Joint probability distribution

When we have two or more random variables, and we have reasons to believe that these variables are related, then we have to study them together. The probability distribution of two or more random variables is called a joint probability distribution.

We will study joint distribution of two random variables only.

Discrete Case:

Joint pmf

Let X and Y be two discrete random variables. The joint probability mass function or joint probability function or joint pmf of X and Y, denoted by $p(x, y)$, is defined as $p(x, y) = P(X = x, Y = y)$. It satisfies the following two conditions:

(i) $p(x, y) \geq 0$

(ii) $\sum_x \sum_y p(x, y) = 1$

Example:

Let $X =$ Number of houses, and $Y =$ Number of cars. Let the joint pmf of X and Y be given as follows:

x	y			
	0	1	2	3
0	0.30	0.15	0.10	0.05
1	0.10	0.10	0.05	0.05
2	0.01	0.03	0.04	0.02

Here,

$p(0, 1) = P(X = 0, Y = 1) = 0.15$,

$p(2, 3) = P(X = 2, Y = 3) = 0.02$,

and so on.

Marginal distributions

The (marginal) distribution of X, denoted by $p_X(x)$, is given by:

$$p_X(x) = P(X = x) = \sum_y p(x, y)$$

The (marginal) distribution of Y, denoted by $p_Y(y)$, is given by:

$$p_Y(y) = P(Y = y) = \sum_x p(x, y)$$

Exercise:

Determine the marginal distribution of X from the above example.

Solution:

x	0	1	2
$p_X(x)$	0.60	0.30	0.10

$p_X(1) = P(X = 1) = 0.30$

and so on.

Note:

Marginal distributions of X and Y are shown in the margins of the following table (in yellow and blue colors, respectively).

x	y				$p_X(x)$
	0	1	2	3	
0	0.30	0.15	0.10	0.05	0.60
1	0.10	0.10	0.05	0.05	0.30
2	0.01	0.03	0.04	0.02	0.10
$p_Y(y)$	0.41	0.28	0.19	0.12	1.00

Conditional distributions:

The conditional distribution of Y given $X = x$, denoted by $p_{Y|X}(y|x)$, is given by:

$$p_{Y|X}(y|x) = P(Y = y \mid X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{p(x, y)}{p_X(x)}$$

Exercise:

For the above joint distribution, determine the conditional distribution of Y given $X = 1$.

Solution:

y	0	1	2	3
$p_{Y \mid X}(y \mid 1)$	1/3	1/3	1/6	1/6

Here,

$$p_{Y \mid X}(0 \mid 1) = \frac{0.10}{0.30} = 1/3$$

and so on.

Continuous Case: **Joint pdf**

Let X and Y be two continuous random variables. The joint probability density function or joint density or joint pdf of X and Y, denoted by $f(x, y)$, is a function that gives density at $X = x$ and $Y = y$. It satisfies the following two conditions:

(i) $f(x, y) \geq 0$

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

We can calculate probabilities by using the following formula:

$$P(a < X < b,\ \ c < Y < d) = \int_a^b \int_c^d f(x, y)\, dy\, dx$$

Example:

Check whether the following function is a joint pdf:

$$f(x, y) = \frac{1}{3}(x + y),\ \ 0 < x < 2,\ \ 0 < y < 1$$

Solution:

i)

$f(x,y) \geq 0$ for $0 < x < 2, 0 < y < 1$.

ii)

$$\int_0^2 \int_0^1 \frac{1}{3}(x+y)\, dy\, dx$$

$$= \int_0^2 \frac{1}{3}\left(xy + \frac{y^2}{2}\right)\Big|_0^1 dx$$

$$= \int_0^2 \frac{1}{3}\left(x + \frac{1}{2}\right) dx$$

$$= 1$$

Therefore, $f(x,y)$ is a joint pdf.

Exercise

Consider the following joint pdf:

$$f(x,y) = \frac{1}{3}(x+y), \;\; 0 < x < 2, \;\; 0 < y < 1.$$

Determine (a) $P(0.6 < X < 1.2, \;0.2 < Y < 0.4)$ (b) $P(0.6 < X < 1.2)$ (c) density at $X = 0.5, Y = 0.4$ (d) $P(X = 0.5, \; Y = 0.4)$ (e) $P(X = 0.5, 0.3 < Y < 0.5)$.

Solution:

(a)

$P(0.6 < X < 1.2, \;\; 0.2 < Y < 0.4)$

$$= \int_{0.6}^{1.2} \int_{0.2}^{0.4} \frac{1}{3}(x+y)\, dy\, dx$$

$$= 0.048$$

(b)

$P(0.6 < X < 1.2)$

$= \int_{0.6}^{1.2} \int_{0}^{1} \frac{1}{3}(x + y)\, dy\, dx$

$= 0.28$

(c)

Density at $X = 0.5, Y = 0.4$

$= f(0.5, 0.4)$

$= \frac{1}{3}(0.5 + 0.4)$

$= 0.3$

(d)

$P(X = 0.5, Y = 0.4) = 0$

(e)

$P(X = 0.5, 0.3 < Y < 0.5) = 0$

Marginal distributions

The (marginal) distribution or density of X, denoted by $f_X(x)$, is given by:

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

The (marginal) distribution or density of Y, denoted by $f_Y(y)$, is given by:

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

Exercise:

$f(x, y) = x + y, \;\; 0 < x < 1, \;\; 0 < y < 1$. Determine the marginal density of X.

Solution:

$f_X(x)$

$$= \int_0^1 (x + y)\, dy$$

$$= x + 0.5, \;\; 0 < x < 1$$

Conditional distributions:

The conditional distribution of Y given $X = x$, denoted by $f_{Y|X}(y|x)$, is given by:

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

Exercise:

For the above joint distribution, determine the conditional distribution of $f_{Y|X}(y|0.25)$.

Solution:

$f_X(x) = x + 0.5$

$\therefore f_X(0.25) = 0.25 + 0.5 = 0.75$

$f_{Y|X}(y|0.25)$

$$= \frac{f(0.25, y)}{f_X(0.25)}$$

$$= \frac{y + 0.25}{0.75}$$

$$= \frac{1}{3}(4y + 1), \;\; 0 < y < 1$$