. Furthermore, the joint marginal probability distribution of X_2 and X_3 is found as follows. The $P(X_2 = x_2, X_3 = x_3)$ is the probability that exactly x_2 trials result in G and that x_3 result in F. The remaining $n - x_2 - x_3$ trials must result in either E or P. Consequently, we can consider each trial in the experiment to result in one of three classes: $\{G\}$, $\{F\}$, and $\{E, P\}$ with probabilities 0.3, 0.08, and $0.6 + 0.02 = 0.62$, respectively. With these new classes, we can consider the trials to comprise a new multinomial experiment. Therefore,

$$f_{X_2 X_3}(x_2, x_3) = P(X_2 = x_2, X_3 = x_3)$$
$$= \frac{n!}{x_2!x_3!(n - x_2 - x_3)!} (0.3)^{x_2} (0.08)^{x_3} (0.62)^{n - x_2 - x_3}$$

The joint probability distribution of other sets of variables can be found similarly.

5-3.2 BIVARIATE NORMAL DISTRIBUTION

An extension of a normal distribution to two random variables is an important bivariate probability distribution. The joint probability distribution can be defined to handle positive, negative, or zero correlation between the random variables.

Example 5-27 **Bivariate Normal Distribution** At the start of this chapter, the length of different dimensions of an injection-molded part were presented as an example of two random variables. If the specifications for X and Y are 2.95 to 3.05 and 7.60 to 7.80 millimeters, respectively, we might be interested in the probability that a part satisfies both specifications; that is, $P(2.95 < X < 3.05,\ 7.60 < Y < 7.80)$. Each length might be modeled by a normal distribution. However, because the measurements are from the same part, the random variables are typically not independent. Therefore, a probability distribution for two normal random variables that are not independent is important in many applications.

Bivariate Normal Probability Density Function

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

$$f_{XY}\left(x, y; \sigma_X, \sigma_Y, \mu_X, \mu_Y, \rho\right) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

$$\exp\left\{\frac{-1}{2\left(1-\rho^2\right)}\left[\frac{\left(x-\mu_X\right)^2}{\sigma_X^2} - \frac{2\rho\left(x-\mu_X\right)\left(y-\mu_Y\right)}{\sigma_X\sigma_Y} + \frac{\left(y-\mu_Y\right)^2}{\sigma_Y^2}\right]\right\} \tag{5-20}$$

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 $-1 < \rho < 1$.

The result that $f_{XY}(x, y; \sigma_X, \sigma_Y, \mu_{\bar{X}}, \mu_{\bar{Y}}, \rho)$ integrates to 1 is left as an exercise. Also, the bivariate normal probability density function is positive over the entire plane of real numbers.

Two examples of bivariate normal distributions along with corresponding **contour plots** are illustrated in Fig. 5-16. Each curve on the contour plots is a set of points for which the probability density function is constant. As seen in the contour plots, the bivariate normal probability density function is constant on ellipses in the (x, y) plane. (We can consider a circle to be a special case of an ellipse.) The center of each ellipse is at the point (μ_X, μ_Y). If $\rho > 0$ $(\rho < 0)$, the major axis of each ellipse has positive (negative) slope, respectively. If $\rho = 0$, the major axis of the ellipse is aligned with either the x or y coordinate axis.

Example 5-28 The joint probability density function

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

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. 5-17. Notice that the contour plot consists of concentric circles about the origin.

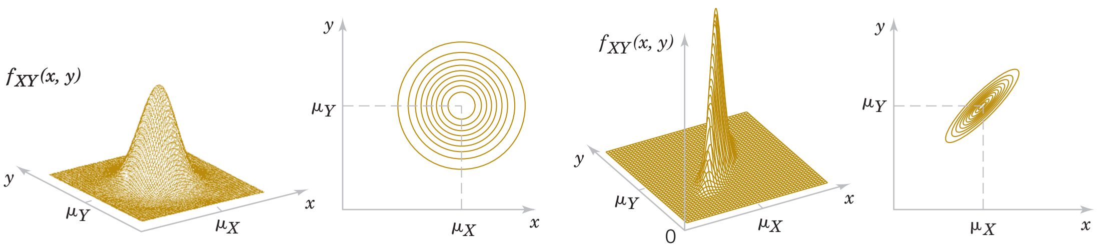

FIGURE 5-16 Examples of bivariate normal distributions.