

## Joint distributions and independence

Random variables related to the same experiment often influence one another. In order to capture this, we introduce the *joint distribution* of two or more random variables. We also discuss the notion of *independence* for random variables, which models the situation where random variables do not influence each other. As with single random variables we treat these topics for discrete and continuous random variables separately.

### 9.1 Joint distributions of discrete random variables

In a census one is usually interested in several variables, such as income, age, and gender. In itself these variables are interesting, but when two (or more) are studied simultaneously, detailed information is obtained on the society where the census is performed. For instance, studying income, age, and gender jointly might give insight to the emancipation of women.

Without mentioning it explicitly, we already encountered several examples of joint distributions of discrete random variables. For example, in Chapter 4 we defined two random variables  $S$  and  $M$ , the sum and the maximum of two independent throws of a die.

**QUICK EXERCISE 9.1** List the elements of the event  $\{S = 7, M = 4\}$  and compute its probability.

In general, the joint distribution of two discrete random variables  $X$  and  $Y$ , defined on the *same* sample space  $\Omega$ , is given by prescribing the probabilities of all possible values of the pair  $(X, Y)$ .

DEFINITION. The *joint probability mass function*  $p$  of two discrete random variables  $X$  and  $Y$  is the function  $p : \mathbb{R}^2 \rightarrow [0, 1]$ , defined by

$$p(a, b) = P(X = a, Y = b) \quad \text{for } -\infty < a, b < \infty.$$

To stress the dependence on  $(X, Y)$ , we sometimes write  $p_{X,Y}$  instead of  $p$ . If  $X$  and  $Y$  take on the values  $a_1, a_2, \dots, a_k$  and  $b_1, b_2, \dots, b_\ell$ , respectively, the joint distribution of  $X$  and  $Y$  can simply be described by listing all the possible values of  $p(a_i, b_j)$ . For example, for the random variables  $S$  and  $M$  from Chapter 4 we obtain Table 9.1.

**Table 9.1.** Joint probability mass function  $p(a, b) = P(S = a, M = b)$ .

		$b$					
$a$	1	2	3	4	5	6	
2	1/36	0	0	0	0	0	
3	0	2/36	0	0	0	0	
4	0	1/36	2/36	0	0	0	
5	0	0	2/36	2/36	0	0	
6	0	0	1/36	2/36	2/36	0	
7	0	0	0	2/36	2/36	2/36	
8	0	0	0	1/36	2/36	2/36	
9	0	0	0	0	2/36	2/36	
10	0	0	0	0	1/36	2/36	
11	0	0	0	0	0	2/36	
12	0	0	0	0	0	1/36	

From this table we can retrieve the distribution of  $S$  and of  $M$ . For example, because

$$\{S = 6\} = \{S = 6, M = 1\} \cup \{S = 6, M = 2\} \cup \dots \cup \{S = 6, M = 6\},$$

and because the six events

$$\{S = 6, M = 1\}, \{S = 6, M = 2\}, \dots, \{S = 6, M = 6\}$$

are mutually exclusive, we find that

$$\begin{aligned} p_S(6) &= P(S = 6) = P(S = 6, M = 1) + \dots + P(S = 6, M = 6) \\ &= p(6, 1) + p(6, 2) + \dots + p(6, 6) \\ &= 0 + 0 + \frac{1}{36} + \frac{2}{36} + \frac{2}{36} + 0 \\ &= \frac{5}{36}. \end{aligned}$$

**Table 9.2.** Joint distribution and marginal distributions of  $S$  and  $M$ .

$a$	$b$						$p_S(a)$
	1	2	3	4	5	6	
2	1/36	0	0	0	0	0	1/36
3	0	2/36	0	0	0	0	2/36
4	0	1/36	2/36	0	0	0	3/36
5	0	0	2/36	2/36	0	0	4/36
6	0	0	1/36	2/36	2/36	0	5/36
7	0	0	0	2/36	2/36	2/36	6/36
8	0	0	0	1/36	2/36	2/36	5/36
9	0	0	0	0	2/36	2/36	4/36
10	0	0	0	0	1/36	2/36	3/36
11	0	0	0	0	0	2/36	2/36
12	0	0	0	0	0	1/36	1/36
$p_M(b)$	1/36	3/36	5/36	7/36	9/36	11/36	1

Thus we see that the probabilities of  $S$  can be obtained by taking the sum of the joint probabilities in the rows of Table 9.1. This yields the probability distribution of  $S$ , i.e., all values of  $p_S(a)$  for  $a = 2, \dots, 12$ . We speak of the *marginal* distribution of  $S$ . In Table 9.2 we have added this distribution in the right “margin” of the table. Similarly, summing over the columns of Table 9.1 yields the marginal distribution of  $M$ , in the bottom margin of Table 9.2.

The joint distribution of two random variables contains a lot more information than the two marginal distributions. This can be illustrated by the fact that in many cases the joint probability mass function of  $X$  and  $Y$  cannot be retrieved from the marginal probability mass functions  $p_X$  and  $p_Y$ . A simple example is given in the following quick exercise.

**QUICK EXERCISE 9.2** Let  $X$  and  $Y$  be two discrete random variables, with joint probability mass function  $p$ , given by the following table, where  $\varepsilon$  is an arbitrary number between  $-1/4$  and  $1/4$ .

$a$	$b$		$p_X(a)$
	0	1	
0	$1/4 - \varepsilon$	$1/4 + \varepsilon$	...
1	$1/4 + \varepsilon$	$1/4 - \varepsilon$	...
$p_Y(b)$	...	...	...

Complete the table, and conclude that we cannot retrieve  $p$  from  $p_X$  and  $p_Y$ .

### The joint distribution function

As in the case of a single random variable, the distribution function enables us to treat pairs of discrete and pairs of continuous random variables in the same way.

**DEFINITION.** The *joint distribution function*  $F$  of two random variables  $X$  and  $Y$  is the function  $F : \mathbb{R}^2 \rightarrow [0, 1]$  defined by

$$F(a, b) = P(X \leq a, Y \leq b) \quad \text{for } -\infty < a, b < \infty.$$

**QUICK EXERCISE 9.3** Compute  $F(5, 3)$  for the joint distribution function  $F$  of the pair  $(S, M)$ .

The distribution functions  $F_X$  and  $F_Y$  can be obtained from the joint distribution function of  $X$  and  $Y$ . As before, we speak of the *marginal* distribution functions. The following rule holds.

**FROM JOINT TO MARGINAL DISTRIBUTION FUNCTION.** Let  $F$  be the joint distribution function of random variables  $X$  and  $Y$ . Then the marginal distribution function of  $X$  is given for each  $a$  by

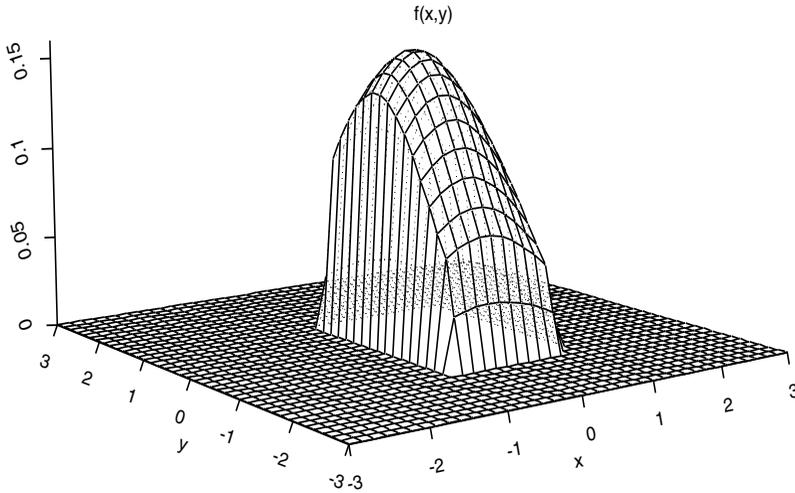
$$F_X(a) = P(X \leq a) = F(a, +\infty) = \lim_{b \rightarrow \infty} F(a, b), \quad (9.1)$$

and the marginal distribution function of  $Y$  is given for each  $b$  by

$$F_Y(b) = P(Y \leq b) = F(+\infty, b) = \lim_{a \rightarrow \infty} F(a, b). \quad (9.2)$$

## 9.2 Joint distributions of continuous random variables

We saw in Chapter 5 that the probability that a single continuous random variable  $X$  lies in an interval  $[a, b]$ , is equal to the area under the probability density function  $f$  of  $X$  over the interval (see also Figure 5.1). For the joint distribution of continuous random variables  $X$  and  $Y$  the situation is analogous: the probability that the pair  $(X, Y)$  falls in the rectangle  $[a_1, b_1] \times [a_2, b_2]$  is equal to the volume under the joint probability density function  $f(x, y)$  of  $(X, Y)$  over the rectangle. This is illustrated in Figure 9.1, where a chunk of a joint probability density function  $f(x, y)$  is displayed for  $x$  between  $-0.5$  and  $1$  and for  $y$  between  $-1.5$  and  $1$ . Its volume represents the probability  $P(-0.5 \leq X \leq 1, -1.5 \leq Y \leq 1)$ . As the volume under  $f$  on  $[-0.5, 1] \times [-1.5, 1]$  is equal to the integral of  $f$  over this rectangle, this motivates the following definition.



**Fig. 9.1.** Volume under a joint probability density function  $f$  on the rectangle  $[-0.5, 1] \times [-1.5, 1]$ .

**DEFINITION.** Random variables  $X$  and  $Y$  have a *joint continuous distribution* if for some function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  and for all numbers  $a_1, a_2$  and  $b_1, b_2$  with  $a_1 \leq b_1$  and  $a_2 \leq b_2$ ,

$$P(a_1 \leq X \leq b_1, a_2 \leq Y \leq b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y) \, dx \, dy.$$

The function  $f$  has to satisfy  $f(x, y) \geq 0$  for all  $x$  and  $y$ , and  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = 1$ . We call  $f$  the *joint probability density function* of  $X$  and  $Y$ .

As in the one-dimensional case there is a simple relation between the joint distribution function  $F$  and the joint probability density function  $f$ :

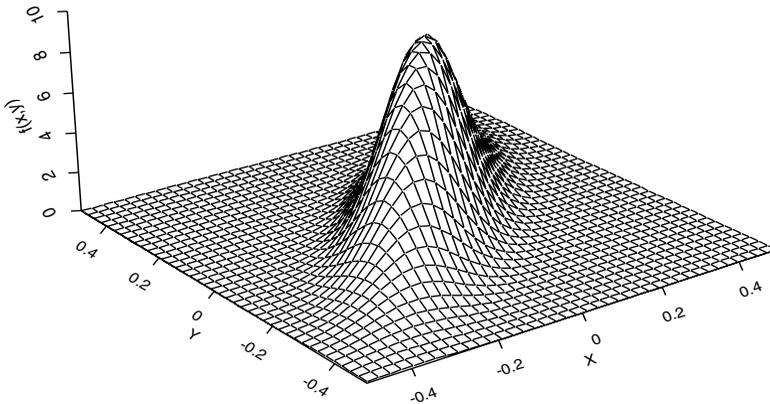
$$F(a, b) = \int_{-\infty}^a \int_{-\infty}^b f(x, y) \, dx \, dy \quad \text{and} \quad f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

A joint probability density function of two random variables is also called a *bivariate probability density*. An explicit example of such a density is the function

$$f(x, y) = \frac{30}{\pi} e^{-50x^2 - 50y^2 + 80xy}$$

for  $-\infty < x < \infty$  and  $-\infty < y < \infty$ ; see Figure 9.2. This is an example of a bivariate normal density (see Remark 11.2 for a full description of bivariate normal distributions).

We illustrate a number of properties of joint continuous distributions by means of the following simple example. Suppose that  $X$  and  $Y$  have joint probability

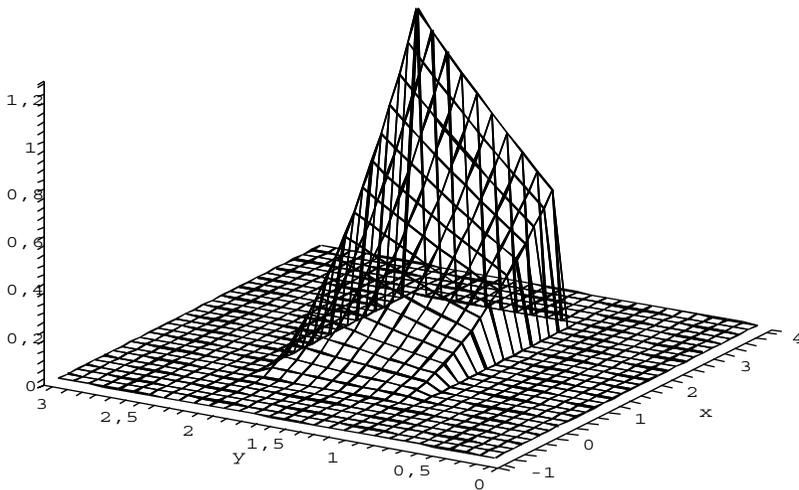


**Fig. 9.2.** A bivariate normal probability density function.

density function

$$f(x, y) = \frac{2}{75}(2x^2y + xy^2) \quad \text{for } 0 \leq x \leq 3 \text{ and } 1 \leq y \leq 2,$$

and  $f(x, y) = 0$  otherwise; see Figure 9.3.



**Fig. 9.3.** The probability density function  $f(x, y) = \frac{2}{75}(2x^2y + xy^2)$ .

As an illustration of how to compute joint probabilities:

$$\begin{aligned} \mathbb{P}\left(1 \leq X \leq 2, \frac{4}{3} \leq Y \leq \frac{5}{3}\right) &= \int_1^2 \int_{\frac{4}{3}}^{\frac{5}{3}} f(x, y) \, dx \, dy \\ &= \frac{2}{75} \int_1^2 \left( \int_{\frac{4}{3}}^{\frac{5}{3}} (2x^2y + xy^2) \, dy \right) dx \\ &= \frac{2}{75} \int_1^2 \left( x^2 + \frac{61}{81}x \right) dx = \frac{187}{2025}. \end{aligned}$$

Next, for  $a$  between 0 and 3 and  $b$  between 1 and 2, we determine the expression of the joint distribution function. Since  $f(x, y) = 0$  for  $x < 0$  or  $y < 1$ ,

$$\begin{aligned} F(a, b) &= \mathbb{P}(X \leq a, Y \leq b) = \int_{-\infty}^a \left( \int_{-\infty}^b f(x, y) \, dy \right) dx \\ &= \frac{2}{75} \int_0^a \left( \int_1^b (2x^2y + xy^2) \, dy \right) dx \\ &= \frac{1}{225} (2a^3b^2 - 2a^3 + a^2b^3 - a^2). \end{aligned}$$

Note that for either  $a$  outside  $[0, 3]$  or  $b$  outside  $[1, 2]$ , the expression for  $F(a, b)$  is different. For example, suppose that  $a$  is between 0 and 3 and  $b$  is larger than 2. Since  $f(x, y) = 0$  for  $y > 2$ , we find for any  $b \geq 2$ :

$$F(a, b) = \mathbb{P}(X \leq a, Y \leq b) = \mathbb{P}(X \leq a, Y \leq 2) = F(a, 2) = \frac{1}{225} (6a^3 + 7a^2).$$

Hence, applying (9.1) one finds the marginal distribution function of  $X$ :

$$F_X(a) = \lim_{b \rightarrow \infty} F(a, b) = \frac{1}{225} (6a^3 + 7a^2)$$

for  $a$  between 0 and 3.

**QUICK EXERCISE 9.4** Show that  $F_Y(b) = \frac{1}{75} (3b^3 + 18b^2 - 21)$  for  $b$  between 1 and 2.

The probability density of  $X$  can be found by differentiating  $F_X$ :

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} \left( \frac{1}{225} (6x^3 + 7x^2) \right) = \frac{2}{225} (9x^2 + 7x)$$

for  $x$  between 0 and 3. It is also possible to obtain the probability density function of  $X$  directly from  $f(x, y)$ . Recall that we determined marginal probabilities of discrete random variables by summing over the joint probabilities (see Table 9.2). In a similar way we can find  $f_X$ . For  $x$  between 0 and 3,

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy = \frac{2}{75} \int_1^2 (2x^2y + xy^2) \, dy = \frac{2}{225}(9x^2 + 7x).$$

This illustrates the following rule.

**FROM JOINT TO MARGINAL PROBABILITY DENSITY FUNCTION.** Let  $f$  be the joint probability density function of random variables  $X$  and  $Y$ . Then the *marginal* probability densities of  $X$  and  $Y$  can be found as follows:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx.$$

Hence the probability density function of each of the random variables  $X$  and  $Y$  can easily be obtained by “integrating out” the other variable.

**QUICK EXERCISE 9.5** Determine  $f_Y(y)$ .

### 9.3 More than two random variables

To determine the joint distribution of  $n$  random variables  $X_1, X_2, \dots, X_n$ , all defined on the *same* sample space  $\Omega$ , we have to describe how the probability mass is distributed over all possible values of  $(X_1, X_2, \dots, X_n)$ . In fact, it suffices to specify the *joint distribution function*  $F$  of  $X_1, X_2, \dots, X_n$ , which is defined by

$$F(a_1, a_2, \dots, a_n) = \text{P}(X_1 \leq a_1, X_2 \leq a_2, \dots, X_n \leq a_n)$$

for  $-\infty < a_1, a_2, \dots, a_n < \infty$ .

In case the random variables  $X_1, X_2, \dots, X_n$  are *discrete*, the joint distribution can also be characterized by specifying the *joint probability mass function*  $p$  of  $X_1, X_2, \dots, X_n$ , defined by

$$p(a_1, a_2, \dots, a_n) = \text{P}(X_1 = a_1, X_2 = a_2, \dots, X_n = a_n)$$

for  $-\infty < a_1, a_2, \dots, a_n < \infty$ .

#### Drawing without replacement

Let us illustrate the use of the joint probability mass function with an example. In the weekly Dutch National Lottery Show, 6 balls are drawn from a vase that contains balls numbered from 1 to 41. Clearly, the first number takes values  $1, 2, \dots, 41$  with equal probabilities. Is this also the case for—say—the third ball?

Let us consider a more general situation. Suppose a vase contains balls numbered  $1, 2, \dots, N$ . We draw  $n$  balls *without replacement* from the vase. Note that  $n$  cannot be larger than  $N$ . Each ball is selected with equal probability, i.e., in the first draw each ball has probability  $1/N$ , in the second draw each of the  $N - 1$  remaining balls has probability  $1/(N - 1)$ , and so on. Let  $X_i$  denote the number on the ball in the  $i$ -th draw, for  $i = 1, 2, \dots, n$ . In order to obtain the marginal probability mass function of  $X_i$ , we first compute the joint probability mass function of  $X_1, X_2, \dots, X_n$ . Since there are  $N(N - 1) \cdots (N - n + 1)$  possible combinations for the values of  $X_1, X_2, \dots, X_n$ , each having the same probability, the joint probability mass function is given by

$$\begin{aligned} p(a_1, a_2, \dots, a_n) &= P(X_1 = a_1, X_2 = a_2, \dots, X_n = a_n) \\ &= \frac{1}{N(N - 1) \cdots (N - n + 1)}, \end{aligned}$$

for all *distinct* values  $a_1, a_2, \dots, a_n$  with  $1 \leq a_j \leq N$ . Clearly  $X_1, X_2, \dots, X_n$  influence each other. Nevertheless, the marginal distribution of each  $X_i$  is the same. This can be seen as follows. Similar to obtaining the marginal probability mass functions in Table 9.2, we can find the marginal probability mass function of  $X_i$  by summing the joint probability mass function over all possible values of  $X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n$ :

$$\begin{aligned} p_{X_i}(k) &= \sum p(a_1, \dots, a_{i-1}, k, a_{i+1}, \dots, a_n) \\ &= \sum \frac{1}{N(N - 1) \cdots (N - n + 1)}, \end{aligned}$$

where the sum runs over all distinct values  $a_1, a_2, \dots, a_n$  with  $1 \leq a_j \leq N$  and  $a_i = k$ . Since there are  $(N - 1)(N - 2) \cdots (N - n + 1)$  such combinations, we conclude that the marginal probability mass function of  $X_i$  is given by

$$p_{X_i}(k) = (N - 1)(N - 2) \cdots (N - n + 1) \cdot \frac{1}{N(N - 1) \cdots (N - n + 1)} = \frac{1}{N},$$

for  $k = 1, 2, \dots, N$ . We see that the marginal probability mass function of each  $X_i$  is the *same*, assigning equal probability  $1/N$  to each possible value.

In case the random variables  $X_1, X_2, \dots, X_n$  are *continuous*, the joint distribution is defined in a similar way as in the case of two variables. We say that the random variables  $X_1, X_2, \dots, X_n$  have a *joint continuous distribution* if for some function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  and for all numbers  $a_1, a_2, \dots, a_n$  and  $b_1, b_2, \dots, b_n$  with  $a_i \leq b_i$ ,

$$\begin{aligned} P(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2, \dots, a_n \leq X_n \leq b_n) \\ = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n. \end{aligned}$$

Again  $f$  has to satisfy  $f(x_1, x_2, \dots, x_n) \geq 0$  and  $f$  has to integrate to 1. We call  $f$  the *joint probability density* of  $X_1, X_2, \dots, X_n$ .

## 9.4 Independent random variables

In earlier chapters we have spoken of independence of random variables, anticipating a formal definition. On page 46 we postulated that the events

$$\{R_1 = a_1\}, \{R_2 = a_2\}, \dots, \{R_{10} = a_{10}\}$$

related to the Bernoulli random variables  $R_1, \dots, R_{10}$  are independent. How should one define independence of random variables? Intuitively, random variables  $X$  and  $Y$  are independent if every event involving only  $X$  is independent of every event involving only  $Y$ . Since for two discrete random variables  $X$  and  $Y$ , any event involving  $X$  and  $Y$  is the union of events of the type  $\{X = a, Y = b\}$ , an adequate definition for independence would be

$$P(X = a, Y = b) = P(X = a)P(Y = b), \quad (9.3)$$

for all possible values  $a$  and  $b$ . However, this definition is useless for continuous random variables. Both the discrete and the continuous case are covered by the following definition.

DEFINITION. The random variables  $X$  and  $Y$ , with joint distribution function  $F$ , are *independent* if

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b),$$

that is,

$$F(a, b) = F_X(a)F_Y(b) \quad (9.4)$$

for all possible values  $a$  and  $b$ . Random variables that are not independent are called *dependent*.

Note that independence of  $X$  and  $Y$  guarantees that the joint probability of  $\{X \leq a, Y \leq b\}$  factorizes. More generally, the following is true: if  $X$  and  $Y$  are independent, then

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B), \quad (9.5)$$

for all suitable  $A$  and  $B$ , such as intervals and points. As a special case we can take  $A = \{a\}$ ,  $B = \{b\}$ , which yields that for independent  $X$  and  $Y$  the probability of  $\{X = a, Y = b\}$  equals the product of the marginal probabilities. In fact, for *discrete* random variables the definition of independence can be reduced—after cumbersome computations—to equality (9.3). For continuous random variables  $X$  and  $Y$  we find, differentiating both sides of (9.4) with respect to  $x$  and  $y$ , that

$$f(x, y) = f_X(x)f_Y(y).$$

**QUICK EXERCISE 9.6** Determine for which value of  $\varepsilon$  the discrete random variables  $X$  and  $Y$  from Quick exercise 9.2 are independent.

More generally, random variables  $X_1, X_2, \dots, X_n$ , with joint distribution function  $F$ , are *independent* if for all values  $a_1, \dots, a_n$ ,

$$F(a_1, a_2, \dots, a_n) = F_{X_1}(a_1)F_{X_2}(a_2) \cdots F_{X_n}(a_n).$$

As in the case of two discrete random variables, the discrete random variables  $X_1, X_2, \dots, X_n$  are independent if

$$P(X_1 = a_1, \dots, X_n = a_n) = P(X_1 = a_1) \cdots P(X_n = a_n),$$

for all possible values  $a_1, \dots, a_n$ . Thus we see that the definition of independence for discrete random variables is in agreement with our intuitive interpretation given earlier in (9.3).

In case of independent continuous random variables  $X_1, X_2, \dots, X_n$  with joint probability density function  $f$ , differentiating the joint distribution function with respect to all the variables gives that

$$f(x_1, x_2, \dots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2) \cdots f_{X_n}(x_n) \quad (9.6)$$

for all values  $x_1, \dots, x_n$ . By integrating both sides over  $(-\infty, a_1] \times (-\infty, a_2] \times \cdots \times (-\infty, a_n]$ , we find the definition of independence. Hence in the continuous case, (9.6) is equivalent to the definition of independence.

## 9.5 Propagation of independence

A natural question is whether transformed independent random variables are again independent. We start with a simple example. Let  $X$  and  $Y$  be two independent random variables with joint distribution function  $F$ . Take an interval  $I = (a, b]$  and define random variables  $U$  and  $V$  as follows:

$$U = \begin{cases} 1 & \text{if } X \in I \\ 0 & \text{if } X \notin I, \end{cases} \quad \text{and} \quad V = \begin{cases} 1 & \text{if } Y \in I \\ 0 & \text{if } Y \notin I. \end{cases}$$

Are  $U$  and  $V$  independent? Yes, they are! By using (9.5) and the independence of  $X$  and  $Y$ , we can write

$$\begin{aligned} P(U = 0, V = 1) &= P(X \in I^c, Y \in I) \\ &= P(X \in I^c)P(Y \in I) \\ &= P(U = 0)P(V = 1). \end{aligned}$$

By a similar reasoning one finds that for *all* values  $a$  and  $b$ ,

$$P(U = a, V = b) = P(U = a)P(V = b).$$

This illustrates the fact that for independent random variables  $X_1, X_2, \dots, X_n$ , the random variables  $Y_1, Y_2, \dots, Y_n$ , where each  $Y_i$  is determined by  $X_i$  only, inherit the independence from the  $X_i$ . The general rule is given here.

**PROPAGATION OF INDEPENDENCE.** Let  $X_1, X_2, \dots, X_n$  be independent random variables. For each  $i$ , let  $h_i : \mathbb{R} \rightarrow \mathbb{R}$  be a function and define the random variable

$$Y_i = h_i(X_i).$$

Then  $Y_1, Y_2, \dots, Y_n$  are also independent.

Often one uses this rule with all functions the same:  $h_i = h$ . For instance, in the preceding example,

$$h(x) = \begin{cases} 1 & \text{if } x \in I \\ 0 & \text{if } x \notin I. \end{cases}$$

The rule is also useful when we need different transformations for different  $X_i$ . We already saw an example of this in Chapter 6. In the single-server queue example in Section 6.4, the  $Exp(0.5)$  random variables  $T_1, T_2, \dots$  and  $U(2, 5)$  random variables  $S_1, S_2, \dots$  are required to be independent. They are generated according to the technique described in Section 6.2. With a sequence  $U_1, U_2, \dots$  of independent  $U(0, 1)$  random variables we can accomplish independence of the  $T_i$  and  $S_i$  as follows:

$$T_i = F^{\text{inv}}(U_{2i-1}) \quad \text{and} \quad S_i = G^{\text{inv}}(U_{2i}),$$

where  $F$  and  $G$  are the distribution functions of the  $Exp(0.5)$  distribution and the  $U(2, 5)$  distribution. The propagation-of-independence rule now guarantees that all random variables  $T_1, S_1, T_2, S_2, \dots$  are independent.

## 9.6 Solutions to the quick exercises

**9.1** The only possibilities with the sum equal to 7 and the maximum equal to 4 are the combinations (3, 4) and (4, 3). They both have probability  $1/36$ , so that  $P(S = 7, M = 4) = 2/36$ .

**9.2** Since  $p_X(0)$ ,  $p_X(1)$ ,  $p_Y(0)$ , and  $p_Y(1)$  are all equal to  $1/2$ , knowing only  $p_X$  and  $p_Y$  yields no information on  $\varepsilon$  whatsoever. You have to be a student at Hogwarts to be able to get the values of  $p$  right!

**9.3** Since  $S$  and  $M$  are discrete random variables,  $F(5, 3)$  is the sum of the probabilities  $P(S = a, M = b)$  of all combinations  $(a, b)$  with  $a \leq 5$  and  $b \leq 3$ . From Table 9.2 we see that this sum is  $8/36$ .

**9.4** For  $a$  between 0 and 3 and for  $b$  between 1 and 2, we have seen that

$$F(a, b) = \frac{1}{225}(2a^3b^2 - 2a^3 + a^2b^3 - a^2).$$

Since  $f(x, y) = 0$  for  $x > 3$ , we find for any  $a \geq 3$  and  $b$  between 1 and 2:

$$\begin{aligned} F(a, b) &= P(X \leq a, Y \leq b) = P(X \leq 3, Y \leq b) \\ &= F(3, b) = \frac{1}{75}(3b^3 + 18b^2 - 21). \end{aligned}$$

As a result, applying (9.2) yields that  $F_Y(b) = \lim_{a \rightarrow \infty} F(a, b) = F(3, b) = \frac{1}{75}(3b^3 + 18b^2 - 21)$ , for  $b$  between 1 and 2.

**9.5** For  $y$  between 1 and 2, we have seen that  $F_Y(y) = \frac{1}{75}(3y^3 + 18y^2 - 21)$ . Differentiating with respect to  $y$  yields that

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{25}(3y^2 + 12y),$$

for  $y$  between 1 and 2 (and  $f_Y(y) = 0$  otherwise). The probability density function of  $Y$  can also be obtained directly from  $f(x, y)$ . For  $y$  between 1 and 2:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x, y) dx = \frac{2}{75} \int_0^3 (2x^2y + xy^2) dx \\ &= \frac{2}{75} \left[ \frac{2}{3}x^3y + \frac{1}{2}x^2y^2 \right]_{x=0}^{x=3} = \frac{1}{25}(3y^2 + 12y). \end{aligned}$$

Since  $f(x, y) = 0$  for values of  $y$  not between 1 and 2, we have that  $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx = 0$  for these  $y$ 's.

**9.6** The number  $\varepsilon$  is between  $-1/4$  and  $1/4$ . Now  $X$  and  $Y$  are independent in case  $p(i, j) = P(X = i, Y = j) = P(X = i)P(Y = j) = p_X(i)p_Y(j)$ , for all  $i, j = 0, 1$ . If  $i = j = 0$ , we should have

$$\frac{1}{4} - \varepsilon = p(0, 0) = p_X(0)p_Y(0) = \frac{1}{4}.$$

This implies that  $\varepsilon = 0$ . Furthermore, for all other combinations  $(i, j)$  one can check that for  $\varepsilon = 0$  also  $p(i, j) = p_X(i)p_Y(j)$ , so that  $X$  and  $Y$  are independent. If  $\varepsilon \neq 0$ , we have  $p(0, 0) \neq p_X(0)p_Y(0)$ , so that  $X$  and  $Y$  are dependent.

## 9.7 Exercises

**9.1** The joint probabilities  $P(X = a, Y = b)$  of discrete random variables  $X$  and  $Y$  are given in the following table (which is based on the *magical square* in Albrecht Dürer's engraving *Melencolia I* in Figure 9.4). Determine the marginal probability distributions of  $X$  and  $Y$ , i.e., determine the probabilities  $P(X = a)$  and  $P(Y = b)$  for  $a, b = 1, 2, 3, 4$ .



**Fig. 9.4.** Albrecht Dürer’s *Melencolia I*.

Albrecht Dürer (German, 1471-1528) *Melencolia I*, 1514. Engraving. Bequest of William P. Chapman, Jr., Class of 1895. Courtesy of the Herbert F. Johnson Museum of Art, Cornell University.

		$a$			
$b$	1	2	3	4	
1	16/136	3/136	2/136	13/136	
2	5/136	10/136	11/136	8/136	
3	9/136	6/136	7/136	12/136	
4	4/136	15/136	14/136	1/136	

**9.2**  $\boxplus$  The joint probability distribution of two discrete random variables  $X$  and  $Y$  is partly given in the following table.

$b$	$a$			$P(Y = b)$
	0	1	2	
-1	...	...	...	1/2
1	...	1/2	...	1/2
$P(X = a)$	1/6	2/3	1/6	1

- Complete the table.
- Are  $X$  and  $Y$  dependent or independent?

**9.3** Let  $X$  and  $Y$  be two random variables, with joint distribution the *Melen-colia* distribution, given by the table in Exercise 9.1. What is

- $P(X = Y)$ ?
- $P(X + Y = 5)$ ?
- $P(1 < X \leq 3, 1 < Y \leq 3)$ ?
- $P((X, Y) \in \{1, 4\} \times \{1, 4\})$ ?

**9.4** This exercise will be easy for those familiar with Japanese puzzles called *nonograms*. The marginal probability distributions of the discrete random variables  $X$  and  $Y$  are given in the following table:

$b$	$a$					$P(Y = b)$
	1	2	3	4	5	
1						5/14
2						4/14
3						2/14
4						2/14
5						1/14
$P(X = a)$	1/14	5/14	4/14	2/14	2/14	1

Moreover, for  $a$  and  $b$  from 1 to 5 the joint probability  $P(X = a, Y = b)$  is either 0 or  $1/14$ . Determine the joint probability distribution of  $X$  and  $Y$ .

**9.5**  $\boxminus$  Let  $\eta$  be an unknown real number, and let the joint probabilities  $P(X = a, Y = b)$  of the discrete random variables  $X$  and  $Y$  be given by the following table:

$b$	$a$		
	-1	0	1
4	$\eta - \frac{1}{16}$	$\frac{1}{4} - \eta$	0
5	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{8}$
6	$\eta + \frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4} - \eta$

- a. Which are the values  $\eta$  can attain?  
b. Is there a value of  $\eta$  for which  $X$  and  $Y$  are independent?

**9.6**  $\square$  Let  $X$  and  $Y$  be two independent  $Ber(\frac{1}{2})$  random variables. Define random variables  $U$  and  $V$  by:

$$U = X + Y \quad \text{and} \quad V = |X - Y|.$$

- a. Determine the joint and marginal probability distributions of  $U$  and  $V$ .  
b. Find out whether  $U$  and  $V$  are dependent or independent.

**9.7** To investigate the relation between hair color and eye color, the hair color and eye color of 5383 persons was recorded. The data are given in the following table:

Eye color	Hair color		
	Fair/red	Medium	Dark/black
Light	1168	825	305
Dark	573	1312	1200

*Source:* B. Everitt and G. Dunn. *Applied multivariate data analysis*. Second edition Hodder Arnold, 2001; Table 4.12. Reproduced by permission of Hodder & Stoughton.

Eye color is encoded by the values 1 (Light) and 2 (Dark), and hair color by 1 (Fair/red), 2 (Medium), and 3 (Dark/black). By dividing the numbers in the table by 5383, the table is turned into a joint probability distribution for random variables  $X$  (hair color) taking values 1 to 3 and  $Y$  (eye color) taking values 1 and 2.

- a. Determine the joint and marginal probability distributions of  $X$  and  $Y$ .  
b. Find out whether  $X$  and  $Y$  are dependent or independent.

**9.8**  $\boxplus$  Let  $X$  and  $Y$  be independent random variables with probability distributions given by

$$P(X = 0) = P(X = 1) = \frac{1}{2} \quad \text{and} \quad P(Y = 0) = P(Y = 2) = \frac{1}{2}.$$

- a. Compute the distribution of  $Z = X + Y$ .
- b. Let  $\tilde{Y}$  and  $\tilde{Z}$  be independent random variables, where  $\tilde{Y}$  has the same distribution as  $Y$ , and  $\tilde{Z}$  the same distribution as  $Z$ . Compute the distribution of  $\tilde{X} = \tilde{Z} - \tilde{Y}$ .

**9.9**  $\boxplus$  Suppose that the joint distribution function of  $X$  and  $Y$  is given by

$$F(x, y) = 1 - e^{-2x} - e^{-y} + e^{-(2x+y)} \quad \text{if } x > 0, y > 0,$$

and  $F(x, y) = 0$  otherwise.

- a. Determine the marginal distribution functions of  $X$  and  $Y$ .
- b. Determine the joint probability density function of  $X$  and  $Y$ .
- c. Determine the marginal probability density functions of  $X$  and  $Y$ .
- d. Find out whether  $X$  and  $Y$  are independent.

**9.10**  $\boxminus$  Let  $X$  and  $Y$  be two continuous random variables with joint probability density function

$$f(x, y) = \frac{12}{5}xy(1+y) \quad \text{for } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1,$$

and  $f(x, y) = 0$  otherwise.

- a. Find the probability  $P(\frac{1}{4} \leq X \leq \frac{1}{2}, \frac{1}{3} \leq Y \leq \frac{2}{3})$ .
- b. Determine the joint distribution function of  $X$  and  $Y$  for  $a$  and  $b$  between 0 and 1.
- c. Use your answer from **b** to find  $F_X(a)$  for  $a$  between 0 and 1.
- d. Apply the rule on page 122 to find the probability density function of  $X$  from the joint probability density function  $f(x, y)$ . Use the result to verify your answer from **c**.
- e. Find out whether  $X$  and  $Y$  are independent.

**9.11**  $\boxplus$  Let  $X$  and  $Y$  be two continuous random variables, with the same joint probability density function as in Exercise 9.10. Find the probability  $P(X < Y)$  that  $X$  is smaller than  $Y$ .

**9.12** The joint probability density function  $f$  of the pair  $(X, Y)$  is given by

$$f(x, y) = K(3x^2 + 8xy) \quad \text{for } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 2,$$

and  $f(x, y) = 0$  for all other values of  $x$  and  $y$ . Here  $K$  is some positive constant.

- a. Find  $K$ .
- b. Determine the probability  $P(2X \leq Y)$ .

**9.13**  $\square$  On a disc with origin  $(0, 0)$  and radius 1, a point  $(X, Y)$  is selected by throwing a dart that hits the disc in an arbitrary place. This is best described by the joint probability density function  $f$  of  $X$  and  $Y$ , given by

$$f(x, y) = \begin{cases} c & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise,} \end{cases}$$

where  $c$  is some positive constant.

- a. Determine  $c$ .
- b. Let  $R = \sqrt{X^2 + Y^2}$  be the distance from  $(X, Y)$  to the origin. Determine the distribution function  $F_R$ .
- c. Determine the marginal density function  $f_X$ . Without doing any calculations, what can you say about  $f_Y$ ?

**9.14** An arbitrary point  $(X, Y)$  is drawn from the square  $[-1, 1] \times [-1, 1]$ . This means that for any region  $G$  in the plane, the probability that  $(X, Y)$  is in  $G$ , is given by the area of  $G \cap \square$  divided by the area of  $\square$ , where  $\square$  denotes the square  $[-1, 1] \times [-1, 1]$ :

$$P((X, Y) \in G) = \frac{\text{area of } G \cap \square}{\text{area of } \square}.$$

- a. Determine the joint probability density function of the pair  $(X, Y)$ .
- b. Check that  $X$  and  $Y$  are two independent,  $U(-1, 1)$  distributed random variables.

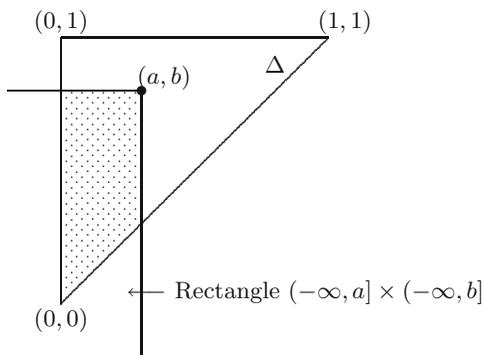
**9.15**  $\triangle$  Let the pair  $(X, Y)$  be drawn arbitrarily from the triangle  $\Delta$  with vertices  $(0, 0)$ ,  $(0, 1)$ , and  $(1, 1)$ .

- a. Use Figure 9.5 to show that the joint distribution function  $F$  of the pair  $(X, Y)$  satisfies

$$F(a, b) = \begin{cases} 0 & \text{for } a \text{ or } b \text{ less than } 0 \\ a(2b - a) & \text{for } (a, b) \text{ in the triangle } \Delta \\ b^2 & \text{for } b \text{ between } 0 \text{ and } 1 \text{ and } a \text{ larger than } b \\ 2a - a^2 & \text{for } a \text{ between } 0 \text{ and } 1 \text{ and } b \text{ larger than } 1 \\ 1 & \text{for } a \text{ and } b \text{ larger than } 1. \end{cases}$$

- b. Determine the joint probability density function  $f$  of the pair  $(X, Y)$ .
- c. Show that  $f_X(x) = 2 - 2x$  for  $x$  between 0 and 1 and that  $f_Y(y) = 2y$  for  $y$  between 0 and 1.

**9.16** (Continuation of Exercise 9.15) An arbitrary point  $(U, V)$  is drawn from the unit square  $[0, 1] \times [0, 1]$ . Let  $X$  and  $Y$  be defined as in Exercise 9.15. Show that  $\min\{U, V\}$  has the same distribution as  $X$  and that  $\max\{U, V\}$  has the same distribution as  $Y$ .



**Fig. 9.5.** Drawing  $(X, Y)$  from  $(-\infty, a] \times (-\infty, b] \cap \Delta$ .

**9.17** Let  $U_1$  and  $U_2$  be two independent random variables, both uniformly distributed over  $[0, a]$ . Let  $V = \min\{U_1, U_2\}$  and  $Z = \max\{U_1, U_2\}$ . Show that the joint distribution function of  $V$  and  $Z$  is given by

$$F(s, t) = P(V \leq s, Z \leq t) = \frac{t^2 - (t - s)^2}{a^2} \quad \text{for } 0 \leq s \leq t \leq a.$$

*Hint:* note that  $V \leq s$  and  $Z \leq t$  happens exactly when both  $U_1 \leq t$  and  $U_2 \leq t$ , but *not* both  $s < U_1 \leq t$  and  $s < U_2 \leq t$ .

**9.18** Suppose a vase contains balls numbered  $1, 2, \dots, N$ . We draw  $n$  balls *without replacement* from the vase. Each ball is selected with equal probability, i.e., in the first draw each ball has probability  $1/N$ , in the second draw each of the  $N - 1$  remaining balls has probability  $1/(N - 1)$ , and so on. For  $i = 1, 2, \dots, n$ , let  $X_i$  denote the number on the ball in the  $i$ th draw. We have shown that the marginal probability mass function of  $X_i$  is given by

$$p_{X_i}(k) = \frac{1}{N}, \quad \text{for } k = 1, 2, \dots, N.$$

a. Show that

$$E[X_i] = \frac{N + 1}{2}.$$

b. Compute the variance of  $X_i$ . You may use the identity

$$1 + 4 + 9 + \dots + N^2 = \frac{1}{6}N(N + 1)(2N + 1).$$

**9.19**  $\square$  Let  $X$  and  $Y$  be two continuous random variables, with joint probability density function

$$f(x, y) = \frac{30}{\pi} e^{-50x^2 - 50y^2 + 80xy}$$

for  $-\infty < x < \infty$  and  $-\infty < y < \infty$ ; see also Figure 9.2.

- a. Determine positive numbers  $a$ ,  $b$ , and  $c$  such that

$$50x^2 - 80xy + 50y^2 = (ay - bx)^2 + cx^2.$$

- b. Setting  $\mu = \frac{4}{5}x$ , and  $\sigma = \frac{1}{10}$ , show that

$$(\sqrt{50}y - \sqrt{32}x)^2 = \frac{1}{2} \left( \frac{y - \mu}{\sigma} \right)^2$$

and use this to show that

$$\int_{-\infty}^{\infty} e^{-(\sqrt{50}y - \sqrt{32}x)^2} dy = \frac{\sqrt{2\pi}}{10}.$$

- c. Use the results from **b** to determine the probability density function  $f_X$  of  $X$ . What kind of distribution does  $X$  have?

**9.20** Suppose we throw a needle on a large sheet of paper, on which horizontal lines are drawn, which are at needle-length apart (see also Exercise 21.16). Choose one of the horizontal lines as  $x$ -axis, and let  $(X, Y)$  be the center of the needle. Furthermore, let  $Z$  be the distance of this center  $(X, Y)$  to the nearest horizontal line under  $(X, Y)$ , and let  $H$  be the angle between the needle and the positive  $x$ -axis.

- a. Assuming that the length of the needle is equal to 1, argue that  $Z$  has a  $U(0, 1)$  distribution. Also argue that  $H$  has a  $U(0, \pi)$  distribution and that  $Z$  and  $H$  are independent.
- b. Show that the needle hits a horizontal line when

$$Z \leq \frac{1}{2} \sin H \quad \text{or} \quad 1 - Z \leq \frac{1}{2} \sin H.$$

- c. Show that the probability that the needle will hit one of the horizontal lines equals  $2/\pi$ .