that independence of X and Z implies

$$\operatorname{Cov}[X, Z] = \rho_{X,Z} = 0. \tag{5.63}$$

In addition, by Theorem 5.17(d),

$$\operatorname{Var}[Y] = \operatorname{Var}[X] + \operatorname{Var}[Z] = \sigma_X^2 + \sigma_Z^2. \tag{5.64}$$

Since $\operatorname{E}[X] = \operatorname{E}[Z] = 0$, Theorem 5.11 tells us that $\operatorname{E}[Y] = \operatorname{E}[X] + \operatorname{E}[Z] = 0$ and Theorem 5.17)(b) says that $\operatorname{E}[XZ] = \operatorname{E}[X]\operatorname{E}[Z] = 0$. This permits us to write

$$\begin{aligned}\operatorname{Cov}[X, Y] &= \operatorname{E}[XY] = \operatorname{E}[X(X+Z)] \\ &= \operatorname{E}\left[X^2 + XZ\right] = \operatorname{E}\left[X^2\right] + \operatorname{E}[XZ] = \operatorname{E}\left[X^2\right] = \sigma_X^2.\end{aligned}$$

This implies

$$\rho_{X,Y} = \frac{\operatorname{Cov}[X, Y]}{\sqrt{\operatorname{Var}[X]\operatorname{Var}[Y]}} = \frac{\sigma_X^2}{\sqrt{\sigma_X^2(\sigma_X^2 + \sigma_Z^2)}} = \sqrt{\frac{\sigma_X^2/\sigma_Z^2}{1 + \sigma_X^2/\sigma_Z^2}}. \tag{5.65}$$

We see in Example 5.18 that the covariance between the transmitted signal X and the received signal Y depends on the ratio σ_X^2/σ_Z^2. This ratio, referred to as the *signal-to-noise ratio*, has a strong effect on communication quality. If $\sigma_X^2/\sigma_Z^2 \ll 1$, the correlation of X and Y is weak and the noise dominates the signal at the receiver. Learning y, a sample of the received signal, is not very helpful in determining the corresponding sample of the transmitted signal, x. On the other hand, if $\sigma_X^2/\sigma_Z^2 \gg 1$, the transmitted signal dominates the noise and $\rho_{X,Y} \approx 1$, an indication of a close relationship between X and Y. When there is strong correlation between X and Y, learning y is very helpful in determining x.

Quiz 5.8

(A) Random variables L and T have joint PMF

$P_{L,T}(l,t)$	$t = 40$sec	$t = 60$sec
$l = 1$ page	0.15	0.1
$l = 2$ pages	0.30	0.2
$l = 3$ pages	0.15	0.1.

Find the following quantities.

(a) $\operatorname{E}[L]$ and $\operatorname{Var}[L]$

(b) $\operatorname{E}[T]$ and $\operatorname{Var}[T]$

(c) The covariance $\operatorname{Cov}[L, T]$

(d) The correlation coefficient $\rho_{L,T}$

(B) The joint probability density function of random variables X and Y is

$$f_{X,Y}(x, y) = \begin{cases} xy & 0 \le x \le 1, 0 \le y \le 2, \\ 0 & \text{otherwise.} \end{cases} \tag{5.66}$$

Find the following quantities.

(a) $\mathrm{E}[X]$ and $\mathrm{Var}[X]$

(b) $\mathrm{E}[Y]$ and $\mathrm{Var}[Y]$

(c) The covariance $\mathrm{Cov}[X, Y]$

(d) The correlation coefficient $\rho_{X,Y}$

5.9 Bivariate Gaussian Random Variables

The *bivariate Gaussian* PDF of X and Y has five parameters: the expected values and standard deviations of X and Y and the correlation coefficient of X and Y. The marginal PDF of X and the marginal PDF of Y are both Gaussian.

For a PDF representing a family of random variables, one or more parameters define a specific PDF. Properties such as $\mathrm{E}[X]$ and $\mathrm{Var}[X]$ depend on the parameters. For example, a continuous uniform (a, b) random variable has expected value $(a+b)/2$ and variance $(b-a)^2/12$. For the bivariate Gaussian PDF, the parameters μ_X, μ_Y, σ_X, σ_Y and $\rho_{X,Y}$ are equal to the expected values, standard deviations, and correlation coefficient of X and Y.

Definition 5.10 Bivariate Gaussian Random Variables

Random variables X and Y have a ***bivariate Gaussian PDF*** *with parameters μ_X, μ_Y, $\sigma_X > 0$, $\sigma_Y > 0$, and $\rho_{X,Y}$ satisfying $-1 < \rho_{X,Y} < 1$ if*

$$f_{X,Y}(x, y) = \frac{\exp\left[-\dfrac{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{2\rho_{X,Y}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}{2(1-\rho_{X,Y}^2)}\right]}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}},$$

Figure 5.6 illustrates the bivariate Gaussian PDF for $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 1$, and three values of $\rho_{X,Y} = \rho$. When $\rho = 0$, the joint PDF has the circular symmetry of a sombrero. When $\rho = 0.9$, the joint PDF forms a ridge over the line $x = y$, and when $\rho = -0.9$ there is a ridge over the line $x = -y$. The ridge becomes increasingly steep as $\rho \to \pm 1$. Adjacent to each PDF, we repeat the graphs in Figure 5.5; each graph shows 200 sample pairs (X, Y) drawn from that bivariate Gaussian PDF. We see that the sample pairs are clustered in the region of the x, y plane where the PDF is large.

To examine mathematically the properties of the bivariate Gaussian PDF, we define

$$\tilde{\mu}_Y(x) = \mu_Y + \rho_{X,Y}\frac{\sigma_Y}{\sigma_X}(x - \mu_X), \qquad \tilde{\sigma}_Y = \sigma_Y\sqrt{1-\rho_{X,Y}^2}, \tag{5.67}$$

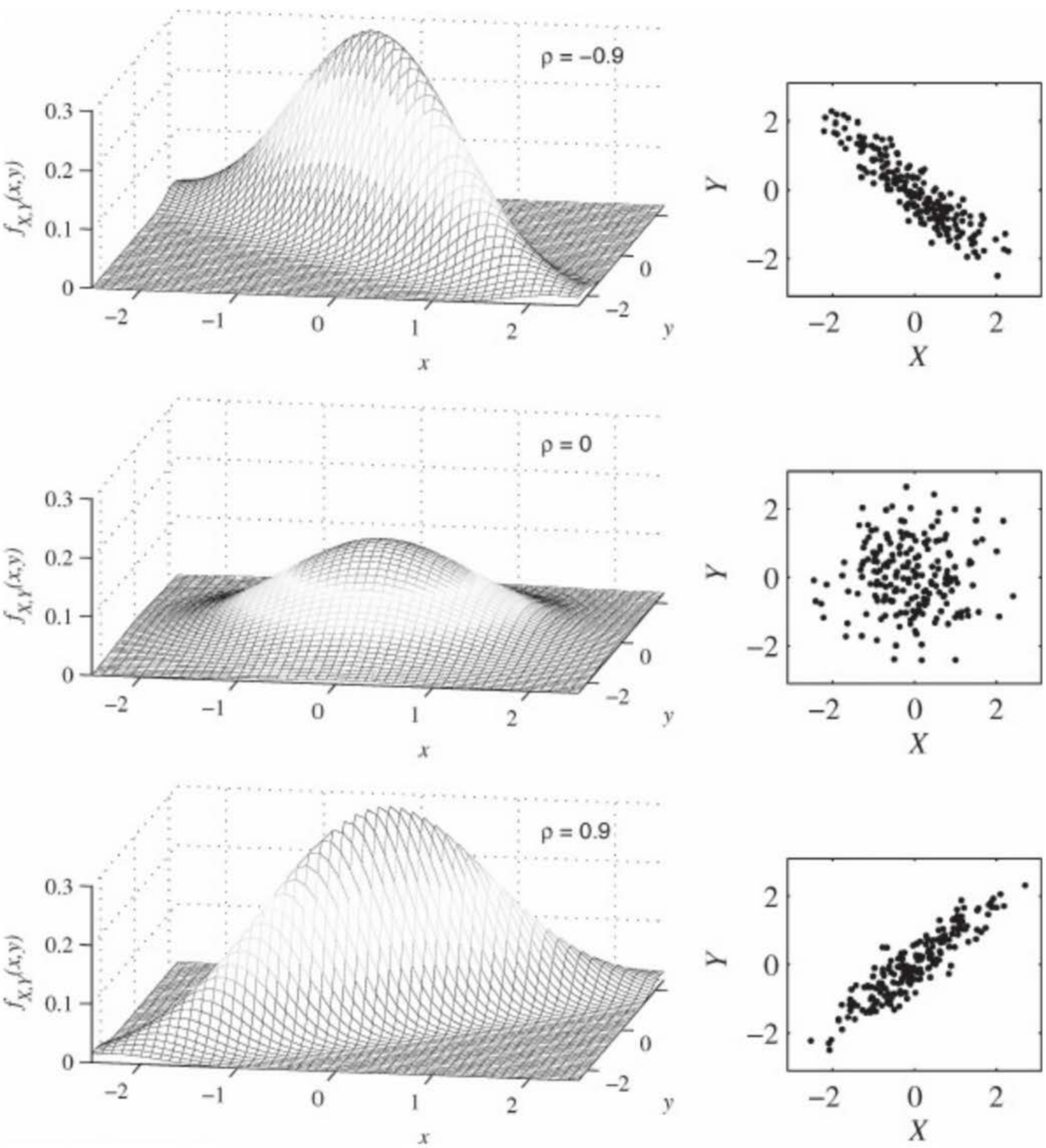

Figure 5.6 The Joint Gaussian PDF $f_{X,Y}(x,y)$ for $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 1$, and three values of $\rho_{X,Y} = \rho$. Next to each PDF, we plot 200 sample pairs (X, Y) generated with that PDF.

and manipulate the formula in Definition 5.10 to obtain the following expression for the joint Gaussian PDF:

$$f_{X,Y}(x,y) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(x-\mu_X)^2/2\sigma_X^2} \frac{1}{\tilde{\sigma}_Y\sqrt{2\pi}} e^{-(y-\tilde{\mu}_Y(x))^2/2\tilde{\sigma}_Y^2}. \tag{5.68}$$

Equation (5.68) expresses $f_{X,Y}(x,y)$ as the product of two Gaussian PDFs, one with parameters μ_X and σ_X and the other with parameters $\tilde{\mu}_Y$ and $\tilde{\sigma}_Y$. This formula plays a key role in the proof of the following theorem.

Theorem 5.18

If X and Y are the bivariate Gaussian random variables in Definition 5.10, X is the Gaussian (μ_X, σ_X) random variable and Y is the Gaussian (μ_Y, σ_Y) random variable:

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(x-\mu_X)^2/2\sigma_X^2}, \qquad f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-(y-\mu_Y)^2/2\sigma_Y^2}.$$

Proof Integrating $f_{X,Y}(x,y)$ in Equation (5.68) over all y, we have

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy \\ &= \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(x-\mu_X)^2/2\sigma_X^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\tilde{\sigma}_Y\sqrt{2\pi}} e^{-(y-\tilde{\mu}_Y(x))^2/2\tilde{\sigma}_Y^2}\,dy}_{1} \end{aligned} \tag{5.69}$$

The integral above the bracket equals 1 because it is the integral of a Gaussian PDF. The remainder of the formula is the PDF of the Gaussian (μ_X, σ_X) random variable. The same reasoning with the roles of X and Y reversed leads to the formula for $f_Y(y)$.

The next theorem identifies $\rho_{X,Y}$ in Definition 5.10 as the correlation coefficient of X and Y.

Theorem 5.19

Bivariate Gaussian random variables X and Y in Definition 5.10 have correlation coefficient $\rho_{X,Y}$.

The proof of Theorem 5.19 involves algebra that is more easily digested with some insight from Chapter 7; see Section 7.6 for the proof.

From Theorem 5.19, we observe that if X and Y are uncorrelated, then $\rho_{X,Y} = 0$ and, by evaluating the PDF in Definition 5.10 with $\rho_{X,Y} = 0$, we have $f_{X,Y}(x,y) = f_X(x)f_Y(y)$. Thus we have the following theorem.

Theorem 5.20

Bivariate Gaussian random variables X and Y are uncorrelated if and only if they are independent.

Another important property of bivariate Gaussian random variables X and Y is that a pair of linear combinations of X and Y forms a pair of bivariate Gaussian random variables.

Theorem 5.21

If X and Y are bivariate Gaussian random variables with PDF given by Definition 5.10, and W_1 and W_2 are given by the linearly independent equations

$$W_1 = a_1X + b_1Y, \qquad W_2 = a_2X + b_2Y,$$

then W_1 and W_2 are bivariate Gaussian random variables such that

$$\begin{aligned} \mathrm{E}[W_i] &= a_i\mu_X + b_i\mu_Y, & i = 1,2, \\ \mathrm{Var}[W_i] &= a_i^2\sigma_X^2 + b_i^2\sigma_Y^2 + 2a_ib_i\rho_{X,Y}\sigma_X\sigma_Y, & i = 1,2, \\ \mathrm{Cov}[W_1,W_2] &= a_1a_2\sigma_X^2 + b_1b_2\sigma_Y^2 + (a_1b_2 + a_2b_1)\rho_{X,Y}\sigma_X\sigma_Y. \end{aligned}$$

Theorem 5.21 is a special case of Theorem 8.11 when we have $n = 2$ jointly Gaussian random variables. We omit the proof since the proof of Theorem 8.11 for n jointly Gaussian random variables is, with some knowledge of linear algebra, simpler. The requirement that the equations for W_1 and W_2 be "linearly independent" is linear algebra terminology that excludes degenerate cases such as $W_1 = X + 2Y$ and $W_2 = 3X + 6Y$ where $W_2 = 3W_1$ is just a scaled replica of W_1.

Theorem 5.21 is powerful. Even the partial result that W_i by itself is Gaussian is a nontrivial conclusion. When an experiment produces linear combinations of Gaussian random variables, knowing that these combinations are Gaussian simplifies the analysis because all we need to do is calculate the expected values, variances, and covariances of the outputs in order to derive probability models.

Example 5.19

For the noisy observation in Example 5.14, find the PDF of $Y = X + Z$.

Since X is Gaussian $(0, \sigma_X)$ and Z is Gaussian $(0, \sigma_Z)$ and X and Z are independent, X and Z are jointly Gaussian. It follows from Theorem 5.21 that Y is Gaussian with $\mathrm{E}[Y] = \mathrm{E}[X] + \mathrm{E}[Z] = 0$ and variance $\sigma_Y^2 = \sigma_X^2 + \sigma_Z^2$. The PDF of Y is

$$f_Y(y) = \frac{1}{\sqrt{2\pi(\sigma_X^2 + \sigma_Z^2)}} e^{-y^2/2(\sigma_X^2+\sigma_Z^2)}. \tag{5.70}$$

Example 5.20

Continuing Example 5.19, find the joint PDF of X and Y when $\sigma_X = 4$ and $\sigma_Z = 3$.

From Theorem 5.21, we know that X and Y are bivariate Gaussian. We also know that $\mu_X = \mu_Y = 0$ and that Y has variance $\sigma_Y^2 = \sigma_X^2 + \sigma_Z^2 = 25$. Substituting $\sigma_X = 4$ and $\sigma_Z = 3$ in the formula for the correlation coefficient derived in Example 5.18, we have

$$\rho_{X,Y} = \sqrt{\frac{\sigma_X^2/\sigma_Z^2}{1 + \sigma_X^2/\sigma_Z^2}} = \frac{4}{5}. \tag{5.71}$$

Applying these parameters to Definition 5.10, we obtain

$$f_{X,Y}(x,y) = \frac{1}{24\pi} e^{-(25x^2/16 - 2xy + y^2)/18}. \tag{5.72}$$

Quiz 5.9

Let X and Y be jointly Gaussian $(0, 1)$ random variables with correlation coefficient $1/2$. What is the joint PDF of X and Y?

5.10 Multivariate Probability Models

The probability model of an experiment that produces n random variables can be represented as an n-dimensional CDF. If all of the random variables are discrete, there is a corresponding n-dimensional PMF. If all of the random variables are continuous, there is an n-dimensional PDF. The PDF is the nth partial derivative of the CDF with respect to all n variables. The probability model (CDF, PMF, or PDF) of n independent random variables is the product of the univariate probability models of the n random variables.

This chapter has emphasized probability models of two random variables X and Y. We now generalize the definitions and theorems to experiments that yield an arbitrary number of random variables $X_1, \ldots, X_n$. This section is heavy on n-dimensional definitions and theorems but relatively light on examples. However, the ideas are straightforward extensions of concepts for a pair of random variables. If you have trouble with a theorem or definition, rewrite it for the special case of $n = 2$ random variables. This will yield a familiar result for a pair of random variables.

To express a complete probability model of $X_1, \ldots, X_n$, we define the joint cumulative distribution function.

Definition 5.11 Multivariate Joint CDF

*The **joint CDF** of $X_1, \ldots, X_n$ is*

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \mathrm{P}\left[X_1 \le x_1, \ldots, X_n \le x_n\right].$$

Definition 5.11 is concise and general. It provides a complete probability model regardless of whether any or all of the X_i are discrete, continuous, or mixed. However, the joint CDF is usually not convenient to use in analyzing practical probability models. Instead, we use the joint PMF or the joint PDF.

Definition 5.12 Multivariate Joint PMF

*The **joint PMF** of the discrete random variables $X_1, \ldots, X_n$ is*

$$P_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \mathrm{P}\left[X_1 = x_1, \ldots, X_n = x_n\right].$$

Definition 5.13 Multivariate Joint PDF

*The **joint PDF** of the continuous random variables $X_1, \ldots, X_n$ is the function*

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \frac{\partial^n F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{\partial x_1 \cdots \partial x_n}.$$

Theorems 5.22 and 5.23 indicate that the joint PMF and the joint PDF have properties that are generalizations of the axioms of probability.

Theorem 5.22

If $X_1, \ldots, X_n$ are discrete random variables with joint PMF $P_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$,

(a) $P_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \geq 0$,

(b) $\displaystyle\sum_{x_1 \in S_{X_1}} \cdots \sum_{x_n \in S_{X_n}} P_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = 1.$

Theorem 5.23

If $X_1, \ldots, X_n$ are continuous random variables with joint PDF $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$,

(a) $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \geq 0$,

(b) $\displaystyle F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1,\ldots,X_n}(u_1,\ldots,u_n)\, du_1 \cdots du_n,$

(c) $\displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\, dx_1 \cdots dx_n = 1.$

Often we consider an event A described in terms of a property of $X_1, \ldots, X_n$, such as $|X_1 + X_2 + \cdots + X_n| \leq 1$, or $\max_i X_i \leq 100$. To find the probability of the event A, we sum the joint PMF or integrate the joint PDF over all $x_1, \ldots, x_n$ that belong to A.

Theorem 5.24

The probability of an event A expressed in terms of the random variables $X_1, \ldots, X_n$ is

$$\textit{Discrete:} \quad \mathrm{P}[A] = \sum_{(x_1,\ldots,x_n) \in A} P_{X_1,\ldots,X_n}(x_1,\ldots,x_n);$$

$$\textit{Continuous:} \quad \mathrm{P}[A] = \int \cdots \int_A f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\, dx_1\, dx_2 \ldots dx_n.$$

Although we have written the discrete version of Theorem 5.24 with a single summation, we must remember that in fact it is a multiple sum over the n variables $x_1, \ldots, x_n$.

x (1 Page)	y (2 Pages)	z (3 Pages)	$P_{X,Y,Z}(x,y,z)$	Total Pages	Events
0	0	4	1/1296	12	B
0	1	3	1/108	11	B
0	2	2	1/24	10	B
0	3	1	1/12	9	B
0	4	0	1/16	8	AB
1	0	3	1/162	10	B
1	1	2	1/18	9	B
1	2	1	1/6	8	AB
1	3	0	1/6	7	B
2	0	2	1/54	8	AB
2	1	1	1/9	7	B
2	2	0	1/6	6	B
3	0	1	2/81	6	
3	1	0	2/27	5	
4	0	0	1/81	4	

Table 5.1 The PMF $P_{X,Y,Z}(x,y,z)$ and the events A and B for Example 5.22.

Example 5.21

Consider a set of n independent trials in which there are r possible outcomes $s_1, \ldots, s_r$ for each trial. In each trial, $P[s_i] = p_i$. Let N_i equal the number of times that outcome s_i occurs over n trials. What is the joint PMF of $N_1, \ldots, N_r$?

The solution to this problem appears in Theorem 2.9 and is repeated here:

$$P_{N_1,\ldots,N_r}(n_1,\ldots,n_r) = \binom{n}{n_1,\ldots,n_r} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}. \tag{5.73}$$

Example 5.22

For each product that a company sells, a company website has a tech support document available for download. The PMF of L, the number of pages in one document, is shown in the table on the right. For a set of four independent information requests, find:

l	1	2	3
$P_L(l)$	1/3	1/2	1/6

(a) the joint PMF of the random variables, X, Y,and Z, the number of 1-page, 2-page, and 3-page downloads, respectively,

(b) $P[A] = P[$total length of four downloads is 8 pages$]$,

(c) $P[B] = P[$at least half of the four downloads have more than 1 page$]$.

The downloads are independent trials, each with three possible outcomes: $L = 1$, $L = 2$, and $L = 3$. Hence, the probability model of the number of downloads of each length in the set of four downloads is the multinomial PMF of Example 5.21:

$$P_{X,Y,Z}(x,y,z) = \binom{4}{x,y,z}\left(\frac{1}{3}\right)^x\left(\frac{1}{2}\right)^y\left(\frac{1}{6}\right)^z. \tag{5.74}$$

The PMF is displayed numerically in Table 5.1. The final column of the table indicates that there are three outcomes in event A and 12 outcomes in event B. Adding the probabilities in the two events, we have $P[A] = 107/432$ and $P[B] = 8/9$.

In analyzing an experiment, we might wish to study some of the random variables and ignore other ones. To accomplish this, we can derive marginal PMFs or marginal PDFs that are probability models for a fraction of the random variables in the complete experiment. Consider an experiment with four random variables W, X, Y, Z. The probability model for the experiment is the joint PMF, $P_{W,X,Y,Z}(w,x,y,z)$ or the joint PDF, $f_{W,X,Y,Z}(w,x,y,z)$. The following theorems give examples of marginal PMFs and PDFs.

Theorem 5.25

For a joint PMF $P_{W,X,Y,Z}(w,x,y,z)$ of discrete random variables W, X, Y, Z, some marginal PMFs are

$$P_{X,Y,Z}(x,y,z) = \sum_{w\in S_W} P_{W,X,Y,Z}(w,x,y,z),$$

$$P_{W,Z}(w,z) = \sum_{x\in S_X}\sum_{y\in S_Y} P_{W,X,Y,Z}(w,x,y,z),$$

Theorem 5.26

For a joint PDF $f_{W,X,Y,Z}(w,x,y,z)$ of continuous random variables W, X, Y, Z, some marginal PDFs are

$$f_{W,X,Y}(w,x,y) = \int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)\,dz,$$

$$f_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)\,dw\,dy\,dz.$$

Theorems 5.25 and 5.26 can be generalized in a straightforward way to any marginal PMF or marginal PDF of an arbitrary number of random variables. For a probability model described by the set of random variables $\{X_1, \ldots, X_n\}$, each nonempty strict subset of those random variables has a marginal probability model. There are 2^n subsets of $\{X_1, \ldots, X_n\}$. After excluding the entire set and the null set $\varnothing$, we find that there are $2^n - 2$ marginal probability models.

Example 5.23

As in Quiz 5.10, the random variables $Y_1, \ldots, Y_4$ have the joint PDF

$$f_{Y_1,\ldots,Y_4}(y_1, \ldots, y_4) = \begin{cases} 4 & 0 \le y_1 \le y_2 \le 1, 0 \le y_3 \le y_4 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.75}$$

Find the marginal PDFs $f_{Y_1,Y_4}(y_1, y_4)$, $f_{Y_2,Y_3}(y_2, y_3)$, and $f_{Y_3}(y_3)$.

$$f_{Y_1,Y_4}(y_1, y_4) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1,\ldots,Y_4}(y_1, \ldots, y_4)\, dy_2\, dy_3. \tag{5.76}$$

In the foregoing integral, the hard part is identifying the correct limits. These limits will depend on y_1 and y_4. For $0 \le y_1 \le 1$ and $0 \le y_4 \le 1$,

$$f_{Y_1,Y_4}(y_1, y_4) = \int_{y_1}^{1} \int_{0}^{y_4} 4\, dy_3\, dy_2 = 4(1 - y_1)y_4. \tag{5.77}$$

The complete expression for $f_{Y_1,Y_4}(y_1, y_4)$ is

$$f_{Y_1,Y_4}(y_1, y_4) = \begin{cases} 4(1 - y_1)y_4 & 0 \le y_1 \le 1, 0 \le y_4 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.78}$$

Similarly, for $0 \le y_2 \le 1$ and $0 \le y_3 \le 1$,

$$f_{Y_2,Y_3}(y_2, y_3) = \int_{0}^{y_2} \int_{y_3}^{1} 4\, dy_4\, dy_1 = 4y_2(1 - y_3). \tag{5.79}$$

The complete expression for $f_{Y_2,Y_3}(y_2, y_3)$ is

$$f_{Y_2,Y_3}(y_2, y_3) = \begin{cases} 4y_2(1 - y_3) & 0 \le y_2 \le 1, 0 \le y_3 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.80}$$

Lastly, for $0 \le y_3 \le 1$,

$$f_{Y_3}(y_3) = \int_{-\infty}^{\infty} f_{Y_2,Y_3}(y_2, y_3)\, dy_2 = \int_{0}^{1} 4y_2(1 - y_3)\, dy_2 = 2(1 - y_3). \tag{5.81}$$

The complete expression is

$$f_{Y_3}(y_3) = \begin{cases} 2(1 - y_3) & 0 \le y_3 \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.82}$$

Example 5.22 demonstrates that a fairly simple experiment can generate a joint PMF that, in table form, is perhaps surprisingly long. In fact, a practical experiment often generates a joint PMF or PDF that is forbiddingly complex. The important exception is an experiment that produces n independent random variables. The following definition extends the definition of independence of two random variables. It states that $X_1, \ldots, X_n$ are independent when the joint PMF or PDF can be factored into a product of n marginal PMFs or PDFs.

Definition 5.14 N Independent Random Variables

Random variables $X_1, \ldots, X_n$ *are* ***independent*** *if for all* $x_1, \ldots, x_n$,

$$\text{Discrete: } P_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = P_{X_1}(x_1)\, P_{X_2}(x_2) \cdots P_{X_N}(x_n);$$

$$\text{Continuous: } f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_{X_1}(x_1)\, f_{X_2}(x_2) \cdots f_{X_n}(x_n).$$

Independence of n random variables is typically a property of an experiment consisting of n independent subexperiments, in which subexperiment i produces the random variable X_i. If all subexperiments follow the same procedure and have the same observation, all of the X_i have the same PMF or PDF. In this case, we say the random variables X_i are *identically distributed.*

Definition 5.15 Independent and Identically Distributed (iid)

$X_1, \ldots, X_n$ *are* ***independent and identically distributed (iid)*** *if*

$$\text{Discrete: } P_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = P_X(x_1)\, P_X(x_2) \cdots P_X(x_n);$$

$$\text{Continuous: } f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_X(x_1)\, f_X(x_2) \cdots f_X(x_n).$$

Example 5.24

The random variables $X_1, \ldots, X_n$ have the joint PDF

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \begin{cases} 1 & 0 \le x_i \le 1, i = 1,\ldots,n, \\ 0 & \text{otherwise.} \end{cases} \tag{5.83}$$

Let A denote the event that $\max_i X_i \le 1/2$. Find $\mathrm{P}[A]$.

We can solve this problem by applying Theorem 5.24:

$$\begin{aligned} \mathrm{P}[A] = \mathrm{P}\left[\max_i X_i \le 1/2\right] &= \mathrm{P}[X_1 \le 1/2, \ldots, X_n \le 1/2] \\ &= \int_0^{1/2} \cdots \int_0^{1/2} 1\, dx_1 \cdots dx_n = \frac{1}{2^n}. \end{aligned} \tag{5.84}$$

As n grows, the probability that the maximum is less than $1/2$ rapidly goes to 0.

We note that inspection of the joint PDF reveals that $X_1, \ldots, X_4$ are iid continuous uniform $(0, 1)$ random variables. The integration in Equation (5.84) is easy because independence implies

$$\begin{aligned} \mathrm{P}[A] &= \mathrm{P}[X_1 \le 1/2, \ldots, X_n \le 1/2] \\ &= \mathrm{P}[X_1 \le 1/2] \times \cdots \times \mathrm{P}[X_n \le 1/2] = (1/2)^n. \end{aligned} \tag{5.85}$$

Quiz 5.10

The random variables $Y_1, \ldots, Y_4$ have the joint PDF

$$f_{Y_1,\ldots,Y_4}(y_1, \ldots, y_4) = \begin{cases} 4 & 0 \leq y_1 \leq y_2 \leq 1, 0 \leq y_3 \leq y_4 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.86}$$

Let C denote the event that $\max_i Y_i \leq 1/2$. Find $\mathrm{P}[C]$.

5.11 MATLAB

It is convenient to use MATLAB to generate pairs of discrete random variables X and Y with an arbitrary joint PMF. There are no generally applicable techniques for generating sample pairs of a continuous random variable. There are techniques tailored to specific joint PDFs, for example, bivariate Gaussian.

MATLAB is a useful tool for studying experiments that produce a pair of random variables X, Y. Simulation experiments often depend on the generation of sample pairs of random variables with specific probability models. That is, given a joint PMF $P_{X,Y}(x, y)$ or PDF $f_{X,Y}(x, y)$, we need to produce a collection of pairs $\{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$. For finite discrete random variables, we are able to develop some general techniques. For continuous random variables, we give some specific examples.

Discrete Random Variables

We start with the case when X and Y are finite random variables with ranges

$$S_X = \{x_1, \ldots, x_n\}, \qquad S_Y = \{y_1, \ldots, y_m\}. \tag{5.87}$$

In this case, we can take advantage of MATLAB techniques for surface plots of $g(x, y)$ over the x, y plane. In MATLAB, we represent S_X and S_Y by the n element vector `sx` and m element vector `sy`. The function `[SX,SY]=ndgrid(sx,sy)` produces the pair of $n \times m$ matrices,

$$\texttt{SX} = \begin{bmatrix} x_1 & \cdots & x_1 \\ \vdots & & \vdots \\ x_n & \cdots & x_n \end{bmatrix}, \qquad \texttt{SY} = \begin{bmatrix} y_1 & \cdots & y_m \\ \vdots & & \vdots \\ y_1 & \cdots & y_m \end{bmatrix}. \tag{5.88}$$

We refer to matrices `SX` and `SY` as a *sample space grid* because they are a grid representation of the joint sample space

$$S_{X,Y} = \{(x, y)|x \in S_X, y \in S_Y\}. \tag{5.89}$$