4. Multiple Random Variables

4.1 Joint and Marginal Distributions

Definition 4.1.1 An n-dimensional random vector is a function from a sample space S into $\mathbb{R}^n$, n-dimensional Euclidean space.

Suppose, for example, that with each point in a sample space we associate an ordered pair of numbers, that is, a point $(x, y) \in \mathbb{R}^2$, where $\mathbb{R}^2$ denotes the plane. Then we have defined a two -dimensional (or bivariate) random vector (X, Y).

Example 4.1.2 (Sample space for dice)

Consider the experiment of tossing two fair dice. The sample space for this experiment has 36 equally likely points. Let

$$X\text{=sum of the two dice} \quad \text{and} \quad Y\text{=|difference of two dice|}.$$

In this way we have defined then bivariate random vector (X, Y).

The random vector (X, Y) defined above is called a discrete random vector because it has only a countable (in this case, finite) number of possible values. The probabilities of events defined in terms of X and Y are just defined in terms of the probabilities of the corresponding events in the sample space S. For example,

$$P(X = 5, Y = 3) = P(\{4, 1\}, \{1, 4\}) = \frac{2}{36} = \frac{1}{18}.$$

Definition 4.1.2 Let (X, Y) be a discrete bivariate random vector. Then the function $f(x, y)$ from $\mathbb{R}^2$ into $\mathbb{R}$ defined by $f(x, y) = P(X = x, Y = y)$ is called the joint probability mass function or joint pmf of (X, Y). If it is necessary to stress the fact that f is the joint pmf of the vector (X, Y) rather than some other vector, the notation $f_{X,Y}(x, y)$ will be used.

The joint pmf can be used to compute the probability of any event defined in terms of (X, Y).

Let A be any subset of $\mathbb{R}^2$. Then

$$P((X,Y) \in A) = \sum_{(x,y)\in A} f(x,y).$$

Expectations of functions of random vectors are computed just as with univariate random variables. Let $g(x,y)$ be a real-valued function defined for all possible values (x,y) of the discrete random vector (X,Y). Then $g(X,Y)$ is itself a random variable and its expected value $Eg(X,Y)$ is given by

$$Eg(X,Y) = \sum_{(x,y)\in\mathbb{R}^2} g(x,y)f(x,y).$$

Example 4.1.2 (Continuation of Example 4.1.2)

For the (X,Y) whose joint pmf is given in the following table

		X										
		2	3	4	5	6	7	8	9	10	11	12
	0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
	1		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$	
Y	2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		
	3				$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$			
	4					$\frac{1}{18}$		$\frac{1}{18}$				
	5						$\frac{1}{18}$					

Letting $g(x,y) = xy$, we have

$$EXY = (2)(0)\frac{1}{36} + \cdots + (7)(5)\frac{1}{18} = 13\frac{11}{18}.$$

The expectation operator continues to have the properties listed in Theorem 2.2.5 (textbook). For example, if $g_1(x,y)$ and $g_2(x,y)$ are two functions and a, b and c are constants, then

$$E(ag_1(X,Y) + bg_2(X,Y) + c) = aEg_1(X,Y) + bEg_2(X,Y) + c.$$

For any (x,y), $f(x,y) \geq 0$ since $f(x,y)$ is a probability. Also, since (X,Y) is certain to be in $\mathbb{R}^2$,

$$\sum_{(x,y)\in\mathbb{R}^2} f(x,y) = P((X,Y) \in \mathbb{R}^2) = 1.$$

Theorem 4.1.6

Let (X, Y) be a discrete bivariate random vector with joint pmf $f_{XY}(x, y)$. Then the marginal pmfs of X and Y, $f_X(x) = P(X = x)$ and $f_Y(y) = P(Y = y)$, are given by

$$f_X(x) = \sum_{y \in \mathbb{R}} f_{X,Y}(x, y) \quad \text{and} \quad f_Y(y) = \sum_{x \in \mathbb{R}} f_{X,Y}(x, y).$$

PROOF: For any $x \in \mathbb{R}$, let $A_x = \{(x, y) : -\infty < y < \infty\}$. That is, A_x is the line in the plane with first coordinate equal to x. Then, for any $x \in \mathbb{R}$,

$$\begin{aligned}
f_X(x) &= P(X = x) \\
&= P(X = x, -\infty < Y < \infty) \qquad (P(-\infty < Y < \infty) = 1) \\
&= P((X, Y) \in A_x) \qquad (\text{definition of } A_x) \\
&= \sum_{(x,y) \in A_x} f_{X,Y}(x, y) \\
&= \sum_{y \in \mathbb{R}} f_{X,Y}(x, y).
\end{aligned}$$

The proof for $f_Y(y)$ is similar. □

Example 4.1.7 (Marginal pmf for dice)

Using the table given in Example 4.1.4, compute the marginal pmf of Y. Using Theorem 4.1.6, we have

$$f_Y(0) = f_{X,Y}(2, 0) + \cdots + f_{X,Y}(12, 0) = \frac{1}{6}.$$

Similarly, we obtain

$$f_Y(1) = \frac{5}{18}, \quad f_Y(2) = \frac{2}{9}, \quad f_Y(3) = \frac{1}{6}, \quad f_Y(4) = \frac{1}{9}, \quad f_Y(5) = \frac{1}{18}.$$

Notice that $\sum_{i=0}^{5} f_Y(i) = 1$.

The marginal distributions of X and Y do not completely describe the joint distribution of X and Y. Indeed, there are many different joint distributions that have the same marginal distribution. Thus, it is hopeless to try to determine

the joint pmf from the knowledge of only the marginal pmfs. The next example illustrates the point.

Example 4.1.9 (Same marginals, different joint pmf)

Considering the following two joint pmfs,

$$f(0,0) = \frac{1}{12}, \quad f(1,0) = \frac{5}{12}, \quad , f(0,1) = f(1,1) = \frac{3}{12}, \quad f(x,y) = 0 \text{ for all other values.}$$

and

$$f(0,0) = f(0,1) = \frac{1}{6}, \quad f(1,0) = f(1,1) = \frac{1}{3}, \quad f(x,y) = 0 \quad \text{for all other values.}$$

It is easy to verify that they have the same marginal distributions. The marginal of X is

$$f_X(0) = \frac{1}{3}, \quad f_X(1) = \frac{2}{3}.$$

The marginal of Y is

$$f_Y(0) = \frac{1}{2}, \quad f_Y(1) = \frac{1}{2}.$$

In the following we consider random vectors whose components are continuous random variables.

Definition 4.1.10A function $f(x,y)$ from $\mathbb{R}^2$ into $\mathbb{R}$ is called a joint probability density function or joint pdf of the continuous bivariate random vector (X,Y) if, for every $A \subset \mathbb{R}^2$,

$$P((X,Y) \in A) = \int\int_A f(x,y)dxdy.$$

If $g(x,y)$ is a real-valued function, then the expected value of $g(X,Y)$ is defined to be

$$Eg(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy.$$

The marginal probability density functions of X and Y are defined as

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

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

Any function $f(x, y)$ satisfying $f(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^2$ and

$$1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy$$

is the joint pdf of some continuous bivariate random vector (X, Y).

Example 4.1.11 (Calculating joint probabilities-I)

Define a joint pdf by

$$f(x, y) = \begin{cases} 6xy^2 & 0 < x < 1 \text{ and } 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Now, consider calculating a probability such as $P(X + Y \geq 1)$. Let $A = \{(x, y) : x + y \geq 1\}$, we can re-express A as

$$A = \{(x, y) : x + y \geq 1, 0 < x < 1, 0 < y < 1\} = \{(x, y) : 1 - y \leq x < 1, 0 < y < 1\}.$$

Thus, we have

$$P(X + Y \geq 1) = \int_A \int f(x, y) dx dy = \int_0^1 \int_{1-y}^1 6xy^2 dx dy = \frac{9}{10}.$$

The joint cdf is the function $F(x, y)$ defined by

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s, t) dt ds.$$

Hence,

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$$

and

$$-\frac{\partial^2 P(X \leq x, Y \geq y)}{\partial x \partial y} = f(x, y)$$