1

Marginal Distribution :

We have noted earlier that if the joint probability function $f(x,y)$ of two random variables X and Y is known, the probability functions of X and Y can be ~~der~~ derived separately.

In the case when the distribution of X is derived from $f(x,y)$, the resulting distribution is called marginal distribution of X.

Similarly the probability density of Y derived from the joint distribution of X and Y is known as the marginal distribution of Y.

Here, we will denote these two distributions by $g(x)$ and $h(x)$, respectively, when they are derived from joint distribution $f(x,y)$.

Thus, given the joint probability $f(x,y)$ of two discrete random variables X and Y, the probability distribution of X alone is

$$g(x) = \sum_y f(x,y)$$

and that of Y alone is

$$h(y) = \sum_x f(x,y)$$

These two are the marginal distributions of two discrete random variables X and Y, respectively.

Now, when X and Y are continuous

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \,; \quad \text{for } -\infty < x < \infty$$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \,; \quad \text{for } -\infty < y < \infty$$

Marginal distributions are indeed probability distributions, since they satisfy all the properties of a probability distribution.

For continuous case,

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

and $P[a < X < b] = P[a < X < b, -\infty < Y < \infty]$

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

Defⁿ :

Let, $F(x, y)$ be the joint CDF of (X, Y). Then the marginal CDF of X is

$$F_1(x) = \lim_{y \to \infty} F(x, y).$$

Similarly, the marginal CDF of Y is

$$F_2(y) = \lim_{x \to \infty} F(x, y).$$

The probability function or PDF of X associated with the marginal CDF of X is called the marginal P.F or marginal PDF of X.

The marginal probability function and marginal PDF can be obtained from joint probability function and joint PDF, respectively.

Example: Suppose X and Y have the following joint probability distribution

	X				
Y	0	1	2	3	Row Sum
0	0	1/8	2/8	1/8	4/8
1	1/8	2/8	1/8	0	4/8
Column Sum	1/8	3/8	3/8	1/8	1

Find the marginal distributions of X and Y.

Solⁿ: For the random variable X

$$g(0) = P(X=0) = \sum_{y=0}^{1} f(0,y) = f(0,0) + f(0,1)$$

$$= 0 + \frac{1}{8} = \frac{1}{8}$$

$$g(1) = P(X=1) = \sum_{y=0}^{1} f(1,y) = f(1,0) + f(1,1)$$

$$= \frac{1}{8} + \frac{2}{8} = 3/8$$

$$g(2) = P(X=2) = \sum_{y=0}^{1} f(2,y) = f(2,0) + f(2,1)$$

$$= \frac{2}{8} + \frac{1}{8} = 3/8$$

$$g(3) = P(X=3) = \sum_{y=0}^{1} f(3,y) = f(3,0) + f(3,1) = \frac{1}{8} + 0 = \frac{1}{8}$$

Similarly for Y

$$h(0) = P(Y=0) = \sum_{x=0}^{3} f(x,0) = f(0,0) + f(1,0) + f(2,0) + f(3,0)$$

$$= 0 + \frac{1}{8} + \frac{2}{8} + \frac{1}{8} = 4/8$$

$$h(1) = P(Y=1) = \sum_{x=0}^{3} f(x,1) = f(0,1) + f(1,1) + f(2,1) + f(3,1)$$

$$= \frac{1}{8} + \frac{2}{8} + \frac{1}{8} + 0 = 4/8$$

The marginal distributions derived above may also be represented as follows:

Marginal distribution of X:

x	0	1	2	3	Sum
$f(x)$	1/8	3/8	3/8	1/8	1

Marginal distribution of Y:

y	0	1	Sum
$f(y)$	4/8	4/8	1

The above two columns show that marginal distributions are necessarily probability distributions.

Example: Find the marginal densities of X and Y from the following joint density function and verify that marginal distributions are also probability distributions.

$$f(x,y) = \frac{1}{8}(6-x-y), \text{ for } 0<x<2,\ 2<y<4$$
$$= 0, \text{ otherwise}$$

Also compute $P(X+Y<3)$ and $P(X<1.5, Y<2.5)$

Solⁿ: By definition, the marginal density of X is

$$g(x) = \frac{1}{8}\int_2^4 f(x,y)\,dy$$

$$= \frac{1}{8}\int_2^4 (6-x-y)\,dy$$

$$= \frac{1}{8}\left[6y - xy - \frac{y^2}{2}\right]_2^4$$

$$= \frac{1}{8}\left[(24-4x-8) - (12-2x-2)\right]$$

$$= \frac{1}{8}(16-4x-10+2x)$$

$$= \frac{1}{4}(3-x), \text{ for } 0<x<2.$$

Similarly, the marginal density of Y is

$$h(y) = \frac{1}{8}\int_0^2 f(x,y)\,dx$$

$$= \frac{1}{8}\int_{x=0}^{2} (6-x-y)\,dx$$

$$= \frac{1}{8}\left[6x - \frac{x^2}{2} - xy\right]_{x=0}^{2} = \frac{1}{8}\left[12 - 2 - 2y\right]$$

$$\Rightarrow h(y) = \frac{1}{4}(5-y), \quad \text{for } 2<y<4$$

Now, we need to verify that $g(x)$ and $h(y)$ are probability distributions.

It is clear that in the given range of the variables X and Y; $g(x) \geq 0$ & $h(y) \geq 0$

Also, $\displaystyle\int_0^2 g(x)\,dx = \frac{1}{4}\int_0^2 (3-x)\,dx = \frac{1}{4}\left[3x - \frac{x^2}{2}\right]_0^2 = 1$

& $\displaystyle\int_2^4 h(y)\,dy = \frac{1}{4}\int_2^4 (5-y)\,dy = \frac{1}{4}\left[5y - \frac{y^2}{2}\right]_2^4 = 1$

The $g(x)$ & $h(y)$ thus satisfy all the conditions of a density function.

Now,

$$P(X+Y<3) = \frac{1}{8}\int_{x=0}^{2}\int_{y=2}^{3-x} (6-x-y)\,dx\,dy$$

$$= \frac{1}{8}\int_{x=0}^{2}\left[\int_{y=2}^{3-x}(6-x-y)\,dy\right]dx$$

$$= \frac{1}{8}\int_{x=0}^{2}\left[6y - xy - \frac{y^2}{2}\right]_{2}^{3-x} dx$$

$$= \frac{1}{8}\int_{x=0}^{2}\left[\left\{6(3-x) - x(3-x) - \frac{(3-x)^2}{2}\right\} - \left\{12 - 2x - 2\right\}\right] dx$$

$$= \frac{1}{8}\int_{x=0}^{2}\left(18 - 6x - 3x + x^2 - \frac{9 - 6x + x^2}{2} - 10 + 2x\right) dx$$

$$= \frac{1}{8}\int_{0}^{2}\left(8 - 7x + x^2 - \frac{9}{2} + 3x - \frac{x^2}{2}\right) dx$$

$$= \frac{1}{8}\int_{0}^{2}\left(\frac{x^2}{2} - 4x + \frac{7}{2}\right) dx = \frac{1}{24}$$

Scanned with CamScanner

and,

$$P\left(X < 3/2,\ Y < 5/2\right) = \frac{1}{8}\int_{x=0}^{3/2}\int_{y=2}^{5/2} (6-x-y)\,dy\,dx$$

$$= \frac{1}{8}\int_{x=0}^{3/2} \left[6y - xy - \frac{y^2}{2}\right]_2^{5/2} dx$$

$$= \frac{1}{8}\int_{x=0}^{3/2} \left(\frac{15}{8} - \frac{x}{2}\right) dx$$

$$= \frac{1}{8}\left[\frac{15x}{8} - \frac{x^2}{4}\right]_0^{3/2}$$

$$= \frac{1}{8}\left(\frac{45}{16} - \frac{9}{16}\right)$$

$$= \frac{1}{\cancel{8}_{2}} \times \frac{\cancel{36}^{9}}{16} = \frac{9}{32}$$

$$0$$

Example: The joint probability function of a discrete bivariate random variable is as follows.

$$f(x,y) = \begin{cases} 1/36 & \text{for } 1 \leq x = y \leq 6 \\ \frac{2}{36} & \text{for } 1 \leq y < x \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

Find the marginal probability functions of X & Y.

Sol^n: The marginal probability function of X is

$$g(x) = \sum_{y} f(x,y)$$

$$= \sum_{y=1}^{6} f(x,y)$$

$$= f(x,x) + \sum_{y<x}^{6} f(x,y) + \sum_{y>x}^{6} f(x,y)$$

$$= 1/36 + (x-1)\frac{2}{36} + 0$$

$$= \frac{2x-1}{36}$$

Thus,

$$g(x) = \begin{cases} \frac{2x-1}{36}; & \text{for } x = 1,2,3,4,5,6 \\ 0; & \text{otherwise} \end{cases}$$

Marginal ~~distribution~~ probability function of Y is

$$h(y) = \sum_{x} f_{x}(x,y)$$

$$= \sum_{x=1}^{6} f(x,y)$$

$$= f(y,y) + \sum_{x<y}^{6} f(x,y) + \sum_{x>y}^{6} f(x,y)$$

$$= \frac{1}{36} + 0 + (6-y)\frac{2}{36}$$

$$= \frac{13-2y}{36}$$

Thus,

$$h(y) = \begin{cases} \frac{13-2y}{36}; & \text{for } y = 1,2,3,4,5,6 \\ 0; & \text{elsewhere} \end{cases}$$

Bivariate cumulative distribution functions:

The joint distribution function or joint cumulative distribution function (joint CDF) F of two dimensional random variable (X, Y) is defined as a function

$$F(x, y) = P(X \leq x \text{ and } Y \leq y); \quad -\infty < x < \infty, \text{ and } -\infty < y < \infty$$

$F(x, y)$ is a monotonic increasing function in x for each fixed y and is a monotonic increasing function in y for each fixed x.

Here it is noted that the CDF of random variable X is $F_1(x) = \lim_{y \to \infty} F(x, y)$

The CDF of random variable Y is

$$F_2(x) = \lim_{x \to \infty} F(x, y)$$

If the bivariate random variable (X, Y) is discrete, then

$$F(x, y) = P(X \leq x \text{ and } Y \leq y)$$

$$= \sum_{r \leq x} \sum_{s \leq y} F(r, s) \text{ for } -\infty < x < \infty \text{ and } -\infty < y < \infty$$

If the bivariate random variable (X, Y) is continuous, then

$$F(x,y) = P(X \leq x \text{ and } Y \leq y)$$

$$= \int_{-\infty}^{y} \int_{-\infty}^{x} f(r,s)\, dr\, ds$$

for $-\infty < x < \infty$ and $-\infty < y < \infty$

If the bivariate random variable (X, Y) is continuous, then

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

Note:

For given numbers $a < b$ and $c < d$

$$P(a \leq X \leq b \text{ and } c \leq Y \leq d)$$

$$= P(a \leq X \leq b \text{ and } Y \leq d) - P(a \leq X \leq b \text{ and } Y \leq c)$$

$$= P(X \leq b \text{ and } Y \leq d) - P(X \leq a \text{ and } Y \leq d)$$
$$- P(X \leq b \text{ and } Y \leq c) + P(X \leq a \text{ and } Y \leq c)$$

$$= F(b,d) - F(a,d) - F(b,c) + F(a,c)$$

Example: The joint PDF of (X, Y) is

$$f(x,y) = \begin{cases} \frac{1}{8}(x+y) \; ; \text{ for } 0<x<2 \text{ and } 0<y<2 \\ 0 \; ; \text{ otherwise} \end{cases}$$

Find its joint CDF.

Sol^n: The joint CDF of (X,Y) is as follows:

For $x<0$ or $y<0$, $F(x,y) = 0$

For $0 \le x < 2$ or $0 \le y < 2$

$$F(x,y) = \int_0^x \int_0^y \frac{1}{8}(x+y)\,dy\,dx$$

$$= \frac{1}{8}\int_0^x \left[xy + \frac{y^2}{2}\right] dx$$

$$= \frac{1}{8}\left[\frac{x^2}{2}y + \frac{xy^2}{2}\right]_0^x$$

$$\Rightarrow F(x,y) = \frac{1}{16}xy(x+y)$$

For $0 \le x < 2$ and $y \geq 2$

$$F(x,y) = \int_0^x \int_0^2 \frac{1}{8}(x+y)\,dy\,dx$$

$$= \frac{1}{8}\int_0^x \left[xy + \frac{y^2}{2}\right]_0^2 dx$$

$$= \frac{1}{4}\int_0^x (x+1)\,dx$$

$$= \frac{1}{4}\left[\frac{x^2}{2} + x\right]_0^x$$

$$= \frac{1}{8}\left(x^2 + 2x\right) = \frac{1}{8}x(x+2)$$

For $x \geq 2$ and $0 \leq y < 2$

$$F(x,y) = \int_0^2 \int_0^y \frac{1}{8}(x+y)\,dy\,dx$$

$$= \frac{1}{8}\int_0^2 \left[xy + \frac{y^2}{2}\right]_0^y dx$$

$$= \frac{1}{8}\int_0^2 \left[xy + \frac{y^2}{2}\right] dx$$

$$= \frac{1}{8}\left[\frac{x^2}{2}y + \frac{xy^2}{2}\right]_0^2$$

$$= \frac{1}{16}\left[4y + 2y^2\right]$$

$$\therefore F(x,y) = \frac{1}{8}y(y+2)$$

For $x \geq 2$ and $y \geq 2$

$$F(x, y) = 1.$$

Thus the complete joint CDF of (X, Y) is given by

$$F(x,y) = \begin{cases} 0; & \text{for } x<0 \text{ or } y<0 \\ \frac{1}{16}xy(x+y); & \text{for } 0 \leq x < 2 \;\&\; 0 \leq y < 2 \\ \frac{1}{8}x(x+2); & \text{for } 0 \leq x < 2 \text{ and } y \geq 2 \\ \frac{1}{8}y(y+2); & \text{for } x \geq 2 \text{ and } 0 \leq y < 2 \\ 1; & \text{for } x \geq 2 \text{ and } y \geq 2 \end{cases}$$

Example: The joint CDF for water and electric demand (X, Y) is

$$F(x,y) = \begin{cases} 0; & \text{for } x<4 \text{ or } y<2 \\ \frac{(x-4)(y-2)}{19208}; & \text{for } 4 \leq x < 200 \text{ and } 2 \leq y < 100 \\ \frac{x-4}{196}; & \text{for } 4 \leq x < 200 \text{ and } y \geq 100 \\ \frac{y-2}{98}; & \text{for } x \geq 200 \text{ and } 2 \leq y < 100 \\ 1; & \text{for } x \geq 200 \text{ and } y \geq 100. \end{cases}$$

(a) Find the joint PDF of (X, Y);

(b) $P(50 \le X \le 100,\ 50 \le Y \le 75)$;

(c) The CDF of X;

(d) The CDF of Y;

Sol^n: (a) For $4 \le x \le 200$ and $2 \le y \le 100$

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

$$= \frac{\partial^2}{\partial x \partial y}\left(\frac{(x-4)(y-2)}{19208}\right)$$

$$= \frac{\partial}{\partial x}\left\{\frac{(x-4)}{19208}\right\}$$

$$= \frac{1}{19208}$$

Thus, $f(x,y) = \begin{cases} \dfrac{1}{19208} & ; \text{ for } 4 \le x \le 200 \text{ and } 2 \le y \le 100 \\ 0, & \text{otherwise} \end{cases}$

Ans

(b) $P(50 \le X \le 100,\ 50 \le Y \le 75)$

$$= F(100,75) - F(50,75) - F(100,50) + F(50,50)$$

$$= \left\{\frac{(100-4)(75-2)}{19208}\right\} - \left\{\frac{(50-4)(75-2)}{19208}\right\}$$

$$- \left\{\frac{(100-4)(50-2)}{19208}\right\} + \left\{\frac{(50-4)(50-2)}{19208}\right\}$$

$$= \frac{625}{9604} \quad \text{Ans}$$

(c) If $4 \le x \le 200$ and $y \ge 100$, then $F(x,y) = F(x,100)$, and it follows that

$$F(x,y) = \frac{x-4}{196}$$

Thus by letting $y \to \infty$, the CDF of X is

$$F_1(x) = \begin{cases} 0; & \text{for } x < 4 \\ \frac{x-4}{196}; & \text{for } 4 \le x < 200 \\ 1; & \text{for } x \ge 200 \end{cases}$$

(d) If $2 \le y < 100$ and $x \geq 200$, then $F(x,y) = F(200, y)$ and it follows that

$$F(x,y) = \frac{y-2}{98}$$

Thus letting $x \to \infty$, the CDF of Y is

$$F_2(y) = \begin{cases} 0; & \text{for } y < 2, \\ \frac{y-2}{98}; & \text{for } 2 \le y < 100, \\ 1; & \text{for } y \geq 100 \end{cases}$$

0

Scanned with CamScanner