first year in college, respectively. Persons in the field of educational testing and measurement are extremely interested in the conditional distribution of Y, given $X = x$, in such situations.

Suppose that we have an application in which we can make the following three assumptions about the conditional distribution of Y, given $X = x$:

(a) It is normal for each real x.
(b) Its mean $E(Y \mid x)$ is a linear function of x.
(c) Its variance is constant; that is, it does not depend upon the given value of x.

Of course, assumption (b), along with a result given in Section 4.3 implies that

MEAN

$$E(Y \mid x) = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

Let us consider the implication of assumption (c). The conditional variance is given by

$$\sigma_{Y|x}^2 = \int_{-\infty}^{\infty} \left[y - \mu_Y - \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) \right]^2 h(y \mid x)\, dy,$$

where $h(y \mid x)$ is the conditional p.d.f. of Y given $X = x$. Multiply each member of this equation by $f_1(x)$ and integrate on x. Since $\sigma_{Y|x}^2$ is a constant, the lefthand member is equal to $\sigma_{Y|x}^2$. Thus we have

$$\sigma_{Y|x}^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[y - \mu_Y - \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) \right]^2 h(y \mid x) f_1(x)\, dy\, dx. \quad (5.6\text{-}1)$$

However, $h(y \mid x) f_1(x) = f(x, y)$; hence the righthand member is just an expectation and Equation 5.6-1 can be written as

$$\sigma_{Y|x}^2 = E\left\{ (Y - \mu_Y)^2 - 2\rho \frac{\sigma_Y}{\sigma_X}(X - \mu_X)(Y - \mu_Y) + \rho^2 \frac{\sigma_Y^2}{\sigma_X^2}(X - \mu_X)^2 \right\}.$$

But using the fact that the expectation E is a linear operator, we have, recalling $E[(X - \mu_X)(Y - \mu_Y)] = \rho \sigma_X \sigma_Y$, that

VARIANCE

$$\begin{aligned} \sigma_{Y|x}^2 &= \sigma_Y^2 - 2\rho \frac{\sigma_Y}{\sigma_X} \rho \sigma_X \sigma_Y + \rho^2 \frac{\sigma_Y^2}{\sigma_X^2} \sigma_X^2 \\ &= \sigma_Y^2 - 2\rho^2 \sigma_Y^2 + \rho^2 \sigma_Y^2 = \sigma_Y^2 (1 - \rho^2). \end{aligned}$$

That is, the conditional variance of Y, for each given x, is $\sigma_Y^2(1 - \rho^2)$. These facts about the conditional mean and variance, along with assumption (a), require that the conditional p.d.f. of Y, given $X = x$, be

$$h(y \mid x) = \frac{1}{\sigma_Y\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left[-\frac{[y-\mu_Y-\rho(\sigma_Y/\sigma_X)(x-\mu_X)]^2}{2\sigma_Y^2(1-\rho^2)}\right],$$

$$-\infty < y < \infty, \text{ for every real } x.$$

Before we make any assumptions about the distribution of X, we give an example and figure to illustrate the implications of our current assumptions.

EXAMPLE 5.6-1 Let $\mu_X = 10$, $\sigma_X^2 = 9$, $\mu_Y = 12$, $\sigma_Y^2 = 16$, and $\rho = 0.6$. We have seen that assumptions (a), (b), and (c) imply that the conditional distribution of Y, given $X = x$, is

$$N\left[12 + (0.6)\left(\frac{4}{3}\right)(x-10),\, 16(1-0.6^2)\right].$$

In Figure 5.6-1 the conditional mean line

$$E(Y \mid x) = 12 + (0.6)\left(\frac{4}{3}\right)(x-10) = 0.8x + 4$$

has been graphed. For each of $x = 5, 10$, and 15, the conditional p.d.f. of Y, given $X = x$, is displayed. ◀

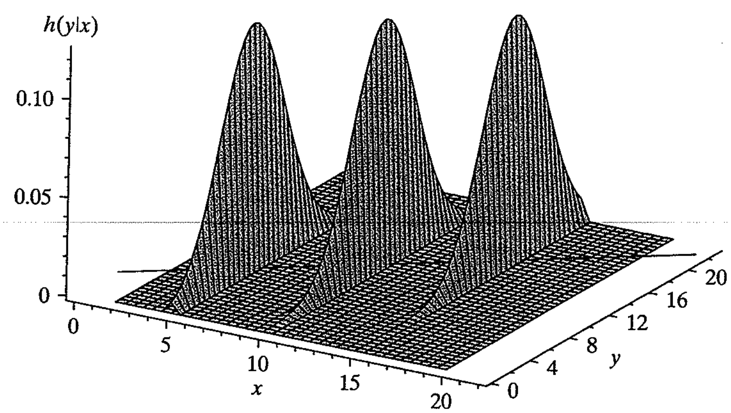

Figure 5.6-1: Conditional p.d.f. of Y, given $x = 5, 10, 15$

Up to this point, nothing has been said about the distribution of X other than that it has mean μ_X and positive variance σ_X^2. Suppose, in addition, we assume that this distribution is also normal; that is, the marginal p.d.f. of X is

$$f_1(x) = \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right], \qquad -\infty < x < \infty.$$

Hence the joint p.d.f. of X and Y is given by the product

JOINT DENSITY

$$f(x, y) = h(y \mid x) f_1(x) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{q(x, y)}{2}\right], \qquad (5.6\text{-}2)$$

where it can be shown (see Exercise 5.6-2) that

$$q(x, y) = \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right].$$

A joint p.d.f. of this form is called a **bivariate normal p.d.f.**

EXAMPLE 5.6-2

Let us assume that in a certain population of college students, the respective grade point averages, say X and Y, in high school and the first year in college have an approximate bivariate normal distribution with parameters $\mu_X = 2.9$, $\mu_Y = 2.4$, $\sigma_X = 0.4$, $\sigma_Y = 0.5$, and $\rho = 0.8$.

Then, for illustration,

$$P(2.1 < Y < 3.3) = P\left(\frac{2.1-2.4}{0.5} < \frac{Y-2.4}{0.5} < \frac{3.3-2.4}{0.5}\right)$$
$$= \Phi(1.8) - \Phi(-0.6) = 0.6898.$$

Since the conditional p.d.f. of Y, given $X = 3.2$, is normal with mean

$$24 + (0.8)\left(\frac{0.5}{0.4}\right)(3.2 - 2.9) = 2.7$$

and standard deviation $(0.5)\sqrt{1-0.64} = 0.3$, we have that

$$P(2.1 < Y < 3.3 \mid X = 3.2)$$
$$= P\left(\frac{2.1-2.7}{0.3} < \frac{Y-2.7}{0.3} < \frac{3.3-2.7}{0.3}\,\middle|\, X = 3.2\right)$$
$$= \Phi(2) - \Phi(-2) = 0.9544.$$

From a practical point of view, however, the reader should be warned that the correlation coefficient of these grade point averages is, in many instances, much smaller than 0.8. ◄

SYMMETRY

Since x and y enter the bivariate normal p.d.f. in a similar manner, the roles of X and Y could have been interchanged. That is, Y could have been assigned the marginal normal p.d.f. $N(\mu_Y, \sigma_Y^2)$, and the conditional p.d.f. of X, given $Y = y$, would have then been normal, with mean $\mu_X + \rho(\sigma_X/\sigma_Y)(y - \mu_Y)$ and variance $\sigma_X^2(1-\rho^2)$. Although this is fairly obvious, we do want to make special note of it.

In order to have a better understanding of the geometry of the bivariate normal distribution, consider the graph of $z = f(x, y)$, where $f(x, y)$ is given by Equation 5.6-2. If we intersect this surface with planes parallel to the yz plane, that is, with $x = x_0$, we have

$$f(x_0, y) = f_1(x_0)h(y \mid x_0).$$

In this equation $f_1(x_0)$ is a constant, and $h(y \mid x_0)$ is a normal p.d.f. Thus $z = f(x_0, y)$ is bell shaped; that is, has the shape of a normal p.d.f. However, note that it is not necessarily a p.d.f. because of the factor $f_1(x_0)$. Similarly, intersections of the surface $z = f(x, y)$ with planes $y = y_0$, parallel to the xz plane will be bell shaped.

If

$$0 < z_0 < \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}},$$

then

$$0 < z_0 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2} < 1.$$

If we intersect $z = f(x, y)$ with the plane $z = z_0$, which is parallel to the xy plane, we have

$$z_0 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2} = \exp\left[\frac{-q(x, y)}{2}\right].$$

Taking the natural logarithm of each side, we obtain

$$\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2$$
$$= -2(1-\rho^2)\ln(z_0 2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}).$$

Thus we see that these intersections are ellipses.

EXAMPLE 5.6-3 With $\mu_X = 10$, $\sigma_X^2 = 9$, $\mu_Y = 12$, $\sigma_Y^2 = 16$, and $\rho = 0.6$, the bivariate normal p.d.f. has been graphed in Figure 5.6-2. For $\rho = 0.6$, level curves or contours are given in Figure 5.6-3. The conditional mean line,

$$E(Y \mid x) = 12 + (0.6)\left(\frac{4}{3}\right)(x - 10) = 0.8x + 4,$$

is also drawn on Figure 5.6-3. Note that this line intersects the level curves at points through which vertical tangents can be drawn to the ellipses. ◀

We close this section by observing another important property of the correlation coefficient ρ if X and Y have a bivariate normal distribution. In Equation 5.6-2 of the product $h(y \mid x)f_1(x)$, let us consider the factor $h(y \mid x)$ if

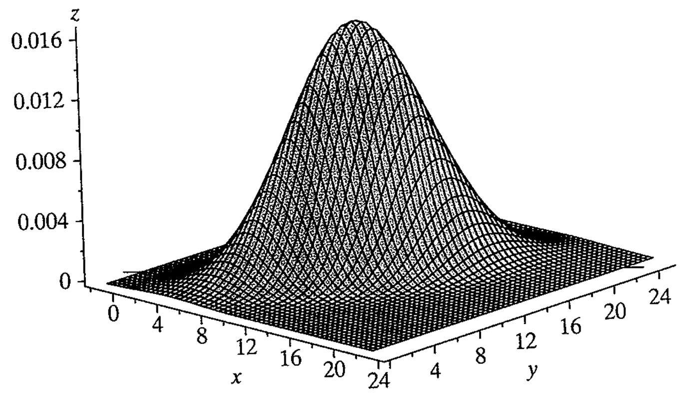

Figure 5.6-2: Bivariate normal p.d.f.

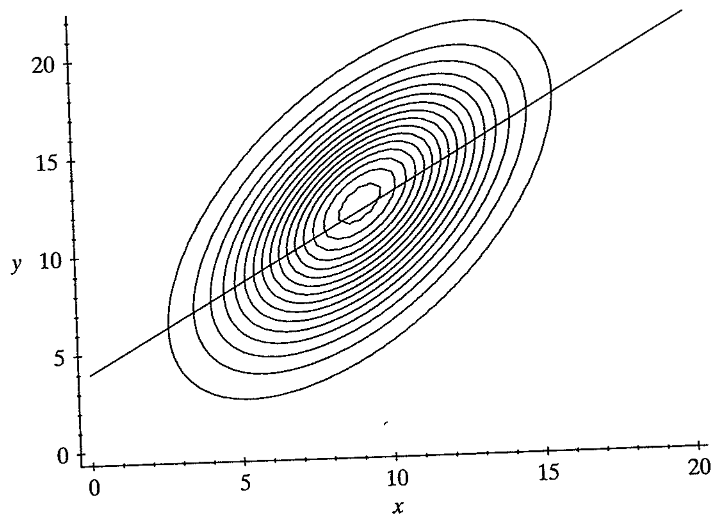

Figure 5.6-3: Contours for the bivariate normal distribution

$\rho = 0$. We see that this product, which is the joint p.d.f. of X and Y, equals $f_1(x)f_2(y)$ because $h(y\,|\,x)$ is, when $\rho = 0$, a normal p.d.f. with mean μ_Y and variance σ_Y^2. That is, if $\rho = 0$, the joint p.d.f. factors into the product of the two marginal probability density functions and hence, X and Y are independent random variables. Of course, if X and Y are any independent random variables (not necessarily normal), we know that ρ, if it exists, is always equal to zero. Thus we have proved the following.

Theorem 5.6-1 If X and Y have a bivariate normal distribution with correlation coefficient ρ, then X and Y are independent if and only if $\rho = 0$.

Thus, in the bivariate normal case, $\rho = 0$ does imply independence of X and Y.

It should be mentioned here than these characteristics of the bivariate normal distribution can be extended to the trivariate normal distribution or, more generally, the multivariate normal distribution. This is done in more advanced texts assuming some knowledge of matrices; for illustration, see Hogg, McKean, and Craig (2005).

EXERCISES

5.6-1 Let X and Y have a bivariate normal distribution with parameters $\mu_X = -3, \mu_Y = 10, \sigma_X^2 = 25, \sigma_Y^2 = 9$, and $\rho = 3/5$. Compute

(a) $P(-5 < X < 5)$.
(b) $P(-5 < X < 5 \mid Y = 13)$.
(c) $P(7 < Y < 16)$.
(d) $P(7 < Y < 16 \mid X = 2)$.

5.6-2 Show that the expression in the exponent of Equation 5.6-2 is equal to the function $q(x, y)$ given in the text.

5.6-3 Let X and Y have a bivariate normal distribution with parameters $\mu_X = 2.8, \mu_Y = 110, \sigma_X^2 = 0.16, \sigma_Y^2 = 100$, and $\rho = 0.6$. Compute

(a) $P(106 < Y < 124)$.
(b) $P(106 < Y < 124 \mid X = 3.2)$.

5.6-4 Let X and Y have a bivariate normal distribution with $\mu_X = 70, \sigma_X^2 = 100, \mu_Y = 80, \sigma_Y^2 = 169$, and $\rho = 5/13$. Find

(a) $E(Y \mid X = 72)$.
(b) $\mathrm{Var}(Y \mid X = 72)$.
(c) $P(Y \le 84, \mid X = 72)$.

5.6-5 Let X denote the height in centimeters and Y the weight in kilograms of male college students. Assume that X and Y have a bivariate normal distribution with parameters $\mu_X = 185, \sigma_X^2 = 100, \mu_Y = 84, \sigma_Y^2 = 64$, and $\rho = 3/5$.

(a) Determine the conditional distribution of Y, given that $X = 190$.
(b) Find $P(86.4 < Y < 95.36 \mid X = 190)$.

5.6-6 For a freshman taking introductory statistics and majoring in psychology, let X equal the student's ACT mathematics score and Y the student's ACT verbal score. Assume that X and Y have a bivariate normal distribution with $\mu_X = 22.7, \sigma_X^2 = 17.64, \mu_Y = 22.7, \sigma_Y^2 = 12.25$, and $\rho = 0.78$. Find

(a) $P(18.5 < Y < 25.5)$.
(b) $E(Y \mid x)$.
(c) $\mathrm{Var}(Y \mid x)$.
(d) $P(18.5 < Y < 25.5 \mid X = 23)$.
(e) $P(18.5 < Y < 25.5 \mid X = 25)$.
(f) For $x = 21, 23$, and 25, draw a graph of $z = h(y \mid x)$ similar to Figure 5.6-1.

5.6-7 For a pair of gallinules, let X equal the weight in grams of the male and Y the weight in grams of the female. Assume that X and Y have a bivariate normal distribution with $\mu_X = 415, \sigma_X^2 = 611, \mu_Y = 347, \sigma_Y^2 = 689$, and $\rho = -0.25$. Find

(a) $P(309.2 < Y < 380.6)$.
(b) $E(Y \mid x)$.
(c) $\mathrm{Var}(Y \mid x)$.
(d) $P(309.2 < Y < 380.6 \mid X = 385.1)$.

5.6-8 Let X and Y have a bivariate normal distribution with parameters $\mu_X = 10, \sigma_X^2 = 9, \mu_Y = 15, \sigma_Y^2 = 16$, and $\rho = 0$. Find