

Joint Probability Distribution

Joint Probability Mass Function for Discrete Random Variable:

The function $f(x, y)$ is a **joint probability distribution** or **probability mass function** of the discrete random variables X and Y if

1. $f(x, y) \geq 0$ for all (x, y) ,
2. $\sum_x \sum_y f(x, y) = 1$,
3. $P(X = x, Y = y) = f(x, y)$.

For any region A in the xy plane, $P[(X, Y) \in A] = \sum_A f(x, y)$.

Joint Density Function for Continuous Random Variable:

The function $f(x, y)$ is a **joint density function** of the continuous random variables X and Y if

1. $f(x, y) \geq 0$, for all (x, y) ,
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$,
3. $P[(X, Y) \in A] = \int \int_A f(x, y) dx dy$, for any region A in the xy plane.

Marginal Probability Distributions:

The **marginal distributions** of X alone and of Y alone are

$$g(x) = \sum_y f(x, y) \quad \text{and} \quad h(y) = \sum_x f(x, y)$$

for the discrete case, and

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

for the continuous case.

Example 3.14:

Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If X is the number of blue pens selected and Y is the number of red pens selected, find

- (a) the joint probability function $f(x, y)$,
- (b) $P[(X, Y) \in A]$, where A is the region $\{(x, y) | x + y \leq 1\}$.

Show that the column and row totals of Table 3.1 give the marginal distribution of X alone and of Y alone.

Solution:

The possible pairs of values (x, y) are $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, $(0, 2)$, and $(2, 0)$.

- (a) Now, $f(0, 1)$, for example, represents the probability that a red and a green pens are selected. The total number of equally likely ways of selecting any 2 pens from the 8 is $\binom{8}{2} = 28$. The number of ways of selecting 1 red from 2 red pens and 1 green from 3 green pens is $\binom{2}{1} \binom{3}{1} = 6$. Hence, $f(0, 1) = 6/28 = 3/14$. Similar calculations yield the probabilities for the other cases, which are presented in Table 3.1. Note that the probabilities sum to 1. In Chapter

5, it will become clear that the joint probability distribution of Table 3.1 can be represented by the formula

$$f(x, y) = \frac{\binom{3}{x} \binom{2}{y} \binom{3}{2-x-y}}{\binom{8}{2}},$$

for $x = 0, 1, 2$; $y = 0, 1, 2$; and $0 \leq x + y \leq 2$.

- (b) The probability that (X, Y) fall in the region A is

$$\begin{aligned} P[(X, Y) \in A] &= P(X + Y \leq 1) = f(0, 0) + f(0, 1) + f(1, 0) \\ &= \frac{3}{28} + \frac{3}{14} + \frac{9}{28} = \frac{9}{14}. \end{aligned}$$

Table 3.1: Joint Probability Distribution for Example 3.14

$f(x, y)$		x			Row Totals
		0	1	2	
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

For the random variable X , we see that

$$g(0) = f(0,0) + f(0,1) + f(0,2) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$

$$g(1) = f(1,0) + f(1,1) + f(1,2) = \frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28},$$

and

$$g(2) = f(2,0) + f(2,1) + f(2,2) = \frac{3}{28} + 0 + 0 = \frac{3}{28},$$

which are just the column totals of Table 3.1. In a similar manner we could show that the values of $h(y)$ are given by the row totals. In tabular form, these marginal distributions may be written as follows:

x	0	1	2
$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$

y	0	1	2
$h(y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$

Example 3.15:

A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day, let X and Y , respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y), & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

- Verify condition 2 of Definition 3.9.
- Find $P[(X, Y) \in A]$, where $A = \{(x, y) \mid 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$.

Find $g(x)$ and $h(y)$ for the joint density function of Example 3.15.

Solution:

- The integration of $f(x, y)$ over the whole region is

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy &= \int_0^1 \int_0^1 \frac{2}{5}(2x + 3y) \, dx \, dy \\ &= \int_0^1 \left(\frac{2x^2}{5} + \frac{6xy}{5} \right) \Big|_{x=0}^{x=1} dy \\ &= \int_0^1 \left(\frac{2}{5} + \frac{6y}{5} \right) dy = \left(\frac{2y}{5} + \frac{3y^2}{5} \right) \Big|_0^1 = \frac{2}{5} + \frac{3}{5} = 1. \end{aligned}$$

(b) To calculate the probability, we use

$$\begin{aligned}
 P[(X, Y) \in A] &= P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right) \\
 &= \int_{1/4}^{1/2} \int_0^{1/2} \frac{2}{5}(2x + 3y) \, dx \, dy \\
 &= \int_{1/4}^{1/2} \left(\frac{2x^2}{5} + \frac{6xy}{5}\right) \Big|_{x=0}^{x=1/2} dy = \int_{1/4}^{1/2} \left(\frac{1}{10} + \frac{3y}{5}\right) dy \\
 &= \left(\frac{y}{10} + \frac{3y^2}{10}\right) \Big|_{1/4}^{1/2} \\
 &= \frac{1}{10} \left[\left(\frac{1}{2} + \frac{3}{4}\right) - \left(\frac{1}{4} + \frac{3}{16}\right)\right] = \frac{13}{160}.
 \end{aligned}$$

By definition,

$$g(x) = \int_{-\infty}^{\infty} f(x, y) \, dy = \int_0^1 \frac{2}{5}(2x + 3y) \, dy = \left(\frac{4xy}{5} + \frac{6y^2}{10}\right) \Big|_{y=0}^{y=1} = \frac{4x + 3}{5},$$

for $0 \leq x \leq 1$, and $g(x) = 0$ elsewhere. Similarly,

$$h(y) = \int_{-\infty}^{\infty} f(x, y) \, dx = \int_0^1 \frac{2}{5}(2x + 3y) \, dx = \frac{2(1 + 3y)}{5},$$

for $0 \leq y \leq 1$, and $h(y) = 0$ elsewhere.

Example 3.16:

Show that the column and row totals of Table 3.1 give the marginal distribution of X alone and of Y alone.

For the random variable X , we see that

$$\begin{aligned}
 g(0) &= f(0, 0) + f(0, 1) + f(0, 2) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14}, \\
 g(1) &= f(1, 0) + f(1, 1) + f(1, 2) = \frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28},
 \end{aligned}$$

and

$$g(2) = f(2, 0) + f(2, 1) + f(2, 2) = \frac{3}{28} + 0 + 0 = \frac{3}{28},$$

which are just the column totals of Table 3.1. In a similar manner we could show that the values of $h(y)$ are given by the row totals. In tabular form, these marginal distributions may be written as follows:

x	0	1	2
$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$

y	0	1	2
$h(y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$

Example 3.17:

Find $g(x)$ and $h(y)$ for the joint density function of Example 3.15.

By definition,

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 \frac{2}{5}(2x + 3y) dy = \left(\frac{4xy}{5} + \frac{6y^2}{10} \right) \Big|_{y=0}^{y=1} = \frac{4x + 3}{5},$$

for $0 \leq x \leq 1$, and $g(x) = 0$ elsewhere. Similarly,

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_0^1 \frac{2}{5}(2x + 3y) dx = \frac{2(1 + 3y)}{5},$$

for $0 \leq y \leq 1$, and $h(y) = 0$ elsewhere. J

Conditional Distributions:

Let X and Y be two random variables, discrete or continuous. The **conditional distribution** of the random variable Y given that $X = x$ is

$$f(y|x) = \frac{f(x, y)}{g(x)}, \text{ provided } g(x) > 0.$$

Similarly, the conditional distribution of X given that $Y = y$ is

$$f(x|y) = \frac{f(x, y)}{h(y)}, \text{ provided } h(y) > 0.$$

Example 3.18:

Referring to Example 3.14, find the conditional distribution of X , given that $Y = 1$, and use it to determine $P(X = 0 | Y = 1)$.

Solution:

We need to find $f(x|y)$, where $y = 1$. First, we find that

$$h(1) = \sum_{x=0}^2 f(x, 1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}.$$

Now

$$f(x|1) = \frac{f(x, 1)}{h(1)} = \left(\frac{7}{3} \right) f(x, 1), \quad x = 0, 1, 2.$$

Therefore,

$$f(0|1) = \left(\frac{7}{3}\right) f(0,1) = \left(\frac{7}{3}\right) \left(\frac{3}{14}\right) = \frac{1}{2}, \quad f(1|1) = \left(\frac{7}{3}\right) f(1,1) = \left(\frac{7}{3}\right) \left(\frac{3}{14}\right) = \frac{1}{2},$$

$$f(2|1) = \left(\frac{7}{3}\right) f(2,1) = \left(\frac{7}{3}\right) (0) = 0,$$

and the conditional distribution of X , given that $Y = 1$, is

x	0	1	2
$f(x 1)$	$\frac{1}{2}$	$\frac{1}{2}$	0

Finally,

$$P(X = 0 | Y = 1) = f(0|1) = \frac{1}{2}.$$

Therefore, if it is known that 1 of the 2 pen refills selected is red, we have a probability equal to $1/2$ that the other refill is not blue. ▮

Example 3.19:

The joint density for the random variables (X, Y) , where X is the unit temperature change and Y is the proportion of spectrum shift that a certain atomic particle produces, is

$$f(x, y) = \begin{cases} 10xy^2, & 0 < x < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

- (a) Find the marginal densities $g(x)$, $h(y)$, and the conditional density $f(y|x)$.
- (b) Find the probability that the spectrum shifts more than half of the total observations, given that the temperature is increased by 0.25 unit.
- (a) By definition,

Solution:

- (a) By definition,

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_x^1 10xy^2 dy$$

$$= \frac{10}{3} xy^3 \Big|_{y=x}^{y=1} = \frac{10}{3} x(1 - x^3), \quad 0 < x < 1,$$

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_0^y 10xy^2 dx = 5x^2 y^2 \Big|_{x=0}^{x=y} = 5y^4, \quad 0 < y < 1.$$

Now

$$f(y|x) = \frac{f(x, y)}{g(x)} = \frac{10xy^2}{\frac{10}{3}x(1 - x^3)} = \frac{3y^2}{1 - x^3}, \quad 0 < x < y < 1.$$

- (b) Therefore,

$$P\left(Y > \frac{1}{2} \mid X = 0.25\right) = \int_{1/2}^1 f(y | x = 0.25) dy = \int_{1/2}^1 \frac{3y^2}{1 - 0.25^3} dy = \frac{8}{9}. \quad \text{▮}$$

Statistically Independent:

Let X and Y be two random variables, discrete or continuous, with joint probability distribution $f(x, y)$ and marginal distributions $g(x)$ and $h(y)$, respectively. The random variables X and Y are said to be **statistically independent** if and only if

$$f(x, y) = g(x)h(y)$$

for all (x, y) within their range.

Example 3.21:

Show that the random variables of Example 3.14 are not statistically independent. Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If X is the number of blue pens selected and Y is the number of red pens selected, find

- (a) the joint probability function $f(x, y)$,
- (b) $P[(X, Y) \in A]$, where A is the region $\{(x, y) | x + y \leq 1\}$.

Solution:

Table 3.1: Joint Probability Distribution for Example 3.14

$f(x, y)$		x			Row
		0	1	2	Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Let us consider the point $(0, 1)$. From Table 3.1 we find the three probabilities $f(0, 1)$, $g(0)$, and $h(1)$ to be

$$f(0, 1) = \frac{3}{14},$$

$$g(0) = \sum_{y=0}^2 f(0, y) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$

$$h(1) = \sum_{x=0}^2 f(x, 1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}.$$

Clearly,

$$f(0, 1) \neq g(0)h(1),$$

and therefore X and Y are not statistically independent.

Practice Problems from Book

3.37 Determine the values of c so that the following functions represent joint probability distributions of the random variables X and Y :

- (a) $f(x, y) = cxy$, for $x = 1, 2, 3$; $y = 1, 2, 3$;
- (b) $f(x, y) = c|x - y|$, for $x = -2, 0, 2$; $y = -2, 3$.

3.38 If the joint probability distribution of X and Y is given by

$$f(x, y) = \frac{x + y}{30}, \quad \text{for } x = 0, 1, 2, 3; \quad y = 0, 1, 2,$$

- find
- (a) $P(X \leq 2, Y = 1)$;
 - (b) $P(X > 2, Y \leq 1)$;
 - (c) $P(X > Y)$;
 - (d) $P(X + Y = 4)$.

3.40 A fast-food restaurant operates both a drive-through facility and a walk-in facility. On a randomly selected day, let X and Y , respectively, be the proportions of the time that the drive-through and walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x, y) = \begin{cases} \frac{2}{3}(x + 2y), & 0 \leq x \leq 1, \quad 0 \leq y \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

- (a) Find the marginal density of X .
- (b) Find the marginal density of Y .
- (c) Find the probability that the drive-through facility is busy less than one-half of the time.

3.41 A candy company distributes boxes of chocolates with a mixture of creams, toffees, and cordials. Suppose that the weight of each box is 1 kilogram, but the individual weights of the creams, toffees, and cordials vary from box to box. For a randomly selected box, let X and Y represent the weights of the creams and the toffees, respectively, and suppose that the joint density function of these variables is

$$f(x, y) = \begin{cases} 24xy, & 0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

- Find the probability that in a given box the cordials account for more than $1/2$ of the weight.
- Find the marginal density for the weight of the creams.
- Find the probability that the weight of the toffees in a box is less than $1/8$ of a kilogram if it is known that creams constitute $3/4$ of the weight.

3.42 Let X and Y denote the lengths of life, in years, of two components in an electronic system. If the joint density function of these variables is

$$f(x, y) = \begin{cases} e^{-(x+y)}, & x > 0, y > 0, \\ 0, & \text{elsewhere,} \end{cases}$$

find $P(0 < X < 1 \mid Y = 2)$.

3.43 Let X denote the reaction time, in seconds, to a certain stimulus and Y denote the temperature ($^{\circ}\text{F}$) at which a certain reaction starts to take place. Suppose that two random variables X and Y have the joint density

$$f(x, y) = \begin{cases} 4xy, & 0 < x < 1, 0 < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find

- $P(0 \leq X \leq \frac{1}{2} \text{ and } \frac{1}{4} \leq Y \leq \frac{1}{2})$;
- $P(X < Y)$.

3.44 Each rear tire on an experimental airplane is supposed to be filled to a pressure of 40 pounds per square inch (psi). Let X denote the actual air pressure for the right tire and Y denote the actual air pressure for the left tire. Suppose that X and Y are random variables with the joint density function

$$f(x, y) = \begin{cases} k(x^2 + y^2), & 30 \leq x < 50, 30 \leq y < 50, \\ 0, & \text{elsewhere.} \end{cases}$$

- (a) Find k .
- (b) Find $P(30 \leq X \leq 40 \text{ and } 40 \leq Y < 50)$.
- (c) Find the probability that both tires are underfilled.

3.49 Let X denote the number of times a certain numerical control machine will malfunction: 1, 2, or 3 times on any given day. Let Y denote the number of times a technician is called on an emergency call. Their joint probability distribution is given as

$f(x, y)$		x		
		1	2	3
y	1	0.05	0.05	0.10
	3	0.05	0.10	0.35
	5	0.00	0.20	0.10

- (a) Evaluate the marginal distribution of X .
- (b) Evaluate the marginal distribution of Y .
- (c) Find $P(Y = 3 \mid X = 2)$.

3.50 Suppose that X and Y have the following joint probability distribution:

$f(x, y)$		x	
		2	4
y	1	0.10	0.15
	3	0.20	0.30
	5	0.10	0.15

- (a) Find the marginal distribution of X .
- (b) Find the marginal distribution of Y .

3.54 Determine whether the two random variables of Exercise 3.49 are dependent or independent.

3.55 Determine whether the two random variables of Exercise 3.50 are dependent or independent.