Random Vector

Let $(\mathcal{S}, \Sigma, P)$ be a probability space. A (univariate) random variable describes a numerical quantity of a typical outcome of a random experiment. In many experiments an observation is expressed as a family of several separate numerical quantities and we may be interested in simultaneously studying all of them together. Consider the following example.

Example 1. *Two distinguishable dice (labelled as D_1 and D_2) are thrown simultaneously. The sample space is $\mathcal{S} = \{(i,j) : i,j \in \{1,2,\ldots,6\}\}$. For $(i,j) \in \mathcal{S}$ define*

$$X_1((i,j)) = i + j = \text{ sum of number of dots on uppermost faces of two dice}$$

and

$$X_2((i,j)) = |i-j| = \text{ absolute difference of number of dots on uppermost faces of two dice.}$$

It may be of interest to study numerical characteristics X_1 and X_2 simultaneously. These considerations lead to the study of the function $\underline{X} = (X_1, X_2) : \mathcal{S} \to \mathbb{R}$

Notations.

- We denote by $\mathbb{R}^n$ the n-dimensional Euclidean space, i.e.,
$$\mathbb{R}^n = \{\underline{x} = (x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R},\ i = 1,2,\ldots,n\}.$$
- For $i = 1,2,\ldots,n$, let $X_i : \mathcal{S} \to \mathbb{R}$ be any functions. Then the function $\underline{X} = (X_1, X_2, \ldots, X_n) : \mathcal{S} \to \mathbb{R}^n$ is defined as
$$\underline{X}(w) = (X_1(w), X_2(w), \ldots, X_n(w)),\ w \in \mathcal{S}.$$
- For $A \subseteq \mathbb{R}^n$,
$$\underline{X}^{-1}(A) = \{w \in \mathcal{S} : \underline{X}(w) \in A\}.$$
- For $\underline{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we denote by $(-\underline{\infty}, \underline{x}]$ the n-dimensional interval
$$(-\underline{\infty}, \underline{x}] = (-\infty, x_1] \times (-\infty, x_2] \times \cdots \times (-\infty, x_n].$$

Definition 2. *A function $\underline{X} : \mathcal{S} \longrightarrow \mathbb{R}^n$ is called an $n-$dimensional random vector (RV) if $\underline{X}^{-1}((-\underline{\infty}, \underline{x}]) \in \Sigma$, for all $\underline{x} \in \mathbb{R}^n$. That is, $\{w \in \mathcal{S} : X_1(w) \le x_1, X_2(w) \le x_2, \ldots, X_n(w) \le x_n\} \in \Sigma$.*

Example 3. *Let $A, B \subseteq \mathcal{S}$. Define $\underline{X} = (X_1, X_2) : \mathcal{S} \to \mathbb{R}^2$ by*

$$X_1(w) = I_A(w) = \begin{cases} 1, \text{if } w \in A, \\ 0, \text{if } w \notin A; \end{cases}$$

and

$$X_2(w) = I_B(w) = \begin{cases} 1, \text{if } w \in B, \\ 0, \text{if } w \notin B. \end{cases}$$

Then $\underline{X}$ is an RV if and only if A and B are events. (Prove!)

Theorem 4. *Let $\underline{X} = (X_1, X_2, \ldots, X_n) : \mathcal{S} \to \mathbb{R}^n$ be a given function. Then $\underline{X}$ is a random vector if and only if $X_1, X_2, \ldots, X_n$ are random variables.*

Proof. Exercise. □

Remark 5. *If $\mathcal{S}$ is finite or countable and $\Sigma = \mathcal{P}(\Sigma)$, then any function $\underline{X} = (X_1, X_2, \ldots, X_n) : \mathcal{S} \to \mathbb{R}^n$ is a random vector.*

Joint Cumulative Distribution Function

Definition 6. *Let $\underline{X} = (X_1, X_2, \ldots, X_n) : \mathcal{S} \to \mathbb{R}^n$ be a random vector. The function $F_{\underline{X}} : \mathbb{R}^n \to \mathbb{R}$, defined by,*

$$F_{\underline{X}}(x_1, x_2, \ldots, x_n) = P(\{w \in \mathcal{S} : X_1(w) \leq x_1, X_2(w) \leq x_2, \ldots, X_n(w) \leq x_n\}),\ \forall\ \underline{x} \in \mathbb{R}^n,$$

is called the ***joint cumulative distribution function*** *(joint c.d.f) or the* ***joint distribution function*** *(d.f) of the random vector $\underline{X}$.*

The joint distribution function of any subset of random variables $X_1, X_2, \ldots, X_n$ is called a marginal distribution function of $F_{\underline{X}}$.

Remark 7. (1) *As in the case of random variables, the set $\{w \in \mathcal{S} : X_1(w) \leq x_1, X_2(w) \leq x_2, \ldots, X_n(w) \leq x_n\}$ will be denoted by $\{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n\}$.*

(2) *In this course, we will mainly study 2- (and sometimes 3-) dimensional random vectors.*

(3) *Let $\underline{X} = (X, Y) : \mathcal{S} \to \mathbb{R}^2$ be a random vector. The joint c.d.f. is a map $F_{\underline{X}} : \mathbb{R}^2 \to \mathbb{R}$, defined by,*

$$F_{\underline{X}}(x, y) = P(\{X \leq x, Y \leq y\}).$$

(4) *The c.d.f. of X and Y are called a marginal c.d.f. of $F_{\underline{X}}$.*

Proposition 8. *Let $\underline{X} = (X, Y) : \mathcal{S} \to \mathbb{R}^2$ be a random vector with joint c.d.f. $F_{\underline{X}}$. Then the marginal c.d.f. of X and Y are given by*

$$F_X(x) = \lim_{y \to \infty} F_{\underline{X}}(x, y) \text{ and } F_Y(y) = \lim_{x \to \infty} F_{\underline{X}}(x, y)$$

Remark 9. *Let $(a_1, b_1), (a_2, b_2) \in \mathbb{R}^2$. Then we know that*

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a).$$

Now,

$$\begin{aligned}
&P(a_1 < X \leq b_1, a_2 < Y \leq b_2) \\
&= P(a_1 < X \leq b_1, Y \leq b_2) - P(a_1 < X \leq b_1, Y \leq a_2) \\
&= [P(X \leq b_1, Y \leq b_2) - P(X \leq a_1, Y \leq b_2)] \\
&\qquad - [P(X \leq b_1, Y \leq a_2) - P(X \leq a_1, Y \leq a_2)] \\
&= F_{\underline{X}}(b_1, b_2) - F_{\underline{X}}(a_1, b_2) - F_{\underline{X}}(b_1, a_2) + F_{\underline{X}}(a_1, a_2).
\end{aligned}$$

Theorem 10. *Let $F_{\underline{X}}$ be the joint cumulative distribution function of a random vector $\underline{X} = (X, Y)$. Then*

(1) $\lim\limits_{\substack{x \to \infty \\ y \to \infty}} F_{\underline{X}}(x, y) = 1.$

(2) $\lim_{y \to -\infty} F_{\underline{X}}(x, y) = 0$ *and* $\lim_{x \to -\infty} F_{\underline{X}}(x, y) = 0.$

(3) *$F_{\underline{X}}(x, y)$ is right continuous and nondecreasing in each argument (keeping other argument fixed).*

(4) *For each $(a_1, b_1] \times (a_2, b_2]$ in $\mathbb{R}^2$,*

$$\Delta = F_{\underline{X}}(b_1, b_2) - F_{\underline{X}}(a_1, b_2) - F_{\underline{X}}(b_1, a_2) + F_{\underline{X}}(a_1, a_2) \geq 0.$$

Theorem 11. *Let $G : \mathbb{R}^2 \to \mathbb{R}$ be a function which satisfies properties $(1) - (4)$ of Theorem 10. Then there exists a probability space $(\mathcal{S}, \Sigma, P)$ and a random vector $\underline{X} = (X_1, X_2, \ldots, X_n)$ defined on $(\mathcal{S}, \Sigma, P)$ such that G is the distribution function of $\underline{X}$.*

Example 12. *Let $G : \mathbb{R}^2 \to \mathbb{R}$ be defined by*

$$G(x,y) = \begin{cases} x, \text{ if } 0 \le x < 1,\ \ y \ge 1, \\ y^2, \text{ if } x \ge 1,\ \ 0 \le y < 1, \\ 1, \text{ if } x \ge 1,\ \ y \ge 1, \\ 0, \text{ otherwise.} \end{cases}$$

Show that G is not a distribution function of any random vector (X, Y).

Solution. Clearly G satisfies properties $(1) - (3)$ of Theorem 10.

For $(a_1, b_1] \times (a_2, b_2]$, where $a_1, a_2 \in [0, 1)$, $b_1, b_2 \in [1, \infty)$ and $a_1 + a_2^2 > 1$. Then

$$G(b_1, b_2) - G(a_1, b_2) - G(b_1, a_2) + G(a_1, a_2) = 1 - a_1 - a_2^2 + 0 < 0.$$

Thus, G is not a joint c.d.f. of any random vector.

Example 13. *Consider the function $G : \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$G(x,y) = \begin{cases} xy^2, \text{ if } 0 \le x < 1,\ \ 0 \le y < 1, \\ x, \text{ if } 0 \le x < 1,\ \ y \ge 1, \\ y^2, \text{ if } x \ge 1,\ \ 0 \le y < 1, \\ 1, \text{ if } x \ge 1,\ \ y \ge 1, \\ 0, \text{ otherwise.} \end{cases}$$

(1) *Show that G is a joint c.d.f. of some random vector (X, Y).*
(2) *Find the marginal c.d.f. of X and Y.*

Solution. Clearly $\lim\limits_{\substack{x \to \infty \\ y \to \infty}} G(x, y) = 1$. For fixed $x \in \mathbb{R}$, $\lim_{y \to -\infty} G(x, y) = 0$ and for fixed $y \in \mathbb{R}$, $\lim_{x \to -\infty} G_(x, y) = 0$.

We note that if $y < 0$, then $G(x, y) = 0$ for all $x \in \mathbb{R}$. Moreover,

$$G(x,y) = \begin{cases} 0, \text{if } x < 0, \\ xy^2, \text{if } 0 \le x < 1,\ \ 0 \le y < 1, \\ y^2, \text{if } x \ge 1, \end{cases}$$

and

$$G(x,y) = \begin{cases} 0, \text{if } x < 0, \\ x, \text{if } 0 \le x < 1,\ \ y \ge 1, \\ 1, \text{if } x \ge 1. \end{cases}$$

One can see that for $y \in \mathbb{R}$, $G(x, y)$ is a continuous (and hence right continuous) function of x. Similarly, for each $x \in \mathbb{R}$, $G(x, y)$ is a continuous function of y

Furthermore, $G(x, y)$ is non-decreasing in each argument keeping other argument fixed.

For $(a_1, b_1] \times (a_2, b_2]$, we need to show that $\Delta = G(b_1, b_2) - G(a_1, b_2) - G(b_1, a_2) + G(a_1, a_2) \ge 0$. We consider the following cases.

(1) $a_1 < 0$. Then $\Delta = G(b_1, b_2) - G(b_1, a_2) \ge 0$ as G is nondecreasing.
(2) $a_2 < 0$.
(3) $0 \le a_1 < 1$, $0 \le a_2 < 1$, $0 \le b_1 < 1$, $0 \le b_2 < 1$.
(4) $0 \le a_1 < 1$, $0 \le a_2 < 1$, $0 \le b_1 < 1$, $b_2 \ge 1$.
(5) $0 \le a_1 < 1$, $0 \le a_2 < 1$, $b_1 \ge 1$, $0 \le b_2 < 1$.
(6) $0 \le a_1 < 1$, $0 \le a_2 < 1$, $b_1 \ge 1$, $b_2 \ge 1$.

(7) $0 \le a_1 < 1$, $a_2 \ge 1$, $0 \le b_1 < 1$, $b_2 \ge 1$.
(8) $0 \le a_1 < 1$, $a_2 \ge 1$, $b_1 \ge 1$, $b_2 \ge 1$.
(9) $a_1 \ge 1$, $0 \le a_2 < 1$, $b_1 \ge 1$, $0 \le b_2 < 1$.
(10) $a_1 \ge 1$, $0 \le a_2 < 1$, $b_1 \ge 1$, $b_2 \ge 1$.
(11) $a_1 \ge 1$, $a_2 \ge 1$, $b_1 \ge 1$, $b_2 \ge 1$.

In all these cases verify that $\Delta \ge 0$.

Therefore, $G(x, y)$ is a distribution function of some random vector (X, Y).

The marginal c.d.f. of X and Y are respectively

$$F_X(x) = \lim_{y \to \infty} G(x, y) = \begin{cases} 0, \text{if } x < 0, \\ x, \text{if } 0 \le x < 1, \\ 1, \text{if } x \ge 1, \end{cases}$$

and

$$F_Y(y) = \lim_{x \to \infty} G(x, y) = \begin{cases} 0, \text{if } y < 0, \\ y^2, \text{if } 0 \le y < 1, \\ 1, \text{if } y \ge 1. \end{cases}$$