Joint Probability distribution

A probability distribution associated with two random variables is reffered as joint probability distribution.

ie If X & Y are two discrete random variables, we define the joint probability function of X & Y by

$$P(X=x,\ Y=y) = f(x,y).$$

Where $f(x,y)$ satisfies the conditions,

$f(x,y) \geq 0$ & $\sum_x \sum_y f(x,y) = 1$ (ie sum of overall the values of x & y is equal to 1)

Suppose $X = \{x_1, x_2, x_3, \dots x_m\}$ & $Y = \{y_1, y_2, y_3, \dots y_n\}$ then

$P(X=x_i,\ Y=y_j) = f(x_i, y_j)$ is denoted by J_{ij}

Also $X \times Y = \{(x_1,y_1), (x_2,y_2), \dots (x_m, y_n)\}$.

f is also reffered as joint probability density function of X & Y in the respective order. The set of values of this function $f(x_i, y_j) = J_{ij}$ for $i = 1, 2, \dots m$, $j = 1, 2, \dots n$ is called the joint probability distribution of X & Y.

These values are presented in the form of two way table called joint probability table:

$X \backslash Y$	y_1	y_2	...	y_n	Sum
x_1	J_{11}	J_{12}	...	J_{1n}	$f(x_1)$
x_2	J_{21}	J_{22}	...	J_{2n}	$f(x_2)$
...	...	...	...	...	...
x_m	J_{m1}	J_{m2}	...	J_{mn}	$f(x_m)$
Sum	$g(y_1)$	$g(y_2)$	...	$g(y_n)$	1

Here $\{f(x_1), f(x_2), \dots f(x_m)\}$ & $\{g(y_1), g(y_2), \dots g(y_m)\}$ are called marginal probability distributions of X & Y respectively.

ie $f(x_1)+f(x_2)+\cdots+f(x_m)=1$ &

$g(y_1)+g(y_2)+\cdots+g(y_n)=1$

this is equivalent to

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f(x_i, y_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} J_{ij} = 1.$$

It means that the total of all the entries in the joint probability table is equal to 1.

Independent random variable :-

The discrete random variables X & Y are said to be independent random variables if

$P(X=x_i, Y=y_j) = P(X=x_i)\cdot P(Y=y_j)$. & conversely.

ie equivalent to $f(x_i)g(y_j) = J_{ij}$ in the joint probability table

That is to say that X & Y are independent if each entry J_{ij} in the table is equal to the product of its marginal entries.

Otherwise X & Y are said to be dependent.

Expectation, Variance and Covariance

If X & Y are two discrete random variables having the joint probability function $f(x,y)$ then

(i) Expections of X & Y are defined as

$$\mu_x = E(X) = \sum x f(x).$$

$$\mu_y = E(Y) = \sum y\, g(y).$$

Also $\mu_{xy} = E(XY) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j f(x,y)$ (or) $\mu_{xy} = E(XY) = \sum x_i y_j J_{ij}$

(ii) The variance of X & Y are defined as

$V(X) = \sigma_x^2 = E(X^2) - \mu_x^2$ where $E(X^2) = \sum x^2 f(x)$ & $\mu_x = E(X)$

$V(Y) = \sigma_y^2 = E(Y^2) - \mu_y^2$ where $E(Y^2) = \sum y^2 f(y)$ & $\mu_y = E(y)$.

(iii) Covariance is defined as

$$COV(X,Y) = E(XY) - \mu_x \mu_y$$

(iv) The correlation of X & Y is defined as

$$\rho(X,Y) = \frac{COV(X,Y)}{\sigma_x \sigma_y}.$$

Note: If X & Y are independent random variable then

(i) $E(XY) = E(X)\cdot E(Y)$ (or) $E(XY) = \mu_x \cdot \mu_y$.

(ii) $COV(X,Y) = 0$ and hence $\rho(X,Y) = 0$.

Problems :

1) The joint probability distribution of two random variables X & Y is as follows

Y \ X	-4	2	7	f(x)
1	1/8	1/4	1/8	= 1/2
5	1/4	1/8	1/8	= 1/2
g(y)	3/8	3/8	1/4	

compute ⓐ $E(X)$ & $E(Y)$ ⓑ $E(XY)$ ⓒ σ_x & σ_y

ⓓ $COV(X,Y)$ ⓔ $\rho(X,Y)$

Sol^n: The marginal distribution of X & Y is as follows.

[This distribution is obtained by adding all the respective row entries & also the respective column entries].

Distribution of X :-

x_i	1	5
$f(x_i)$	1/2	1/2

Distribution of Y :

y_j	-4	2	7
$g(y_j)$	3/8	3/8	1/4

(a)

(i) $E(X) = \sum x_i f(x_i) = 1(1/2) + 5(1/2) = \frac{1}{2} + \frac{5}{2} = \frac{6}{2} = 3.$

$\boxed{E(X) = 3}$

(ii) $E(Y) = \sum y_j g(y_j) = -4(3/8) + 2(3/8) + 7(1/4) = \frac{-12}{8} + \frac{6}{8} + \frac{7}{4} = \frac{-12+6+14}{8}$

$\boxed{E(Y) = 1}$

(b) $E(XY) = \sum x_i y_j J_{ij}$

$= (1)(-4)(1/8) + (1)(2)(1/4) + (1)(7)(1/8) + (5)(-4)(1/4) + (5)(2)(1/8) + (5)(7)(1/8)$

$= \frac{-1}{2} + \frac{1}{2} + \frac{7}{8} - 5 + \frac{5}{4} + \frac{35}{8}$

$\boxed{E(XY) = \frac{3}{2}}$

c) $\sigma_x^2 = E(x^2) - \mu_x^2$

$E(X^2) = \sum x_i^2 f(x_i)$

$= 1(1/2) + (25)(1/2)$

$E(X^2) = 13$

$\therefore \sigma_x^2 = 13 - (3)^2$

$= 13 - 9$

$\sigma_x^2 = 4$

$\boxed{\sigma_x = 2}$

$\sigma_y^2 = E(Y^2) - \mu_y^2$

$E(Y^2) = \sum y_j^2 g(y_j)$

$= 16(3/8) + 4(3/8) + 49(1/4)$

$E(Y^2) = 79/4$

$\sigma_y^2 = \left(\frac{79}{4}\right) - (1)^2$

$\sigma_y^2 = \frac{75}{4}$

$\sigma_y = \sqrt{\frac{75}{4}} = 4.33$

$\boxed{\sigma_y = 4.33}$

d) $\text{COV}(X,Y) = E(XY) - \mu_x \mu_y$

$= 3/2 - (3)(1)$

$= \frac{3}{2} - 3$

$\boxed{\text{COV}(X,Y) = \frac{-3}{2}}$

e) $\rho(X,Y) = \frac{\text{COV}(X,Y)}{\sigma_x \sigma_y}$

$= \frac{-3/2}{(2)\left(\sqrt{\frac{75}{4}}\right)}$

$\boxed{\rho(X,Y) = -0.1732}$

2) Find the covariance & correlation of the random variable X & Y as:

X \ Y	1	3	9	$f(x_i)$
2	1/8	1/24	1/12	= 1/4
4	1/4	1/4	0	= 1/2
6	1/8	1/24	1/12	= 1/4
$g(y_j)$	1/2	1/3	1/6	

Solⁿ: The marginal distribution for X & Y is as follows:

Distribution for X:

x_i	2	4	6
$f(x_i)$	1/4	1/2	1/4

Distribution for Y

y_i	1	3	9
$g(y_i)$	1/2	1/3	1/6

Covariance of $(X,Y) = \text{cov}(X,Y) = E(XY) - E(X)E(Y)$

$$\mu_x = E(x) = \sum x_i f(x_i) = (2\times 1/4) + (4\times \tfrac{1}{2}) + (6\times 1/4)$$

$$E(x) = 4$$

$$\mu_y = E(Y) = \sum y_j\, g(y_j) = (1\times 1/2) + (3\times 1/3) + (9\times 1/6)$$

$$E(Y) = 3$$

$$\therefore E(XY) = \sum x_i y_j J_{ij} = (2\times 1\times 1/8) + (2\times 3\times 1/24) + (2\times 9\times 1/12) + (4(1)\times 1/4)$$
$$+ (4\times 3\times 1/4) + (4\times 9\times 0) + (6\times 1\times 1/8) + (6\times 3\times 1/24) + (6\times 9\times \tfrac{1}{12})$$

$$E(XY) = 12$$

$$\therefore \text{Cov}(X,Y) = 12 - (4)(3)$$

$$\text{cov}(X,Y) = 0$$

$$\therefore \rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_x \sigma_y} = 0.$$

If $\text{cov}(X,Y) = 0 \Rightarrow \rho(X,Y) = 0$ then X & Y are independent.

3) The joint probability distribution table for two random variables X & Y is as follows.

X \ Y	-2	-1	4	5	$f(x_i)$
1	0.1	0.2	0	0.3	0.6
2	0.2	0.1	0.1	0	0.4
$g(y_j)$	0.3	0.3	0.1	0.3	

Determine the marginal probability distributions of X & Y. Also compute (a) Expectation of X, Y & XY (b) S.D's. of X, Y (c) Covariance of X & Y (d) Correlation of X & Y e) Further verify that X & Y are dependent random variables.

Soln.: Distribution for X:

x_i	1	2
$f(x_i)$	0.6	0.4

Distribution for Y

y_i	-2	-1	4	5
$g(y_i)$	0.3	0.3	0.1	0.3

(a) $\mu_x = E(X) = \sum x_i f(x_i)$
$= (1)(0.6) + 2(0.4)$
$= 1.4$

$\mu_y = E(Y) = \sum y_i g(y_i)$
$= (-2)(0.3) + (-1)(0.3) + 4(0.1) + 5(0.3)$
$= 1$

$E(XY) = \sum_{i,j} x_i y_j J_{ij}$
$= (1)(-2)(0.1) + (1)(-1)(0.2) + (1)(4)(0) + (1)(5)(0.3) + (2)(-2)(0.2) + (2)(-1)(0.1) + (2)(4)(0.1) + (2)(5)(0)$
$= -0.2 - 0.2 + 0 + 1.5 - 0.8 - 0.2 + 0.8 + 0$

$E(XY) = 0.9$

$\therefore E(X) = 1.4,\ E(Y) = 1,\ E(XY) = 0.9.$

b) $\sigma_x^2 = E(X^2) - \mu_x^2$

but $E(X^2) = \sum x_i^2 f(x_i)$
$= (1)(0.6) + 4(0.4) = 2.2$

$\therefore \sigma_x^2 = 2.2 - (1.4)^2 = 0.24$

$\therefore \sigma_x = \sqrt{0.24} = 0.49$

$\sigma_y^2 = E(Y^2) - \mu_y^2$

$E(Y^2) = \sum y_i^2 f(y_i)$
$= 4(0.3) + 1(0.3) + (16)(0.1) + 25(0.3)$
$= 10.6$

$\sigma_y^2 = 10.6 - (1)^2$

$\sigma_y^2 = 9.6$

$\therefore \sigma_y = \sqrt{9.6} = 3.1$

c) $\mathrm{cov}(X,Y) = E(XY) - E(X)E(Y)$

$= 0.9 - (1.4)(1)$

$\boxed{\mathrm{cov}(X,Y) = -0.5}$

d) correlation of X & Y $= \rho(X,Y) = \dfrac{\mathrm{cov}(X,Y)}{\sigma_x \sigma_y}$

$= \dfrac{-0.5}{(0.49)(3.1)}$

$\boxed{\rho(X,Y) = -0.3}$

e) If X & Y are independent random variables we must have $f(x_i)\,g(y_j) = J_{ij}$

It can be seen that $f(x_1)\,g(x_1) = (0.6)(0.3) = 0.18$

but $J_{11} = .0.1$

ie $f(x_1) \cdot g(y_1) \neq J_{11}$

Similarly for others also the condition is not satisfied.

Hence we conclude that X & Y are dependent random variables.

Scanned with CamScanner