

Correlation

The correlation of two random variables X and Y , denoted by $\rho(X, Y)$, is defined by

$$\rho(X, Y) = \text{cov}(x, y) / \sigma_X \sigma_Y$$

- 1) If $\sigma_X = 0$ or If $\sigma_Y = 0$ then $\rho(X, Y) = 0$.

The random variables X and Y are said to be uncorrelated if and only if $\rho(X, Y) = 0$; otherwise they are said to be correlated.

- 2) Note that if X and Y are independent random variables, then they are uncorrelated. But the opposite is not true (i.e. we can find two random variable that are uncorrelated but not independent).

Finally, it is worth noting the following properties of ρ :

- (i) $\rho(X, Y) = \rho(Y, X)$,
- (ii) $-1 \leq \rho \leq 1$,
- (iii) $\rho(X, X) = 1, \rho(X, -X) = -1$.

Example:-

Let the joint density function

$$f(x, y) = 8xy \quad , 0 \leq x \leq 1, , 0 \leq y \leq x$$

Find the correlation of x and y .

Sol:-

Since $\text{cov}(x, y) = 4/225$

To find correlation, we first need to find variances of x and of y .

$$E(x^2) = \int_0^1 4x^5 dx = 2/3, E(y^2) = \int_0^1 4y^3(1 - y^2) dy = 1/3$$

$$\text{Thus } \text{var}(x) = 2/3 - (4/5)^2 = 2/75$$

$$\text{var}(y) = 1/3 - (8/15)^2 = 11/225$$

$$\rho(X, Y) = \text{cov}(X, Y) / \sigma_X \sigma_Y$$

$$(4/225)/(2/75)(11/225)=4/\sqrt{66}.$$

Example

A pair of dice are tossed. The sample space Ω consists of the 36 ordered pairs (a,b) where a and b can be any integers between 1 and 6.

Let X assign to each point (a, b) the maximum of its numbers, that is,

$$X(a, b) = \max(a, b).$$

Now let Y assign to each point (a, b) the sum of its numbers, that is,

$$Y(a, b) = a + b.$$

So, for example $X(1, 1) = 1$, $X(3, 4) = 4$, $X(5, 2) = 5$, $X(6, 6) = 6$;

and in the case of the random variable Y, $Y(1, 1) = 2$, $Y(3, 4) = 7$, $Y(6, 3) = 9$,

$$Y(6, 6) = 12.$$

Then X is a random variable where any number between 1 and 6 could occur, and no other number can occur.

Thus, the range space R_x of X is as follows:

$R_x = \{1, 2, 3, 4, 5, 6\}$ And, Y is a random variable where any number between 2 and 12 could occur, and no other number can occur.

Thus, the range space R_y of Y is as follows:

$R_y = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ The joint distribution appears in the following table:

The entry $h(3, 5) = 2/36$ comes from the fact that (3, 2) and (2, 3) are the only points in Ω whose maximum number is 3 and whose sum is 5, that is,

$$h(3, 5) \equiv P(X = 3, Y = 5) = P\{(3, 2), (2, 3)\} = 2/36$$

x \ y												
	2	3	4	5	6	7	8	9	10	11	12	P _x
1	1/36	0	0	0	0	0	0	0	0	0	0	1/36
2	0	2/36	1/36	0	0	0	0	0	0	0	0	3/36
3	0	0	2/36	2/36	1/36	0	0	0	0	0	0	5/36
4	0	0	0	2/36	2/36	2/36	1/36	0	0	0	0	7/36
5	0	0	0	0	2/36	2/36	2/36	2/36	1/36	0	0	9/36
6	0	0	0	0	0	2/36	2/36	2/36	2/36	2/36	1/36	11/36
P _y	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36	1

The other entries are obtained in a similar manner.

First we compute the expectation of X and Y as follows:

$$E(X) = 1(1/36) + 2(3/36) + 3(5/36) + 4(7/36) + 5(9/36) + 6(11/36) \\ = 161/36 \approx 4.47$$

$$E(Y) = 2(1/36) + 3(2/36) + 4(3/36) + \dots + 12(1/36) = 252/36 = 7$$

Next, we compute σ_x and σ_y as follows:

$$E(X^2) = \sum x^2 f(x) = 1^2(1/36) + 2^2(3/36) + 3^2(5/36) + 4^2(7/36) + 5^2(9/36) \\ + 6^2(11/36) = 791/36 \approx 21.97$$

$$\text{Hence } \text{var}(X) = E(X^2) - \mu_x^2 = 21.97 - 19.98 \\ = 1.99 \text{ and } \sigma_x = \sqrt{1.99} \approx 1.4$$

Similarly

$$E(Y^2) = \sum y^2 g(y) = 2^2(1/36) + 3^2(2/36) + 4^2(3/36) + \dots + 12^2(1/36) \\ = 1974/36 \approx 54.8$$

$$\text{Hence } \text{var}(Y) = E(Y^2) - \mu_y^2 = 54.8 - 49 = 5.8$$

$$\text{and } \sigma_y = \sqrt{5.8} \approx 2.4$$

Now we compute $E(XY)$ as follows:

$$\begin{aligned} E(XY) &= \sum \sum xy h(x, y) \\ &= 1(2)(1/36) + 2(3)(2/36) + 2(4)(1/36) + \dots + 6(12)(1/36) \\ &= 1232/36 \approx 34.2 \end{aligned}$$

So, the covariance of X and Y is computed as:

$$\text{cov}(X, Y) = E(XY) - \mu_X \mu_Y = 34.2 - (4.47)(7) \approx 2.9$$

and

$$\rho(X, Y) = \text{cov}(X, Y) / \sigma_X \sigma_Y = 2.9 / (1.4)(2.4) \approx 0.86.$$

$$h(2, 3) \equiv P(X = 2, Y = 3) = P\{(1, 2), (2, 1)\} = 2/36$$

$$h(4, 6) \equiv P(X = 4, Y = 6) = P\{(2, 4), (4, 2)\}$$

$$\begin{aligned} h(4, 5) \equiv P(X = 4, Y = 5) &= P\{(4, 1), (1, 4)\} = 2/36 \\ &= 2/36 \end{aligned}$$

$$h(5, 6) \equiv P(X = 5, Y = 6) = P\{(1, 5), (5, 1)\} = 2/36$$