Most important to statisticians are intersections that occur in the course of sampling. Suppose that $Y_1, Y_2, \ldots, Y_n$ denote the outcomes of $n$ successive trials of an experiment. For example, this sequence could represent the weights of $n$ people or the measurements of $n$ physical characteristics for a single person. A specific set of outcomes, or sample measurements, may be expressed in terms of the intersection of the $n$ events $(Y_1 = y_1), (Y_2 = y_2), \ldots, (Y_n = y_n)$, which we will denote as $(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n)$, or, more compactly, as $(y_1, y_2, \ldots, y_n)$. Calculation of the probability of this intersection is essential in making inferences about the population from which the sample was drawn and is a major reason for studying multivariate probability distributions.

## 5.2 Bivariate and Multivariate Probability Distributions

Many random variables can be defined over the same sample space. For example, consider the experiment of tossing a pair of dice. The sample space contains 36 sample points, corresponding to the $mn = (6)(6) = 36$ ways in which numbers may appear on the faces of the dice. Any one of the following random variables could be defined over the sample space and might be of interest to the experimenter:

$Y_1$: The number of dots appearing on die 1.
$Y_2$: The number of dots appearing on die 2.
$Y_3$: The sum of the number of dots on the dice.
$Y_4$: The product of the number of dots appearing on the dice.

The 36 sample points associated with the experiment are equiprobable and correspond to the 36 numerical events $(y_1, y_2)$. Thus, throwing a pair of 1s is the simple event (1, 1). Throwing a 2 on die 1 and a 3 on die 2 is the simple event (2, 3). Because all pairs $(y_1, y_2)$ occur with the same relative frequency, we assign probability $1/36$ to each sample point. For this simple example, the intersection $(y_1, y_2)$ contains at most one sample point. Hence, the bivariate probability function is

$$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2) = 1/36, \qquad y_1 = 1, 2, \ldots, 6, y_2 = 1, 2, \ldots, 6.$$

A graph of the bivariate probability function for the die-tossing experiment is shown in Figure 5.1. Notice that a nonzero probability is assigned to a point $(y_1, y_2)$ in the plane if and only if $y_1 = 1, 2, \ldots, 6$ and $y_2 = 1, 2, \ldots, 6$. Thus, exactly 36 points in the plane are assigned nonzero probabilities. Further, the probabilities are assigned in such a way that the sum of the nonzero probabilities is equal to 1. In Figure 5.1 the points assigned nonzero probabilities are represented in the $(y_1, y_2)$ plane, whereas the probabilities associated with these points are given by the lengths of the lines above them. Figure 5.1 may be viewed as a theoretical, three-dimensional relative frequency histogram for the pairs of observations $(y_1, y_2)$. As in the single-variable discrete case, the theoretical histogram provides a model for the sample histogram that would be obtained if the die-tossing experiment were repeated a large number of times.

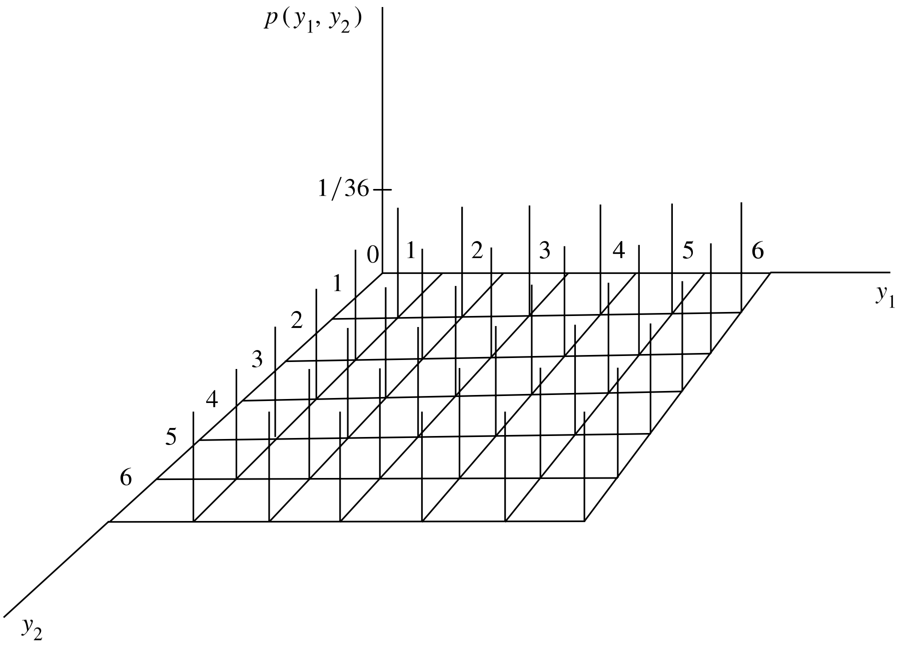


FIGURE 5.1 Bivariate probability function; $y_1$ = number of dots on die 1, $y_2$ = number of dots on die 2

**DEFINITION 5.1**

Let $Y_1$ and $Y_2$ be discrete random variables. The *joint* (or bivariate) *probability function* for $Y_1$ and $Y_2$ is given by

$$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2), \qquad -\infty < y_1 < \infty, -\infty < y_2 < \infty.$$

In the single-variable case discussed in Chapter 3, we saw that the probability function for a discrete random variable $Y$ assigns nonzero probabilities to a finite or countable number of distinct values of $Y$ in such a way that the sum of the probabilities is equal to 1. Similarly, in the bivariate case the joint probability function $p(y_1, y_2)$ assigns nonzero probabilities to only a finite or countable number of pairs of values $(y_1, y_2)$. Further, the nonzero probabilities must sum to 1.

**THEOREM 5.1**

If $Y_1$ and $Y_2$ are discrete random variables with joint probability function $p(y_1, y_2)$, then

1. $p(y_1, y_2) \geq 0$ for all $y_1, y_2$.
2. $\sum_{y_1, y_2} p(y_1, y_2) = 1$, where the sum is over all values $(y_1, y_2)$ that are assigned nonzero probabilities.

As in the univariate discrete case, the joint probability function for discrete random variables is sometimes called the *joint probability mass function* because it specifies the probability (mass) associated with each of the possible pairs of values for the random variables. Once the joint probability function has been determined for discrete random variables $Y_1$ and $Y_2$, calculating joint probabilities involving $Y_1$ and $Y_2$ is straightforward. For the die-tossing experiment, $P(2 \leq Y_1 \leq 3, 1 \leq Y_2 \leq 2)$ is

$$\begin{aligned} P(2 \leq Y_1 \leq 3, 1 \leq Y_2 \leq 2) &= p(2, 1) + p(2, 2) + p(3, 1) + p(3, 2) \\ &= 4/36 = 1/9. \end{aligned}$$

---

EXAMPLE **5.1** A local supermarket has three checkout counters. Two customers arrive at the counters at different times when the counters are serving no other customers. Each customer chooses a counter at random, independently of the other. Let $Y_1$ denote the number of customers who choose counter 1 and $Y_2$, the number who select counter 2. Find the joint probability function of $Y_1$ and $Y_2$.

**Solution** We might proceed with the derivation in many ways. The most direct is to consider the sample space associated with the experiment. Let the pair $\{i, j\}$ denote the simple event that the first customer chose counter $i$ and the second customer chose counter $j$, where $i, j = 1, 2$, and 3. Using the $mn$ rule, the sample space consists of $3 \times 3 = 9$ sample points. Under the assumptions given earlier, each sample point is equally likely and has probability $1/9$. The sample space associated with the experiment is

$$S = [\{1,\ 1\},\ \{1,\ 2\},\ \{1,\ 3\},\ \{2,\ 1\},\ \{2,\ 2\},\ \{2,\ 3\},\ \{3,\ 1\},\ \{3,\ 2\},\ \{3,\ 3\}].$$

Notice that sample point $\{1,\ 1\}$ is the only sample point corresponding to $(Y_1 = 2, Y_2 = 0)$ and hence $P(Y_1 = 2, Y_2 = 0) = 1/9$. Similarly, $P(Y_1 = 1, Y_2 = 1) = P(\{1, 2\} \text{ or } \{2, 1\}) = 2/9$. Table 5.1 contains the probabilities associated with each possible pair of values for $Y_1$ and $Y_2$—that is, the joint probability function for $Y_1$ and $Y_2$. As always, the results of Theorem 5.1 hold for this example.

**Table 5.1 Probability function for $Y_1$ and $Y_2$, Example 5.1**

| | $y_1$ | | |
|---|---|---|---|
| $y_2$ | 0 | 1 | 2 |
| 0 | 1/9 | 2/9 | 1/9 |
| 1 | 2/9 | 2/9 | 0 |
| 2 | 1/9 | 0 | 0 |

■

---

As in the case of univariate random variables, the distinction between jointly discrete and jointly continuous random variables may be characterized in terms of their (joint) distribution functions.

DEFINITION **5.2** For any random variables $Y_1$ and $Y_2$, the joint (bivariate) distribution function $F(y_1, y_2)$ is

$$F(y_1, y_2) = P(Y_1 \leq y_1, Y_2 \leq y_2), \qquad -\infty < y_1 < \infty, -\infty < y_2 < \infty.$$