Chapter 2

Construction of probability measures

In the previous chapter we assumed the existence of measures. In this chapter we consider their existence and construction.

2.1 π systems, Dynkin systems, semi-rings and σ-algebras

We start by introducing two additional systems that may be associated with a set $\mathbb{X}$.

Definition 2.1. *1. A system $\mathcal{P}$ associated with $\mathbb{X}$ is called a π system if $A, B \in \mathcal{P} \implies A \cap B \in \mathcal{P}$.*

2. A system $\mathcal{D}$ associated with $\mathbb{X}$ is called a Dynkin[1] system if:

a) $\mathbb{X} \in \mathcal{D}$

b) $A \in \mathcal{D} \implies A^c \in \mathcal{D}$

c) $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{D}$ *and* $A_i \cup A_j = \emptyset \; \forall \; i \neq j, \; i, j \in \mathbb{N} \implies \bigcup_{j\in\mathbb{N}} A_j \in \mathcal{D}$.

It is evident from part 2. of this definition that a σ-algebra associated with $\mathbb{X}$ is also a Dynkin system associated with $\mathbb{X}$. As in the case of σ-algebras, there exist smallest Dynkin systems generated by subsets of $\mathbb{X}$.

[1] Eugene Borisovich Dynkin was a Russian mathematician that made important contributions to algebra and probability. He was a student of Andrei Kolmogorov.

Theorem 2.1. *Let $\mathcal{C} \subset 2^{\mathbb{X}}$. There exists a smallest Dynkin system $\delta(\mathcal{C})$ such that $\mathcal{C} \subset \delta(\mathcal{C})$. It is called the Dynkin system generated by $\mathcal{C}$. In addition, $\delta(\mathcal{C}) \subset \sigma(\mathcal{C})$.*

Proof. Existence and characterization of $\delta(\mathcal{C})$ is proved as in Theorem 1.2. Since $\sigma(\mathcal{C})$ is a Dynkin system $\delta(\sigma(\mathcal{C})) = \sigma(\mathcal{C})$. Since $\mathcal{C} \subset \sigma(\mathcal{C})$, $\delta(\mathcal{C}) \subset \delta(\sigma(\mathcal{C})) = \sigma(\mathcal{C})$ as in Theorem 1.3. ■

The next theorem shows that a Dynkin system is a σ-algebra if, and only if, it is a π system.

Theorem 2.2. *A Dynkin system $\mathcal{D}$ is a σ-algebra $\iff A, B \in \mathcal{D} \implies A \cap B \in \mathcal{D}$.*

Proof. ($\implies$) If $\mathcal{D}$ is a σ-algebra, then $A, B \in \mathcal{D} \implies A \cap B \in \mathcal{D}$, since σ-algebras are closed under countable intersections.

($\Longleftarrow$) If $\mathcal{D}$ is a Dynkin system it satisfies requirements 1 and 2 for σ-algebras in Definition 1.1. Let $A_i \in \mathcal{D}$ for $i \in \mathbb{N}$, we must show that $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{D}$. Define $B_1 := A_1$, $B_2 := A_2 - B_1 = A_2 \cap B_1^c$, $B_3 := A_3 - \cup_{i=1}^{2} B_i = A_3 \cap (\cup_{i=1}^{2} B_i)^c \cdots$ $B_n := A_n - \cup_{i=1}^{n-1} B_i = A_n \cap (\cup_{i=1}^{n-1} B_i)^c$. The collection $\{B_i\}_{i\in\mathbb{N}}$ is pairwise disjoint, and since each B_i is the intersection of two sets in $\mathcal{D}$, using closeness under finite intersections, $\bigcup_{i\in\mathbb{N}} B_i = \bigcup_{i\in\mathbb{N}} A_i \in \mathcal{D}$. ■

Theorem 2.3. *If $\mathcal{P}$ is a π system associated with $\mathbb{X}$, then $\delta(\mathcal{P}) = \sigma(\mathcal{P})$.*

Proof. From Theorem 2.1, $\delta(\mathcal{P}) \subset \sigma(\mathcal{P})$ and from Theorem 2.2 if $\delta(\mathcal{P})$ is a π system it is a σ-algebra. Since $\sigma(\mathcal{P})$ is the smallest σ-algebra generated by $\mathcal{P}$, it must be that $\delta(\mathcal{P}) = \sigma(\mathcal{P})$, so it suffices to show that $\delta(\mathcal{P})$ is a π-system. For any $D \in \delta(\mathcal{P})$, let $\mathcal{D}_D = \{A \subset \mathbb{X} : A \cap D \in \delta(\mathcal{P})\}$. First, we show that $\mathcal{D}_D$ is a Dynkin system. We verify conditions a), b) and c) in Definition 2.1.

a) Note that $\mathbb{X} \cap D = D \in \delta(\mathcal{P})$, hence $\mathbb{X} \in \mathcal{D}_D$.

b) If $A \in \mathcal{D}_D$, then $A \cap D \in \delta(\mathcal{P})$. Now, $A^c \cap D = (A^c \cup D^c) \cap D = (A \cap D)^c \cap D = ((A \cap D) \cup D^c)^c$ where $A \cap D$ and D^c are disjoint. Also, since $D \in \delta(\mathcal{P})$, we have $D^c \in \delta(\mathcal{P})$, and $A \cap D \in \delta(\mathcal{P})$ by assumption, so $((A \cap D) \cup D^c)^c \in \delta(\mathcal{P})$. Thus $A^c \in \mathcal{D}_D$.

c) Let A_i for $i \in \mathbb{N}$ be pairwise disjoint with $A_i \cap D \in \delta(\mathcal{P})$ and note that $\{(A_i \cap D)\}_{i \in \mathbb{N}}$ forms a disjoint collection. Thus, $\bigcup\limits_{i \in \mathbb{N}} (A_i \cap D) = D \cap \bigcup\limits_{i \in \mathbb{N}} A_i$ and $\bigcup\limits_{i \in \mathbb{N}} A_i \in \mathcal{D}_D$. Thus, $\mathcal{D}_D$ is a Dynkin system.

Fix $G \in \mathcal{P}$. Then, $G \in \delta(\mathcal{P})$ and we can define $\mathcal{D}_G = \{A \subset \mathbb{X} : A \cap G \in \delta(\mathcal{P})\}$. Now, consider $G' \in \mathcal{P}$. Since, $\mathcal{P}$ is a π-system, $G' \cap G \in \mathcal{P} \subset \delta(\mathcal{P})$. Hence, $G' \in \mathcal{D}_G$, showing that $\mathcal{P} \subset \mathcal{D}_G$ for all $G \in \mathcal{P}$. But $\mathcal{D}_G$ is a Dynkin system and consequently, by definition $\delta(\mathcal{P}) \subset \mathcal{D}_G$, $\forall\, G \in \mathcal{P}$.

Thus, we have that if $D \in \delta(\mathcal{P})$ and $G \in \mathcal{P}$, then $G \cap D \in \delta(\mathcal{P})$ and $\mathcal{P} \subset \mathcal{D}_D$ (by definition of $\mathcal{D}_D$). Then, $\delta(\mathcal{P}) \subset \mathcal{D}_D$ for all $D \in \delta(\mathcal{P})$ implying that $\delta(\mathcal{P})$ is a π system by definition of $\mathcal{D}_D$. ■

Definition 2.2. *A semi-ring, denoted by $\mathcal{S}$, is a non-empty system associated with $\mathbb{X}$ having the following properties:*

1. *$\emptyset \in \mathcal{S}$,*

2. *$A, B \in \mathcal{S} \implies A \cap B \in \mathcal{S}$,*

3. *for all $A, B \in \mathcal{S}$ there exists $m \in \mathbb{N}$ and $\{S_j\}_{j=1}^m \subset \mathcal{S}$ that is pairwise disjoint such that $B - A = \cup_{j=1}^m S_j$.*

Remark 2.1. 1. *A semi-ring is a π system in view of condition 2.*

2. *Property 3 in Definition 2.2 is equivalent to the following:*

   3'. if $A, B \in \mathcal{S}$ and $A \subset B$, then $B = A \cup \left(\bigcup_{j=1}^m S_j \right)$ where the collection $\{A, S_1, \cdots S_m\} \subset \mathcal{S}$ is pairwise disjoint.

To verify that 3 $\implies$ 3' note that $A \subset B \implies B = A \cup (B - A) = A \cup \left(\bigcup_{j=1}^{m} S_j\right)$ by 3, where $\{A, S_1, \cdots S_m\} \subset \mathcal{S}$ is pairwise disjoint. Now, to verify that 3' $\implies$ 3 note that $B = (B \cap A) \cup (B - A)$. Since $(B \cap A) \subset B$, by 3' $B = (B \cap A) \cup \left(\bigcup_{j=1}^{m} S_j\right)$. Thus, $(B \cap A) \cup (B - A) = (B \cap A) \cup \left(\bigcup_{j=1}^{m} S_j\right)$ which implies that $B - A = \bigcup_{j=1}^{m} S_j$ where $\{S_j\}_{j=1}^{m} \subset \mathcal{S}$ is pairwise disjoint.

3. *A ring $\mathcal{R}$ is a non-empty system of sets associated with $\mathbb{X}$ such that $A, B \in \mathcal{R} \implies A \cup B \in \mathcal{R}$ and $A - B \in \mathcal{R}$. If $A \in \mathcal{R}$ then $A - A = \emptyset \in \mathcal{R}$. Also, if $A, B \in \mathcal{R}$, and noting that $A \cap B = A - (A - B)$, we have that $A \cap B \in \mathcal{R}$. Now let $A \subset B$, $A, B \in \mathcal{R}$. Since $B = A \cup (B - A)$ and $(B - A) \in \mathcal{R}$, we conclude that every ring is a semi-ring using property 3'.*

4. *If $\mathcal{A}$ is an algebra, then for $A, B \in \mathcal{A}$ we have that $A \cup B, A \cap B, B^c \in \mathcal{A}$, and since $A - B = A \cap B^c \in \mathcal{A}$, an algebra is a ring.*

It follows from these remarks that we have the following hierarchy of systems: $\mathcal{A}$ (algebras) are $\mathcal{R}$ (rings) are $\mathcal{S}$ (semi-rings) are π systems.

2.2 Uniqueness of measures

The following theorem shows that under some conditions, measures that coincide on some generating class $\mathcal{G}$ coincide on $\sigma(\mathcal{G})$.

Theorem 2.4. *Let $(\mathbb{X}, \sigma(\mathcal{P}))$ be a measurable space and $\mathcal{P}$ a collection of subsets of $\mathbb{X}$, such that:*

1. *$\mathcal{P}$ is a π system,*

2. *there exists $\{P_j\}_{j\in\mathbb{N}} \subset \mathcal{P}$ with $P_1 \subset P_2 \subset \cdots$ such that $\bigcup_{j\in\mathbb{N}} P_j := \lim_{j\to\infty} P_j = \mathbb{X}$ (the sequence $\{P_j\}_{j\in\mathbb{N}}$ is exhausting).*

If μ and v are measures that coincide on $\mathcal{P}$ and, are finite for all P_j, then $\mu(A) = v(A)$ for all $A \in \sigma(\mathcal{P})$.

Proof. For $j \in \mathbb{N}$ let $\mathcal{D}_j = \{A \in \sigma(\mathcal{P}) : \mu(A \cap P_j) = v(A \cap P_j)\}$. First, we show that $\mathcal{D}_j$ is a Dynkin system.

1. $\mathbb{X} \in \mathcal{D}_j$ since $\mu(\mathbb{X} \cap P_j) = \mu(P_j) = v(P_j) = v(\mathbb{X} \cap P_j)$.

2. Let $A \in \mathcal{D}_j$. Note that $P_j = (A \cap P_j) \cup (A^c \cap P_j)$ and note that the two sets in the union are disjoint. Since μ is a measure $\mu(P_j) = \mu(A \cap P_j) + \mu(A^c \cap P_j)$. Hence, $\mu(A^c \cap P_j) = \mu(P_j) - \mu(A \cap P_j)$. Since μ and v coincide in $\mathcal{P}$ we have that $v(P_j) = \mu(P_j)$ and since $A \in \mathcal{D}_j$ we have that $\mu(A \cap P_j) = v(A \cap P_j)$. Hence,

$$\mu(A^c \cap P_j) = \mu(P_j) - \mu(A \cap P_j) = v(P_j) - v(A \cap P_j) = v(A^c \cap P_j).$$

   Thus, $A^c \in \mathcal{D}_j$.

3. Let $A_1, A_2, \cdots$ be a pairwise disjoint collection in $\mathcal{D}_j$.

$$\begin{aligned}
\mu\left(\left(\bigcup_{i \in \mathbb{N}} A_i\right) \cap P_j\right) = \mu\left(\bigcup_{i \in \mathbb{N}} (A_i \cap P_j)\right) &= \sum_{i=1}^{\infty} \mu(A_i \cap P_j) \\
&= \sum_{i=1}^{\infty} v(A_i \cap P_j) \text{ since } A_i \in \mathcal{D}_j \\
&= v\left(\bigcup_{i \in \mathbb{N}} (P_j \cap A_i)\right) = v\left(P_j \cap \left(\bigcup_{i \in \mathbb{N}} A_i\right)\right)
\end{aligned}$$

   and consequently, $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{D}_j$.

Since $\mathcal{P}$ is a π-system, by Theorem 2.3 and the definition $\delta(\mathcal{P})$, we have $\mathcal{P} \subset \delta(\mathcal{P}) = \sigma(\mathcal{P}) \subset \mathcal{D}_j$. But by construction $\mathcal{D}_j \subset \sigma(\mathcal{P})$ and we conclude that $\mathcal{D}_j = \sigma(\mathcal{P})$. So, for all $A \in \sigma(\mathcal{P})$ and $j = 1, 2, \cdots$,

$$\mu(A \cap P_j) = v(A \cap P_j). \tag{2.1}$$

By continuity of measures from below and noting that $(A \cap P_1) \subset (A \cap P_2) \subset \cdots$, letting $j \to \infty$ in (2.1) we have for all $A \in \sigma(\mathcal{P})$,

$$\begin{aligned}\lim_{j\to\infty} \mu(A \cap P_j) &= \mu\left(\lim_{j\to\infty} (A \cap P_j)\right) = \mu\left(\underset{j\in\mathbb{N}}{\cup} (A \cap P_j)\right) \\ &= \mu\left(A \cap \left(\underset{j\in\mathbb{N}}{\cup} P_j\right)\right) = \mu\left(A \cap \mathbb{X}\right) \\ &= \mu(A)\end{aligned}$$

Similarly, $\lim_{j\to\infty} v(A \cap P_j) = v(A)$ and we conclude that $\mu(A) = v(A)$. ■

2.3 Existence of measures - Carathéodory's Extension Theorem

We take the following path to construct a measure on $\mathcal{F}$. We start with a class of subsets $\mathcal{S}$ of $\mathbb{X}$, such that $\mathcal{F} = \sigma(\mathcal{S})$, and define a pre-measure μ on $\mathcal{S}$. If $\mathcal{S}$ and μ satisfy the requirements of Theorem 2.4, then μ will extend uniquely to $\mathcal{F}$, provided we are able to extend it from $\mathcal{S}$ to $\mathcal{F}$. The result that provides the conditions and possibility for such an extension is known as Carathéodory's Extension Theorem.

Theorem 2.5. *(Carathéodory) Let $\mathcal{S}$ be a semi-ring of subsets of $\mathbb{X}$ and $\mu : \mathcal{S} \to [0, \infty]$ be a pre-measure. Then, μ has an extension to a measure μ on $\sigma(\mathcal{S})$. If there exists $\{E_j\}_{j\in\mathbb{N}} \in \mathcal{S}$ with $E_1 \subset E_2 \cdots$ such that $\lim_{j\to\infty} E_j \to \mathbb{X}$ and $\mu(E_j) < \infty$ for all j, then the extension is unique.*

Proof. **Step 1.** We start by defining the set function $\mu^* : 2^{\mathbb{X}} \to [0, \infty]$. For any $A \subset \mathbb{X}$ define the collection of countable covers for A that are composed of sets in $\mathcal{S}$ by

$$C(A) = \{\{S_j\}_{j\in\mathbb{N}} \subset \mathcal{S} : A \subset \underset{j\in\mathbb{N}}{\cup} S_j\}.$$

If A cannot be covered by some $\underset{j\in\mathbb{N}}{\cup} S_j$, then $C(A) = \emptyset$. Now, define

$$\mu^*(A) := \inf \left\{ \sum_{j\in\mathbb{N}} \mu(S_j) : \{S_j\}_{j\in\mathbb{N}} \in C(A) \right\},$$

where $\inf \emptyset := \infty$. Note that,

a) $\mu^*(\emptyset) = 0$, by taking $S_1 = S_2 = \cdots = \emptyset$

b) $A \subset B$ implies that every cover for B is also a cover for A, i.e., $C(B) \subset C(A)$. Therefore,

$$\mu^*(A) = \inf\left\{\sum_{j\in\mathbb{N}} \mu(S_j) : \{S_j\}_{j\in\mathbb{N}} \in C(A)\right\} \leq \inf\left\{\sum_{j\in\mathbb{N}} \mu(T_j) : \{T_j\}_{j\in\mathbb{N}} \in C(B)\right\} = \mu^*(B).$$

c) Let $A_n \subset \mathbb{X}$ for $n \in \mathbb{N}$ and, without loss of generality, assume that $\mu^*(A_n) < \infty$ (that is $C(A_n) \neq \emptyset$). Choose $\epsilon > 0$ and let $\{S_{nk}\}_{k\in\mathbb{N}} \in C(A_n)$ be such that

$$\sum_{k\in\mathbb{N}} \mu(S_{nk}) \leq \mu^*(A_n) + \epsilon/2^n.$$

Now, $\underset{n\in\mathbb{N}}{\cup} A_n \subset \underset{n\in\mathbb{N}}{\cup}\,\underset{k\in\mathbb{N}}{\cup} S_{nk}$ and by the definition of infimum and sub-additivity of pre-measures

$$\begin{aligned}\mu^*\left(\underset{n\in\mathbb{N}}{\cup} A_n\right) &\leq \sum_{n\in\mathbb{N}}\sum_{k\in\mathbb{N}} \mu(S_{nk}) \\ &\leq \sum_{n\in\mathbb{N}} (\mu^*(A_n) + \epsilon/2^n) = \sum_{n\in\mathbb{N}} \mu^*(A_n) + \epsilon.\end{aligned}$$

Hence, $\mu^*\left(\underset{n\in\mathbb{N}}{\cup} A_n\right) \leq \sum_{n\in\mathbb{N}} \mu^*(A_n)$. If $\mu^*(A_n) = \infty$ for some n, then the last inequality holds trivially.

Since μ^* satisfies properties a)-c), it is called an outer-measure on $2^{\mathbb{X}}$.

Step 2. We now show that μ^* extends μ (defined on $\mathcal{S}$) to $2^{\mathbb{X}}$. By this we mean that $\mu^*(S) = \mu(S)$ for $S \in \mathcal{S}$.

First, let $\mathcal{S}_U = \{S : S = \cup_{j=1}^m S_j, S_j \in \mathcal{S}, S_i \cap S_j = \emptyset\ \forall i \neq j \text{ and } m \in \mathbb{N}\}$ be the collection of sets that can be written as disjoint finite unions of elements of $\mathcal{S}$ and let $\bar{\mu}(S) = \sum_{j=1}^m \mu(S_j)$ for $S \in \mathcal{S}_U$. Note that $\bar{\mu}(S)$ is invariant to the pairwise disjoint finite union used to represent S. To see this, suppose $S = \cup_{j=1}^m S_j$ and $S = \cup_{k=1}^n T_k$ for $m, n \in \mathbb{N}$.

Then, $\cup_{j=1}^m S_j = \cup_{k=1}^n T_k$ and $S_j = S_j \cap (\cup_{k=1}^n T_k) = \cup_{k=1}^n (S_j \cap T_k)$ and $S_j \cap T_k \in \mathcal{S}$, since a semi-ring is a π-system. Since μ is a pre-measure on $\mathcal{S}$, and $\{T_k\}_{k=1}^n$ is a pairwise disjoint collection, $\mu(S_j) = \sum_{k=1}^n \mu(T_k \cap S_j)$. Then,

$$\bar{\mu}(S) = \sum_{j=1}^m \mu(S_j) = \sum_{k=1}^n \sum_{j=1}^m \mu(T_k \cap S_j) = \sum_{k=1}^n \mu(T_k).$$

We now show that $\mathcal{S}_U$ is closed under (arbitrary) finite intersections and unions. If $A, B \in \mathcal{S}_U$ then $A \cap B = (\cup_{j=1}^m S_j) \cap (\cup_{k=1}^n T_k)$ where the two unions are over pairwise disjoint sets. Then, $A \cap B = \cup_{j=1}^m \cup_{k=1}^n (S_j \cap T_k) \in \mathcal{S}_U$ since $S_j \cap T_k \in \mathcal{S}$ for all j, k and $\{S_j \cap T_k\}_{j=1,k=1}^{m,n}$ is pairwise disjoint.

Also, since $S_j, T_k \in \mathcal{S}$, their difference can be written as a finite union of pairwise disjoint elements of $\mathcal{S}$. Hence, $S_j - T_k \in \mathcal{S}_U$. Now,

$$A - B = \cup_{j=1}^m S_j - \cup_{k=1}^n T_k = \cup_{j=1}^m \cap_{k=1}^n (S_j \cap T_k^c) = \cup_{j=1}^m \cap_{k=1}^n (S_j - T_k).$$

Since, $S_j - T_k \in \mathcal{S}_U$ and given that we have shown that $\mathcal{S}_U$ is closed under finite intersections, $\cap_{k=1}^n (S_j - T_k) \in \mathcal{S}_U$. Hence, $A - B$ is the finite union of pairwise disjoint elements in $\mathcal{S}_U$ and we conclude that $A - B \in \mathcal{S}_U$, since $\mathcal{S}_U$ is closed under pairwise disjoint unions. Lastly, since $A \cup B = (A - B) \cup (A \cap B) \cup (B - A)$ and all sets in the union are disjoint and in $\mathcal{S}_U$, we conclude that $A \cup B \in \mathcal{S}_U$.

We now show that $\bar{\mu}$ is σ-additive on $\mathcal{S}_U$, i.e., a pre-measure. Let $\{T_k\}_{k \in \mathbb{N}} \subset \mathcal{S}_U$ such that $\{T_k\}_{k \in \mathbb{N}}$ is pairwise disjoint and such that $T := \underset{k \in \mathbb{N}}{\cup} T_k \in \mathcal{S}_U$. Since $T_k \in \mathcal{S}_U$, by definition there exist $\{S_j\}_{j \in \mathbb{N}} \in \mathcal{S}$ and a sequence of $0 = n_0 \leq n_1 \leq \cdots$ of integers such that

$$T_k = S_{n_{(k-1)}+1} \cup S_{n_{(k-1)}+2} \cup \cdots \cup S_{n_k} \text{ for } k \in \mathbb{N},$$

where the collection $\{S_{n_{(k-1)}+1}, S_{n_{(k-1)}+2}, \cdots, S_{n_k}\}$ is pairwise disjoint and

$$T = \bigcup_{k \in \mathbb{N}} \bigcup_{j=n_{(k-1)}+1}^{n_k} S_j.$$

Also, since $T \in \mathcal{S}_U$, it can be written as $T = \bigcup_{l=1}^{N} U_l$ where $N \in \mathbb{N}$ with $U_l \in \mathcal{S}$ and $\{U_l\}_{l=1}^{N}$ a pairwise disjoint collection. Hence,

$$\bigcup_{l=1}^{N} U_l = \bigcup_{k\in\mathbb{N}} \bigcup_{j=n_{(k-1)}+1}^{n_k} S_j.$$

Defining disjoint subsets $J_1, \cdots, J_N$ of $\mathbb{N}$ such that $\cup_{l=1}^{N} J_l = \mathbb{N}$ we write $U_l = \bigcup_{j\in J_l} S_j$ and note that $U_l \in \mathcal{S}$. Now, $T = \bigcup_{k\in\mathbb{N}} T_k = \cup_{l=1}^{N} U_l$ and

$$\begin{aligned}
\bar{\mu}(T) &= \sum_{l=1}^{N} \mu(U_l) \text{ by definition of } \bar{\mu} \\
&= \sum_{l=1}^{N} \sum_{j\in J_l} \mu(S_j) \text{ by } \mu \text{ being a pre-measure on } \mathcal{S} \\
&= \sum_{k\in\mathbb{N}} \sum_{j=n_{(k-1)}+1}^{n_k} \mu(S_j) = \sum_{k\in\mathbb{N}} \bar{\mu}(T_k).
\end{aligned}$$

Now, for any $S \in \mathcal{S}$ and any $\mathcal{S}$-covering of S, i.e., $\{S_j\}_{j\in\mathbb{N}} \in C(S)$

$$\begin{aligned}
\mu(S) = \bar{\mu}(S) = \bar{\mu}\left(\bigcup_{j\in\mathbb{N}} S_j \cap S\right) & \text{ since } S \in \mathcal{S} \implies S \in \mathcal{S}_U \\
&\leq \sum_{j\in\mathbb{N}} \bar{\mu}(S_j \cap S) \text{ since } \bar{\mu} \text{ is a pre-measure and sub-additive} \\
&= \sum_{j\in\mathbb{N}} \mu(S_j \cap S) \leq \sum_{j\in\mathbb{N}} \mu(S_j).
\end{aligned}$$

Taking the infimum over $C(S)$, we have $\mu(S) \leq \mu^*(S)$. Now, taking $(S, \emptyset, \cdots) \in C(S)$ gives $\mu^*(S) \leq \mu(S)$. Combining the two inequalities, we have

$$\mu^*(S) = \mu(S) \text{ for all } S \in \mathcal{S}.$$

Step 3. We will show that $\mathcal{S} \subset \mathcal{A}^*$ where

$$\mathcal{A}^* = \{A \subset \mathbb{X} : \mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \cap A^c), \forall\, Q \subset \mathbb{X}\}. \tag{2.2}$$

Let $S, T \in \mathcal{S}$ and note that $T = (T \cap S) \cup (T \cap S^c) = (T \cap S) \cup (T - S) = (T \cap S) \cup (\cup_{j=1}^{m} S_j)$ with $\{S_j\}_{j=1}^{m}$ disjoint, $m \in \mathbb{N}$ and where the last equality follows from the third defining property of semi-rings. Since μ is a pre-measure on $\mathcal{S}$ we have

$$\mu(T) = \mu(T \cap S) + \sum_{j=1}^{m} \mu(S_j).$$

Since μ^* and μ coincide on $\mathcal{S}$ and $T \cap S \in \mathcal{S}$, and since μ^* is sub-additive, from c) in Step 1, we have $\mu^*(T - S) = \mu^*(\cup_{j=1}^{m} S_j) \leq \sum_{j=1}^{m} \mu^*(S_j) = \sum_{j=1}^{m} \mu(S_j)$. Consequently,

$$\mu(T) = \mu(T \cap S) + \sum_{j=1}^{m} \mu(S_j) \geq \mu^*(T \cap S) + \mu^*(T - S). \tag{2.3}$$

Take $Q \subset \mathbb{X}$ and $\{T_j\}_{j \in \mathbb{N}} \in C(Q)$. Using $\mu^*(T_j) = \mu(T_j)$ and summing (2.3) over j taking $T = T_j$

$$\sum_{j \in \mathbb{N}} \mu^*(S \cap T_j) + \sum_{j \in \mathbb{N}} \mu^*(T_j - S) \leq \sum_{j \in \mathbb{N}} \mu^*(T_j).$$

Sub-additivity and monotonicity of μ^* together with $Q \subset \bigcup_{j \in \mathbb{N}} T_j$ give

$$\begin{aligned} \mu^*(Q \cap S) + \mu^*(Q - S) &\leq \mu^*(\cup_{j \in \mathbb{N}}(T_j \cap S)) + \mu^*(\cup_{j \in \mathbb{N}}(T_j - S)) \\ &\leq \sum_{j \in \mathbb{N}} \mu^*(T_j) = \sum_{j \in \mathbb{N}} \mu(T_j). \end{aligned}$$

Taking the infimum over $C(Q)$, $\mu^*(Q \cap S) + \mu^*(Q - S) \leq \mu^*(Q)$. The reverse inequality follows easily from sub-additivity of μ^*. Consequently, if $S \in \mathcal{S}$ we have that $S \in \mathcal{A}^*$.

Step 4. We show that $\mathcal{A}^*$ is a σ-algebra and μ^* is a measure on $(\mathbb{X}, \mathcal{A}^*)$.

1. For all $Q \subset \mathbb{X}$, $Q \cap \mathbb{X} = Q$ and $Q \cap \mathbb{X}^c = \emptyset$. Since $\mu^*(\emptyset) = 0$ we have that $\mathbb{X} \in \mathcal{A}^*$.

2. For all $Q \subset \mathbb{X}$ suppose $A \in \mathcal{A}^*$, i.e.

$$\mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \cap A^c).$$

But by symmetry of the right hand side of the equality due to $(A^c)^c = A$, we have $A^c \in \mathcal{A}^*$.

3. If $A, A' \in \mathcal{A}^*$, for all $Q \subset \mathbb{X}$

$$
\begin{aligned}
\mu^*(Q \cap (A \cup A')) + \mu^*(Q - (A \cup A'))& \\
&= \mu^*(Q \cap (A \cup (A' - A))) + \mu^*(Q - (A \cup A')) \\
&= \mu^*((Q \cap A) \cup [Q \cap (A' - A)]) + \mu^*(Q - (A \cup A')) \\
&\leq \mu^*(Q \cap A) + \mu^*(Q \cap (A' - A)) + \mu^*(Q - (A \cup A')) \\
&\text{using subadditivity of } \mu^* \\
&= \mu^*(Q \cap A) + \mu^*((Q - A) \cap A') + \mu^*((Q - A) - A') \\
&= \mu^*(Q \cap A) + \mu^*(Q - A) = \mu^*(Q) \\
&\text{using the defining expression for } \mathcal{A}^* \text{ twice, once for } Q - A \text{ and once for } Q.
\end{aligned}
$$

Thus,

$$\mu^*(Q \cap (A \cup A')) + \mu^*(Q - (A \cup A')) \leq \mu^*(Q). \tag{2.4}$$

Now, $Q = \{Q \cap (A \cup A')\} \cup \{Q \cap (A \cup A')^c\}$. By sub-additivity of μ^*

$$\mu^*(Q) \leq \mu^*(Q \cap (A \cup A')) + \mu^*(Q - (A \cup A')). \tag{2.5}$$

Combining inequalities (2.4) and (2.5) we conclude that $\mu^*(Q) = \mu^*(Q \cap (A \cup A')) + \mu^*(Q - (A \cup A'))$ and consequently $\mathcal{A}^*$ is closed under finite unions.

If A, $A' \in \mathcal{A}^*$ such that $A \cap A' = \emptyset$, then for $Q = (A \cup A') \cap P$ with $P \subset \mathbb{X}$ the equality $\mu^*(Q \cap A) + \mu^*(Q - A) = \mu^*(Q)$ becomes

$$\mu^*((A \cup A') \cap P) = \mu^*(P \cap A) + \mu^*(P \cap A'), \forall\, P \subset \mathbb{X}.$$

For a disjoint collection $\{A_j\}_{j=1}^m \in \mathcal{A}^*$,

$$\mu^*((\cup_{j=1}^m A_j) \cap P) = \sum_{j=1}^m \mu^*(P \cap A_j).$$

If $A = \cup_{j \in \mathbb{N}} A_j$, where $\{A_j\}$ is a disjoint collection,

$$\mu^*(P \cap A) \geq \mu^*(P \cap (\cup_{j=1}^m A_j)) = \sum_{j=1}^m \mu^*(P \cap A_j).$$

Since $\cup_{j=1}^{m} A_j \in \mathcal{A}^*$ we have that

$$\begin{aligned}\mu^*(P) &= \mu^*(P \cap (\cup_{j=1}^{m} A_j)) + \mu^*(P - \cup_{j=1}^{m} A_j) \\ &\geq \mu^*(P \cap (\cup_{j=1}^{m} A_j)) + \mu^*(P - A) \\ &= \sum_{j=1}^{m} \mu^*(P \cap A_j) + \mu^*(P - A).\end{aligned}$$

Let $m \to \infty$, to conclude

$$\mu^*(P) \geq \sum_{j=1}^{\infty} \mu^*(P \cap A_j) + \mu^*(P - A) \geq \mu^*(P \cap A) + \mu^*(P - A)$$

The reverse inequality follows directly from sub-additivity of μ^*. Thus,

$$\mu^*(P) = \mu^*(P \cap A) + \mu^*(P - A), \forall P \subset \mathbb{X}.$$

Consequently, $A = \cup_{j \in \mathbb{N}} A_j$ where the collection $\{A_j\}_{j \in \mathbb{N}}$ is pairwise disjoint is in $\mathcal{A}^*$. Consequently, $\mathcal{A}^*$ is a Dynkin system that is closed under finite unions. By DeMorgan Laws, $\mathcal{A}^*$ is closed under finite intersections, and by Theorem 2.2, $\mathcal{A}^*$ is a σ-algebra.

Now, we show that μ^* is a measure on $\sigma(\mathcal{S})$. From above, $\mathcal{S} \subset \mathcal{A}^*$, so $\sigma(\mathcal{S}) \subset \mathcal{A}^*$. Also, μ^* is a measure on $\mathcal{A}^*$ and on $\sigma(\mathcal{S})$, which extends μ on $\mathcal{S}$. By Theorem 2.4, and under the conditions in the enunciation of this theorem, any two extensions μ^* and v^* of μ coincide on $\sigma(\mathcal{S})$. ■

Remark 2.2. $(\mathbb{X}, \mathcal{A}^*, \mu^*)$ *is a complete measure space. To verify completeness, let* $E \in \mathcal{A}^*$ *such that* $\mu^*(E) = 0$, *and consider* $B \subset E$. *We must verify that* $B \in \mathcal{A}^*$, *i.e., for any* $Q \subset \mathbb{X}$, *it must be that*

$$\mu^*(Q) = \mu^*(Q \cap B) + \mu^*(Q \cap B^c).$$

Now, $Q \cap B \subset Q \cap E \subset E \implies \mu^*(Q \cap B) \leq \mu^*(E) = 0$ *and, consequently* $\mu^*(Q \cap B) = 0$. *Also,* $Q \cap B^c \subset Q \implies \mu^*(Q \cap B^c) \leq \mu^*(Q)$. *Hence,*

$$\mu^*(Q) \geq \mu^*(Q \cap B^c) + \mu^*(Q \cap B). \tag{2.6}$$

By sub-additivity

$$\mu^*(Q) \leq \mu^*(Q \cap B^c) + \mu(Q \cap B) \tag{2.7}$$

Given (2.6) *and* (2.7) *we have* $\mu^*(Q) = \mu^*(Q \cap B^c) + \mu^*(Q \cap B)$*. In addition,* $\mu^*(B) = 0$ *follows from monotonicity of measures.*

2.4 Lebesgue measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

In this section, we use Carathéodory's Theorem to construct a measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. The first step is to show that $\mathcal{I}^{n,h}$ is a semi-ring.

Theorem 2.6. *Let* $R^{n,h} = \times_{i=1}^{n}[a_i, b_i)$ *for* $n \in \mathbb{N}$ *be a half-open rectangle in* $\mathbb{R}^n$ *and* $\mathcal{I}^{n,h}$ *be the collection formed by all such rectangles with real endpoints.* $\mathcal{I}^{n,h}$ *is a semi-ring.*

Proof. Let $\mathcal{I}^{1,h} = \{[a_i, b_i) : a_i \leq b_i \text{ where } a_i,\ b_i \in \mathbb{R}\}$ and note that:

1. if $b_i = a_i$, $[a_i, b_i) = \emptyset$,
2. if $[a_i, b_i)$, $[a_j, b_j) \in \mathcal{I}^{n,h}$ then $[a_i, b_i) \cap [a_j, b_j) = \begin{cases} \emptyset & \in \mathcal{I}^{1,h} \\ [a_j, b_i) & \in \mathcal{I}^{1,h} \\ [a_i, b_j) & \in \mathcal{I}^{1,h} \\ [a_i, b_i) & \in \mathcal{I}^{1,h} \end{cases}$,
3. if $[a_1, b_1) \subset [a_2, b_2)$ then $[a_2, b_2) = [a_2, a_1) \cup [a_1, b_1) \cup [b_1, b_2)$, where the members in the union are all disjoint.

Hence, $\mathcal{I}^{1,h}$ is a semi-ring.

Now, suppose $\mathcal{I}^{n,h}$ is a semi-ring. We will verify that $\mathcal{I}^{n+1,h}$ is a semi-ring. First, note that $\mathcal{I}^{n+1,h} = \mathcal{I}^{n,h} \times \mathcal{I}^{1,h}$ and since $\emptyset \in \mathcal{I}^{n,h}$ we immediately conclude that $\emptyset \in \mathcal{I}^{n+1,h}$. The intersection of two rectangles in $\mathcal{I}^{n+1,h}$ is given by

$$(R^{n,h} \times R^{1,h}) \cap (I^{n,h} \times I^{1,h}) = (R^{n,h} \cap I^{n,h}) \times (R^{1,h} \cap I^{1,h})$$

where $I^{n,h}$ is a half-open rectangle in $\mathbb{R}^n$ and the righthand side of the equality is an element of $\mathcal{I}^{n+1,h}$. Also, $(R^{n,h} \times R^{1,h}) - (I^{n,h} \times I^{1,h}) = (R^{n,h} \times R^{1,h}) \cap (I^{n,h} \times I^{1,h})^c$ and note that

$$\begin{aligned}(I^{n,h} \times I^{1,h})^c &= \{(x,y) : x \notin I^{n,h}, y \notin I^{1,h}, \text{ or } x \in I^{n,h} \text{ and } y \notin I^{1,h}, \text{ or } x \notin I^{n,h} \text{ and } y \in I^{1,h}\} \\ &= ((I^{n,h})^c \times (I^{1,h})^c) \cup (I^{n,h} \times (I^{1,h})^c) \cup ((I^{n,h})^c \times I^{1,h})\end{aligned}$$

where the components of the union are disjoint. Thus,

$$\begin{aligned}(R^{n,h} \times R^{1,h}) - (I^{n,h} \times I^{1,h}) &= [(R^{n,h} \times R^{1,h}) \cap ((I^{n,h})^c \times (I^{1,h})^c)] \cup [(R^{n,h} \times R^{1,h}) \cap (I^{n,h} \times (I^{1,h})^c)] \\ &\quad \cup [(R^{n,h} \times R^{1,h}) \cap ((I^{n,h})^c \times I^{1,h})] \\ &= [(R^{n,h} - I^{n,h}) \times (R^{1,h} - I^{1,h})] \cup [(R^{n,h} \cap I^{n,h}) \times (R^{1,h} - I^{1,h})] \\ &\quad \cup [(R^{n,h} - I^{n,h}) \times (R^{1,h} \cap I^{1,h})].\end{aligned}$$

By the induction assumption, $R^{n,h} - I^{n,h}$ and $R^{1,h} - I^{1,h}$ can be expressed as finite unions of disjoint rectangles, which completes the proof. ■

Definition 2.3. *Let* $\lambda^n : \mathcal{I}^{n,h} \to [0, \infty)$ *be defined as* $\lambda^n(R^{n,h}) = \prod_{j=1}^{n}(b_j - a_j)$ *whenever* $b_j > a_j$ *for* $j = 1, \cdots, n$ *and* $\lambda^n(R^{n,h}) = 0$ *if* $b_j \leq a_j$ *for some* j.

Theorem 2.7. λ^n *is a pre-measure on* $\mathcal{I}^{n,h}$.

Proof. We start by showing that λ^1 is a pre-measure on $\mathcal{I}^{1,h}$. Let $[a, b) \in \mathcal{I}^{1,h}$ and $[a, b) = \cup_{i=1}^{n}[a_i, b_i)$ with $a_1 = a$, $a_2 = b_1$, $a_3 = b_2, \cdots, a_n = b_{n-1}, b_n = b$. Then,

$$\begin{aligned}\sum_{i=1}^{n} \lambda^1([a_i, b_i)) &= (b_1 - a_1) + (b_2 - a_2) + \cdots + (b_{n-1} - a_{n-1}) + (b_n - a_n) \\ &= (a_2 - a) + (a_3 - a_2) + \cdots + (a_n - a_{n-1}) + (b - a_n) = b - a \\ &= \lambda^1([a, b)) = \lambda^1\left(\cup_{i=1}^{n}[a_i, b_i)\right).\end{aligned}$$

Therefore, λ^1 is finitely additive. For σ-additivity, we need to show that for $[a, b) = \bigcup_{i \in \mathbb{N}} [a_i, b_i)$, where $\{[a_i, b_i)\}_{i \in \mathbb{N}}$ is a pairwise disjoint collection we have $b - a = \sum_{i=1}^{\infty}(b_i - a_i)$.

For any $n \in \mathbb{N}$, let $\{[a_i, b_i)\}_{i=1}^n$ be a pairwise disjoint collection. Then, since $\mathcal{I}^{1,h}$ is a semi-ring, we can write

$$[a, b) - \cup_{i=1}^n [a_i, b_i) = \cup_{j=1}^m I_j,$$

where the last set is the finite union of pairwise disjoint half-open rectangles. Thus, since λ^1 is finitely additive on $\mathcal{I}^{1,h}$

$$\lambda^1([a, b)) = \sum_{i=1}^n \lambda^1([a_i, b_i)) + \sum_{j=1}^m \lambda^1(I_j) \geq \sum_{i=1}^n \lambda^1([a_i, b_i)).$$

Thus, $\lambda^1([a, b)) = b - a \geq \lim_{n\to\infty} \sum_{i=1}^n \lambda^1([a_i, b_i)) = \sum_{i=1}^\infty \lambda^1([a_i, b_i))$. We need only show that $b - a \leq \sum_{i=1}^\infty \lambda^1([a_i, b_i))$ to complete the proof.

Let $0 < \epsilon < b - a$ and consider a pair-wise disjoint collection $\{[a_i, b_i)\}_{i\in\mathbb{N}}$ such that $[a, b) = \bigcup_{i\in\mathbb{N}} [a_i, b_i)$. Note that

$$\begin{aligned}
[a, b - \epsilon) \subset [a, b - \epsilon] &\subset \cup_{i=1}^\infty (a_i - 2^{-i}\epsilon, b_i) \\
&\subset \cup_{i=1}^n (a_i - 2^{-i}\epsilon, b_i) \text{ for some } n \in \mathbb{N}, \text{ by the Heine-Borel Theorem} \\
&\subset \cup_{i=1}^n [a_i - 2^{-i}\epsilon, b_i).
\end{aligned}$$

But $\lambda^1([a_i, b_i)) = \lambda^1([a_i - 2^{-i}\epsilon, b_i)) - \frac{1}{2^i}\epsilon$. Hence,

$$\begin{aligned}
\lambda^1([a, b - \epsilon)) &\leq \sum_{i=1}^n \lambda^1\left([a_i - \frac{1}{2^i}\epsilon, b_i)\right) \text{ by subadditivity} \\
&= \sum_{i=1}^n (b_i - a_i + \frac{1}{2^i}\epsilon) \\
b - a - \epsilon &\leq \sum_{i=1}^n (b_i - a_i) + \epsilon \sum_{i=1}^n \frac{1}{2^i} \text{ or} \\
b - a &\leq \sum_{i=1}^n (b_i - a_i) + \epsilon \left(1 + \sum_{i=1}^n \frac{1}{2^i}\right).
\end{aligned}$$

Taking limits as $n \to \infty$ on both sides of the last inequality gives $b - a \leq \sum_{i=1}^\infty (b_i - a_i)$, which combined with the previously obtained reverse inequality gives $b - a = \sum_{i=1}^\infty (b_i - a_i)$. Hence, λ^1 is a pre-measure on $\mathcal{I}^{1,h}$.

Clearly, $\lambda^n(\emptyset) = 0$. The proof is completed by using induction on n, the dimension of the space. Hence, we assume that λ^n is σ-additive on $\mathcal{I}^{n,h}$ for some n and show that λ^{n+1} is σ-additive on $\mathcal{I}^{n+1,h}$. This final step is left as an exercise. ■

Theorem 2.8. *There exists a unique extension of λ^n from $\mathcal{I}^{n,h}$ to a measure on the Borel sets $\mathcal{B}(\mathbb{R}^n)$. This extension is denoted by λ^n and is called Lebesgue measure.*

Proof. We know that $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{I}^{n,h})$ from Theorem 1.5. Since, $[-k,k)^n = [-k,k) \times [-k,k) \cdots \times [-k,k) \uparrow \mathbb{R}^n$ as $k \to \infty$ is an exhausting sequence of n-rectangles, and since $\lambda^n([-k,k)^n) = (2k)^n < \infty$, all conditions of Carathéodory's Theorem are fulfilled. ■

Remark 2.3. *Let $(\mathbb{R}, \sigma(\mathcal{I}^{1,h}) = \mathcal{B}(\mathbb{R}))$ be a measurable space. From Theorem 1.4 if we set $S = [0,1)$ and consider $\mathcal{I} = \mathcal{I}^{1,h} \cap S = \left\{[0,1) \cap A : A \in \mathcal{I}^{1,h}\right\}$ then $\sigma(\mathcal{I}^{1,h} \cap [0,1)) = \mathcal{B}(\mathbb{R}) \cap [0,1)$ is a σ-algebra associated with $[0,1)$. Thus, we define $\mathcal{B}_{[0,1)} := \sigma(\mathcal{I}^{1,h} \cap [0,1))$ and note that*

$$([0,1), \mathcal{B}_{[0,1)} := \sigma(\mathcal{I}))$$

is a measurable space where $\mathcal{I} = \{[a,b) : 0 \leq a \leq b \leq 1\}$. Define the set function $\lambda : \mathcal{I} \to [0,1]$ such that $\lambda(\emptyset) = 0$ and $\lambda([a,b)) = b - a$. Since λ is σ-additive (pre-measure) on $\mathcal{I}$ (a semi-ring), using Carathéodory's Theorem, we can state that

$$([0,1), \mathcal{B}_{[0,1)} := \sigma(\mathcal{I}), \lambda^*)$$

is a measure space, where λ^* is the unique extension of λ from $\mathcal{I}$ to $\sigma(\mathcal{I})$. In addition, $\lambda^*([0,1)) = 1$. Thus, we have constructed a specific probability space.

2.5 Distribution functions and probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

We will now construct probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. This will be done using distribution functions.

Definition 2.4. *Let* $F : \mathbb{R} \to [0,1]$ *be a function with the following properties:*

1. $\lim_{h \downarrow 0} F(x+h) := F(x+) = F(x)$ *for all* $x \in \mathbb{R}$ *and* $h > 0$,

2. $x < y \implies F(x) \le F(y)$,

3. $\lim_{x \to \infty} F(x) = 1, \; \lim_{x \to -\infty} F(x) = 0.$

F is called a (proper or non-defective) distribution function (df). If only conditions 1 and 2 are met, F is called a defective df.

Remark 2.4. *1. Let* $F(x-) := \lim_{h \downarrow 0} F(x-h)$ *for* $h > 0$*. The left jump of F at x is defined as* $LJ_F(x) = F(x) - F(x-)$ *and the right jump of F at x is defined as* $RJ_F(x) = F(x+) - F(x)$*. The jump of F at x is defined as* $J_F(x) = LJ_F(x) + RJ_F(x) = F(x+) - F(x-)$*. If F is a df,* $RJ_F(x) = 0$ *for all* $x \in \mathbb{R}$ *and* $J_F(x) = F(x) - F(x-)$*. In addition, since F is nondecreasing* $J_F(x) \ge 0$*. If* $J_F(x) = 0$ *then F is continuous at x.*
2. For any two $x \le y \in \mathbb{R}$ *we have that* $0 \le F(y) - F(x) \le 1$.

Definition 2.5. *Let* $S(p) = \{x \in \mathbb{R} : F(x) \ge p\}$ *for* $p \in [0,1]$*. The left (generalized) inverse of a df F, denoted by F^-, is defined as* $F^-(p) := \inf S(p)$*, for* $p \in [0,1]$.

Remark 2.5. *Note that when* $p = 0$, $S(0) = \mathbb{R}$*, which is nonempty but not bounded below. As such,* $\inf S(0)$ *is not defined as a real number. In this case, we put* $F^-(0) := -\infty$*. Also, when* $p = 1$*, either* $S(1) = [a, \infty)$*, which is nonempty and bounded below by a, in which case* $F^-(1) = a \in \mathbb{R}$ *or* $S(1) = \emptyset$*, in which case we put* $F^-(1) = \infty$*. Hence,* $F^- : [0,1] \to \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$.

Theorem 2.9. *Let* $S(p) = \{x : F(x) \ge p\}$ *for* $p \in (0,1]$*. Then,*

1. *$S(p)$ is a closed set.*

2. $t < F^-(p) \iff F(t) < p$ *or* $F^-(p) \le t \iff p \le F(t)$.

Proof. 1. If $s_n \in S(p)$ and $s_n \downarrow s$, bp right continuity of F we have $p \leq F(s_n) \downarrow F(s)$. Thus, $p \leq F(s)$ and $s \in S(p)$. If $s_n \in S(p)$ and $s_n \uparrow s$, we have $p \leq F(s_n) \uparrow F(s-) \leq F(s)$. Thus, $p \leq F(s)$ which implies that $s \in S(p)$. Consequently, by a characterization of closed sets, $S(p)$ is closed.

2. Since $S(p)$ is closed, its infimum $F^-(p) \in S(p)$ and therefore $F(F^-(p)) \geq p$. $t < F^-(p) \implies t \notin S(p) \implies F(t) < p$. The reverse implications all apply. ■

Theorem 2.10. *Let $A \subset \mathbb{R}$, $\mathcal{S}_F(A) = \{p \in (0,1] : F^-(p) \in A\}$ and $\mathcal{I}^1 = \{(a,b] : -\infty \leq a < b < \infty\}$. If $A \in \mathcal{B}(\mathbb{R})$, then $\mathcal{S}_F(A) \in \mathcal{B}_{(0,1]} = \sigma(\mathcal{I}^1) \cap (0,1]$.*

Proof. Let $\mathcal{G} = \{A \subset \mathbb{R} : \mathcal{S}_F(A) \in \mathcal{B}_{(0,1]}\}$. Note that

$$\begin{aligned}
\mathcal{S}_F((a,b]) &= \{p \in (0,1] : F^-(p) \in (a,b]\} = \{p \in (0,1] : a < F^-(p) \leq b\} \\
&= \{p \in (0,1] : F(a) < p \leq F(b)\} \text{ by Theorem 2.9} \\
&= (F(a), F(b)] \in \mathcal{B}_{(0,1]}.
\end{aligned}$$

Hence, $(a,b] \in \mathcal{G}$ and $\mathcal{I}^1 \subset \mathcal{G}$. If $\mathcal{G}$ is a σ-algebra, $\sigma(\mathcal{I}^1) = \mathcal{B}(\mathbb{R}) \subset \mathcal{G}$. Hence, $A \in \mathcal{B}(\mathbb{R})$ implies $S_F(A) \in \mathcal{B}_{(0,1]}$. Consequently, we need only show that $\mathcal{G}$ is a σ-algebra associated with $\mathbb{R}$.

1. $\mathcal{S}_F(\mathbb{R}) = \{p \in (0,1] : F^-(p) \in \mathbb{R}\} = (0,1) = \bigcup_{n \in \mathbb{N}} (0, 1 - n^{-1}] \in \mathcal{B}_{(0,1]}$, thus $\mathbb{R} \in \mathcal{G}$.

2. By definition of $\mathcal{S}_F$

$$\begin{aligned}
\mathcal{S}_F(A^c) &= \{p \in (0,1] : F^-(p) \in A^c\} = \{p \in (0,1] : F^-(p) \notin A\} \\
&= (\mathcal{S}_F(A))^c \in \mathcal{B}_{(0,1]}
\end{aligned}$$

where the last inclusion statement follows if $A \in \mathcal{G}$ and the fact that $\mathcal{B}_{(0,1]}$ is a σ-algebra.

3. If $\{A_n\}_{n\in\mathbb{N}} \in \mathcal{G}$ we have by definition of $\mathcal{S}_F$

$$\begin{aligned}\mathcal{S}_F\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= \left\{p \in (0,1] : F^-(p) \in \bigcup_{n\in\mathbb{N}} A_n\right\} = \{p \in (0,1] : F^-(p) \in A_n \text{ for some } n\} \\ &= \bigcup_{n\in\mathbb{N}} \{p \in (0,1] : F^-(p) \in A_n\} = \bigcup_{n\in\mathbb{N}} \mathcal{S}_F(A_n) \in \mathcal{B}_{(0,1]} \qquad (2.8)\end{aligned}$$

where the last inclusion statement follows since $A_n \in \mathcal{G}$ and the fact that $\mathcal{B}_{(0,1]}$ is a σ-algebra.

■

Definition 2.6. *Let $A \in \mathcal{B}(\mathbb{R})$ and define $P_F(A) = \lambda^1(\mathcal{S}_F(A))$ where λ^1 is the Lebesgue measure on $\mathcal{B}_{(0,1]}$.*

Theorem 2.11. *Let P_F be given in Definition 2.6. Then, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_F)$ is a probability space.*

Proof. First, note that

$$P_F(\emptyset) = \lambda^1(\mathcal{S}_F(\emptyset)) = \lambda^1(\{p \in (0,1] : F^-(p) \in \emptyset\}) = \lambda^1(\emptyset) = 0.$$

Second, if $\{A_n\}_{n\in\mathbb{N}}$ is a pairwise disjoint collection of sets in $\mathcal{B}(\mathbb{R})$ then

$$\begin{aligned}P_F\left(\cup_{n\in\mathbb{N}} A_n\right) &= \lambda^1\left(\mathcal{S}_F\left(\cup_{n\in\mathbb{N}} A_n\right)\right) = \lambda^1\left(\cup_{n\in\mathbb{N}} \mathcal{S}_F(A_n)\right) \text{ by (2.8)} \\ &= \sum_{n=1}^{\infty} \lambda^1(\mathcal{S}_F(A_n)) = \sum_{n=1}^{\infty} P_F(A_n).\end{aligned}$$

where the next to last equality follows from the fact that λ^1 is a measure and $\{\mathcal{S}_F(A_n)\}_{n\in\mathbb{N}}$ is a pairwise disjoint collection.

Lastly,

$$\begin{aligned}P_F(\mathbb{R}) &= \lambda^1(\mathcal{S}_F(\mathbb{R})) = \lambda^1(\{p \in (0,1] : F^-(p) \in \mathbb{R}\}) = \lambda^1((0,1)) \\ &= \lambda^1\left(\bigcup_{n\in\mathbb{N}} (0, 1 - n^{-1}]\right) = \lambda^1\left((0,1/2] \cup (1/2,2/3] \cup (2/3,3/4] \cup \cdots\right) \\ &= 1/2 + (2/3 - 1/2) + (3/4 - 2/3) + \cdots = 1.\end{aligned}$$

■

Remark 2.6. *Note that*

$$\begin{aligned} P_F\left((-\infty,x]\right) &= \lambda^1\left(\mathcal{S}_F((-\infty,x])\right) = \lambda^1(\{p\in(0,1] : F^-(p)\in(-\infty,x]\}) \\ &= \lambda^1(\{p\in(0,1] : p\le F(x)\}) = \lambda^1((0,F(x)]) = F(x). \end{aligned}$$

2.6 Exercises

1. Let μ be a measure on $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ such that $\mu([-n,n))<\infty$ for all $n\in\mathbb{N}$. Define,

$$F_\mu(x) := \begin{cases} \mu([0,x)) & \text{if } x>0, \\ 0 & \text{if } x=0, \\ -\mu([x,0)) & \text{if } x<0. \end{cases}$$

   Show that $F_\mu : \mathbb{R}\to\mathbb{R}$ is monotonically increasing and left continuous.

2. Let F_μ be defined as in question 1 and let $\nu_{F_\mu}(([a,b)) = F_\mu(b)-F_\mu(a)$ for all $a\le b$, $a,b\in\mathbb{R}$. Show that ν_{F_μ} extends uniquely to a measure on $\mathcal{B}(\mathbb{R})$ and $\nu_{F_\mu}=\mu$.

3. If F is a distribution function, show that it can have an infinite number of jump discontinuities, but at most countably many.

4. Show that $\lambda^1((a,b)) = b-a$ for all $a,b\in\mathbb{R}$, $a\le b$. State and prove the same for λ^n.

5. Consider the measure space $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n),\lambda^n)$. Show that for every $B\in\mathcal{B}(\mathbb{R}^n)$ and $x\in\mathbb{R}^n$, $x+B\in\mathcal{B}(\mathbb{R}^n)$ and that $\lambda^n(x+B)=\lambda^n(B)$. Note: $x+B := \{z : z=x+b,\ b\in B\}$.