MODULE 6

RANDOM VECTOR AND ITS JOINT DISTRIBUTION

LECTURE 28

Topics

6.3 CONDITIONAL DISTRIBUTIONS
6.4 INDEPENDENT RANDOM VARIABLES

6.3 CONDITIONAL DISTRIBUTIONS

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\underline{X} = (X_1, \dots, X_p): \Omega \to \mathbb{R}^p$ be a p-dimensional $(p \geq 2)$ random vector with distribution function $F_{\underline{X}}(\cdot)$.

Definition 3.1

Let $D \in \mathcal{B}_p$ be such that $P(\{\underline{X} \in D\}) > 0$. Then the conditional distribution function of $\underline{X}$ given that $\underline{X} \in D$ is defined by

$$\begin{aligned} F_{\underline{X}|D}(\underline{x}) &= P(\{\underline{X} \in (-\underline{\infty}, \underline{x}]\} | \{\underline{X} \in D\}) \\ &= \frac{P(\{\underline{X} \in (-\underline{\infty}, \underline{x}] \cap D\})}{P(\{\underline{X} \in D\})} \\ &= \frac{P(\{X_1 \leq x_1, \dots, X_p \leq x_p, \underline{X} \in D\})}{P(\{\underline{X} \in D\})}, \quad \underline{x} \in \mathbb{R}^p. \blacksquare \end{aligned}$$

For a given $D \in \mathcal{B}_p$ it can be verified that $F_{\underline{X}|D}(\cdot)$ is a distribution function, i.e., it satisfies properties (i) $-$ (iv) of Theorem 1.3. For a fixed $k \in \{1, \dots, p-1\}$, let $\underline{Y} = (X_1, \dots, X_k) (= (Y_1, \dots, Y_k)$, say) and $\underline{Z} = (X_{k+1}, \dots, X_p)$ $(= (Z_1, \dots, Z_{p-k})$, say$)$, so that $\underline{X} = (\underline{Y}, \underline{Z})$. In many situations it may be of interest to study the conditional probability distribution of numerical characteristic $\underline{Y}$ given a fixed value of numerical characteristic $\underline{Z}$. For example if X_1 and X_2 denote respectively the heights and weights of newly born babies in a community then it may be of interest to study the probability distribution of heights of babies having weight of 3Kg (i.e., conditional distribution of X_1 given that $\{X_2 = 3\}$).

To make the above discussion precise, first suppose that $\underline{X} = \left(\underline{Y}, \underline{Z}\right)$ is of discrete type so that $\underline{Y}$ and $\underline{Z}$ are also of discrete type (see Theorem 2.1 (i)). Let $S_{\underline{X}}$, $S_{\underline{Y}}$ and $S_{\underline{Z}}$ denote the supports of $\underline{X}, \underline{Y}$ and $\underline{Z}$ respectively. Further let $f_{\underline{X}}(\cdot) \doteqdot f_{\underline{Y},\underline{Z}}(\cdot)$ and $f_{\underline{Z}}(\cdot)$denote the joint p.m.f.s of $\underline{X} = \left(\underline{Y}, \underline{Z}\right)$ and $\underline{Z}$, respectively. Let $\underline{z} \in S_{\underline{Z}}$ be fixed such that $f_{\underline{Z}}\left(\underline{z}\right) = P\left(\left\{\underline{Z} = \underline{z}\right\}\right) > 0$. Define$S_{\underline{Y}|\underline{Z}=\underline{z}} = \left\{\underline{y} \in \mathbb{R}^k : \left(\underline{y}, \underline{z}\right) \in S_{\underline{X}}\right\}$. Then $S_{\underline{Y}|\underline{Z}=\underline{z}} \subseteq S_{\underline{Y}} = \left\{\underline{y} \in \mathbb{R}^p : \left(\underline{y}, \underline{t}\right) \in S_{\underline{X}},\ \text{for some } \underline{t} \in \mathbb{R}^{p-k}\right\}$ and, using Definition 3.1, the conditional distribution function of $\underline{Y}$ given $\left\{\underline{Z} = \underline{z}\right\} \left(= \left\{\underline{Z} \in \left\{\underline{z}\right\}\right\}\right)$ is given by

$$F_{\underline{Y}|\underline{Z}}\left(\underline{y}\,|\underline{z}\right) = \frac{P\left(\left\{Y_1 \le y_1, \dots, Y_k \le y_k,\ \ \underline{Z} = \underline{z}\right\}\right)}{P\left(\left\{\underline{Z} = \underline{z}\right\}\right)}, \quad \underline{y} \in \mathbb{R}^k \tag{3.1}$$

$$= \frac{\sum_{\underline{x} \in S_{\underline{Y}|\underline{Z}=\underline{z}} \cap \left(\left(-\underline{\infty}, \underline{y}\right]\right)} f_{\underline{X}}\left(\underline{x}, \underline{z}\right)}{f_{\underline{Z}}\left(\underline{z}\right)}$$

$$= \sum_{\underline{x} \in S_{\underline{Y}|\underline{Z}=\underline{z}} \cap \left(\left(-\underline{\infty}, \underline{y}\right]\right)} \frac{f_{\underline{X}}\left(\underline{x}, \underline{z}\right)}{f_{\underline{Z}}\left(\underline{z}\right)}. \tag{3.2}$$

Clearly the p.m.f. corresponding to distribution function $F_{\underline{Y}|\underline{Z}}\left(\cdot\,|\underline{z}\right)$ is (see Remark 2.1 (xi))

$$f_{\underline{Y}|\underline{Z}}\left(\underline{y}\,|\underline{z}\right) = \begin{cases} \dfrac{f_{\underline{Y},\underline{Z}}\left(\underline{y}, \underline{z}\right)}{f_{\underline{Z}}\left(\underline{z}\right)}, & \text{if } \underline{y} \in S_{\underline{Y}|\underline{Z}=\underline{z}} \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

$$= \frac{f_{\underline{Y}\,\underline{Z}}\left(\underline{y}, \underline{z}\right)}{f_{\underline{Z}}\left(\underline{z}\right)}, \quad \underline{y} \in \mathbb{R}^k \tag{3.4}$$

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

The above discussion leads to the following definition.

Definition 3.2

Let $\underline{X} = (X_1, \ldots, X_p)$ be a discrete type random vector. Then, under the above notation,

(i) the conditional p.m.f. of $\underline{Y}$ given $\underline{Z} = \underline{z}$ (where $\underline{z} \in S_{\underline{Z}}$ is fixed) is defined by (3.3) (or (3.4));

(ii) the conditional distribution function of $\underline{Y}$ given $\underline{Z} = \underline{z}$ (where $\underline{z} \in S_{\underline{Z}}$ is fixed) is defined by (3.1) (or (3.2)); ■

Now suppose that $\underline{X} = (\underline{Y}, \underline{Z})$ is of absolutely continuous type so that $\underline{Y}$ and $\underline{Z}$ are also of absolutely continuous type (see Theorem 2.1 (ii)). Let $f_{\underline{X}}(\cdot) \doteqdot f_{\underline{Y},\underline{Z}}(\cdot)$, $f_{\underline{Y}}(\cdot)$ and $f_{\underline{Z}}(\cdot)$denote the p.d.f.s. of $\underline{X}$, $\underline{Y}$ and $\underline{Z}$ respectively. Then we have $P(\{\underline{Z} = \underline{z}\}) = 0, \forall \underline{z} \in \mathbb{R}^{p-k}$ (Remark 2.1 (viii)) and therefore conditional distribution function of $\underline{Y}$ given $\{\underline{Z} = \underline{z}\}$ cannot be defined by (3.1). For $\underline{z} \in \mathbb{R}^{p-k}$, note that

$$\{\underline{Z} = \underline{z}\} = \bigcap_{n_1=1}^{\infty} \cdots \bigcap_{n_{p-k}=1}^{\infty} \left\{z_i - \frac{1}{n_i} < Z_i \le z_i, i = 1, \ldots, p-k\right\},$$

and therefore, using continuity of probability measures,

$$\begin{aligned} P(\{\underline{Z} = \underline{z}\}) &= \lim_{\substack{n_i \to \infty \\ i=1,\ldots,p-k}} P\left(\left\{z_i - \frac{1}{n_i} < Z_i \le z_i, i = 1, \ldots, p-k\right\}\right) \\ &= \lim_{\substack{h_i \downarrow 0 \\ i=1,\cdots,p-k}} P(\{z_i - h_i < Z_i \le z_i, i = 1, \ldots, p-k\}). \end{aligned}$$

Thus if $\underline{z} \in \mathbb{R}^{p-k}$ is such that

$$P(\{z_i - \delta_i < Z_i \le z_i, i = 1, \ldots, p-k\}) > 0, \forall \underline{\delta} = (\delta_1, \ldots, \delta_{p-k}) \in (0, \infty)^{p-k}, \qquad (3.5)$$

then the conditional distribution function of $\underline{Y}$ given $\underline{Z} = \underline{z}$ may be defined by

$$\begin{aligned} F_{\underline{Y}|\underline{Z}}(\underline{y}|\underline{z}) &= \lim_{\substack{h_i \downarrow 0 \\ i=1,\cdots,p-k}} P(\{Y_i \le y_i, i = 1, \ldots, k\} | \{z_i - h_i < Z_i \le z_i, i = 1, \ldots, p-k\}) \quad (3.6) \\ &= \lim_{\substack{h_i \downarrow 0 \\ i=1,\ldots,p-k}} \frac{P(\{Y_i \le y_i, i = 1, \ldots, k, z_i - h_i < Z_i \le z_i, i = 1, \ldots, p-k\})}{P(\{z_i - h_i < Z_i \le z_i, i = 1, \cdots, p-k\})} \\ &= \lim_{\substack{h_i \downarrow 0 \\ i=1,\ldots,p-k}} \frac{\int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_k} \int_{z_1-h_1}^{z_1} \cdots \int_{z_{p-k}-h_{p-k}}^{z_{p-k}} f_{\underline{Y},\underline{Z}}(\underline{s}, \underline{t}) d\underline{t} d\underline{s}}{\int_{z_1-h_1}^{z_1} \cdots \int_{z_{p-k}-h_{p-k}}^{z_{p-k}} f_{\underline{Z}}(\underline{t}) d\underline{t}} \end{aligned}$$

$$= \frac{\int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_k} \left\{ \lim_{\substack{h_i \downarrow 0 \\ i=1,\ldots,p-k}} \frac{1}{h_1 \cdots h_{p-k}} \int_{z_1-h_1}^{z_1} \cdots \int_{z_{p-k}-h_{p-k}}^{z_{p-k}} f_{\underline{Y},\underline{Z}}(\underline{s},\underline{t}) d\underline{t} \right\} d\underline{s}}{\lim_{\substack{h_i \downarrow 0 \\ i=1,\ldots,p-k}} \frac{1}{h_1 \cdots h_{p-k}} \int_{z_1-h_1}^{z_1} \cdots \int_{z_{p-k}-h_{p-k}}^{z_{p-k}} f_{\underline{Z}}(\underline{t}) d\underline{t}}$$

$$= \frac{\int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_k} f_{\underline{Y},\underline{Z}}(\underline{s}\,,\underline{z}) d\underline{s}}{f_{\underline{Z}}(\underline{z})}$$

$$= \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_k} \frac{f_{\underline{Y},\underline{Z}}(\underline{s}\,,\underline{z})}{f_{\underline{Z}}(\underline{z})} d\underline{s}, \quad \underline{y} \in \mathbb{R}^k, \tag{3.7}$$

provided $f_{\underline{Z}}(\underline{z}) > 0$ and $\underline{z}$ is such that (3.5) is satisfied. In that case the p.d.f corresponding to distribution function $F_{\underline{Y}|\underline{Z}}(\cdot\,|\underline{z})$ is given by

$$f_{\underline{Y}|\underline{Z}}\left(\underline{y}|\underline{z}\right) = \frac{f_{\underline{Y}\underline{Z}}\left(\underline{y},\underline{z}\right)}{f_{\underline{Z}}(\underline{z})}, \quad \underline{y} \in \mathbb{R}^k\,. \tag{3.8}$$

The above discussion is summarized in the following definition.

Definition 3.3

Let $\underline{X} = (X_1, \ldots, X_p)$ be a random vector of absolutely continuous type. Let $\underline{z} \in \mathbb{R}^k$ be such that $f_{\underline{Z}}(\underline{z}) > 0$ and it satisfies (3.5). Then

(i) the conditional p.d.f. of $\underline{Y}$ given $\underline{Z} = \underline{z}$ is defined by (3.8);

(ii) the conditional distribution function of $\underline{Y}$ given $\underline{Z} = \underline{z}$ is defined by (3.6) (or (3.7)). ■

Remark 3.1

Using (3.4) and (3.8), for fixed $\underline{z} \in D = \{\underline{t} \in \mathbb{R}^{p-k} : f_{\underline{Y}|\underline{Z}}(\cdot\,|\underline{t}) \text{ is defined}\}$, the conditional p.m.f./p.d.f. of $\underline{Y}$ given $\underline{Z} = \underline{z}$ is given by

$$f_{\underline{Y}|\underline{Z}}\left(\underline{y}|\underline{z}\right) = c(\underline{z}) f_{\underline{Y},\underline{Z}}\left(\underline{y},\underline{z}\right), \ \underline{y} \in \mathbb{R}^k,$$

where $c(\underline{z})$ is the normalizing constant. ■

Example 3.1

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

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

(i) Find the conditional p.m.f. of X_1 given that $(X_2, X_3) = (2, 1)$;

(ii) Find the conditional p.m.f. of (X_1, X_3) given that $X_2 = 3$.

Solution.

(i) We have

$$f_{X_1|(X_2,X_3)}\big(x_1 \big| (2, 1)\big) = \frac{P(\{X_1 = x_1, X_2 = 2, X_3 = 1\})}{P(\{(X_2, X_3) = (2, 1)\})}$$

$$= \begin{cases} \frac{2x_1}{72\, P(\{X_2 = 2,\ X_3 = 1\})}, & \text{if } x_1 \in \{1, 2\} \\ 0, & \text{otherwise} \end{cases},$$

$$P(\{X_2 = 2, X_3 = 1\}) = \sum_{x_1=1}^{2} P(\{X_1 = x_1, X_2 = 2, X_3 = 1\})$$

$$= \frac{2}{72}(1 + 2)$$

$$= \frac{1}{12}.$$

Therefore

$$f_{X_1|(X_2,X_3)}(x_1|(2, 1)) = \begin{cases} \frac{x_1}{3}, & \text{if } x_1 \in \{1, 2\} \\ 0, & \text{otherwise} \end{cases}.$$

(ii) We have

$$f_{X_1,X_3|X_2}(x_1, x_3|3) = \frac{P(\{X_1 = x_1, X_2 = 3, X_3 = x_3\})}{P(\{X_2 = 3\})}.$$

Using Example 2.2, $P(\{X_2 = 3\}) = \frac{1}{2}$ and therefore

$$f_{X_1,X_3|X_2}(x_1, x_3|3) = \begin{cases} \frac{x_1 x_3}{12}, & \text{if } (x_1, x_3) \in \{1, 2\} \times \{1, 3\} \\ 0, & \text{otherwise} \end{cases}. \blacksquare$$

Example 3.2

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

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

(i) For $0 < x_3 < x_2 < 1$, find the conditional p.d.f. of X_1 given $(X_2, X_3) = (x_2, x_3)$;

(ii) For $0 < x_2 < 1$, find the conditional p.d.f. of (X_1, X_3) given $X_2 = x_2$.

Solution.

(i) For $0 < x_3 < x_2 < 1$

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

Using Example 2.3 (ii), for $0 < x_3 < x_2 < 1$, we have

$$f_{X_2,X_3}(x_2, x_3) = -\frac{\ln x_2}{x_2}.$$

Therefore,

$$f_{X_1|(X_2,X_3)}(x_1|x_2, x_3) = \begin{cases} -\frac{1}{x_1 \ln x_2}, & \text{if } x_2 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases}.$$

Alternatively $f_{X_1|(X_2,X_3)}(x_1|x_2, x_3)$ can be found by using Remark 3.1.

(ii) For $0 < x_2 < 1$,

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

Using Example 2.3 (iii) we have, for $0 < x_2 < 1$,

$$f_{X_2}(x_2) = -\ln x_2.$$

Therefore, for $0 < x_2 < 1$,

$$f_{X_1,X_3|X_2}(x_1,x_3|x_2) = \begin{cases} -\dfrac{1}{x_1x_2\ln x_2}, & \text{if } x_2 < x_1 < 1,\;\; 0 < x_3 < x_2 \\ 0, & \text{otherwise} \end{cases}.$$

Alternatively $f_{X_1,X_3}(x_1,x_3|x_2)$ can be found using Remark 3.1. ■

6.4 INDEPENDENT RANDOM VARIABLES

Let $(\Omega,\mathcal{F},P)$ be a probability space and let $\{X_\lambda : \lambda \in \Lambda\}$ be a collection of random variables, where $\Lambda \subseteq \mathbb{R}$ is a non-empty index set.

Definition 4.1

The random variables $\{X_\lambda : \lambda \in \Lambda\}$ are said to be (statistically) independent if for any finite sub collection $\{\lambda_1,\ldots,\lambda_p\} \subseteq \Lambda$ we have

$$F_{X_{\lambda_1},\ldots,X_{\lambda_p}}(x_1,\ldots,x_p) = \prod_{i=1}^{p} F_{X_{\lambda_i}}(x_i),\;\; \forall \underline{x} = (x_1,\ldots,x_p) \in \mathbb{R}^p. \blacksquare$$

The observations made in the following remark are immediate from Definition 4.1.

Remark 4.1

(i) The random variables $\{X_\lambda : \lambda \in \Lambda\}$ are independent if, and only if, every finite sub collection $\{X_{\lambda_1},\ldots,X_{\lambda_p}\} \subseteq \{X_\lambda : \lambda \in \Lambda\}$ constitutes a collection of independent random variables;

(ii) Suppose that $\Lambda_1 \subseteq \Lambda_2 \subseteq \mathbb{R}$ and $\Lambda_1 \neq \phi$. Then

$$\{X_\lambda : \lambda \in \Lambda_2\} \text{ are independent} \Rightarrow \{X_\lambda : \lambda \in \Lambda_1\} \text{ are independent;}$$

(iii) It can be shown that (see Theorem 5.3 (ii) in the sequel) $X_1,\ldots,X_p$ are independent if, and only if, for any $A_i \in \mathcal{B}_1, i = 1,\ldots,p$,

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

Theorem 4.1

Let $X = (X_1, \dots, X_p)$ be a p-dimentsional $(p \geq 2)$ random vector with joint distribution function $F_{X_1,\dots,X_p}(\cdot)$. Let $F_{X_i}(\cdot)$ denote the marginal distribution function of X_i, $i = 1, \dots, p$. Then the random variables $X_1, \dots, X_p$ are independent if, and only if,

$$F_{X_1,\dots,X_p}(x_1, \dots, x_p) = \prod_{i=1}^{p} F_{X_i}(x_i), \quad \forall \underline{x} = (x_1, \dots, x_p) \in \mathbb{R}^p. \qquad (4.1)$$

Proof. First suppose that $X_1, \dots, X_p$ are independent. Then, by definition, (4.1) obviously holds. Conversely suppose that (4.1) holds. Then, for any $\underline{y} \in \mathbb{R}^p$ and any permutation $(\beta_1, \dots, \beta_p)$ of $(1, \dots, p)$,

$$P(\{X_i \leq y_i, i = 1, \dots, p\}) = \prod_{i=1}^{p} P(\{X_i \leq y_i\})$$

$$\Rightarrow P\left(\{X_{\beta_i} \leq y_{\beta_i}, i = 1, \dots, p\}\right) = \prod_{i=1}^{p} P(\{X_{\beta_i} \leq y_{\beta_i}\})$$

$$\Rightarrow F_{X_{\beta_1},\dots,X_{\beta_p}}\left(y_{\beta_1}, \dots, y_{\beta_p}\right) = \prod_{i=1}^{p} F_{X_{\beta_i}}(y_{\beta_i}), \forall \underline{y} = (y_1, \dots, y_p) \in \mathbb{R}^p, \underline{\beta} = (\beta_1, \dots, \beta_p) \in S_p,$$

where S_p denotes the set of all permutations of $(1, \dots, p)$. It follows that, for any $(\beta_1, \dots, \beta_p) \in S_p$ and any $\underline{x} \in \mathbb{R}^p$,

$$F_{X_{\beta_1},\dots,X_{\beta_p}}(x_1, \dots, x_p) = \prod_{i=1}^{p} F_{X_{\beta_i}}(x_i). \qquad (4.2)$$

Let $q \in \{2, \dots, p\}$ and let $\{\lambda_1, \dots, \lambda_q\} \subseteq \{1, \dots, p\} = \Lambda$, say. Let $\lambda_{q+1}, \dots, \lambda_p$ be such that $\Lambda - \{\lambda_1, \dots, \lambda_q\} = \{\lambda_{q+1}, \dots, \lambda_p\}$. Then $(\lambda_1, \dots, \lambda_q, \lambda_{q+1}, \dots, \lambda_p) \in S_p$ and by Lemma 1.2

$$F_{X_{\lambda_1},\dots,X_{\lambda_q}}(x_1, \dots, x_q) = \lim_{\substack{x_j \to \infty \\ j = q+1,\dots,p}} F_{X_{\lambda_1},\dots,X_{\lambda_p}}(x_1, \dots, x_p)$$

$$= \lim_{\substack{x_j \to \infty \\ j = q+1,\cdots,p}} \prod_{l=1}^{p} F_{X_{\lambda_l}}(x_l) \qquad \text{(using (4.2))}$$

$$= \prod_{l=1}^{q} F_{X_{\lambda_l}}(x_l), \;\; \forall \underline{x} = (x_1, \dots, x_q) \in \mathbb{R}^q.$$

Hence the result follows. ■

The following remark is immediate from the above theorem and Remark 1.2(ii).

Remark 4.2

Random variables $X_1, \dots, X_p$ are independent if, and only if, for any $\underline{\beta} = (\beta_1, \dots, \beta_p) \in S_p$ the random variables $X_{\beta_1}, \dots, X_{\beta_p}$ are independent. ■

Theorem 4.2

Let $\underline{X} = (X_1, \dots, X_p)$ be a p-dimensional ($p \geq 2$) random vector of either discrete type or of absolutely continuous type. Let $f_{X_1,\dots,X_p}(\cdot)$ denote the joint p.m.f. (or p.d.f.) of $\underline{X}$ and let $f_{X_i}(\cdot)$ denote the marginal p.m.f. (or p.d.f.) of $X_i, \; i = 1, \dots, p$. Then

(i) $X_1, \dots, X_p$ are independent if, and only if,

$$f_{X_1,\dots,X_p}(x_1, \dots, x_p) = \prod_{i=1}^{p} f_{X_i}(x_i), \;\; \forall \underline{x} = (x_1, \dots, x_p) \in \mathbb{R}^p. \qquad (4.3)$$

(ii) $X_1, \dots, X_p$ are independent if, and only if,

$$f_{X_1,\dots,X_p}(x_1, \dots, x_p) = \prod_{i=1}^{p} g_i(x_i), \;\; \forall \underline{x} = (x_1, \dots, x_p) \in \mathbb{R}^p, \qquad (4.4)$$

for some non-negative functions $g_1(\cdot), \dots, g_p(\cdot)$. In that case $f_{X_i}(x_i) = d_i g_i(x), \; x \in \mathbb{R}, i = 1, \dots, p$ for some positive constants $d_1, \dots, d_p$.

(iii) $X_1, X_2, \dots, X_p$ are independent $\Rightarrow S_{\underline{X}} = \prod_{i=1}^{p} S_{X_i}$, where, for a random variable $\underline{Y}, \; S_{\underline{Y}} = \left\{\underline{y} \in \mathbb{R}^p : f_{\underline{Y}}\left(\underline{y}\right) > 0\right\}$.

Proof.

(i) For notational simplicity we will provide the proof for $p = 2$.

Case I. $\underline{X}$ is of discrete type

Let $S_{\underline{X}}$ be the support of $X = (X_1, X_2)$ and let S_{X_i} be the support of $X_i, i = 1, 2$ .... First suppose that (4.3) holds. Then clearly $S_{\underline{X}} = S_{X_1} \times S_{X_2}$(see (iii) proved in the sequel). Therefore, for $\underline{x} = (x_1, x_2) \in \mathbb{R}^2$,

$$F_{X_1,X_2}(x_1,x_2) = \sum_{\underline{y}\in S_{\underline{X}}\cap((-\underline{\infty},\underline{x}])} f_{X_1,X_2}(y_1,y_2)$$

$$= \sum_{y_1\in S_{X_1}\cap(-\infty,x_1]}\ \sum_{y_2\in S_{X_2}\cap(-\infty,x_2]} f_{X_1}(y_1)f_{X_2}(y_2) \qquad (S_{\underline{X}} = S_{X_1}\times S_{X_2})$$

$$= \left(\sum_{y_1\in S_{X_1}\cap(-\infty,x_1]} f_{X_1}(y_1)\right)\left(\sum_{y_2\in S_{X_2}\cap(-\infty,x_2]} f_{X_2}(y_2)\right)$$

$$= F_{X_1}(x_1)F_{X_2}(x_2).$$

Using Theorem 4.1 it follows that X_1 and X_2 are independent.

Conversely suppose that X_1 and X_2 are independent.Then, by Theorem 4.1,

$$F_{X_1,X_2}(z_1,z_2) = F_{X_1}(z_1)F_{X_2}(z_2),\ \ \forall \underline{z} = (z_1,z_2)\in\mathbb{R}^2.$$

Let $\underline{x} = (x_1,x_2)\in\mathbb{R}^2$. Define $\underline{x_n} = \left(x_1 - \frac{1}{n}, x_2 - \frac{1}{n}\right)$, $n = 1, 2, \ldots.$ Then, by Remark 2.1 (v),

$$f_{X_1,X_2}(x_1,x_2) = P(\{X_1 = x_1, X_2 = x_2\})$$

$$= \lim_{n\to\infty}\sum_{k=0}^{2}\ \sum_{\underline{z}\in\Delta_{k,2}((\underline{x}_n,\underline{x}])} F_{X_1,X_2}(z_1,z_2)$$

$$= \lim_{n\to\infty}[F_{X_1,X_2}(x_1,x_2) - F_{X_1,X_2}\left(x_1-\frac{1}{n},x_2\right) - F_{X_1,X_2}\left(x_1,x_2-\frac{1}{n}\right)$$

$$+F_{X_1,X_2}\left(x_1-\frac{1}{n},x_2-\frac{1}{n}\right)]$$

$$= \lim_{n\to\infty}[F_{X_1}(x_1)F_{X_2}(x_2) - F_{X_1}\left(x_1-\frac{1}{n}\right)F_{X_2}(x_2) - F_{X_1}(x_1)F_{X_2}\left(x_2-\frac{1}{n}\right)$$

$$+F_{X_1}\left(x_1-\frac{1}{n}\right)F_{X_2}\left(x_2-\frac{1}{n}\right)]$$

$$= F_{X_1}(x_1)F_{X_2}(x_2) - F_{X_1}(x_1-)F_{X_2}(x_2) - F_{X_1}(x_1)F_{X_2}(x_2-) + F_{X_1}(x_1-)F_{X_2}(x_2-)$$

$$= F_{X_2}(x_2)\left[F_{X_1}(x_1) - F_{X_1}(x_1-)\right] - F_{X_2}(x_2-)[F_{X_1}(x_1) - F_{X_1}(x_1-)]$$

$$= \left[F_{X_1}(x_1) - F_{X_1}(x_1-)\right]\left[F_{X_2}(x_2) - F_{X_2}(x_2-)\right]$$

$$= P(\{X_1 = x_1\})\,P(\{X_2 = x_2\})$$

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

i.e., (4.3) holds.

Case II. $\underline{X}$ is of absolutely continuous type

First suppose that (4.3) holds. Then, for $\underline{x} = (x_1, x_2) \in \mathbb{R}$,

$$F_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1,X_2}(y_1, y_2)\, dy_2 dy_1$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1}(y_1)\, f_{X_2}(y_2)\, dy_2 dy_1$$

$$= \left(\int_{-\infty}^{x_1} f_{X_1}(y_1) dy_1 \right) \left(\int_{-\infty}^{x_2} f_{X_2}(y_2) dy_2 \right)$$

$$= F_{X_1}(x_1) F_{X_2}(x_2).$$

Using Theorem 4.1 it follows that X_1 and X_2 are independent.

Conversely suppose that X_1 and X_2 are independent.Then, by Theorem 4.1,

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

For simplicity assume that $f_{X_1,X_2}(x_1, x_2)$ is continuous everywhere. Then, by Remark 2.1 (xiii)

$$f_{X_1,X_2}(x_1, x_2) = \frac{\partial^2 F_{X_1,X_2}(x_1, x_2)}{\partial x_1 \partial x_2}$$

$$= \frac{\partial^2}{\partial x_1 \partial x_2} \left(F_{X_1}(x_1) F_{X_2}(x_2) \right)$$

$$= \left(\frac{\partial F_{X_1}(x_1)}{\partial x_1} \right) \left(\frac{\partial F_{X_2}(x_2)}{\partial x_2} \right)$$

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

(ii) First suppose that X_1 and X_2 are independent. Then clearly (4.4) holds with the choice $g_i(x_i) = f_{X_i}(x_i),\ x_i \in \mathbb{R}, i = 1, 2$.Conversely suppose that (4.4) holds. Let

$$c_i = \int_{-\infty}^{\infty} g_i(x)\, dx, \;\; i = 1, 2,$$

so that $c_1 \geq 0$, $c_2 \geq 0$ and

$$c_1 c_2 = \left(\int_{-\infty}^{\infty} g_1(x_1) dx_1 \right)\left(\int_{-\infty}^{\infty} g_2(x_2) dx_2 \right)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x_1) g_2(x_2) dx_2\, dx_1$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) dx_2\, dx_1$$

$$= 1.$$

It follows that $c_1 > 0$, $c_2 > 0$ and $c_1 c_2 = 1$. Also

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) dx_2$$

$$= \int_{-\infty}^{\infty} g_1(x_1) g_2(x_2) dx_2$$

$$= c_2 g_1(x_1), \;\; x_1 \in \mathbb{R}.$$

Similarly

$$f_{X_2}(x_2) = c_1 g_2(x_2)\,, \;\; x_2 \in \mathbb{R}\,.$$

Thus we have

$$f_{X_1,X_2}(x_1, x_2) = g_1(x_1) g_2(x_2)$$

$$= (c_1 g_1(x_1))(c_2 g_2(x_2)) \quad (c_1 c_2 = 1)$$

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

Using (i) it follows that X_1 and X_2 are independent.

(iii) Since X_1 and X_2 are independent by (i), $f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) \forall \underline{x} \in \mathbb{R}^2$. Therefore

$$S_{\underline{X}} = \{(x_1, x_2): f_{X_1,X_2}(x_1, x_2) > 0\}$$
$$= \{(x_1, x_2): f_{X_1}(x_1) f_{X_2}(x_2) > 0\}$$
$$= \{x: f_{X_1}(x) > 0\} \times \{y: f_{X_2}(y) > 0\}$$
$$= S_{X_1} \times S_{X_2}. \blacksquare$$

Remark 4.3

(i) Let $\underline{X} = (X_1, X_2)$ be a bivariate vector of either discrete type or of absolutely continuous type. Let $D = \{x_2 \in \mathbb{R}: f_{X_1|X_2}(\cdot\,|x_2\,) \text{ is defined}\}$.Then by Theorem 4.2 (i)

$$X_1 \text{ and } X_2 \text{ are independent} \iff f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2), \forall \underline{x} = (x_1, x_2) \in \mathbb{R}^2$$
$$\iff \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_2}(x_2)} = f_{X_1}(x_1), \forall x_1 \in \mathbb{R}, x_2 \in D$$
$$\iff f_{X_1|X_2}(x_1|x_2) = f_{X_1}(x_1)\,, \quad \forall x_1 \in \mathbb{R}, x_2 \in D.$$

It follows that X_1 and X_2 are independent if, and only if, for every $x_2 \in D$ the conditional distribution of X_1 given $X_2 = x_2$ is the same as unconditional distribution ofX_1. Similarly, by symmetry, X_1 and X_2 are independent if, and only if, for every $x_1 \in E = \{t \in \mathbb{R}: f_{X_2|X_1}(\cdot\,|t) \text{ is defined}\}$ the conditional distribution of X_2 given $X_1 = x_1$ is the same as the unconditional distribution of X_2.

(ii) Let $\Lambda \subseteq \mathbb{R}$ be an arbitrary non-empty index set, and let $\{\underline{X}_\lambda: \lambda \in \Lambda\}$ be a collection of random vectors defined on a probability space $(\Omega, \mathcal{F}, P)$, where $\underline{X}_\lambda$ may be of different dimensions. One can define the independence of random vectors $\{\underline{X}_\lambda: \lambda \in \Lambda\}$ by extending Definition 4.1 in an obvious manner. We say that the random vectors $\{\underline{X}_\lambda: \lambda \in \Lambda\}$ are independent if for any finite subcollection $\{\lambda_1, \dots, \lambda_p\} \subseteq \Lambda$, we have

$$F_{\underline{X}_{\lambda_1}, \dots, \underline{X}_{\lambda_p}}(\underline{x}_1, \dots, \underline{x}_p) = P(\{\underline{X}_{\lambda_i} \in (-\underline{\infty}, \underline{x}_i], i = 1, \dots, p\})$$
$$= \prod_{i=1}^{p} P(\{\underline{X}_{\lambda_i} \in (-\underline{\infty}, \underline{x}_i]\})$$
$$= \prod_{i=1}^{p} F_{\underline{X}_{\lambda_i}}(x_i), \quad \forall \underline{x}_1, \dots, \underline{x}_p.$$

With above definition of independence of random vectors $\{\underline{X}_\lambda : \lambda \in \Lambda\}$ the results stated in Theorem 4.1 and 4.2 hold with random variables $X_1, \dots, X_p$ replaced by random vectors $\underline{X}_1, \cdots, \underline{X}_p$. Morever, Remarks 4.1, 4.2 and 4.3 (i) also hold with random variables $X_\lambda s$ replaced by random vectors $\underline{X}_\lambda s$.

(iii) Let $\underline{X} = (X_1, \dots, X_p)$ be a random vector and let $k_1, \dots, k_r$ be positive integers such that $\sum_{i=1}^r k_i = p$. Define $\underline{Y_1} = (X_1, \dots, X_{k_1})$, $\underline{Y_2} = (X_{k_1+1}, \dots, X_{k_1+k_2})$ and $\underline{Y_i} = \left(X_{\sum_{j=1}^{i-1} k_j+1}, \dots, X_{\sum_{j=1}^{i} k_j}\right), i = 2, 3, \dots, r$. Suppose that $X_1, \dots, X_p$ are independent random variables. Then, on using the analog of Theorem 4.1 for random vectors, it follows that $\underline{Y_1}, \dots, \underline{Y_r}$ are independent random vectors. ■