MODULE 6

RANDOM VECTOR AND ITS JOINT DISTRIBUTION

LECTURE 29

Topics

6.5 EXPECTATIONS AND MOMENTS

6.5.1 Cauchy- Schwarz Inequality for Random Variables

Theorem 4.3

Let $\underline{X}_1, \dots, \underline{X}_p$ be independent random vectors such that $\underline{X}_i$ is q_i-dimensional, $i = 1, \dots, p$. Let $\psi_i: \mathbb{R}^{q_i} \to \mathbb{R}^{r_i}$, $i = 1, \dots, p$, be Borel functions. Then $\psi_1(\underline{X}_1), \dots, \psi_p(\underline{X}_p)$ are independent.

Proof. Let $\underline{X} = (\underline{X}_1, \dots, \underline{X}_p)$ and let $\underline{Y}_i = \psi_i(\underline{X}_i)$, $i = 1, \dots, p$. For fixed $\underline{y}_i \in \mathbb{R}^{r_i}$ define $A_i = \left\{\underline{x} \in \mathbb{R}^{q_i}: \psi_i(\underline{x}) \leq \underline{y}_i\right\}, i = 1, \dots, p$ (where, for $\underline{x}, \underline{y} \in \mathbb{R}^r$, $\underline{x} \leq \underline{y}$ means $x_i \leq y_i, i = 1, \dots, r$) . Then, for $\underline{y}_i \in \mathbb{R}^{r_i}$, $i = 1, \dots, p$, the joint distribution function of $\underline{Y}_1 = \psi_1(\underline{X}_1), \dots, \underline{Y}_p = \psi_p(\underline{X}_p)$ is given by

$$F_{\underline{Y}_1,\dots,\underline{Y}_p}\left(\underline{y}_1, \dots, \underline{y}_p\right) = P(\{\underline{Y}_1 \in (-\underline{\infty}, y_1], \dots, \underline{Y}_p \in (-\underline{\infty}, y_p]\})$$

$$= P(\{\underline{X}_1 \in A_1, \dots, \underline{X}_p \in A_p\})$$

$$= \prod_{j=1}^{p} P(\{\underline{X}_j \in A_j\}) \qquad \text{(using Remark 4.1 (iii))}$$

$$= \prod_{j=1}^{p} P(\{\underline{Y}_j \leq \underline{y}_j\})$$

$$= \prod_{j=1}^{p} F_{\underline{Y}_j}(\underline{y}_j),$$

where $F_{Y_j}(\cdot)$ denotes the marginal distribution function of $\underline{Y}_j, j = 1, 2, \dots, p$. Now, using the analog of Theorem 4.1 for random vectors, it follows that $\underline{Y}_1, \dots, \underline{Y}_p$ are independent. ■

Example 4.1

Let $\underline{X} = (X_1, X_2, X_3)$ be a discrete type random vector with joint p.m.f.

$$f_{\underline{X}}(x_1, x_2, x_3) = \begin{cases} \frac{x_1 x_2 x_3}{72}, & \text{if } (x_1, x_2, x_3) \in \{1,2\} \times \{1,2,3\} \times \{1,3\} \\ 0, & \text{otherwise} \end{cases}.$$

(i) Are X_1, X_2 and X_3 independent random variables?
(ii) Are X_1 and X_3 independent random variables?

Solution. (i) From Example 2.2 (ii) we have

$$f_{X_1}(x_1) = \begin{cases} \frac{x_1}{3}, & \text{if } x_1 \in \{1,2\} \\ 0, & \text{otherwise} \end{cases}; \quad f_{X_2}(x_2) = \begin{cases} \frac{x_2}{6}, & \text{if } x_2 \in \{1,2,3\} \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_{X_3}(x_3) = \begin{cases} \frac{x_3}{4}, & \text{if } x_3 \in \{1,3\} \\ 0, & \text{otherwise} \end{cases}.$$

Clearly

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3), \ \ \forall \underline{x} = (x_1, x_2, x_3) \in \mathbb{R}^3.$$

Now using Theorem 4.2 (i) it follows that X_1, X_2 and X_3 are independent.

One can also directly infer the independence of X_1, X_2 and X_3 from Theorem 4.2 (ii) by nothing that

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = g_1(x_1) g_2(x_2) g_3(x_3), \ \ \forall \underline{x} = (x_1, x_2, x_3) \in \mathbb{R}^3,$$

where

$$g_1(x_1) = \begin{cases} \frac{x_1}{72}, & \text{if } x_1 \in \{1,2\} \\ 0, & \text{otherwise} \end{cases}, \quad g_2(x_2) = \begin{cases} x_2, & \text{if } x_2 \in \{1,2,3\} \\ 0, & \text{otherwise} \end{cases}$$

and

$$g_3(x_3) = \begin{cases} x_3, & \text{if } x_1 \in \{1,3\} \\ 0, & \text{otherwise} \end{cases}.$$

(ii) From Example 2.2 (iii) we have

$$f_{X_1,X_3}(x_1, x_3) = \begin{cases} \frac{x_1 x_3}{12}, & \text{if } (x_1, x_3) \in \{1,2\} \times \{1,3\} \\ 0, & \text{otherwise} \end{cases}.$$

Clearly

$$f_{X_1,X_3}(x_1,x_3) = f_{X_1}(x_1)f_{X_3}(x_3), \;\; \forall (x_1,x_3) \in \mathbb{R}^2.$$

Therefore X_1 and X_3 are independent. ■

Example 4.2

Let $\underline{X} = (X_1, X_2, X_3)$ be a random vector of absolutely continuous type with p.d.f.

$$f_{\underline{X}}(x_1,x_2,x_3) = \begin{cases} \frac{1}{x_1 x_2}, & \text{if } 0 < x_3 < x_2 < x_1 < 1 \\ 0\;, & \text{otherwise} \end{cases}.$$

(i) Are X_1, X_2 and X_3 independent random variables?
(ii) Let $x_2 \in (0,1)$ be fixed. Are X_1 and X_3 independent given $X_2 = x_2$?

Solution. (i) We have

$$f_{X_1}(x_1) = \begin{cases} \int_0^{x_1}\int_0^{x_2} \frac{1}{x_1 x_2} dx_3 dx_2, & \text{if } 0 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} 1, & \text{if } 0 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases},$$

$$f_{X_2}(x_2) = \begin{cases} -\ln x_2, & \text{if } 0 < x_2 < 1 \\ 0, & \text{otherwise} \end{cases} \qquad \text{(See Example 2.3 (iii))}$$

and

$$f_{X_3}(x_3) = \begin{cases} \int_{x_3}^{1}\int_{x_2}^{1} \frac{1}{x_1 x_2} dx_1 dx_2, & \text{if } 0 < x_3 < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{(\ln x_3)^2}{2}, & \text{if } 0 < x_3 < 1 \\ 0, & \text{otherwise} \end{cases}.$$

Clearly

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) \neq f_{X_1}(x_1)f_{X_2}(x_2)f_{X_3}(x_3), \;\; \forall (x_1,x_2,x_3) \in \mathbb{R}^3,$$

and therefore X_1, X_2 and X_3 are not independent.

Note that $S_{\underline{X}} = \{(x_1, x_2, x_3) : f_{\underline{X}}(x_1, x_2, x_3) > 0\} = \{(x_1, x_2, x_3):\ 0 < x_3 < x_2 < x_1 < 1\}$, $S_{X_1} = \{x_1 : f_{X_1}(x_1) > 0\} = (0, 1) = S_{X_2} = S_{X_3}$. Since $S_{\underline{X}} \neq S_{X_1} \times S_{X_2} \times S_{X_3}$ one can also infer the non-independence of X_1, X_2 and X_3 from Theorem 4.2 (iii).

(ii) Fix $x_2 \in (0, 1)$. From Example 3.2 (ii) we have

$$f_{X_1,X_3|X_2}(x_1, x_3|x_2) = \frac{f_{X_1,X_2,X_3}(x_1, x_2, x_3)}{f_{X_2}(x_2)}$$

$$= \begin{cases} -\dfrac{1}{x_1 x_2 \ln x_2}, & \text{if } x_2 < x_1 < 1, 0 < x_3 < x_2 \\ 0, & \text{otherwise} \end{cases}.$$

Also it is easy to see that

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_2}(x_2)} = \begin{cases} -\dfrac{1}{x_1 \ln x_2}, & \text{if } x_2 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_{X_3|X_2}(x_3|x_2) = \frac{f_{X_2,X_3}(x_2, x_3)}{f_{X_2}(x_2)} = \begin{cases} \dfrac{1}{x_2}, & \text{if } 0 < x_3 < x_2 \\ 0, & \text{otherwise} \end{cases}.$$

Clearly, for fixed $x_2 \in (0, 1)$,

$$f_{X_1,X_3|X_2}(x_1, x_3|x_2) = f_{X_1|X_2}(x_1|x_2) f_{X_3|X_2}(x_3|x_2), \ \ \forall (x_1, x_3) \in \mathbb{R}^2.$$

Now using Theorem 4.2 (i) on conditional p.d.f. of (X_1, X_3) given $X_2 = x_2$ it follows that, given $X_2 = x_2$, the random variables X_1 and X_3 are conditionally independent.

One can also infer the conditional independence of X_1 and X_3 given $X_2 = x_2$ directly from Theorem 4.2 (ii) by nothing that, for a fixed $x_2 \in (0,1)$,

$$f_{X_1,X_3|X_2}(x_1, x_3|x_2) = \frac{f_{X_1,X_2,X_3}(x_1, x_2, x_3)}{f_{X_2}(x_2)}$$

$$= c(x_2) f_{X_1,X_2,X_3}(x_1, x_2, x_3)$$

$$= g_{x_2}^{(1)}(x_1) g_{x_2}^{(2)}(x_3), \quad (x_1, x_3) \in \mathbb{R}^2,$$

where, for a fixed $x_2 \in (0, 1)$

$$g_{x_2}^{(1)}(x_1) = \begin{cases} \frac{c(x_2)}{x_2 x_1}, & \text{if } x_2 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases} \text{ and } g_{x_2}^{(2)}(x_3) = \begin{cases} 1\ , & \text{if } 0 < x_3 < x_2 \\ 0, & \text{otherwise} \end{cases}. \blacksquare$$

6.5 EXPECTATIONS AND MOMENTS

Let $\underline{X} = (X_1, \dots, X_p)$ be a p-dimensional random vector of either discrete type or of absolutely continuous type. Let $f_X(\cdot)$ and $S_{\underline{X}} = \{\underline{x} \in \mathbb{R}^p : f_{\underline{X}}(\underline{x}) > 0\}$ denote respectively the p.m.f. (or p.d.f.) of $\underline{X}$ (or $f_{\underline{X}}$). Further let $f_{X_i}(\cdot)$ and $S_{X_i} = \{x \in \mathbb{R} : f_{X_i}(\underline{x}) > 0\}$ denote respectively the p.m.f. (or p.d.f.) and support of X_i (or $f_{X_i}(\cdot)$), $i = 1, \dots, p$.

The proof of the following theorem, being similar to that of Theorem 3.2, Module 3, is omitted.

Theorem 5.1

Let $\psi : \mathbb{R}^p \to \mathbb{R}$ be a Borel function such that $E(\psi(\underline{X}))$ is finite.

(i) If $\underline{X}$ is of discrete type then

$$E\left(\psi(\underline{X})\right) = \sum_{\underline{x} \in S_{\underline{X}}} \psi(\underline{x}) f_X(x).$$

(ii) If $\underline{X}$ is of absolutely continuous type then

$$E\left(\psi(\underline{X})\right) = \int_{\mathbb{R}^p} \psi(\underline{x}) f_{\underline{X}}(\underline{x}) d\underline{x}. \blacksquare$$

Definition 5.1

Some special kind of expectations are defined below:

(i) For non-negative integers $k_1, \dots, k_p$, let $\psi(\underline{x}) = x_1^{k_1} \cdots x_p^{k_p}$. Then

$$\mu'_{k_1,\dots,k_p} = E\left(X_1^{k_1} \cdots X_p^{k_p}\right),$$

provided it is finite, is called a *joint moment of order* $k_1 + \cdots + k_p$ of $\underline{X}$;

(ii) For non-negative integers $k_1, \dots, k_p$, let $\psi(\underline{x}) = (x_1 - E(X_1))^{k_1} \cdots (x_p - E(X_p))^{k_p}$. Then

$$\mu_{k_1,\dots,k_p} = E\left((X_1 - E(X_1))^{k_1} \cdots (X_p - E(X_p))^{k_p}\right),$$

provided it is finite, is called a *joint central moment of order* $k_1 + \cdots + k_p$ of $\underline{X}$;

(iii) Let $\psi(\underline{x}) = (x_i - E(X_i))(x_j - E(X_j))$, $i, j = 1, \dots, p$. Then

$$\text{Cov}(X_i, X_j) = E\big((X_i - E(X_i))(X_j - E(X_j))\big),$$

provided it is finite, is called the covariance between X_i and X_j. ■

Note that

$$\text{Cov}(X_i, X_i) = E((X_i - E(X_i))^2) = \text{Var}(X_i), \quad i = 1, \dots, p,$$

and, for $i, j \in \{1, \dots, p\}$, $i \neq j$,

$$\text{Cov}(X_i, X_j) = E\left(\left(X_i - E(X_i)\right)\left(X_j - E(X_j)\right)\right)$$

$$= E\left(\left(X_j - E(X_j)\right)\left(X_i - E(X_i)\right)\right)$$

$$= \text{Cov}(X_j, X_i).$$

Also, for $i, j \in \{1, \dots, p\}$,

$$\text{Cov}(X_i, X_j) = E\left(\left(X_i - E(X_i)\right)\left(X_j - E(X_j)\right)\right)$$

$$= E(X_i X_j) - E(X_i)E(X_j).$$

Theorem 5.2

Let $\underline{X} = (X_1, \dots, X_{p_1})$ and $\underline{Y} = (Y_1, \dots, Y_{p_2})$ be random vectors and let $a_1, \dots, a_{p_1}$, $b_1, \dots, b_{p_2}$ be real constants. Then, provided the involved expectations are finite,

(i) $E\left(\sum_{i=1}^{p_1} a_i X_i\right) = \sum_{i=1}^{p_1} a_i \, E(X_i);$

(ii) $\text{Cov}\left(\sum_{i=1}^{p_1} a_i X_i, \sum_{j=1}^{p_2} b_j Y_j\right) = \sum_{i=1}^{p_1} \sum_{j=1}^{p_2} a_i b_j \, \text{Cov}(X_i, Y_j).$

In particular

$$\text{Var}\left(\sum_{i=1}^{p_1} a_i X_i\right) = \sum_{i=1}^{p_1} a_i^2 \, \text{Var}(X_i) + \sum_{i=1}^{p_1} \sum_{\substack{j=1 \\ j \neq i}}^{p_2} a_i a_j \, \text{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{p_1} a_i^2 \, \text{Var}(X_i) + 2 \sum_{1 \leq i < j} \sum_{\leq p_1} a_i a_j \, \text{Cov}(X_i, X_j).$$

Proof. We will provide the proof for the absolutely continuous case. The proof for the discrete case follows similarly.

(i) Let $f_{\underline{X}}(\cdot)$ denote the joint p.d.f. of $\underline{X} = (X_1, \dots, X_{p_1})$.Then

$$E\left(\sum_{i=1}^{p_1} a_i X_i\right) = \int_{\mathbb{R}^{p_1}} \left(\sum_{i=1}^{p_1} a_i x_i\right) f_{\underline{X}}(\underline{x}) d\underline{x}$$

$$= \sum_{i=1}^{p_1} a_i \int_{\mathbb{R}^{p_1}} x_i \, f_{\underline{X}}(\underline{x}) d\underline{x}$$

$$= \sum_{i=1}^{p_1} a_i E(X_i). \qquad \text{(using Theorem 5.1)}$$

(ii) We have

$$\mathrm{Cov}\left(\sum_{i=1}^{p_1} a_i X_i, \sum_{j=1}^{p_2} b_j Y_j\right) = E\left(\left(\sum_{i=1}^{p_1} a_i X_i - E\left(\sum_{i=1}^{p_1} a_i X_i\right)\right)\left(\sum_{j=1}^{p_2} b_j Y_j - E\left(\sum_{j=1}^{p_2} b_j Y_j\right)\right)\right)$$

$$= E\left(\left(\sum_{i=1}^{p_1} a_i X_i - \sum_{i=1}^{p_1} a_i E\,(X_i)\right)\left(\sum_{j=1}^{p_2} b_j Y_j - \sum_{j=1}^{p_2} b_j E\,(Y_j)\right)\right) \text{ (using (i))}$$

$$= E\left(\left(\sum_{i=1}^{p_1} a_i \left(X_i - E(X_{\mathrm{i}})\right)\right)\left(\sum_{j=1}^{p_2} b_j \left(Y_j - E(Y_j)\right)\right)\right)$$

$$= E\left(\sum_{i=1}^{p_1}\sum_{j=1}^{p_2} a_i b_j \left(X_i - E(X_i)\right)\left(Y_j - E(Y_j)\right)\right)$$

$$= \sum_{i=1}^{p_1}\sum_{j=1}^{p_2} a_i b_j E\left(\left(X_i - E(X_i)\right)\left(Y_j - E(Y_j)\right)\right) \qquad \text{(again using (i))}$$

$$= \sum_{i=1}^{p_1}\sum_{j=1}^{p_2} a_i b_j \;\mathrm{Cov}(X_i, X_j).$$

Also,

$$\operatorname{Var}\left(\sum_{i=1}^{p} a_i X_i\right) = \operatorname{Cov}\left(\sum_{i=1}^{p_1} a_i X_i\,, \sum_{j=1}^{p_1} a_j X_j\right)$$

$$= \sum_{i=1}^{p_1}\sum_{j=1}^{p_1} a_i a_j \operatorname{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{p_1} a_i^2 \operatorname{Cov}(X_i, X_i) + \sum_{i=1}^{p_1}\sum_{\substack{j=1\\ i\neq j}}^{p_1} a_i a_j \operatorname{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{p_1} a_i^2 \operatorname{Var}(X_i) + \sum_{i=1}^{p_1}\sum_{\substack{j=1\\ i\neq j}}^{p_1} a_i a_j \operatorname{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{p_1} a_i^2 \operatorname{Var}(X_i) + 2 \sum_{1\le i<j}\sum_{\le p_1} a_i a_j \operatorname{Cov}(X_i, X_j).$$

$$\left(\text{since } \operatorname{Cov}(X_i, X_j) = \operatorname{Cov}(X_j, X_i)\right). \blacksquare$$

Theorem 5.3

Let $\underline{X}_1, \dots, \underline{X}_p$ be independent random vectors, where $\underline{X}_i$ is r_i- dimensional, $i = 1, \dots, p$.

(i) Let $\psi_i \colon \mathbb{R}^{r_i} \to \mathbb{R}$, $i = 1, 2, \dots, p$, be Borel functions. Then

$$E\left(\prod_{i=1}^{p} \psi_i\left(\underline{X}_i\right)\right) = \prod_{i=1}^{p} E\left(\psi_i(\underline{X}_i)\right),$$

provided the involved expectations are finite.

(ii) For $A_i \in \mathcal{B}_{r_i}$, $i = 1, \dots, p$,

$$P\left(\{\underline{X}_i \in A_i, i = 1, \dots, p\}\right) = \prod_{i=1}^{p} P\left(\{\underline{X}_i \in A_i\}\right).$$

Proof. We will provide the proof for the absolutely continuous case. The proof for the discrete case follows similarly and is left as an exercise.

(i) Let $\underline{X} = (\underline{X}_1, \dots, \underline{X}_p)$. Since $\underline{X}_1, \dots, \underline{X}_p$ are independent. We have

$$f_{\underline{X}}\left(\underline{x}_1, \dots, \underline{x}_p\right) = \prod_{i=1}^{p} f_{\underline{X}_i}\left(\underline{x}_i\right), \quad \forall\, (\underline{x}_1, \dots, \underline{x}_p) \in \mathbb{R}^r,$$

where $r = \sum_{i=1}^{p} r_i$. Therefore,

$$\begin{aligned}
E\left(\prod_{i=1}^{p} \psi_i\left(\underline{X}_i\right)\right) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{i=1}^{p} \psi_i\left(\underline{x}_i\right)\right)\left(\prod_{i=1}^{p} f_{\underline{X}_i}\left(\underline{x}_i\right)\right) d\underline{x}_1 \cdots d\underline{x}_p \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{p} \left(\psi_i\left(\underline{x}_i\right) f_{\underline{X}_i}\left(\underline{x}_i\right)\right) d\underline{x}_1 \cdots d\underline{x}_p, \\
&= \left(\int_{\mathbb{R}^{r_1}}^{\infty} \psi_1\left(\underline{x}_1\right) f_{\underline{X}_1}\left(\underline{x}_1\right) d\underline{x}_1\right) \cdots \left(\int_{\mathbb{R}^{r_p}} \psi_p\left(\underline{x}_p\right) f_{\underline{X}_p}\left(\underline{x}_p\right) d\underline{x}_p\right) \\
&= E\left(\psi_1\left(\underline{X}_1\right)\right) \cdots E\left(\psi_p\left(\underline{X}_p\right)\right).
\end{aligned}$$

(ii) Let

$$\psi_i\left(\underline{X}_i\right) = \begin{cases} 1, & \text{if } \underline{X}_i \in A_i \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, \dots, p,$$

so that

$$\prod_{i=1}^{p} \psi_i\left(\underline{X}_i\right) = \begin{cases} 1, & \text{if } \underline{X}_i \in A_i, i = 1, \dots, p \\ 0, & \text{otherwise} \end{cases}.$$

Now using (i) we get

$$E\left(\prod_{i=1}^{p} \psi_i\left(\underline{X}_i\right)\right) = \prod_{i=1}^{p} E\left(\psi_i\left(\underline{X}_i\right)\right)$$

$$\Rightarrow P(\{X_i \in A_i, i = 1, \dots, p\}) = \prod_{i=1}^{p} P(\{X_i \in A_i\}). \blacksquare$$

Corollary 5.1

Let $X_1, \dots, X_p$ be independent random variables. Then

$$\operatorname{Cov}\left(X_i, X_j\right) = 0, \quad \forall i \neq j,$$

and, for real constants $a_1, \dots, a_p$,

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

provided the involved expectations are finite.

Proof. Fix $i, j \in \{1, \dots, p\},\ i \neq j$. Using Theorem 5.3 (i), we have

$$E(X_i X_j) = E(X_i)E(X_j)$$

$$\Longrightarrow \text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = 0.$$

By Theorem 5.2 we have

$$\text{Var}\left(\sum_{i=1}^{p} a_i X_i\right) = \sum_{i=1}^{p} a_i^2 \,\text{Var}(X_i) + \sum_{\substack{i=1 \\ i \neq j}}^{p} \sum_{j=1}^{p} a_i a_j \,\text{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{p} a_i^2 \,\text{Var}(X_i). \qquad \left(\text{since } \text{Cov}(X_i, X_j) = 0, i \neq j\right). \blacksquare$$

Definition 5.2

(i) The correlation coefficient between random variables X and Y is defined by

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

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

(ii) Random variables X and Y are said to be uncorrelated if $\text{Cov}(X, Y) = 0$. $\blacksquare$

Note that $\rho(X, Y) = \rho(Y, X)$. Also from Corollary 5.1 it is clear that if X and Y are independent random variables then they are uncorrelated. However, as the following examples illustrates, the converse may not be true (i.e., uncorrected random variables may not be independent).

Example 5.1

Let (X, Y)be a bivariate random vector of discrete type with p.m.f. given by

(x, y)	$(-1, 1)$	$(0, 0)$	$(1, 1)$
$f_{X,Y}(x, y)$	p_1	p_2	p_1

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

Clearly

$$E(XY) = (-1)p_1 + (0)p_2 + (1)p_1 = 0$$

$$E(X) = (-1)p_1 + (0)p_2 + (1)p_1 = 0$$

$$E(Y) = (1)p_1 + (0)p_2 + (1)p_1 = 2p_1$$

$$\Longrightarrow \mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0$$

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

However

$$P(\{(X,Y) = (-1,1)\}) = p_1 \neq 2{p_1}^2 = P(\{X = -1\})P(\{Y = 1\}),$$

implying that X and Y are not independent. ■

Example 5.2

Let $\underline{X} = (X_1, X_2)$ be a bivariate random vector of absolutely continuous type with p.d.f. given by

$$f_{\underline{X}}(x_1,x_2) = \begin{cases} 1, & \text{if } 0 < |x_2| \leq x_1 < 1 \\ 0, & \text{otherwise} \end{cases}.$$

Then

$$E(X_1X_2) = \int_0^1 \int_{-x_1}^{x_1} x_1x_2\,dx_2\,dx_1 = 0$$

$$E(X_1) = \int_0^1 \int_{-x_1}^{x_1} x_1\,dx_2\,dx_1 = \frac{2}{3}$$

$$E(X_2) = \int_0^1 \int_{-x_1}^{x_1} x_2\,dx_2\,dx_1 = 0$$

and

$$\mathrm{Cov}(X_1,X_2) = E(X_1X_2) - E(X_1)E(X_2) = 0.$$

Therefore,

$$\rho(X_1,X_2) = 0,$$

i.e., X_1 and X_2 are uncorrelated. Also

$$f_{X_1}(x_1) = \begin{cases} \int_{-x_1}^{x_1} dx_1\,, & \text{if } 0 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 2x_1, & \text{if } 0 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_{X_2}(x_2) = \begin{cases} \int_{|x_2|}^{1} dx_1\,, & \text{if} -1 < x_2 < 1 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 1 - |x_2|, & \text{if} -1 < x_2 < 1 \\ 0, & \text{otherwise} \end{cases}.$$

Clearly

$$f_{X_1,X_2}(x_1, x_2) \neq f_{X_1}(x_1) f_{X_2}(x_2), \quad \forall \underline{x} = (x_1, x_2) \in \mathbb{R}^2,$$

and therefore X_1 and X_2 are not independent.

One can also infer that X_1 and X_2 are not independent by directly observing from the joint p.d.f. $f_{\underline{X}}(\cdot)$ that $S_{\underline{X}} = \left\{\underline{x} \in \mathbb{R}^2 : f_X\left(\underline{x}\right) > 0\right\} = \{(x_1, x_2) : 0 < |x_2| \le x_1 < 1\}, S_{X_1} = \left\{x_1 \in \mathbb{R}^1 : f_{X_1}(x_1) > 0\right\} = (0, 1),\ S_{X_2} = \left\{x_2 \in \mathbb{R}^1 : f_{X_2}(x_2) > 0\right\} = (-1, 1)$ and that $S_{\underline{X}} \neq S_{X_1} \times S_{X_2}$. ■

Theorem 5.4

6.5.1 Cauchy- Schwarz Inequality for Random Variables

Let (X, Y) be a bivariate random vector. Then, provided the involved expectations are finite,

$$(E(XY))^2 \le E(X^2) E(Y^2). \tag{5.1}$$

The equality in (5.1) is attained if, and only if, $P(\{Y = cX\}) = 1$ or $P(\{X = cY\}) = 1$, for some real constant c.

Proof. Consider the following two cases.

Case I. $E(X^2) = 0$.

In this case $P(\{X = 0\}) = 1$ (see Theorem 3.3 (iii), Module 3) and hence $P(\{XY = 0\}) = 1$. It follows that $E(XY) = 0,\ E(X) = 0, P(\{X = cY\}) = 1$, (for $c = 0$) and the equality in (5.1) is attained.

Case II. $E(X)^2 > 0$.

Then,

$$0 \leq E((Y-\lambda X)^2) = E(X^2)\lambda^2 - 2E(XY)\lambda + E(Y^2)$$

i.e., $$E(X^2)\lambda^2 - 2E(XY)\lambda + E(Y^2) \geq 0, \ \forall \lambda \in \mathbb{R}.$$

This implies that the discriminant of the quadratic equation $E(X^2)\lambda^2 - 2E(XY)\lambda + E(Y^2) = 0$ is non- negative, i.e.,

$$4(E(XY))^2 \leq 4E(X^2)E(Y^2)$$

$$\Rightarrow (E(XY))^2 \leq E(X^2)E(Y^2),$$

and the equality is attained if, and only if,

$$E((Y-cX)^2) = 0, \ \text{ for some } c \in \mathbb{R}$$

$$\Leftrightarrow P(\{Y = cX\}) = 1, \ \text{ for some } c \in \mathbb{R}. \blacksquare$$

Corollary 5.2

Let (X_1, X_2) be a bivariate random vector with $E\,(X_i) = \mu_i \in (-\infty, \infty)$ and $\mathrm{Var}(X_i) = \sigma_i^2 \in (0, \infty), i = 1, 2$. Then

(i) $|\rho(X_1, X_2)| \leq 1$;

(ii) $\rho(X_1, X_2) = \pm 1$ if, and only if, $\frac{X_1 - \mu_1}{\sigma_1} = d\,\frac{X_2 - \mu_2}{\sigma_2}$, for some real constant d; here $\mu_i = E(X_i), \ i = 1,2$.

Proof. Taking $X = X_1 - \mu_1$ and $Y = X_2 - \mu_2$ in Theorem 5.4, we get

$$(E\big((X_1-\mu_1)(X_2-\mu_2)\big))^2 \leq E((X_1-\mu_1)^2)E((X_2-\mu_2)^2)$$

$$\Leftrightarrow \rho^2(X_1, X_2) \leq 1$$

$$\Leftrightarrow |\rho(X_1, X_2)| \leq 1,$$

and the equality is attained if and only if,

$$P((X_1-\mu_1) = c(X_2-\mu_2)) = 1 \text{ , for some } c \in \mathbb{R}$$

$$\Leftrightarrow P\left(\frac{X_1-\mu_1}{\sigma_1} = d\,\frac{X_2-\mu_2}{\sigma_2}\right) = 1 \text{ , } \ \text{ for some } d \in \mathbb{R}. \blacksquare$$

Let $\underline{X} = (\underline{Y}, \underline{Z})$ be a p-dimensional random vector of either discrete type or of absolutely continuous type and let $\underline{Y}$ and $\underline{Z}$, respectively, be p_1 and p_2 dimensional, so that $p = p_1 +$

p_2. For a given $\underline{z} \in S_{\underline{Z}}$ (or $\underline{z}$ satisfying (3.5) and $f_{\underline{Z}}(\underline{z}) > 0$) the conditional p.m.f. (or p.d.f.) of $\underline{Y}$ given $\underline{Z} = \underline{z}$ is given by

$$f_{\underline{Y}|\underline{Z}}\left(\underline{y}|\underline{z}\right) = \frac{f_{\underline{Y},\underline{Z}}\left(\underline{y},\underline{z}\right)}{f_{\underline{Z}}(\underline{z})}, \quad \underline{y} \in \mathbb{R}^{p_1}.$$

Let $\psi: \mathbb{R}^{p_1} \to \mathbb{R}$ be a Borel function and let $\underline{z} \in S_{\underline{Z}}$ (or $\underline{z}$ satisfies (3.5) with $f_{\underline{Z}}(\underline{z}) > 0$) . Then the conditional expectation of $\psi(\underline{Y})$ given that $\underline{Z} = \underline{z}$ may be defined by

$$E\left(\psi(\underline{Y})\middle|\underline{Z} = \underline{z}\right) = \int_{\mathbb{R}^{p_1}} \psi\left(\underline{y}\right) f_{\underline{Y}|\underline{Z}}\left(\underline{y}|\underline{z}\right) \mathrm{d}\underline{y}\,,$$

provided the expectation is finite.

Similarly the conditional variance of $\psi(\underline{Y})$, given that $\underline{Z} = \underline{z}$, may be defined by

$$\mathrm{Var}\left(\psi(\underline{Y})\middle|\underline{Z} = \underline{z}\right) = E\left(\left(\psi(\underline{Y}) - E(\psi(\underline{Y})|\underline{Z} = \underline{z})\right)^2\middle|\underline{Z} = \underline{z}\right).$$

Throughout we will use the following notation

$$E\left(\psi(\underline{Y})\middle|\underline{Z}\right) = \psi^*(\underline{Z}), \tag{5.2}$$

where ψ^* is defined by

$$\psi^*(\underline{z}) = E\left(\psi(\underline{Y})\middle|\underline{Z} = \underline{z}\right), \tag{5.3}$$

for all $\underline{z} \in S_{\underline{z}}$ (or all $\underline{z}$ satisfying (3.5) with $f_{\underline{Z}}(\underline{z}) > 0$) .