

Moments, Covariance and Correlation Coefficient

Let $\underline{X} = (X_1, X_2, \dots, X_n)$ be a n -dimensional ($n \geq 2$) random vector and $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function such that $\psi^{-1}(A) \in \mathbb{B}_{\mathbb{R}^n}$, for all $A \in \mathbb{B}_{\mathbb{R}}$. Suppose $E(\psi(\underline{X}))$ is finite.

(1) If \underline{X} is of discrete type with joint p.m.f. $f_{\underline{X}}$ and support $E_{\underline{X}}$, then

$$E(\psi(\underline{X})) = \sum_{(x_1, x_2, \dots, x_n) \in E_{\underline{X}}} \psi(x_1, x_2, \dots, x_n) f_{\underline{X}}(x_1, x_2, \dots, x_n).$$

(2) If \underline{X} is of continuous type with joint p.d.f. $f_{\underline{X}}$, then

$$E(\psi(\underline{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \psi(x_1, x_2, \dots, x_n) f_{\underline{X}}(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n.$$

(3) For nonnegative integers k_1, k_2, \dots, k_n , let $\psi(x_1, x_2, \dots, x_n) = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$. Then

$$\mu'_{k_1, k_2, \dots, k_n} = E(\psi(\underline{X})) = E(X_1^{k_1} X_2^{k_2} \cdots X_n^{k_n}),$$

provided it is finite, is called the joint moment of order $k_1 + k_2 + \cdots + k_n$ of $\underline{X} = (X_1, X_2, \dots, X_n)$.

(4) For $n = 2$, let $\psi(x_1, x_2) = (x_1 - E(X_1))(x_2 - E(X_2))$. Then

$$Cov(X_1, X_2) = E\left((X_1 - E(X_1))(X_2 - E(X_2))\right),$$

provided it is finite, is called the covariance between X_1 and X_2 .

Note: By the definition of covariance, it is easy to see

$$Cov(X_1, X_1) = Var(X_1);$$

$$Cov(X_1, X_2) = Cov(X_2, X_1);$$

$$Cov(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2).$$

Theorem 1. Let $\underline{X} = (X_1, X_2)$ and $\underline{Y} = (Y_1, Y_2)$ be two random vectors and a_1, a_2, b_1, b_2 be real constants. Then, provided the involved expectations are finite,

$$(1) E(a_1 X_1 + a_2 X_2) = a_1 E(X_1) + a_2 E(X_2);$$

$$(2) Cov(a_1 X_1 + a_2 X_2, b_1 Y_1 + b_2 Y_2) = a_1 b_1 Cov(X_1, Y_1) + a_1 b_2 Cov(X_1, Y_2) + a_2 b_1 Cov(X_2, Y_1) + a_2 b_2 Cov(X_2, Y_2) = \sum_{i=1}^2 \sum_{j=1}^2 a_i b_j Cov(X_i, Y_j).$$

In particular,

$$Var(a_1 X_1 + a_2 X_2) = Cov(a_1 X_1 + a_2 X_2, a_1 X_1 + a_2 X_2) = a_1^2 Var(X_1) + a_2^2 Var(X_2) + 2a_1 a_2 Cov(X_1, X_2).$$

Proof. (1) Suppose \underline{X} is continuous type with joint p.d.f. $f_{\underline{X}}$. Let $\psi(x_1, x_2) = a_1 x_1 + a_2 x_2$. Then

$$\begin{aligned} E(a_1 X_1 + a_2 X_2) &= E(\psi(\underline{X})) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (a_1 x_1 + a_2 x_2) f_{\underline{X}}(x_1, x_2) dx_1 dx_2 \\ &= a_1 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 f_{\underline{X}}(x_1, x_2) dx_1 dx_2 + a_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2 f_{\underline{X}}(x_1, x_2) dx_1 dx_2 \end{aligned}$$

By taking $\psi_1(x_1, x_2) = x_1$ and $\psi_2(x_1, x_2) = x_2$, we have

$$E(X_1) = E(\psi_1(\underline{X})) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 f_{\underline{X}}(x_1, x_2) dx_1 dx_2$$

and

$$E(X_2) = E(\psi_2(\underline{X})) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2 f_{\underline{X}}(x_1, x_2) dx_1 dx_2.$$

Thus,

$$E(a_1X_1 + a_2X_2) = a_1E(X_1) + a_2E(X_2).$$

Similarly, we can prove for discrete type random vector.

(2)

$$Cov(a_1X_1 + a_2X_2, b_1Y_1 + b_2Y_2)$$

$$\begin{aligned} &= Cov\left(\sum_{i=1}^2 a_i X_i, \sum_{j=1}^2 b_j Y_j\right) \\ &= E\left(\left(\sum_{i=1}^2 a_i X_i - E\left(\sum_{i=1}^2 a_i X_i\right)\right)\left(\sum_{j=1}^2 b_j Y_j - E\left(\sum_{j=1}^2 b_j Y_j\right)\right)\right) \\ &= E\left(\left(\sum_{i=1}^2 a_i (X_i - E(X_i))\right)\left(\sum_{j=1}^2 b_j (Y_j - E(Y_j))\right)\right) \text{ (by (1))} \\ &= E\left(\sum_{i=1}^2 \sum_{j=1}^2 a_i b_j (X_i - E(X_i))(Y_j - E(Y_j))\right) \\ &= \sum_{i=1}^2 \sum_{j=1}^2 a_i b_j E\left((X_i - E(X_i))(Y_j - E(Y_j))\right) \\ &= \sum_{i=1}^2 \sum_{j=1}^2 a_i b_j Cov(X_i, Y_j). \end{aligned}$$

□

Remark 2. In general, we have

$$(1) E(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + \cdots + a_nE(X_n);$$

$$(2) Cov\left(\sum_{i=1}^{n_1} a_i X_i, \sum_{j=1}^{n_2} b_j Y_j\right) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} a_i b_j Cov(X_i, Y_j).$$

In particular,

$$Var\left(\sum_{i=1}^{n_1} a_i X_i\right) = \sum_{i=1}^{n_1} a_i^2 Var(X_i) + 2 \sum_{1 \leq i < j \leq n_1} a_i a_j Cov(X_i, X_j)$$

Theorem 3. Let X_1, X_2, \dots, X_n be the independent random variables. Let $\psi_i : \mathbb{R} \rightarrow \mathbb{R}$ be a function such that $\psi_i^{-1}(A) \in \mathbb{B}_{\mathbb{R}}$, for all $A \in \mathbb{B}_{\mathbb{R}}$, for $i = 1, 2, \dots, n$. Then

$$E\left(\prod_{i=1}^n \psi_i(X_i)\right) = \prod_{i=1}^n E\left(\psi_i(X_i)\right),$$

provided the involved expectations are finite.

Proof. We will prove the theorem for $n = 2$ and continuous random vector. Suppose $\underline{X} = (X_1, X_2)$ is a continuous type random vector with joint p.d.f. $f_{\underline{X}}$. Consider the function

$\psi(x_1, x_2) = \psi_1(x_1)\psi_2(x_2)$. Then

$$\begin{aligned}
E(\psi_1(X_1)\psi_2(X_2)) &= E(\psi(\underline{X})) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_1(x_1)\psi_2(x_2)f_{\underline{X}}(x_1, x_2) dx_1 dx_2 \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_1(x_1)\psi_2(x_2)f_{X_1}(x_1)f_{X_2}(x_2) dx_1 dx_2 \text{ (since } X_1 \text{ and } X_2 \text{ are independent)} \\
&= \left(\int_{-\infty}^{\infty} \psi_1(x_1)f_{X_1}(x_1) dx_1 \right) \left(\int_{-\infty}^{\infty} \psi_2(x_2)f_{X_2}(x_2) dx_2 \right) \\
&= E(\psi_1(X_1))E(\psi_2(X_2))
\end{aligned}$$

□

Corollary 4. Let X_1, X_2, \dots, X_n be the independent random variables. Then

$$Cov(X_i, X_j) = 0, \forall i \neq j$$

and for real constants a_1, a_2, \dots, a_n ,

$$Var\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 Var(X_i),$$

provided the involved expectations are finite.

Proof. Fix $i, j \in \{1, 2, \dots, n\}, i \neq j$. Then by Theorem 3, we have

$$\begin{aligned}
E(X_i X_j) &= E(X_i)E(X_j) \\
\Rightarrow Cov(X_i, X_j) &= E(X_i X_j) - E(X_i)E(X_j) = 0
\end{aligned}$$

Since $Cov(X_i, X_j) = 0, \forall i \neq j$, by Remark 2,

$$Var\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 Var(X_i).$$

□

Definition 5. (1) The correlation coefficient between random variables X and Y is defined by

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}},$$

provided $0 < Var(X), Var(Y) < \infty$.

(2) The random variables X and Y are said to be uncorrelated if $Cov(X, Y) = 0$.

Note: By definition, it is clear that if X and Y are independent random variables, then they are uncorrelated but converse need not be true.

Theorem 6. Let X and Y be two random variables. Then, provided the involved expectations are finite,

(1) $(E(XY))^2 \leq E(X^2)E(Y^2)$. Moreover, $(E(XY))^2 = E(X^2)E(Y^2)$ if and only if $P(Y = cX) = 1$ or $P(X = cY) = 1$, for some $c \in \mathbb{R}$.

This inequality is known as Cauchy-Schwarz inequality for random variables.

(2) $|\rho(X, Y)| \leq 1$. To prove it, apply (1) on random variables $X' = X - E(X)$ and $Y' = Y - E(Y)$.

Example 7. Let $\underline{Z} = (X, Y)$ be a random vector of discrete type with joint p.m.f.

$$f(x, y) = \begin{cases} p_1, & \text{if } (x, y) = (-1, 1) \\ p_2, & \text{if } (x, y) = (0, 0) \\ p_1, & \text{if } (x, y) = (1, 1) \\ 0, & \text{otherwise} \end{cases}$$

where $p_1, p_2 \in (0, 1)$ and $2p_1 + p_2 = 1$.

Then the support of \underline{Z} , X and Y are

$$E_{\underline{Z}} = \{(-1, 1), (0, 0), (1, 1)\}$$

$$E_X = \{-1, 0, 1\}$$

and

$$E_Y = \{0, 1\},$$

respectively. Clearly $E_{\underline{Z}} \neq E_X \times E_Y$. So, X and Y are not independent.

Now,

$$E(XY) = \sum_{(x,y) \in E_{\underline{Z}}} xyf(x, y) = 0;$$

$$E(X) = \sum_{(x,y) \in E_{\underline{Z}}} xf(x, y) = 0;$$

$$E(Y) = \sum_{(x,y) \in E_{\underline{Z}}} yf(x, y) = 2p_1;$$

$$\Rightarrow \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0 \Rightarrow \rho(X, Y) = 0$$

This shows that X and Y are uncorrelated but not independent.

We can also show that X and Y are not independent by another way.

The marginal p.m.f. of X is

$$f_X(x) = \begin{cases} \sum_{y \in R_x} f(x, y), & \text{if } x \in \{-1, 0, 1\} \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} p_1, & \text{if } x = -1 \\ p_2, & \text{if } x = 0 \\ p_1, & \text{if } x = 1 \\ 0, & \text{otherwise} \end{cases}$$

Similarly, the marginal p.m.f. of Y is

$$f_Y(y) = \begin{cases} \sum_{x \in R_y} f(x, y), & \text{if } y \in \{0, 1\} \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} p_2, & \text{if } y = 0 \\ 2p_1, & \text{if } y = 1 \\ 0, & \text{otherwise} \end{cases}$$

Since $f(-1, 1) \neq f_X(-1)f_Y(1)$, X and Y are not independent.

Example 8. Let $\underline{Z} = (X, Y)$ be a random vector of continuous type with joint p.d.f.

$$f(x, y) = \begin{cases} 1, & \text{if } 0 < |y| \leq x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Now,

$$\begin{aligned}
 E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x,y) dx dy = \int_0^1 \int_{-x}^x xy dy dx = 0; \\
 E(X) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x,y) dx dy = \int_0^1 \int_{-x}^x x dy dx = \frac{2}{3}; \\
 E(Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x,y) dx dy = \int_0^1 \int_{-x}^x y dy dx = 0; \\
 &\Rightarrow \text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0 \Rightarrow \rho(X,Y) = 0
 \end{aligned}$$

Thus X and Y are uncorrelated.

The marginal p.d.f. of X is

$$\begin{aligned}
 f_X(x) &= \int_{-\infty}^{\infty} f(x,y) dy \\
 &= \begin{cases} \int_{-x}^x dy, & \text{if } 0 < x < 1 \\ 0, & \text{otherwise} \end{cases} \\
 &= \begin{cases} 2x, & \text{if } 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}
 \end{aligned}$$

Similarly, the marginal p.d.f. of Y is

$$\begin{aligned}
 f_Y(y) &= \int_{-\infty}^{\infty} f(x,y) dx \\
 &= \begin{cases} \int_{|y|}^1 dx, & \text{if } -1 < y < 1 \\ 0, & \text{otherwise} \end{cases} \\
 &= \begin{cases} 1 - |y|, & \text{if } -1 < y < 1 \\ 0, & \text{otherwise} \end{cases}
 \end{aligned}$$

Since $f(x,y) \neq f_X(x)f_Y(y)$, X and Y are not independent.

We can also show that X and Y are not independent by another way. Then the support of \underline{Z} , X and Y are

$$\begin{aligned}
 E_{\underline{Z}} &= \{(x,y) \in \mathbb{R}^2 \mid 0 < |y| \leq x < 1\} \\
 E_X &= (0,1)
 \end{aligned}$$

and

$$E_Y = (-1,1),$$

respectively. Clearly $E_{\underline{Z}} \neq E_X \times E_Y$. So, X and Y are not independent.

This example also shows that X and Y are uncorrelated but not independent.