Joint Probability

Let X and Y be two discrete random variables. Let $p(x, y)$ be a function such that $p(x, y) = P(X = x, Y = y)$, then $p(x, y)$ is called joint probability function of X and Y, if the following conditions are satisfied

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

(ii) $\sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) = 1$

Marginal distribution of X

In the bivariate probability distribution, if the probability mass function of only X is taken then it is called Marginal distribution of X, denoted by P_i or $P(x_i)$ or $P(x)$.

$$P(x_i) = P_i = \sum_{j=1}^{m} P(x_i, y_j)$$

Marginal distribution of Y

In the bivariate probability distribution, if the probability mass function of only Y is taken then it is called Marginal distribution of Y, denoted by P_j or $P(y_j)$ or $P(y)$.

$$P(y_j) = P_j = \sum_{i=1}^{n} P(x_i, y_j)$$

Joint probability distribution of X and Y

The set of values of the $P(x_i, y_j) = P_{ij}$ for $i = 1,2, \dots n, \quad j = 1,2, \dots m$ is called the joint probability distribution of X and Y

X \ Y	y_1	y_1	…	y_j	…	y_m	$\sum_i y_i$
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	…	$p(x_1, y_j)$	…	$p(x_1, y_m)$	P_1
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	…	$p(x_2, y_j)$	…	$p(x_2, y_m)$	P_2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
x_i	$p(x_i, y_1)$	$p(x_i, y_2)$	…	$p(x_i, y_j)$	…	$p(x_i, y_m)$	P_i
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
x_n	$p(x_n, y_1)$	$p(x_n, y_2)$	…	$p(x_n, y_j)$	…	$p(x_n, y_m)$	P_n
$\sum_i x_i$	Q_1	Q_2	…	Q_j	…	Q_m	1

Independent Random variables

Two random variables X and Y are said to be independent, if their joint probability mass function equal to the product of their marginal distribution

$$P(x_i, y_j) = P(x_i)P(y_j)$$

OR

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

OR

$$E(XY) = E(X)E(Y)$$

Expectation (mean), Variance and Covariance

If X and Y are two discrete random variables having the joint probability $P(x, y)$ then the expectations of X and Y are defined as follows

$$\mu_X = E(X) = \sum_{j=1}^{m} x_i P(x_i) = \sum xp(x)$$

$$\mu_Y = E(Y) = \sum_{i=1}^{n} y_j P(y_j) = \sum yp(y)$$

$$E(XY) = \sum xyp(x, y)$$

Variance

$$V(X) = \sigma_X^2 = E(X^2) - [E(X)]^2$$

$$V(Y) = \sigma_Y^2 = E(Y^2) - [E(Y)]^2$$

Standard deviation

$$SD(X) = \sigma_X = \sqrt{E(X^2) - [E(X)]^2}$$

$$SD(Y) = \sigma_Y = \sqrt{E(Y^2) - [E(Y)]^2}$$

Coefficient of correlation $\boldsymbol{r(X, Y)}$ or $\boldsymbol{\rho}$

$$r(X, Y) = \frac{covariance(X, Y)}{\sigma_X \sigma_Y}$$

$$covariance(X, Y) = E(XY) - E(X)E(Y)$$

Note:

$E(X+Y) = E(X) + E(Y)$

If X and Y are independent random variables then $E(XY) = E(X)E(Y)$ or $covariance(X,Y) = 0$

Problems:

1. In the Joint probability distribution, find the correlation coefficient.

X \ Y	1	2	3
2	0.2	0.1	0
4	0.1	0.1	0.1
6	0.2	0.2	0

Solution: Given

X \ Y	1	2	3	P(X)
2	0.2	0.1	0	0.3
4	0.1	0.1	0.1	0.3
6	0.2	0.2	0	0.4
P(Y)	0.5	0.4	0.1	1

Marginal distribution of X is

X	2	4	6
P(X)	0.3	0.3	0.4

Marginal distribution of Y is

Y	1	2	3
P(Y)	0.5	0.4	0.1

$E(X) = \sum xp(x) = (2 \times 0.3) + (4 \times 0.3) + (6 \times 0.4) = 4.2$

$E(Y) = \sum yp(y) = (1 \times 0.5) + (2 \times 0.4) + (3 \times 0.1) = 1.6$

$E(X^2) = \sum x^2 p(x) = (2^2 \times 0.3) + (4^2 \times 0.3) + (6^2 \times 0.4) = 20.4$

$$E(Y^2) = \sum y^2 p(y) = (1^2 \times 0.5) + (2^2 \times 0.4) + (3^2 \times 0.1) = 3$$

$$V(X) = E(X^2) - [E(X)]^2$$

$$V(X) = 20.4 - 4.2^2 = 2.76$$

$$\sigma_X = \sqrt{V(X)} = 1.6613$$

$$V(Y) = E(Y^2) - [E(Y)]^2$$

$$V(Y) = 3 - 1.6^2 = 0.44$$

$$\sigma_Y = \sqrt{V(Y)} = 0.6633$$

$$E(XY) = \sum xyp(x, y)$$

X \ Y	1	2	3
2	0.2	0.1	0
4	0.1	0.1	0.1
6	0.2	0.2	0

$$\begin{aligned} E(XY) &= (2 \times 1 \times 0.2) + (2 \times 2 \times 0.1) + (2 \times 3 \times 0) \\ &\quad + (4 \times 1 \times 0.1) + (4 \times 2 \times 0.1) + (4 \times 3 \times 0.1) \\ &\quad + (6 \times 1 \times 0.2) + (6 \times 2 \times 0.2) + (6 \times 3 \times 0) \end{aligned}$$

$$E(XY) = 6.8$$

$$cov(X, Y) = E(XY) - E(X)E(Y) = 6.8 - (4.2 \times 1.6) = 0.08$$

$$r = \frac{cov(X, Y)}{\sigma_X \sigma_y} = \frac{0.08}{1.6613 \times 0.6633} = 0.0725$$

2. A joint probability distribution is given by the following table

X \ Y	-3	2	4
1	0.1	0.2	0.2
3	0.3	0.1	0.1

Find the (i) Marginal distribution of X and Y

(ii) $\mu_x, \mu_y, \sigma_x, \sigma_y$

(iii) Correlation coefficient

Solution:

X \ Y	-3	2	4	P(X)
1	0.1	0.2	0.2	0.5
3	0.3	0.1	0.1	0.5
P(Y)	0.4	0.3	0.3	1

Marginal distribution of X

X	1	3
P(X)	0.5	0.5

Marginal distribution of Y

Y	-3	2	4
P(Y)	0.4	0.3	0.3

$$\mu_X = E(X) = \sum xp(x) = (1 \times 0.5) + (3 \times 0.5) = 2$$

$$\mu_Y = E(Y) = \sum yp(y) = (-3 \times 0.4) + (2 \times 0.3) + (4 \times 0.3) = 0.6$$

$$E(X^2) = \sum x^2 p(x) = (1^2 \times 0.5) + (3^2 \times 0.5) = 5$$

$$E(Y^2) = \sum y^2 p(y) = ((-3)^2 \times 0.4) + (2^2 \times 0.3) + (4^2 \times 0.3) = 9.6$$

$$V(X) = E(X^2) - [E(X)]^2$$

$$V(X) = 5 - 2^2 = 1$$

$V(Y) = E(Y^2) - [E(Y)]^2$

$V(Y) = 9.6 - 0.6^2 = 9.24$

$\sigma_X = \sqrt{V(X)} = 1$

$\sigma_Y = \sqrt{V(Y)} = 3.0397$

$E(XY) = \sum xyp(x,y) = -0.3 + 0.4 + 0.8 - 2.7 + 0.6 + 1.2 = 0$

$COV(X,Y) = E(XY) - E(X)E(Y) = 0 - (2 \times 0.6) = -1.2$

$$r = \frac{COV(X,Y)}{\sigma_X \sigma_Y} = \frac{-1.2}{3.0397} = -0.3947$$

3. A coin is tossed three times. Let X be equal to 0 or 1 according as a head or a tail occurs on the first toss. Let Y be equal to the total number of heads which occurs. Determine (i) the marginal distributions of X and Y, and (ii) the joint distribution of X and Y, (iii) expected values of $X, Y, X+Y$ and XY, (iv) σ_X and σ_Y, (v) $Cov(X,Y)$ and $\rho(X,Y)$.

Solution: Here the sample space is given by

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

(i) The distribution of the random variable X is given by the following table

X (First toss Head or Tail)	0 (First toss Head)	1 (First toss Tail)
P(X) (Probability of random variable X)	$\frac{4}{8}$	$\frac{4}{8}$

which is the marginal distribution of the random variable X.

The distribution of the random variable Y is given by the following table

Y (Total number of Heads)	0 (zero Heads)	1 (one Head)	2 (two Head)	3 (three Head)
P(Y) (Probability of random variable Y)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

which is the marginal distribution of the random variable Y.

(ii) The joint distribution of the random variables X and Y is given by the following table

X \ Y	0 (zero Heads)	1 (one Head)	2 (two Head)	3 (three Head)
0 (First toss Head)	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$
1 (First toss Tail)	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0

$$E[X] = \mu_X = \sum x_i P(x_i) = \left(0 \times \frac{4}{8}\right) + \left(1 \times \frac{4}{8}\right) = \frac{4}{8} = 0.5$$

$$E[Y] = \mu_Y = \sum y_j P(y_j) = \left(0 \times \frac{1}{8}\right) + \left(1 \times \frac{3}{8}\right) + \left(2 \times \frac{3}{8}\right) + \left(3 \times \frac{1}{8}\right) = \frac{12}{8} = 1.5$$

$$E(X + Y) = E(X) + E(Y) = 0.5 + 1.5 = 2$$

OR

$$\begin{aligned} E[X + Y] &= \sum\sum P_{ij}(x_i + y_j) \\ &= P_{11}(x_1 + y_1) + P_{12}(x_1 + y_2) + P_{13}(x_1 + y_3) + P_{14}(x_1 + y_4) + P_{21}(x_2 + y_1) \\ &\quad + P_{22}(x_2 + y_2) + P_{23}(x_2 + y_3) + P_{24}(x_2 + y_4) \\ &= 0(0 + 0) + \frac{1}{8}(0 + 1) + \frac{2}{8}(0 + 2) + \frac{2}{8}(1 + 1) + \frac{1}{8}(1 + 2) + 0(1 + 3) = \frac{16}{8} = 2. \end{aligned}$$

$$\begin{aligned} E(XY) &= \sum\sum P_{ij}(x_i y_j) \\ &= P_{11}(x_1y_1) + P_{12}(x_1y_2) + P_{13}(x_1y_3) + P_{14}(x_1y_4) + P_{21}(x_2y_1) + P_{22}(x_2y_2) + P_{23}(x_2y_3) \\ &\quad + P_{24}(x_2y_4) \\ &= 0(0 \times 0) + \frac{1}{8}(0 \times 1) + \frac{2}{8}(0 \times 2) + \frac{2}{8}(1 \times 1) + \frac{1}{8}(1 \times 2) + 0(1 \times 3) = \frac{1}{2} = 0.5 \end{aligned}$$

$$\sigma_X^2 = E[X^2] - [E(X)]^2 = \left(0^2 \times \frac{4}{8}\right) + (1^2 \times \frac{4}{8}) - \left(\frac{4}{8}\right)^2 = \frac{1}{4}$$

$$\sigma_Y^2 = E[Y^2] - [E(Y)]^2 = \left(0^2 \times \frac{1}{8}\right) + \left(1^2 \times \frac{3}{8}\right) + \left(2^2 \times \frac{3}{8}\right) + \left(3^2 \times \frac{1}{8}\right) - \left(\frac{2}{8}\right)^2 = \frac{3}{4}$$

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{2} - \frac{1}{2} \times \frac{3}{2} = -\frac{1}{4}$$

$$r = \rho(X, Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{-1/4}{(1/2)(\sqrt{3}/2)} = -\frac{1}{\sqrt{3}}.$$

4. A joint probability distribution is given by the following table

Y / X	2	3	4
1	0.06	0.15	0.09
2	0.14	0.35	0.21

Determine the marginal distributions of X and Y. Also verify that X and Y are independent.

Solution:

Y / X	2	3	4	P(X)
1	0.06	0.15	0.09	0.3
2	0.14	0.35	0.21	0.7
P(Y)	0.2	0.5	0.3	1

Here

$P_1 = 0.3, P_2 = 0.7$

$Q_1 = 0.2, Q_2 = 0.5, Q_3 = 0.3$

$P_{11} = P_1Q_1 = 0.06 = P(x_1, y_1)$

$P_{12} = P_1Q_2 = 0.15 = P(x_1, y_2)$

$P_{13} = P_1Q_3 = 0.09 = P(x_1, y_3)$

$P_{21} = P_2Q_1 = 0.14 = P(x_2, y_1)$

$P_{22} = P_2Q_2 = 0.35 = P(x_2, y_2)$

$P_{23} = P_2Q_3 = 0.2 = P(x_2, y_3)$

Thus, $P_iQ_j = P_{ij}$ for all values of i and j. Accordingly, X and Y are stochastically independent.

OR

Alternate method:

Marginal distribution of X

X	1	2
P(X)	0.3	0.7

$E(X) = 1.7$

Y	2	3	4
P(Y)	0.2	0.5	0.3

$E(Y) = 3.1$

$E(XY) = 0.12 + 0.45 + 0.36 + 0.56 + 2.1 + 1.68 = 5.27$

$E(XY) = E(X)E(Y) = (1.7 \times 3.1) = 5.27$

Thus, X and Y are stochastically independent.

5. A probability distributions of two stochastically independent random variables X and Y are given by the following table.

X	0	1
P(X)	0.2	0.8

Y	1	2	3
P(Y)	0.1	0.4	0.5

Find the joint probability distribution. Also compute E(X) and E(Y).

Solution:

$P_{11} = P(X = 0) \times P(Y = 1) = 0.2 \times 0.1 = 0.02$

$P_{12} = P(X = 0) \times P(Y = 2) = 0.2 \times 0.4 = 0.08$

$P_{13} = P(X = 0) \times P(Y = 3) = 0.2 \times 0.5 = 0.1$

$P_{21} = P(X = 1) \times P(Y = 1) = 0.8 \times 0.1 = 0.08$

$P_{22} = P(X = 1) \times P(Y = 2) = 0.8 \times 0.4 = 0.32$

$P_{23} = P(X = 1) \times P(Y = 3) = 0.8 \times 0.5 = 0.4$

Hence the joint probability distribution is

X \Y	1	2	3	P(X)
0	0.02	0.08	0.1	0.2
1	0.08	0.32	0.4	0.8
P(Y)	0.1	0.4	0.5	1

$E(X) = 0.8$

$E(Y) = 2.4$

$E(XY) = \sum xyp(x, y) = 0 + 0 + 0 + 0.08 + 0.64 + 1.2 = 1.92$

$E(XY) = E(X)E(Y) = (0.8)(2.4) = 1.92$

6. The joint probability distribution of two random variables X and Y is given by the following table.

X \ Y	1	3	9
2	1/8	1/24	1/12
4	1/4	1/4	0
6	1/8	1/24	1/12

Find the marginal distribution of X and Y, and evaluate $cov(X, Y)$.

Solution: From the table, we note that

$$P_1 = \frac{1}{8} + \frac{1}{24} + \frac{1}{12} = \frac{1}{4}$$

$$P_2 = \frac{1}{4} + \frac{1}{4} + \ 0 = \frac{1}{2}$$

$$P_3 = \frac{1}{8} + \frac{1}{24} + \frac{1}{12} = \frac{1}{4}$$

$$Q_1 = \frac{1}{8} + \frac{1}{4} + \frac{1}{8} = \frac{1}{2}$$

$$Q_2 = \frac{1}{24} + \frac{1}{4} + \frac{1}{24} \ = \frac{1}{3}$$

$$Q_3 = \frac{1}{12} + 0 + \frac{1}{12} = \frac{1}{6}$$

The marginal distribution of X is given by the table:

x_i	2	4	6
P_i	1/4	1/2	1/4

And the marginal distribution of Y is given by the table:

y_j	1	3	9
Q_j	1/2	1/3	1/6

Therefore, the means of these distributions are respectively,

$$\mu_X = \sum x_i \, P(x_i) = \left(2 \times \frac{1}{4}\right) + \left(4 \times \frac{1}{2}\right) + \left(6 \times \frac{1}{4}\right) = 4$$

$$\mu_Y = \sum y_j \, P(y_j) = \left(1 \times \frac{1}{2}\right) + \left(3 \times \frac{1}{3}\right) + \left(9 \times \frac{1}{6}\right) = 3$$

$\mathrm{E}[XY] = \sum_i \sum_j P_{ij} \ x_i y_j$

$= \left(2 \times \frac{1}{8}\right) + \left(6 \times \frac{1}{24}\right) + \left(18 \ \times \frac{1}{12}\right) + \left(4 \times \frac{1}{4}\right) + \left(12 \times \frac{1}{4}\right) + 36 \times 0 + \left(6 \times \frac{1}{8}\right) + \left(18 \times \frac{1}{24}\right) + \left(54 \times \frac{1}{12}\right)$

$= 2 + 4 + 6 = 12$

$Cov\ (X, Y) = E[XY] - \mu_X \mu_Y \ = 12 - 12 \ = 0$

$\rho(X, Y) = 0.$

7. For the following bivariate probability distribution of X and Y find
(i) $P(X \leq 1, Y = 2)$ (ii) $P(X \leq 1)$ (iii) $P(Y = 3)$ (iv) $P(Y \leq 3)$ (v) $P(X < 3, Y \leq 4)$

X \ Y	1	2	3	4	5	6
0	0	0	1/32	2/32	2/32	3/32
1	1/16	1/16	1/8	1/8	1/8	1/8
2	1/32	1/32	1/64	1/64	0	2/64

Solution:

X \ Y	1	2	3	4	5	6	P(X)
0	0	0	1/32	2/32	2/32	3/32	8/32
1	1/16	1/16	1/8	1/8	1/8	1/8	10/16
2	1/32	1/32	1/64	1/64	0	2/64	8/64
P(Y)	3/32	3/32	11/64	13/64	6/32	16/64	1

$$P(X \leq 1, Y = 2) = P(X = 0, Y = 2) + P(X = 1, Y = 2) = 0 + \frac{1}{16} = \frac{1}{16}$$

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{8}{32} + \frac{10}{16} = \frac{7}{8}$$

$$P(Y = 3) = \frac{11}{64}$$

$$P(Y \leq 3) = \frac{3}{32} + \frac{3}{32} + \frac{11}{64} = \frac{23}{64}$$

$$P(X < 3, Y \leq 4) = P(X = 0, Y \leq 4) + P(X = 1, Y \leq 4) + P(X = 2, Y \leq 4)$$

$$= P(X = 0, Y = 1) + P(X = 0, Y = 2) + P(X = 0, Y = 3) + P(X = 0, Y = 4)$$

$$+P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 1, Y = 3) + P(X = 1, Y = 4)$$

$$+P(X = 2, Y = 1) + P(X = 2, Y = 2) + P(X = 2, Y = 3) + P(X = 2, Y = 4)$$

$$P(X < 3, Y \leq 4) = \left(0 + 0 + \frac{1}{32} + \frac{2}{32}\right) + \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{8} + \frac{1}{8}\right) + \left(\frac{1}{32} + \frac{1}{32} + \frac{1}{64} + \frac{1}{64}\right) = \frac{9}{16}$$

8. For the following bivariate probability distribution, find the value of k.

X \ Y	1	2	3
-5	0	0.1	0.1
0	0.1	k	0.2
5	0.2	0.1	0

Solution:

X \ Y	1	2	3	P(X)
-5	0	0.1	0.1	0.2
0	0.1	k	0.2	0.3+k
5	0.2	0.1	0	0.3
P(Y)	0.3	0.2+k	0.3	1

$$0.8 + k = 1$$

$$k = 0.2$$