

# Random Variable

Let  $(\mathcal{S}, \Sigma, P)$  be a probability space. Since  $P$  is a set function, it is not very easy to handle. Also in many situations, one may not be interested in the sample space rather one may be interested in some numerical characteristics of the sample space. For example, when a coin is tossed  $n$ -times, which replication resulted in heads is not of much interest. Rather, one is interested in the number of heads, and consequently, the number of tails, that appear out of  $n$  toss.

It is therefore desirable to introduce a point function on the sample space so that we can use the theory of calculus or real analysis to study properties of  $P$ .

**Definition 1.** A function  $X : \mathcal{S} \rightarrow \mathbb{R}$  is called a random variable (RV) if  $X^{-1}(B) \in \Sigma$ , for all  $B \in \mathbb{B}_{\mathbb{R}}$ , that is,  $X^{-1}(B) = \{w \in \mathcal{S} : X(w) \in B\}$  is an event.

## Notations.

We will use the following notations throughout the course.

- For  $B \in \mathbb{B}_{\mathbb{R}}$ ,  $\{X \in B\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : X(w) \in B\} \stackrel{\text{def}}{=} X^{-1}(B)$ ;
- $\{a < X \leq b\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : a < X(w) \leq b\} \stackrel{\text{def}}{=} X^{-1}((a, b])$ ;
- $\{a \leq X \leq b\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : a \leq X(w) \leq b\} \stackrel{\text{def}}{=} X^{-1}([a, b])$ ;
- $\{a < X < b\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : a < X(w) < b\} \stackrel{\text{def}}{=} X^{-1}((a, b))$ ;
- $\{a \leq X < b\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : a \leq X(w) < b\} \stackrel{\text{def}}{=} X^{-1}([a, b))$ ;
- $\{X = a\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : X(w) = a\} \stackrel{\text{def}}{=} X^{-1}(\{a\})$ ;
- $\{X \leq a\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : X(w) \leq a\} \stackrel{\text{def}}{=} X^{-1}((-\infty, a])$ ;
- $\{X < a\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : X(w) < a\} \stackrel{\text{def}}{=} X^{-1}((-\infty, a))$ ;
- $\{X \geq a\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : X(w) \geq a\} \stackrel{\text{def}}{=} X^{-1}([a, \infty))$ ;
- $\{X > a\} \stackrel{\text{def}}{=} \{w \in \mathcal{S} : X(w) > a\} \stackrel{\text{def}}{=} X^{-1}(a, \infty)$ .

**Remark 2.** (1)  $X$  is a random variable if and only if for each  $x \in \mathbb{R}$ ,  $\{X \leq x\} \in \Sigma$ .

(2) If  $\Sigma = \mathcal{P}(\mathcal{S})$ , then any function  $X : \mathcal{S} \rightarrow \mathbb{R}$  is a random variable.

(3) Let  $(\mathcal{S}, \Sigma, P)$  be a probability space and  $X : \mathcal{S} \rightarrow \mathbb{R}$  be a random variable. Then the random variable  $X$  induces a probability space  $(\mathbb{R}, \mathbb{B}_{\mathbb{R}}, P_X)$ , where  $P_X(B) = P(\{w \in \mathcal{S} : X(w) \in B\})$ , for all  $B \in \mathbb{B}_{\mathbb{R}}$ .

**Example 3.** Suppose that a fair coin is independently flipped thrice. Then

$\mathcal{S} = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$  and

$P(E) = \frac{\text{number of elements in } E}{8}$ , for every  $E \in \mathcal{P}(\mathcal{S})$ . Define  $X : \mathcal{S} \rightarrow \mathbb{R}$  by  $X(w) =$  number of heads, i.e.,

$$X(w) = \begin{cases} 0, & w = TTT \\ 1, & w \in \{HTT, TTH, THT\} \\ 2, & w \in \{HHT, THH, HTH\} \\ 3, & w = HHH. \end{cases}$$

Clearly  $X$  is a random variable. The induced probability space is  $(\mathbb{R}, \mathbb{B}_{\mathbb{R}}, P_X)$ , where  $P_X(\{0\}) = P_X(\{3\}) = \frac{1}{8}$ ,  $P_X(\{1\}) = P_X(\{2\}) = \frac{3}{8}$ , and  $P_X(B) = \sum_{i \in \{0,1,2,3\} \cap B} P_X(\{i\})$ , for

all  $B \in \mathbb{B}_{\mathbb{R}}$ .

**Definition 4.** Let  $(\mathcal{S}, \Sigma, P)$  be a probability space and  $X : \mathcal{S} \rightarrow \mathbb{R}$  be a random variable. The function  $F_X : \mathbb{R} \rightarrow \mathbb{R}$ , defined by,

$$F_X(x) = P(\{X \leq x\}), \quad \forall x \in \mathbb{R},$$

is called the **cumulative distribution function** (c.d.f) or the **distribution function** (d.f) of the random variable  $X$ .

**Theorem 5.** Let  $F_X$  be the cumulative distribution function of a random variable  $X$ . Then

- (1)  $F_X$  is increasing;
- (2)  $F_X$  is right continuous;
- (3)  $F_X(-\infty) \stackrel{\text{def}}{=} \lim_{x \rightarrow -\infty} F_X(x) = 0$  and  $F_X(\infty) \stackrel{\text{def}}{=} \lim_{x \rightarrow \infty} F_X(x) = 1$ .

*Proof.* (1) Let  $x_1 < x_2$ . Then  $(-\infty, x_1] \subset (-\infty, x_2]$ . Then by the properties of the probability function, we have

$$F_X(x_1) = P(\{X \leq x_1\}) \leq P(\{X \leq x_2\}) = F_X(x_2).$$

- (2) Fix  $a \in \mathbb{R}$ . Since  $F_X$  is non-decreasing,  $F_X(a+) = \lim_{x \rightarrow a+} F_X(x)$  exists. Therefore

$$F_X(a+) = \lim_{n \rightarrow \infty} F_X(a + \frac{1}{n}) = \lim_{n \rightarrow \infty} P(\{X \leq a + \frac{1}{n}\}).$$

Let  $E_n = \{w \in \mathcal{S} : X(w) \in (-\infty, a + \frac{1}{n}]\}$ . Then  $E_n$  is an decreasing sequence of events and  $\text{Lim}_{n \rightarrow \infty} E_n = \cap_{n=1}^{\infty} E_n = \{w \in \mathcal{S} : X(w) \in (-\infty, a]\}$ . Now by using continuity of probability, we have

$$\begin{aligned} F_X(a+) &= \lim_{n \rightarrow \infty} P(\{X \leq a + \frac{1}{n}\}) \\ &= \lim_{n \rightarrow \infty} P(E_n) \\ &= P(\text{Lim}_{n \rightarrow \infty} E_n) \\ &= P(\{X \in (-\infty, a]\}) \\ &= P(\{X \leq a\}) \\ &= F_X(a) \end{aligned}$$

- (3) Let  $A_n = \{w \in \mathcal{S} : X(w) \in (-\infty, -n]\}$  and  $B_n = \{w \in \mathcal{S} : X(w) \in (-\infty, n]\}$ . Then  $A_n$  and  $B_n$  are decreasing and increasing sequence of events, respectively. Also  $\text{Lim}_{n \rightarrow \infty} A_n = \cap_{n=1}^{\infty} A_n = \emptyset$  and  $\text{Lim}_{n \rightarrow \infty} B_n = \cup_{n=1}^{\infty} B_n = \{w \in \mathcal{S} : X(w) \in \mathbb{R}\} = \mathcal{S}$ . Therefore, by using continuity of probability, we have

$$\begin{aligned} F_X(-\infty) &= \lim_{n \rightarrow \infty} F_X(-n) \\ &= \lim_{n \rightarrow \infty} P(\{X \in (-\infty, -n]\}) \\ &= \lim_{n \rightarrow \infty} P(A_n) \\ &= P(\text{Lim}_{n \rightarrow \infty} A_n) \\ &= P(\emptyset) = 0, \end{aligned}$$

and

$$\begin{aligned}
F_X(\infty) &= \lim_{n \rightarrow \infty} F_X(n) \\
&= \lim_{n \rightarrow \infty} P(\{X \in (-\infty, n]\}) \\
&= \lim_{n \rightarrow \infty} P(B_n) \\
&= P(\text{Lim}_{n \rightarrow \infty} B_n) \\
&= P(\mathcal{S}) = 1.
\end{aligned}$$

□

**Remark 6.** (1) Let  $E_n = \{w \in \mathcal{S} : X(w) \in (-\infty, a - \frac{1}{n}]\} = \{X \leq a - \frac{1}{n}\}$ . Then  $E_n$  is an increasing sequence of events and  $\text{Lim}_{n \rightarrow \infty} E_n = \cup_{n=1}^{\infty} E_n = \{w \in \mathcal{S} : X(w) \in (-\infty, a)\} = \{X < a\}$ . Now by using continuity of probability, we have

$$\begin{aligned}
P(\{X < a\}) &= P(\text{Lim}_{n \rightarrow \infty} E_n) \\
&= \lim_{n \rightarrow \infty} P(E_n) \\
&= \lim_{n \rightarrow \infty} P(\{X \leq a - \frac{1}{n}\}) \\
&= \lim_{n \rightarrow \infty} F_X(a - \frac{1}{n}) \\
&= F_X(a-).
\end{aligned}$$

Therefore,  $P(\{X < a\}) = F_X(a-)$ ,  $\forall x \in \mathbb{R}$ .

(2) For  $-\infty < a < b < \infty$ , we have

$$(a) P(\{a < X \leq b\}) = P(\{X \in ((-\infty, b] - (-\infty, a])\}) = P(\{X \leq b\}) - P(\{X \leq a\}) = F_X(b) - F_X(a).$$

$$(b) P(\{a < X < b\}) = P(\{X \in ((-\infty, b) - (-\infty, a])\}) = P(\{X < b\}) - P(\{X \leq a\}) = F_X(b-) - F_X(a).$$

$$(c) P(\{a \leq X < b\}) = P(\{X \in ((-\infty, b) - (-\infty, a))\}) = P(\{X < b\}) - P(\{X < a\}) = F_X(b-) - F_X(a-).$$

$$(d) P(\{a \leq X \leq b\}) = P(\{X \in ((-\infty, b] - (-\infty, a))\}) = P(\{X \leq b\}) - P(\{X < a\}) = F_X(b) - F_X(a-).$$

(3) For  $-\infty < a < \infty$ , we have

$$(a) P(\{X \geq a\}) = P(\{X \in (\mathbb{R} - (-\infty, a))\}) = P(\{X \in \mathbb{R}\}) - P(\{X < a\}) = 1 - F_X(a-).$$

$$(b) P(\{X > a\}) = P(\{X \in (\mathbb{R} - (-\infty, a])\}) = P(\{X \in \mathbb{R}\}) - P(\{X \leq a\}) = 1 - F_X(a).$$

(4) The distribution function  $F_X$  has atmost countable number of discontinuities.

**Example 7.** Let  $(\mathcal{S}, \Sigma, P)$  be a probability space. Define  $X : \mathcal{S} \rightarrow \mathbb{R}$  by  $X(w) = c$ , for all  $w \in \mathcal{S}$ , where  $c$  is a fixed real number. Clearly,  $X$  is a random variable and the the cumulative distribution function of  $X$  is

$$F_X(x) = P(\{X \leq x\}) = \begin{cases} 0, & x < c \\ 1, & x \geq c. \end{cases}$$

**Example 8.** Let  $(\mathcal{S}, \Sigma, P)$  be a probability space and  $E$  be an event. Define  $I_E : \mathcal{S} \rightarrow \mathbb{R}$  by  $I_E(w) = 1$ , if  $w \in E$  and  $I_E(w) = 0$ , if  $w \notin E$ . The function  $I_E$  is called the indicator function or characteristic function of  $E$  and is sometimes denoted by  $1_E$  or  $\chi_E$ . We have

$$\{I_E \leq a\} = I_E^{-1}((-\infty, a]) = \begin{cases} \emptyset, & a < 0 \\ E^c, & 0 \leq a < 1 \\ \mathcal{S}, & a \geq 1. \end{cases}$$

Clearly,  $I_E$  is a random variable and the cumulative distribution function of  $I_E$  is

$$F_{I_E}(a) = P(\{I_E \leq a\}) = \begin{cases} 0, & a < 0 \\ P(E^c), & 0 \leq a < 1 \\ 1, & a \geq 1. \end{cases}$$

**Example 9.** Let  $\mathcal{S} = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$  with  $P(E) = \frac{\text{number of elements in } E}{8}$ , for every  $E \in \mathcal{P}(\mathcal{S})$ . Let  $X : \mathcal{S} \rightarrow \mathbb{R}$  be a random variable, defined by  $X(w) = \text{number of heads}$ . Then the cumulative distribution function of  $X$  is

$$F_X(x) = P(\{X \leq x\}) = \sum_{i \in \{0,1,2,3\} \cap (-\infty, x]} P(\{X = i\}) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3. \end{cases}$$

**Example 10.** Consider the probability space  $(\mathbb{R}, \mathbb{B}_{\mathbb{R}}, P)$  with  $P(B) = \int_0^{\infty} e^{-t} I_B(t) dt$ , where  $I_B$  is the indicator function of  $B$ . Define  $X : \mathbb{R} \rightarrow \mathbb{R}$  by

$$X(w) = \begin{cases} 0, & w \leq 0 \\ \sqrt{w}, & w > 0. \end{cases}$$

We have

$$\{X \leq a\} = X^{-1}((-\infty, a]) = \begin{cases} \emptyset, & a < 0 \\ (-\infty, a^2], & a \geq 0. \end{cases}$$

Thus  $X$  is a random variable. Now, the cumulative distribution function of  $X$  is

$$F_X(x) = P(\{X \leq x\}) = \begin{cases} P(\emptyset), & x < 0 \\ P((-\infty, x^2]), & x \geq 0. \end{cases}$$

Thus

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \int_0^{x^2} e^{-t} dt, & x \geq 0 \end{cases} = \begin{cases} 0, & x < 0 \\ 1 - e^{-x^2}, & x \geq 0. \end{cases}$$

**Definition 11.** A real-valued function  $F : \mathbb{R} \rightarrow \mathbb{R}$  that is increasing, right continuous and satisfies

$$F(-\infty) = 0 \text{ and } F(\infty) = 1$$

is called a distribution function.

**Theorem 12.** Every distribution function is the distribution function of a random variable on some probability space.