MODULE 6

RANDOM VECTOR AND ITS JOINT DISTRIBUTION

LECTURE 27

Topics

6.2 TYPES OF RANDOM VARIABLES

Theorem 2.1

Let $\underline{X} = (X_1, \dots, X_p)$ be a p- dimensional $(p \geq 2)$ random vector with distribution function $F_{\underline{X}}(\cdot)$. For a fixed positive integer $k \in \{1, \dots, p-1\}$, let $\underline{Y} = (X_1, \dots, X_k)$ and $\underline{Z} = (X_{k+1}, \dots, X_p)$ so that $\underline{X} = (\underline{Y}, \underline{Z})$.

(i) Suppose that $\underline{X}$ is of discrete type with support $S_{\underline{X}}$ and p.m.f. $f_{\underline{X}}(\cdot)$. For $\underline{y} \in \mathbb{R}^k$, define $R_{\underline{y}} = \left\{\underline{z} \in \mathbb{R}^{p-k} : \left(\underline{y}, \underline{z}\right) \in S_{\underline{X}}\right\}$ (note that, for each $\underline{y} \in \mathbb{R}^k, R_{\underline{y}}$ is a countable set. Then the random vector $\underline{Y} = (X_1, \dots, X_k)$ is of discrete type with support $S_{\underline{Y}} = \left\{\underline{y} \in \mathbb{R}^k : \left(\underline{y}, \underline{z}\right) \in S_{\underline{X}}, \text{for some } \underline{z} \in \mathbb{R}^{p-k}\right\}$ and joint p.m.f. (called the *marginal p.m.f.* of $\underline{Y}$)

$$f_{\underline{Y}}\left(\underline{y}\right) = \begin{cases} \sum_{\underline{z} \in R_{\underline{y}}} f_{\underline{X}}\left(\underline{y}, \underline{z}\right), & \text{if } \underline{y} \in S_{\underline{Y}} \\ 0, & \text{otherwise} \end{cases}.$$

(ii) Suppose that $\underline{X}$ is of absolutely continuous type with joint p.d.f. $f_{\underline{X}}(\cdot)$. Then the random vector $\underline{Y} = (X_1, \dots, X_k)$ is of absolutely continuous type with p.d.f. (called the *marginal p.d.f.* of $\underline{Y}$)

$$f_{\underline{Y}}\left(\underline{y}\right) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\underline{X}}\left(\underline{y}, \underline{z}\right) dz_{p-k} \cdots dz_1,\ \underline{y} \in \mathbb{R}^k,$$

where $\underline{z} = \left(z_1, \dots, z_{p-k}\right)$.

Proof.

(i) Note that $\{\underline{X} \in S_{\underline{X}}\} = \{(\underline{Y}, \underline{Z}) \in S_{\underline{X}}\} \subseteq \{\underline{Y} \in S_{\underline{Y}}\}$. Therefore

$$P(\{\underline{Y} \in S_{\underline{Y}}\}) \geq P(\{\underline{X} \in S_{\underline{X}}\}) = 1,$$

i.e., $$P(\{\underline{Y} \in S_{\underline{Y}}\}) = 1.$$

Also $S_{\underline{Y}}$ is countable and, for $\underline{y} \in S_{\underline{Y}}$,

$$P\left(\left\{\underline{Y} = \underline{y}\right\}\right) = P\left(\left\{\underline{Y} = \underline{y}\right\} \cap \{\underline{X} \in S_{\underline{X}}\}\right) \qquad (\text{since } P(\{\underline{X} \in S_{\underline{X}}\}) = 1)$$

$$= P\left(\left\{\underline{Y} = \underline{y}\right\} \cap \{(\underline{Y}, \underline{Z}) \in S_{\underline{X}}\}\right)$$

$$= P\left(\left\{\underline{Y} = \underline{y}\right\} \cap \left\{\left(\underline{y}, \underline{Z}\right) \in S_{\underline{X}}\right\}\right)$$

$$= P\left(\left\{\underline{Y} = \underline{y}\right\} \cap \left\{\underline{Z} \in R_{\underline{y}}\right\}\right)$$

$$= P\left(\bigcup_{\underline{z} \in R_{\underline{y}}} \left\{(\underline{Y}, \underline{Z}) = \left(\underline{y}, \underline{z}\right)\right\}\right)$$

$$= \sum_{\underline{z} \in R_{\underline{y}}} P\left(\left\{(\underline{Y}, \underline{Z}) = \left(\underline{y}, \underline{z}\right)\right\}\right)$$

$$= \sum_{\underline{z} \in R_{\underline{y}}} P\left(\left\{\underline{X} = \left(\underline{y}, \underline{z}\right)\right\}\right)$$

$$= \sum_{\underline{z} \in R_{\underline{y}}} f_{\underline{X}}\left(\underline{y}, \underline{z}\right).$$

Note that, for $\underline{y} \in S_{\underline{Y}}$, $R_{\underline{y}} \neq \phi$, and for $\underline{z} \in R_{\underline{y}}, \left(\underline{y}, \underline{z}\right) \in S_{\underline{X}}$. Therefore we have $f_{\underline{X}}\left(\underline{y}, \underline{z}\right) > 0, \forall \underline{y} \in S_{\underline{Y}}$ and $\underline{z} \in R_{\underline{y}}$. If follows that $P(\{\underline{Y} \in S_{\underline{Y}}\}) = 1, P\left(\left\{\underline{Y} = \underline{y}\right\}\right) > 0$, $\forall \underline{y} \in S_{\underline{Y}}$. Hence the assertion follows.

(ii) Note that, for $\underline{y} \in \mathbb{R}^k$,

$$F_{\underline{Y}}\left(\underline{y}\right) = \lim_{\substack{z_i \to \infty \\ i=1,\ldots,,p-k}} F_{\underline{X}}\left(\underline{y}, \underline{z}\right)$$

$$= \lim_{\substack{z_i \to \infty \\ i=1,\ldots,p-k}} \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_k} \cdots \int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_{p-k}} f_{\underline{X}}\left(\underline{s}, \underline{t}\right) d\underline{t}\, d\underline{s}$$

$$= \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_k} \cdots \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\underline{X}} \left(\underline{s}, \underline{t}\right) d\underline{t} \right] d\underline{s},$$

$$= \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_k} h\left(\underline{s}\right) d\underline{s}, \qquad (2.3)$$

where $\underline{s} = (s_1, \dots, s_k)$, $\underline{t} = \left(t_1, \dots, t_{p-k}\right), d\underline{t} = dt_{p-k} \cdots dt_1, d\underline{s} = ds_k \cdots ds_1$ and

$$h\left(\underline{s}\right) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\underline{X}} \left(\underline{s}, \underline{t}\right) d\underline{t}, \quad \underline{s} \in \mathbb{R}^k.$$

Clearly $h\left(\underline{s}\right) \geq 0$, $\forall \underline{s} \in \mathbb{R}^k$ and

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h\left(\underline{s}\right) ds_k \cdots ds_1 = \int_{\mathbb{R}^p} f_{\underline{X}}\left(\underline{s}, \underline{t}\right) d\underline{t} d\underline{s} = 1.$$

Now, using (2.3) and the above properties of $h(\cdot)$, it follows that $\underline{Y}$ is of absolutely continuous type with p.d.f.

$$f_{\underline{Y}}\left(\underline{y}\right) = h\left(\underline{y}\right) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\underline{X}}\left(\underline{y}, \underline{t}\right) d\underline{t}, \quad \underline{y} \in \mathbb{R}^k. \blacksquare$$

Example 2.1

Let $\underline{Z} = (X, Y)$ be a bivariate random vector with p.m.f.

$$f_{\underline{Z}}(x, y) = P(\{X = x, Y = y\}) = \begin{cases} cy, & \text{if } (x, y) \in R \\ 0, & \text{otherwise} \end{cases},$$

where $R = \{(s, t) \in \mathbb{R}^2 : s, t \in \{1, \dots, n\}, s \leq t\}, n\ (\geq 2)$ is fixed positive integer and c is a fixed real constant.

(i) Find the value of constant c;
(ii) Find marginal p.m.f.s of X and Y;
(iii) Find $P(\{X > Y\})$, $P(\{X = Y\})$ and $P(\{X < Y\})$.

Solution.

(i) Clearly we must have $c > 0$. Then the support of $\underline{Z}$ is $S_{\underline{Z}} = R = \{(s,t) \in \mathbb{R}^2: s,t \in \{1, \dots, n\}, s \leq t\}$ and therefore

$$\sum_{(x,y)\in S_{\underline{Z}}} f_{\underline{Z}}(x,y) = 1$$

$$\Rightarrow c \sum_{y=1}^{n} \sum_{x=1}^{y} y = 1$$

$$\Rightarrow c \sum_{y=1}^{n} y^2 = 1$$

$$\Rightarrow c = \frac{6}{n(n+1)(2n+1)}.$$

(ii) By Theorem 2.1 (i) the support of X is $S_X = \left\{x \in \mathbb{R}: (x,y) \in S_{\underline{Z}} \text{ for some } y \in \mathbb{R}\right\} = \{1,2, \dots, n\}$, and the support of Y is $S_Y = \left\{y \in \mathbb{R}: (x,y) \in S_{\underline{Z}} \text{ for some } x \in \mathbb{R}\right\} = \{1, 2, \dots, n\}$. For $x \in S_X$, define $R_x = \left\{y \in \mathbb{R}: (x,y) \in S_{\underline{Z}}\right\}$. Then, by Theorem 2.1, the marginal p.m.f. of X is

$$f_X(x) = \begin{cases} \displaystyle\sum_{y\in R_x} f_{\underline{Z}}(x,y), & \text{if } x \in S_X \\ 0, & \text{otherwise} \end{cases}.$$

For $x \in S_X$, we have $R_x = \{x, x+1, \dots, n\}$

$$\sum_{y\in R_x} f_{\underline{Z}}(x,y) = c\sum_{y=x}^{n} y = c\left[\frac{n(n+1)}{2} - \frac{(x-1)x}{2}\right].$$

Therefore the marginal p.m.f. X is

$$f_X(x) = \begin{cases} \dfrac{3[n(n+1) - (x-1)x]}{n(n+1)(2n+1)}, & \text{if } x \in S_X \\ 0, & \text{otherwise} \end{cases},$$

where $S_X = \{1, \dots, n\}$.

For $y \in S_Y$, define $R_y^* = \left\{x \in \mathbb{R}: (x,y) \in S_{\underline{Z}}\right\} = \{1, 2, \dots, y\}$. Then, by Theorem 2.1, the marginal p.m.f. of Y is

$$f_Y(y) = \begin{cases} \sum_{x \in R_y^*} f_{\underline{Z}}(x, y), & \text{if } y \in S_Y \\ 0, & \text{otherwise} \end{cases}.$$

For $y \in S_Y$, we have

$$\sum_{x \in R_y^*} f_{\underline{Z}}(x, y) = c \sum_{x=1}^{y} y = cy^2.$$

Therefore the marginal p.m.f. of Y is

$$f_Y(y) = \begin{cases} \frac{6y^2}{n(n+1)(2n+1)}, & \text{if } y \in S_Y \\ 0, & \text{otherwise} \end{cases},$$

where $S_Y = \{1, 2, \dots, n\}$.

(iii) Let $A = \{(s, t): s > t\}$ and $B = \{(s, t): s = t\}$. Then by Remark 2.1 (ix)

$$\begin{aligned} P(\{X > Y\}) &= P\{\underline{Z} \in A\}) \\ &= \sum_{(x,y) \in S_{\underline{Z}} \cap A} f_{\underline{Z}}(x, y) \\ &= 0 \qquad \left(\text{since } S_{\underline{Z}} \cap A = \phi\right). \end{aligned}$$

$$\begin{aligned} P(\{X = Y\}) &= P\{\underline{Z} \in B\}) \\ &= \sum_{(x,y) \in S_{\underline{Z}} \cap B} f_{\underline{Z}}(x, y) \\ &= c \sum_{y=1}^{n} y \\ &= \frac{3}{2n+1}. \end{aligned}$$

Therefore

$$P(\{X < Y\}) = 1 - P(\{X = Y\}) - P(\{X > Y\})$$

$$= 1 - \frac{3}{2n+1}$$

$$= \frac{2(n-1)}{2n+1}. \blacksquare$$

Example 2.2

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

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

where c is a real constant.

(i) Find the value of c;
(ii) Find the marginal p.m.f.s. of X_1; of X_2; of X_3;
(iii) Find the marginal p.m.f. of $\underline{Y} = (X_1, X_3)$;
(iv) Find $P(\{X_1 = X_2 = X_3\})$.

Solution.
(i) Clearly we must have $c > 0$. Then the support of $\underline{X}$ is $S_{\underline{X}} = \{(x_1, x_2, x_3): x_1 \in \{1, 2\}, x_2 \in \{1, 2, 3\}, x_3 \in \{1, 3\}\}$. Therefore

$$\sum_{\underline{x} \in S_{\underline{X}}} f_{\underline{X}}(x_1, x_{2,} x_3) = 1$$

$$\Rightarrow c \sum_{x_1 \in \{1,2\}} \sum_{x_2 \in \{1,2,3\}} \sum_{x_3 \in \{1,3\}} x_1 x_2 x_3 = 1$$

$$\Rightarrow c = \frac{1}{72}.$$

(ii) The supports of X_1, X_2 and X_3 are

$$S_{X_1} = \{x_1 \in \mathbb{R}^1: (x_1, x_2, x_3) \in S_{\underline{X}} \text{ for some } (x_2, x_3) \in \mathbb{R}^2\} = \{1, 2\},$$

$$S_{X_2} = \{x_2 \in \mathbb{R}^1: (x_1, x_2, x_3) \in S_{\underline{X}} \text{ for some } (x_1, x_3) \in \mathbb{R}^2\} = \{1, 2, 3\}$$

and

$$S_{X_3} = \{x_3 \in \mathbb{R}^1: (x_1, x_2, x_3) \in S_{\underline{X}} \text{ for some } (x_1, x_2) \in \mathbb{R}^2\} = \{1, 3\},$$

respectively.

For $x_1 \in S_{X_1}$, $R_{x_1} = \{(x_2, x_3): (x_1, x_2, x_3) \in S_{\underline{Z}}\} = \{1,2,3\} \times \{1,3\}$. Then, for $x_1 \in S_{X_1}$

$$f_{X_1}(x_1) = P(\{X_1 = x_1\})$$

$$= \sum_{(x_2,x_3)\in R_{x_1}} f_{\underline{X}}(x_1, x_2, x_3)$$

$$= \sum_{x_2\in\{1,2,3\}} \sum_{x_3\in\{1,3\}} x_1 x_2 x_3$$

$$= \frac{x_1}{3}.$$

Therefore the marginal p.m.f. of X_1 is

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

Similarly the p.m.f.s of X_2 and X_3 are

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

and

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

respectively.

(iii) The support of $\underline{Y} = (X_1, X_3)$ is

$$S_{\underline{Y}} = \{(y_1, y_2): (y_1, s, y_2) \in S_{\underline{Z}} \text{ for some } s \in \mathbb{R}\}$$

$$= \{1,2\} \times \{1,3\}$$

$$= \{(1,1), (1,3), (2,1), (2,3)\}.$$

For $\underline{y} = (y_1, y_2) \in S_{\underline{Y}}, R_{\underline{y}} = \{s \in \mathbb{R}: (y_1, s, y_3) \in S_Z\} = \{1,2,3\}$. Therefore, for $\underline{y} = (y_1, y_2) \in S_{\underline{Y}}$,

$$f_{\underline{Y}}(\underline{y}) = (\{\underline{Y} = \underline{y}\}) = \sum_{s\in\{1,2,3\}} c y_1 s y_2$$

$$= \frac{y_1 y_2}{12},$$

and the marginal p.m.f. of $\underline{Y} = (Y_1, Y_2)$ is

$$f_{\underline{Y}}(y_1, y_2) = \begin{cases} \frac{y_1 y_2}{12}, & \text{if } (y_1, y_2) \in \{(1,1), (1,3), (2,1), (2,3)\} \\ 0, & \text{otherwise} \end{cases}.$$

(iv) Let $A = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2 = x_3\}$. Then $S_{\underline{X}} \cap A = \{(1,1,1)\}$ and therefore

$$P(\{X_1 = X_2 = X_3\}) = \sum_{\underline{x} \in S_{\underline{X}} \cap A} f_{\underline{X}}(\underline{x})$$

$$= c$$

$$= \frac{1}{72}. \blacksquare$$

Example 2.3

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

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

where c is a real constant.

(i) Find the value of constant c;
(ii) Find the marginal p.d.f. of $\underline{Y} = (X_2, X_3)$;
(iii) Find the marginal p.d.f. of X_2;
(iv) Find $P(\{X_1 > 2X_2\})$.

Solution.

(i) Clearly we have $c > 0$. Also

$$\int_{\mathbb{R}^3} f_{\underline{X}}(\underline{x}) d\underline{x} = 1$$

$$\Rightarrow \int_0^1 \int_0^{x_1} \int_0^{x_2} \frac{c}{x_1 x_2} dx_3 dx_2\, dx_1 = 1$$

$$\Rightarrow c = 1.$$

(ii) The marginal p.d.f. of $\underline{Y} = (X_2, X_3)$ is

$$f_{\underline{Y}}(y_1, y_2) = \int_{-\infty}^{\infty} f_{\underline{X}}\,(x_1, y_1, y_2)\, dx_1$$

$$= \begin{cases} \displaystyle\int_{y_1}^{1} \frac{1}{x_1 y_1}\, dx_1\,, & \text{if } 0 < y_2 < y_1 < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \dfrac{-\ln y_1}{y_1}, & \text{if } 0 < y_2 < y_1 < 1 \\ 0, & \text{otherwise} \end{cases}.$$

(iii) The marginal p.d.f. of X_2 is

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\underline{X}}\,(x_1, x_2, x_3)\, dx_1\, dx_3$$

$$= \begin{cases} \displaystyle\int_{0}^{x_2}\int_{x_2}^{1} \frac{1}{x_1 x_2}\, dx_1\, dx_3, & \text{if } 0 < x_2 < 1 \\ 0, & \text{otherwise} \end{cases}$$

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

(iv) Let $A = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 > 2x_2\}$. Then

$$P(\{X_1 > 2\, X_2\}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\underline{X}}(\underline{x}) I_A(\underline{x})\, d\underline{x}$$

$$= \int\int\int_{0<x_3<x_2<x_1<1} \frac{1}{x_1 x_2} I_A(\underline{x})\, d\underline{x}$$

$$= \int_{0}^{1}\int_{0}^{\frac{x_1}{2}}\int_{0}^{x_2} \frac{1}{x_1 x_2}\, dx_3 dx_2\, dx_1$$

$$= \frac{1}{2}.$$

We conclude this section with the following remark. ■

Remark 2.2

(i) There are random vectors that are neither of discrete type nor of continuous type (and hence also nor of absolutely continuous type). To see this let $\underline{X} = (X_1, X_2)$ have the joint distribution function

$$F_{X_1,X_2}(x_1,x_2) = \begin{cases} \frac{1}{2} + \frac{x_1 x_2}{2}, & \text{if } 0 \le x_1 < 1, 0 \le x_2 < 1 \\ \frac{1}{2} + \frac{x_1}{2}, & \text{if } 0 \le x_1 < 1, x_2 \ge 1 \\ \frac{1}{2} + \frac{x_2}{2}, & \text{if } x_1 \ge 1, 0 \le x_2 < 1 \\ 1, & \text{if } x_1 \ge 1, x_2 \ge 1 \\ 0, & \text{otherwise} \end{cases}.$$

It is easy to verify that $F_{X_1,X_2}(\cdot)$ is a distribution function (i.e., it satisfies properties (i)-(iv) of Theorem 1.3). The marginal distribution functions of X_1 and X_2 are

$$F_{X_1}(x_1) = \lim_{x_2 \to \infty} F_{X_1,X_2}(x_1,x_2) = \begin{cases} 0, & \text{if } x_1 < 0 \\ \frac{1}{2} + \frac{x_1}{2}, & \text{if } 0 \le x_1 < 1 \\ 1, & \text{if } x_1 \ge 1 \end{cases}$$

and

$$F_{X_2}(x_2) = \lim_{x_1 \to \infty} F_{X_1,X_2}(x_1,x_2) = \begin{cases} 0, & \text{if } x_2 < 0 \\ \frac{1}{2} + \frac{x_2}{2}, & \text{if } 0 \le x_2 < 1. \\ 1, & \text{if } x_2 \ge 1 \end{cases}$$

Clearly the set of discontinuity points of F_{X_1} $(= F_{X_2})$ is $D = \{0\}$ and

$$\sum_{x \in D} \left[F_{X_1}(x) - F_{X_1}(x-)\right] = \sum_{x \in D} \left[F_{X_2}(x) - F_{X_2}(x-)\right] = \frac{1}{2} \ne 1.$$

It follows that X_1 and X_2 are not of discrete type and therefore using Theorem 2.1 (i) it follows that (X_1, X_2) is not of discrete type.

Note that

$$\left|F_{X_1,X_2}(x_1,x_2) - F_{X_1,X_2}(0,0)\right| = \left|F_{X_1,X_2}(x_1,x_2) - \tfrac{1}{2}\right|$$

$$= \begin{cases} \dfrac{1}{2}, & \text{if } x_1 < 0 \text{ or } x_2 < 0 \\ \dfrac{x_1 x_2}{2}, & \text{if } 0 \le x_1 < 1, 0 \le x_2 < 1 \end{cases}$$

$$\nrightarrow 0, \quad \text{as } (x_1,x_2) \rightarrow (0,0),$$

i.e., $F_{X_1,X_2}(\cdot)$ is not continuous at $(0,0)$. Therefore (X_1,X_2) is also not of continuous type.

(ii) There are random vectors which are of continuous type but not of absolutely continuous type. These random vectors are normally difficult to study.