6.1-6.2 Joint Distribution Function

For any two random variables X and Y, the *joint cumulative probability distribution function* of X and Y by

$$F(a, b) = P\{X \le a, Y \le b\}, \quad -\infty < a, b < \infty$$

The distribution of X (and Y) can be obtained from the joint distribution of X and Y as follows:

$$\begin{aligned} F_X(a) &= P\{X \le a\} \\ &= P\{X \le a, Y < \infty\} \\ &= P\left(\lim_{b\to\infty} \{X \le a, Y \le b\}\right) \\ &= \lim_{b\to\infty} P\{X \le a, Y \le b\} \\ &= \lim_{b\to\infty} F(a, b) \\ &\equiv F(a, \infty) \end{aligned}$$

$$\begin{aligned} F_Y(b) &= P\{Y \le b\} \\ &= \lim_{a\to\infty} F(a, b) \\ &\equiv F(\infty, b) \end{aligned}$$

All joint probability statements about X and Y can, in theory, be answered in terms of their joint distribution function. For instance,

$$\begin{aligned} &P\{a_1 < X \le a_2, b_1 < Y \le b_2\} \\ &\quad = F(a_2, b_2) + F(a_1, b_1) - F(a_1, b_2) - F(a_2, b_1) \end{aligned} \tag{1.2}$$

whenever $a_1 < a_2, b_1 < b_2$.

Discrete Random Variables

In the case when X and Y are both discrete random variables, it is convenient to define the *joint probability mass function* of X and Y by

$$p(x,y) = P\{X = x, Y = y\}$$

The probability mass function of X can be obtained from $p(x,y)$ by

$$\begin{aligned} p_X(x) &= P\{X = x\} \\ &= \sum_{y:p(x,y)>0} p(x,y) \end{aligned}$$

Similarly,

$$p_Y(y) = \sum_{x:p(x,y)>0} p(x,y)$$

EXAMPLE 1b

Suppose that 15 percent of the families in a certain community have no children, 20 percent have 1 child, 35 percent have 2 children, and 30 percent have 3. Suppose further that in each family each child is equally likely (independently) to be a boy or a girl. If a family is chosen at random from this community, then B, the number of boys, and G, the number of girls, in this family will have the joint probability mass function shown in Table 6.2.

TABLE 6.2: $P\{B = i, G = j\}$

i \ j	0	1	2	3	Row sum $= P\{B = i\}$
0	.15	.10	.0875	.0375	.3750
1	.10	.175	.1125	0	.3875
2	.0875	.1125	0	0	.2000
3	.0375	0	0	0	.0375
Columnsum $= P\{G = j\}$	.3750	.3875	.2000	.0375	

Continuous Random Variables

We say that X and Y are *jointly continuous* if there exists a function $f(x, y)$, defined for all real x and y, having the property that, for every set C of pairs of real numbers (that is, C is a set in the two-dimensional plane),

$$P\{(X, Y) \in C\} = \iint_{(x,y)\in C} f(x, y)\, dx\, dy \tag{1.3}$$

The function $f(x, y)$ is called the *joint probability density function* of X and Y.

PROPERTIES

1. $$F(a,b) = P\{X \in (-\infty, a], Y \in (-\infty, b]\}$$
$$= \int_{-\infty}^{b} \int_{-\infty}^{a} f(x,y)\, dx\, dy$$

2. $$f(a,b) = \frac{\partial^2}{\partial a\, \partial b} F(a,b)$$

3. $$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

4. $$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

EXAMPLE 1c

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty,\ 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Compute (a) $P\{X > 1, Y < 1\}$, (b) $P\{X < Y\}$, and (c) $P\{X < a\}$.

Solution. (a)

$$P\{X > 1, Y < 1\} = \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}\,dx\,dy = \int_0^1 2e^{-2y}\left(-e^{-x}\Big|_1^\infty\right)dy = e^{-1}\int_0^1 2e^{-2y}\,dy = e^{-1}(1 - e^{-2})$$

(b)

$$P\{X < Y\} = \iint\limits_{(x,y):x<y} 2e^{-x}e^{-2y}\,dx\,dy = \int_0^\infty \int_0^y 2e^{-x}e^{-2y}\,dx\,dy = \int_0^\infty 2e^{-2y}(1 - e^{-y})dy = \int_0^\infty 2e^{-2y}dy - \int_0^\infty 2e^{-3y}dy = 1 - \frac{2}{3} = \frac{1}{3}$$

(c)

$$P\{X < a\} = \int_0^a \int_0^\infty 2e^{-2y}e^{-x}\,dy\,dx = \int_0^a e^{-x}\,dx = 1 - e^{-a}$$

We can also define joint probability distributions for n random variables in exactly the same manner as we did for $n = 2$. For instance, the joint cumulative probability distribution function $F(a_1, a_2, \ldots, a_n)$ of the n random variables $X_1, X_2, \ldots, X_n$ is defined by

$$F(a_1, a_2, \ldots, a_n) = P\{X_1 \le a_1, X_2 \le a_2, \ldots, X_n \le a_n\}$$

Further, the n random variables are said to be *jointly continuous* if there exists a function $f(x_1, x_2, \ldots, x_n)$, called the *joint probability density function*, such that, for any set C in n-space,

$$P\{(X_1, X_2, \ldots, X_n) \in C\} = \iint\limits_{(x_1,\ldots,x_n)\in C} \cdots \int f(x_1, \ldots, x_n)dx_1dx_2\cdots dx_n$$

EXAMPLE 1f The multinomial distribution

One of the most important joint distributions is the multinomial distribution, which arises when a sequence of n independent and identical experiments is performed. Suppose that each experiment can result in any one of r possible outcomes, with respective probabilities $p_1, p_2, \ldots, p_r$, $\sum_{i=1}^{r} p_i = 1$. If we let X_i denote the number of the n experiments that result in outcome number i, then

$$P\{X_1 = n_1, X_2 = n_2, \ldots, X_r = n_r\} = \frac{n!}{n_1!n_2!\cdots n_r!} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r} \tag{1.5}$$

whenever $\sum_{i=1}^{r} n_i = n$.

Example As an application of the multinomial distribution, suppose that a fair die is rolled 9 times. The probability that 1 appears three times, 2 and 3 twice each, 4 and 5 once each, and 6 not at all is

$$\frac{9!}{3!2!2!1!1!0!}\left(\frac{1}{6}\right)^3\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^0 = \frac{9!}{3!2!2!}\left(\frac{1}{6}\right)^9$$

■

Independent Random Variables

The random variables X and Y are said to be *independent* if, for any two sets of real numbers A and B,

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\} \tag{2.1}$$

Hence, in terms of the joint distribution function F of X and Y, X and Y are independent if

$$F(a, b) = F_X(a)F_Y(b) \quad \text{for all } a, b$$

When X and Y are discrete random variables, the condition of independence (2.1) is equivalent to

$$p(x, y) = p_X(x)p_Y(y) \quad \text{for all } x, y \tag{2.2}$$

In the jointly continuous case, the condition of independence is equivalent to

$$f(x, y) = f_X(x)f_Y(y) \quad \text{for all } x, y$$

(these statements require proves)

EXAMPLE 2c

A man and a woman decide to meet at a certain location. If each of them independently arrives at a time uniformly distributed between 12 noon and 1 P.M., find the probability that the first to arrive has to wait longer than 10 minutes.

Solution: Since X and Y are independent, $f(x, y) = f_X(x) f_Y(x) = (\frac{1}{60})^2, \quad 0 < x, y < 60,$ and 0 otherwise. Hence

$$P\{|X - Y| < 10\} = P\{Y < X - 10\} + P\{Y > X + 10\} = 2\int_{10}^{60}\int_{0}^{y-10}\left(\frac{1}{60}\right)^2 dx\,dy \quad = 25/36$$

EXAMPLE 2f

If the joint density function of X and Y is

$$f(x, y) = 6e^{-2x}e^{-3y} \qquad 0 < x < \infty,\ 0 < y < \infty$$

and is equal to 0 outside this region, are the random variables independent? What if the joint density function is

$$f(x, y) = 24xy \qquad 0 < x < 1, 0 < y < 1,\ 0 < x + y < 1$$

and is equal to 0 otherwise?

Solution:

a) $f_X(x) = 2\,e^{-2x},\ x > 0;\quad f_Y(y) = 3\,e^{-3y},\ y > 0;$

$f_X(x)f_Y(y) = 6\,e^{-2x}e^{-3y},\ x, y > 0.$ Hence $f_X(x)f_Y(y) = f(x, y)$, INDEPENDENT

b) $f_X(x) = \int_0^{1-x} 24xy\,dy = 12\,x\,(1 - x^2),\ 0 < x < 1;\quad f_Y(y) = 12\,y\,(1 - y^2),\ 0 < y < 1$

$f_X(x)f_Y(y) \neq f(x, y)$, DEPENDENT

Proposition 2.1. The continuous (discrete) random variables X and Y are independent if and only if their joint probability density (mass) function can be expressed as

$$f_{X,Y}(x, y) = h(x)g(y) \qquad -\infty < x < \infty, -\infty < y < \infty$$

The concept of independence may, of course, be defined for more than two random variables. In general, the n random variables $X_1, X_2, \ldots, X_n$ are said to be independent if, for all sets of real numbers $A_1, A_2, \ldots, A_n$,

$$P\{X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n\} = \prod_{i=1}^{n} P\{X_i \in A_i\}$$

As before, it can be shown that this condition is equivalent to

$$\begin{aligned} P\{X_1 \le a_1, X_2 \le a_2, \ldots, X_n \le a_n\} \\ = \prod_{i=1}^{n} P\{X_i \le a_i\} \quad \text{for all } a_1, a_2, \ldots, a_n \end{aligned}$$

Finally, we say that an infinite collection of random variables is independent if every finite subcollection of them is independent.

Exercises

6.22. The joint density function of X and Y is

$$f(x, y) = \begin{cases} x + y & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Are X and Y independent?
(b) Find the density function of X.
(c) Find $P\{X + Y < 1\}$.

6.27. If X_1 and X_2 are independent exponential random variables with respective parameters λ_1 and λ_2, find the distribution of $Z = X_1/X_2$. Also compute $P\{X_1 < X_2\}$.