function $X : \Omega \to \mathbb{R}$ given by

$$X(H) := 1, \quad X(TH) := 2, \quad X(TTH) := 3, \quad X(TTTH) := 4, \cdots$$

Now, we focus on analysis of such functions X defined on the sample space Ω of some random experiment $\mathcal{E}$.

Notation 1.92 (Pre-image of a set under a function)**.** Let Ω be a non-empty set and let $X : \Omega \to \mathbb{R}$ be a function. Given any subset A of $\mathbb{R}$, we consider the subset $X^{-1}(A)$ of Ω defined by

$$X^{-1}(A) := \{\omega \in \Omega : X(\omega) \in A\}.$$

The set $X^{-1}(A)$ shall be referred to as the pre-image of A under the function X.

Remark 1.93. In Notation 1.92, we do not know whether the function X is bijective. As such, we cannot identify X^{-1} as the 'inverse' function of X. To avoid any confusion, treat $X^{-1}(A)$ as one symbol referring to the set as defined above and do not identify it as a combination of symbols X^{-1} and A.

Remark 1.94 (Shorthand notation for Pre-images). In the setting of Notation 1.92, we shall suppress the symbols ω and use the following notation for convenience, viz.

$$X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\} = (X \in A).$$

For specific sets A, other notations, again for convenience, may be used. For example for

(a) If $A = (-\infty, x]$, then we write

$$X^{-1}(A) = (X \in A) = \{\omega \in \Omega : X(\omega) \in (-\infty, x]\} = \{\omega \in \Omega : X(\omega) \leq x\} = (X \leq x).$$

For $A = (-\infty, x), (x, \infty), [x, \infty)$, we shall write $X^{-1}(A)$ to be equal to $(X < x), (X > x), (X \geq x)$ respectively.

(b) If $A = \{x\}$, then we write

$$X^{-1}(A) = (X \in A) = \{\omega \in \Omega : X(\omega) \in \{x\}\} = \{\omega \in \Omega : X(\omega) = x\} = (X = x).$$

Remark 1.95 (Properties of pre-images). Let $X : \Omega \to \mathbb{R}$ be a function. The following are some properties of the pre-images under X, which follow from the fact that X is a function.

(a) $X^{-1}(\mathbb{R}) = \Omega$.

(b) $X^{-1}(\emptyset_{\mathbb{R}}) = \emptyset_{\Omega}$, where $\emptyset_{\mathbb{R}}$ and $\emptyset_{\Omega}$ denote the empty sets under $\mathbb{R}$ and Ω, respectively. When there is no chance of confusion, we simply write $X^{-1}(\emptyset) = \emptyset$.

(c) For any two subsets A, B of $\mathbb{R}$ with $A \cap B = \emptyset$, we have $X^{-1}(A) \cap X^{-1}(B) = \emptyset$.

(d) For any subset A of $\mathbb{R}$, we have $X^{-1}(A^c) = \left(X^{-1}(A)\right)^c$.

(e) Let $\mathcal{I}$ be an indexing set. For any collection $\{A_i : i \in \mathcal{I}\}$ of subsets of $\mathbb{R}$, we have

$$X^{-1}\left(\bigcup_{i\in\mathcal{I}} A_i\right) = \bigcup_{i\in\mathcal{I}} X^{-1}(A_i), \quad X^{-1}\left(\bigcap_{i\in\mathcal{I}} A_i\right) = \bigcap_{i\in\mathcal{I}} X^{-1}(A_i).$$

The above properties shall be used frequently throughout the course.

Note 1.96. As discussed in Note 1.91, we now look at real valued functions defined on Ω, where Ω is the sample space of a random experiment $\mathcal{E}$. We shall also assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. The event space $\mathcal{F}$ shall be taken as the power set 2^{Ω}, unless stated otherwise.

Definition 1.97 (Random variable or RV)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Any real valued function $X : \Omega \to \mathbb{R}$ shall be referred to as a random variable or simply, an RV. In this case, we shall say that X is an RV defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Note 1.98. Since $\mathcal{F}$ is taken to be 2^{Ω}, we immediately have

$$X^{-1}(A) = (X \in A) \in \mathcal{F}$$

for any subset A of $\mathbb{R}$. If $\mathcal{F}$ is taken to be a smaller collection of subsets of Ω, then the above observation may not hold for any arbitrary function X. Given such $\mathcal{F}$, we then restrict our attention to the class of functions X satisfying the above property and refer to them as RVs. It is therefore important to specify $\mathcal{F}$ before we discuss RVs X. As mentioned earlier, $\mathcal{F}$ shall be taken as 2^{Ω}, unless stated otherwise.

Note 1.99. The probability function/measure $\mathbb{P}$ has not been used in the definition of an RV X. We now discuss the role of $\mathbb{P}$ in analysis of RVs X.

Notation 1.100. We write $\mathbb{B}$ to denote the power set of $\mathbb{R}$.

Notation 1.101. Let X be an RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then for all $A \in \mathbb{B}$, we have $X^{-1}(A) \in \mathcal{F}$ and hence $\mathbb{P}(X^{-1}(A))$ is well defined. We denote this in terms of a set function $\mathbb{P} \circ X^{-1} : \mathbb{B} \to [0, 1]$ given by $\mathbb{P} \circ X^{-1}(A) := \mathbb{P}(X^{-1}(A)) = \mathbb{P}(X \in A), \forall A \in \mathbb{B}$. A shorthand notation $\mathbb{P}_X$ shall also be used to refer to $\mathbb{P} \circ X^{-1}$.

Notation 1.102. Similar to the discussion in Remark 1.94, we shall write $\mathbb{P}(X \leq x), \mathbb{P}(X = x)$ etc. for $\mathbb{P} \circ X^{-1}(A)$ where $A = (-\infty, x], \{x\}$ etc. respectively.

Proposition 1.103. *Let X be an RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, the set function $\mathbb{P} \circ X^{-1}$ is a probability function/measure defined on the collection $\mathbb{B}$.*

Proof. We verify the axioms/properties of a probability function/measure as mentioned in Definition 1.33.

We have $\mathbb{P} \circ X^{-1}(\mathbb{R}) = \mathbb{P}(X^{-1}(\mathbb{R})) = \mathbb{P}(\Omega) = 1$. Since $\mathbb{P}$ is a probability measure on $\mathcal{F}$, we also have $\mathbb{P} \circ X^{-1}(A) = \mathbb{P}(X^{-1}(A)) \geq 0, \forall A \in \mathbb{B}$.

If $\{A_n\}_n$ is a sequence of pairwise disjoint sets in $\mathbb{B}$, then $\{X^{-1}(A_n)\}_n$ is a sequence of pairwise disjoint events in $\mathcal{F}$. Hence,

$$\mathbb{P} \circ X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} X^{-1}(A_n)\right) = \sum_{n=1}^{\infty} \mathbb{P}(X^{-1}(A_n)) = \sum_{n=1}^{\infty} \mathbb{P} \circ X^{-1}(A_n).$$

This proves countable additivity property for $\mathbb{P} \circ X^{-1}$ and the proof is complete. $\square$

Definition 1.104 (Induced Probability Space and Induced Probability Measure)**.** If X is an RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then the probability function/measure $\mathbb{P} \circ X^{-1}$ on $\mathbb{B}$ is referred to as the induced probability function/measure induced by X. In this case, $(\mathbb{R}, \mathbb{B}, \mathbb{P} \circ X^{-1})$ is referred to as the induced probability space induced by X.

Example 1.105. Recall from Remark 1.90, that if we toss a fair coin twice independently, then the sample space is $\Omega = \{HH, HT, TH, TT\}$ with $\mathbb{P}(\{HH\}) = \mathbb{P}(\{HT\}) = \mathbb{P}(\{TH\}) = \mathbb{P}(\{TT\}) = \frac{1}{4}$. Consider the RV $X : \Omega \to \mathbb{R}$ which denotes the number of heads. Here,

$$X(HH) = 2, \quad X(HT) = X(TH) = 1, \quad X(TT) = 0.$$

Consider the induced probability measure $\mathbb{P} \circ X^{-1}$ on $\mathbb{B}$. We have

$$\mathbb{P} \circ X^{-1}(\{0\}) = \mathbb{P}(X^{-1}(\{0\})) = \mathbb{P}(\{TT\}) = \frac{1}{4},$$

$$\mathbb{P} \circ X^{-1}(\{1\}) = \mathbb{P}(X^{-1}(\{1\})) = \mathbb{P}(\{HT, TH\}) = \frac{1}{2},$$

$$\mathbb{P} \circ X^{-1}(\{2\}) = \mathbb{P}(X^{-1}(\{2\})) = \mathbb{P}(\{HH\}) = \frac{1}{4}.$$

More generally, for any $A \in \mathbb{B}$, we have

$$\mathbb{P} \circ X^{-1}(A) = \mathbb{P}(\{\omega : X(\omega) \in A\}) = \sum_{i \in \{0,1,2\} \cap A} \mathbb{P} \circ X^{-1}(\{i\}).$$

Remark 1.106. If we know the probability function/measure $\mathbb{P}_X = \mathbb{P} \circ X^{-1}$ for any RV X, then we get the information about all the probabilities $\mathbb{P}(X \in A), A \in \mathbb{B}$ for events $X^{-1}(A) = (X \in A), A \in \mathbb{B}$ involving the RV X. In what follows, our analysis of RV X shall be through the understanding of probability function/measure $\mathbb{P}_X = \mathbb{P} \circ X^{-1}$ on $\mathbb{B}$.

Definition 1.107 (Law/Distribution of an RV)**.** If X is an RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then the probability function/measure $\mathbb{P}_X$ on $\mathbb{B}$ is referred to as the law or distribution of the RV X.

We now discuss some properties of a probability function/measure. To do this, we first introduce a concept involving sequence of sets.

Definition 1.108 (Increasing and decreasing sequence of sets)**.** Let $\{A_n\}_n$ be a sequence of subsets of a non-empty set Ω.

(a) If $A_n \subseteq A_{n+1}, \forall n = 1, 2, \cdots$, we say that the sequence $\{A_n\}_n$ is increasing. In this case, we say A_n increases to A, denoted by $A_n \uparrow A$, where $A = \bigcup_{n=1}^{\infty} A_n$.

(b) If $A_n \supseteq A_{n+1}, \forall n = 1, 2, \cdots$, we say that the sequence $\{A_n\}_n$ is decreasing. In this case, we say A_n decreases to A, denoted by $A_n \downarrow A$, where $A = \bigcap_{n=1}^{\infty} A_n$.

Remark 1.109. $A_n \uparrow A$ if and only if $A_n^c \downarrow A^c$.

Proposition 1.110 (Continuity of a probability measure)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.*

(a) (Continuity from below) Let $\{A_n\}_n$ be sequence in $\mathcal{F}$, such that $A_n \uparrow A$. Then

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A).$$

(b) (Continuity from above) Let $\{A_n\}_n$ be sequence in $\mathcal{F}$, such that $A_n \downarrow A$. Then

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A).$$

Proof. To prove the first statement. Since $\{A_n\}_n$ is an increasing sequence of sets, we have

$$A_n \cap (A_1 \cup A_2 \cup \cdots \cup A_{n-1})^c = A_n \cap A_{n-1}^c, \forall n \geq 2.$$

Then using a hint from practice problem set 1, we have

$$\bigcup_{n=1}^{\infty} A_n = A_1 \cup \left(\bigcup_{n=2}^{\infty} (A_n \cap A_{n-1}^c) \right).$$

Since the sets $A_1, A_2 \cap A_1^c, A_3 \cap A_2^c, \cdots$ are pairwise disjoint, using the countable additivity of $\mathbb{P}$, we have

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mathbb{P}(A_1) + \sum_{n=2}^{\infty} \mathbb{P}(A_n \cap A_{n-1}^c) = \mathbb{P}(A_1) + \lim_{k\to\infty} \sum_{n=2}^{k} \mathbb{P}(A_n \cap A_{n-1}^c) \\
&= \mathbb{P}(A_1) + \lim_{k\to\infty} \sum_{n=2}^{k} [\mathbb{P}(A_n) - \mathbb{P}(A_{n-1})] \\
&= \mathbb{P}(A_1) + \lim_{k\to\infty} [\mathbb{P}(A_k) - \mathbb{P}(A_1)] = \lim_{k\to\infty} \mathbb{P}(A_k).
\end{aligned}$$

This completes the proof of the first statement.

To prove the second statement. First observe that $A_n^c \uparrow A^c$ with $A = \bigcap_{n=1}^{\infty} A_n$. Using the first statement, we have

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \lim_{n\to\infty} \mathbb{P}(A_n^c) = \lim_{n\to\infty} [1 - \mathbb{P}(A_n^c)] = \lim_{n\to\infty} \mathbb{P}(A_n).$$

The proof is complete. $\square$

Definition 1.111 (Distribution function of an RV). Let X be an RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with law/distribution $\mathbb{P}_X$. Consider the function $F_X : \mathbb{R} \to \mathbb{R}$ defined by $F_X(x) :=$

$\mathbb{P}_X((-\infty, x]) = \mathbb{P}(X \le x), \forall x \in \mathbb{R}$. The function F_X is called the cumulative distribution function (CDF) or simply, the distribution function (DF) of the RV X.

Remark 1.112 (RVs equal in law/distribution). Let X be an RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let Y be an RV defined on a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$. If $\mathbb{P} \circ X^{-1} = \mathbb{P}' \circ Y^{-1}$, i.e. $\mathbb{P} \circ X^{-1}(A) = \mathbb{P}' \circ Y^{-1}(A), \forall A \in \mathbb{B}$, then we say that X and Y are equal in law/distribution. In this case, $F_X = F_Y$, i.e. $F_X(x) = F_Y(x), \forall x \in \mathbb{R}$.

Remark 1.113. Let X and Y be two RVs, possibly defined on different probability spaces. If $F_X = F_Y$, then it can be shown that X and Y are equal in law/distribution. The proof of this statement is beyond the scope of this course. This statement is often restated as 'the DF of an RV uniquely determines the law/distribution of the RV'.

Example 1.114. Consider X as in Example 1.105. Then for all $x \in \mathbb{R}$, we have

$$F_X(x) = \mathbb{P}_X((-\infty, x])) = \sum_{i \in \{0,1,2\} \cap (-\infty, x]} \mathbb{P}_X(\{i\}) = \begin{cases} 0, \text{ if } x < 0, \\ \mathbb{P}_X(\{0\}), \text{ if } 0 \le x < 1, \\ \mathbb{P}_X(\{0\}) + \mathbb{P}_X(\{1\}), \text{ if } 1 \le x < 2, \\ \mathbb{P}_X(\{0\}) + \mathbb{P}_X(\{1\}) + \mathbb{P}_X(\{2\}), \text{ if } x \ge 2. \end{cases}$$

Therefore,

$$F_X(x) = \begin{cases} 0, \text{ if } x < 0, \\ \frac{1}{4}, \text{ if } 0 \le x < 1, \\ \frac{3}{4}, \text{ if } 1 \le x < 2, \\ 1, \text{ if } x \ge 2. \end{cases}$$

Theorem 1.115. *Let X be an RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with law $\mathbb{P}_X$ and DF F_X. Then*

(a) F_X is non-decreasing, i.e. $F_X(x) \le F_Y(y), \forall x < y$.

(b) F_X is right continuous, i.e. $\lim_{h \downarrow 0} F_X(x+h) = F_X(x), \forall x \in \mathbb{R}$.

(c) $F_X(-\infty) := \lim_{x \to -\infty} F_X(x) = 0$ and $F_X(\infty) := \lim_{x \to \infty} F_X(x) = 1$.

Proof. For all $x < y$, observe that $(-\infty, x] \subsetneq (-\infty, y]$. Since $\mathbb{P}_X$ is a probability measure, we have $\mathbb{P}_X((-\infty, x]) \leq \mathbb{P}_X((-\infty, y])$. The statement (a) follows.

By definition, F_X takes values in $[0,1]$ and hence it is bounded. Since F_X is non-decreasing, the limit $F_X(x+) = \lim_{h \downarrow 0} F_X(x+h)$ exists for all $x \in \mathbb{R}$. Using the non-decreasing property, we use the following fact from real analysis that $F_X(x+) = \lim_{n\to\infty} F_X(x + \frac{1}{n})$. By Proposition 1.110, we have

$$F_X(x+) = \lim_{n\to\infty} F_X\left(x + \frac{1}{n}\right) = \lim_{n\to\infty} \mathbb{P}_X(\left(-\infty, x + \frac{1}{n}\right]) = \mathbb{P}_X((-\infty, x]) = F_X(x).$$

This proves statement (b). Here, we use the fact that $(-\infty, x + \frac{1}{n}] \downarrow (-\infty, x]$.

Similar to the proof of statement (b), we have

$$F_X(-\infty) = \lim_{n\to\infty} F_X(-n) = \lim_{n\to\infty} \mathbb{P}_X((-\infty, -n]) = \mathbb{P}_X(\emptyset) = 0,$$

and

$$F_X(\infty) = \lim_{n\to\infty} F_X(n) = \lim_{n\to\infty} \mathbb{P}_X((-\infty, n]) = \mathbb{P}_X(\mathbb{R}) = 1.$$

Here, we use that facts that $(-\infty, -n] \downarrow \emptyset$ and $(-\infty, n] \uparrow \mathbb{R}$. This proves statement (c). $\square$

The next theorem is stated without proof. The arguments required to prove this statement is beyond the scope of this course.

Theorem 1.116. *Let $F : \mathbb{R} \to \mathbb{R}$ be a non-decreasing and right continuous function such that $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$. Then there exists an RV X defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $F = F_X$, i.e. $F(x) = F_X(x), \forall x$.*

Remark 1.117. Given any function $F : \mathbb{R} \to \mathbb{R}$, as soon as we check the relevant conditions, we can claim that it is the DF of some RV by Theorem 1.116.

Example 1.118. Consider the function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(x) = \begin{cases} 0, & \text{if } x < 0, \\ x, & \text{if } 0 \leq x < 1, \\ 1, & \text{if } x \geq 1. \end{cases}$$

The function is a constant on $(-\infty, 0)$ and on $[1, \infty)$. Moreover, it is non-decreasing in the interval $[0, 1)$. Further for $x < 0, y \in (0, 1), z > 1$, we have

$$F(x) = F(0) < F(y) < F(1) = F(z).$$

Hence, F in non-decreasing over $\mathbb{R}$. Again, by definition F is continuous on the intervals $(-\infty, 0)$, $(0, 1)$ and $(1, \infty)$. We check for right continuity at the points 0 and 1. We have

$$\lim_{h \downarrow 0} F(0 + h) = \lim_{h \downarrow 0} h = 0 = F(0), \quad \lim_{h \downarrow 0} F(1 + h) = \lim_{h \downarrow 0} 1 = 1 = F(1).$$

Hence, F is right continuous on $\mathbb{R}$. Finally, $\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} 0 = 0$ and $\lim_{x \to \infty} F(x) = \lim_{x \to \infty} 1 = 1$. Hence, F is the DF of some RV. Later on, we shall identify the corresponding RV.

Proposition 1.119 (Further properties of a DF)**.** *Let X be an RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with law $\mathbb{P}_X$ and DF F_X.*

(a) For all $x \in \mathbb{R}$, the limit $F_X(x-) = \lim_{h \downarrow 0} F_X(x - h)$ exists and equals $\mathbb{P}_X((-\infty, x)) = \mathbb{P}(X < x)$.

Proof. Since F_X in non-decreasing and bounded, as argued in Theorem 1.115, the limit $F_X(x-) = \lim_{h \downarrow 0} F_X(x - h)$ exists and moreover, by Proposition 1.110 we have

$$F_X(x-) = \lim_{n \to \infty} F_X\left(x - \frac{1}{n}\right) = \lim_{n \to \infty} \mathbb{P}_X(\left(-\infty, x - \frac{1}{n}\right]) = \mathbb{P}_X((-\infty, x)) = \mathbb{P}(X < x).$$

Here, we use the fact that $(-\infty, x - \frac{1}{n}] \uparrow (-\infty, x)$. $\square$

(b) For all $x \in \mathbb{R}$, $\mathbb{P}(X \geq x) = 1 - F_X(x-)$.

Proof. We have, $\mathbb{P}(X \geq x) = \mathbb{P}_X([x, \infty)) = \mathbb{P}_X((-\infty, x)^c) = 1 - \mathbb{P}_X((-\infty, x)) = 1 - F_X(x-)$. $\square$

(c) For any $x \in \mathbb{R}$, $F_X(x-) \leq F_X(x+)$.

Proof. By the non-decreasing property of F_X, for all $x \in \mathbb{R}$ and positive integers n, we have, $F_X(x - \frac{1}{n}) \leq F_X(x + \frac{1}{n})$. Letting n go to infinity in this inequality, we get the result. $\square$

(d) F_X is continuous at x if and only if $F_X(x) = F_X(x-)$.

Proof. A real valued function is continuous at a point x if and only if the function is both right continuous and left continuous at the point x. Now, by construction, F_X is right continuous on $\mathbb{R}$. Hence, F_X is continuous at x if and only if F_X is left continuous at x. The last statement is exactly the statement to be proved. $\square$

(e) *Only possible discontinuities of F_X are jump discontinuities.*

Proof. As discussed in Theorem 1.115 and in part (a), for any $x \in \mathbb{R}$, both the limits $F_X(x+)$ and $F_X(x-)$ exist and $F_X(x+) = F_X(x)$. Since $F_X(x-) \leq F_X(x+)$, the only possible discontinuity appears if and only if $F_X(x-) < F_X(x+)$. These discontinuities are jump discontinuities. This completes the proof. $\square$

(f) *For all $x \in \mathbb{R}$, we have $F_X(x+) - F_X(x-) = \mathbb{P}(X = x)$.*

Proof. By the finite additivity of $\mathbb{P}_X$, we have $F_X(x+) - F_X(x-) = \mathbb{P}(X \leq x) - \mathbb{P}(X < x) = \mathbb{P}_X((-\infty, x]) - \mathbb{P}_X((-\infty, x)) = \mathbb{P}_X(\{x\}) = \mathbb{P}(X = x)$. $\square$

(g) *If F_X has a jump at x, then the jump is given by $F_X(x+) - F_X(x-) = \mathbb{P}(X = x)$.*

Proof. If F_X has a jump at x, then the jump is given by $F_X(x+) - F_X(x-)$. The result follows from statement (f). $\square$

(h) *F_X is continuous at x if and only if $\mathbb{P}(X = x) = 0$.*

Proof. Recall that $F_X(x+) = F_X(x)$. Then by statement (d) and (f), we have F_X is continuous at x if and only if $F_X(x+) = F_X(x-)$ and hence, if and only if $\mathbb{P}(X = x) = 0$. $\square$

(i) *Consider the set $D := \{x \in \mathbb{R} : F_X \text{ is discontinuous at } x\} = \{x \in \mathbb{R} : F_X(x-) < F_X(x+)\} = \{x \in \mathbb{R} : \mathbb{P}(X = x) > 0\}$. Then D is either finite or countably infinite. (Note that if F_X is continuous on $\mathbb{R}$, then $D = \emptyset$.)*

Proof. Left as an exercise in practice problem set 3. $\square$

(j) *For all $x < y$, we have*

$$\mathbb{P}(x < X \leq y) = F_X(y) - F_X(x),$$

$$\mathbb{P}(x < X < y) = F_X(y-) - F_X(x),$$

$$\mathbb{P}(x \leq X < y) = F_X(y-) - F_X(x-),$$

$$\mathbb{P}(x \le X \le y) = F_X(y) - F_X(x-).$$

Proof. We prove the first two equalities. Proof of the last two equalities are similar.

By the finite additivity of $\mathbb{P}_X$, we have $F_X(y) - F_X(x) = \mathbb{P}_X((-\infty, y]) - \mathbb{P}_X((-\infty, x]) = \mathbb{P}_X((x, y]) = \mathbb{P}(x < X \le y)$.

Again, $F_X(y-) - F_X(x) = \mathbb{P}_X((-\infty, y)) - \mathbb{P}_X((-\infty, x]) = \mathbb{P}_X((x, y)) = \mathbb{P}(x < X < y)$.

This completes the proof. $\square$

Example 1.120. Consider the function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(x) = \begin{cases} 0, \text{ if } x < 0, \\ \frac{1}{4} + \frac{x}{2}, \text{ if } 0 \le x \le 1, \\ \frac{1}{2} + \frac{x}{4}, \text{ if } 1 < x < 2, \\ 1, \text{ if } x \ge 2. \end{cases}$$

Assume that F is the DF of some RV X (left as an exercise in practice problem set 3). Since F is continuous on the intervals $(-\infty, 0), (0, 1), (1, 2)$ and $(2, \infty)$, discontinuities may arise only at the points $0, 1, 2$.

We have $F(0-) = \lim_{h \downarrow 0} F(0 - h) = 0$ and $F(0) = \frac{1}{4}$. Therefore F is discontinuous at 0 with jump $F(0) - F(0-) = \frac{1}{4}$.

We have $F(1-) = \lim_{h \downarrow 0} F(1 - h) = \lim_{h \downarrow 0} [\frac{1}{4} + \frac{1-h}{2}] = \frac{3}{4}$ and $F(1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$. Therefore F is continuous at 1.

We have $F(2-) = \lim_{h \downarrow 0} F(2 - h) = \lim_{h \downarrow 0} [\frac{1}{2} + \frac{2-h}{4}] = 1$ and $F(2) = 1$. Therefore F is continuous at 2.

Only discontinuity of F is at the point 0. In particular, $\mathbb{P}(X = 0) = F(0) - F(0-) = \frac{1}{4}$. At all other points F is continuous and hence $\mathbb{P}(X = x) = 0, \forall x \ne 0$.

Observe that $\mathbb{P}(0 \le X < 1) = F(1-) - F(0-) = \frac{3}{4}$. Again, $\mathbb{P}(\frac{3}{2} < X \le 2) = F(2) - F(\frac{3}{2}) = 1 - [\frac{1}{2} + \frac{3}{8}] = \frac{1}{8}$.

We now discuss special classes of RVs defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Recall that $\mathbb{P}_X$ and F_X denote the law/distribution and the distribution function (DF) of an RV X, respectively.

Definition 1.121 (Discrete RV)**.** An RV X is said to be a discrete RV if there exists a finite or countably infinite set $S \subsetneq \mathbb{R}$ such that

$$1 = \mathbb{P}_X(S) = \mathbb{P}(X \in S) = \sum_{x \in S} \mathbb{P}_X(\{x\}) = \sum_{x \in S} \mathbb{P}(X = x)$$

and $\mathbb{P}(X = x) > 0, \forall x \in S$. In this situation, we refer to the set S as the support of the discrete RV X.

Remark 1.122. Let X be a discrete RV with DF F_X and support S. Then we have the following observations.

(a) $\mathbb{P}_X(S^c) = 1 - \mathbb{P}_X(S) = 0$. In particular, for any $x \in S^c$, $0 \leq \mathbb{P}(X = x) = \mathbb{P}_X(\{x\}) \leq \mathbb{P}_X(S^c) = 0$ and hence $\mathbb{P}(X = x) = 0, \forall x \in S^c$.

(b) Since $\mathbb{P}_X(S) = 1$, for any $A \subseteq \mathbb{R}$, we have $\mathbb{P}_X(A) = \mathbb{P}_X(A \cap S)$ (see problem set 1). Moreover,

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}_X(A \cap S) = \sum_{x \in A \cap S} \mathbb{P}(X = x).$$

(c) Recall that F_X is right continuous, i.e. $F_X(x+) = F_X(x), \forall x \in \mathbb{R}$. Moreover, $F_X(x) - F_X(x-) = \mathbb{P}(X = x)$. From the discussion above, we conclude that

$$F_X(x) - F_X(x-) = \mathbb{P}(X = x) \begin{cases} > 0, & \text{if } x \in S, \\ = 0, & \text{if } x \in S^c. \end{cases}$$

Hence, the set of discontinuities of F_X is exactly the support S.

(d) Note that

$$1 = \sum_{x \in S} \mathbb{P}(X = x) = \sum_{x \in S} [F_X(x) - F_X(x-)].$$

Hence, the sum of the jumps of F_X is exactly 1.

Example 1.123. Consider the DF $F : \mathbb{R} \to [0,1]$ considered in Example 1.120 given by

$$F(x) = \begin{cases} 0, \text{ if } x < 0, \\ \frac{1}{4} + \frac{x}{2}, \text{ if } 0 \le x \le 1, \\ \frac{1}{2} + \frac{x}{4}, \text{ if } 1 < x < 2, \\ 1, \text{ if } x \ge 2. \end{cases}$$

As discussed earlier, F only has a discontinuity at the point 0. If an RV X has this F as the DF, then

$$\sum_{x \in D} \mathbb{P}(X = x) = \mathbb{P}(X = 0) = \frac{1}{4} \neq 1,$$

with $D = \{0\}$ as the set of discontinuities of F. This RV X is not discrete.

Example 1.124. Let X denote the number of heads in tossing a fair coin twice independently. As computed earlier in Example 1.114, the DF F_X is given by

$$F_X(x) = \begin{cases} 0, \text{ if } x < 0, \\ \frac{1}{4}, \text{ if } 0 \le x < 1, \\ \frac{3}{4}, \text{ if } 1 \le x < 2, \\ 1, \text{ if } x \ge 2. \end{cases}$$

Clearly, the set D of discontinuities of F_X is $\{0, 1, 2\}$ with

$$\mathbb{P}(X = x) = F_X(x) - F_X(x-) = \begin{cases} \frac{1}{4} - 0 = \frac{1}{4}, \text{ if } x = 0, \\ \frac{3}{4} - \frac{1}{4} = \frac{1}{2}, \text{ if } x = 1, \\ 1 - \frac{3}{4} = \frac{1}{4}, \text{ if } x = 2. \end{cases}$$

Since $\sum_{x \in D} \mathbb{P}(X = x) = 1$, the RV X is discrete with support D.