Chapter 2

Construction of probability measures

We have revealed a number of properties of measures, but we have not discussed their existence (in general) or how to construct them.

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

2. A class of subsets $\mathcal{D}$ of $\mathbb{X}$ 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) If $A_1, A_2, \dots \in \mathcal{D}$ and are pairwise disjoint, then $\cup_{i=1}^{\infty} A_i \in \mathcal{D}$.

It is clear from the definition that a σ-algebra associated with $\mathbb{X}$ is also a Dynkin system.

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

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 $\mathcal{C} \subseteq \sigma(\mathcal{C})$, $\delta(\mathcal{C}) \subseteq \delta(\sigma(\mathcal{C})) = \sigma(\mathcal{C})$ as in Theorem 1.3.

■

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

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

Proof. If $\mathcal{D}$ is a σ-algebra, $A, B \in \mathcal{D} \implies A \cap B = (A^c \cup B^c)^c \in \mathcal{D}$.

If $\mathcal{D}$ is a Dynkin system it satisfies requirements 1 and 2 for σ-algebras in Definition 1.1. Now let A_i, $i \in \mathbb{N}$ be such that $A_i \in \mathcal{D}$. We must show that $\cup_{i\in\mathbb{N}} A_i \in \mathcal{D}$. 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, $\cup_{i\in\mathbb{N}} B_i = \cup_{i\in\mathbb{N}} A_i \in \mathcal{D}$. ■

Another may to state the previous theorem is to say that a Dynkin system is a σ-algebra if, and only if, it is a π-system.

Theorem 2.3. *If $\mathcal{G} \subseteq 2^{\mathbb{X}}$ is a π-system, then $\delta(\mathcal{G}) = \sigma(\mathcal{G})$.*

Proof. From Theorem 2.1, $\delta(\mathcal{G}) \subseteq \sigma(\mathcal{G})$ and from Theorem 2.2 if $\delta(\mathcal{G})$ is a π-system it is a σ-algebra. Since $\sigma(\mathcal{G})$ is the smallest σ-algebra it must be that $\delta(\mathcal{G}) = \sigma(\mathcal{G})$, so it suffices to show that $\delta(\mathcal{G})$ is a π-system.

For any $D \in \delta(\mathcal{G})$, let $\mathcal{D}_D = \{A \subseteq \mathbb{X} : A \cap D \in \delta(\mathcal{G})\}$. First, we show that $\mathcal{D}_D$ is a Dynkin system. Note that $\mathbb{X} \cap D = D$, so $\mathbb{X} \in \mathcal{D}_D$. If $A \in \mathcal{D}_D$, then $A \cap D \in \delta(\mathcal{G})$. $A^c \cap D = (A^c \cup D^c) \cap D = (A \cap D)^c \cap D = ((A \cap D) \cup D^c)^c$ where $A \cap D$ and D^c are disjoint. Also, since $D \in \delta(\mathcal{G})$ so is D^c and $A \cap D \in \delta(\mathcal{G})$ by assumption, so $((A \cap D) \cup D^c)^c \in \delta(\mathcal{G})$. Thus $A^c \in \mathcal{D}_D$. Now, let A_i for $i \in \mathbb{N}$ be pairwise disjoint with $A_i \cap D \in \delta(\mathcal{G})$ and note that $\{(A_i \cap D)\}_{i\in\mathbb{N}}$ forms a disjoint collection. Thus $U = \cup_{i\in\mathbb{N}}(A_i \cap D) = D \cap \cup_{i\in\mathbb{N}} A_i$ and $\cup_{i\in\mathbb{N}} A_i \in \mathcal{D}_D$. Thus, $\mathcal{D}_D$ is a Dynkin system.

Now, since $\mathcal{G} \subseteq \delta(\mathcal{G})$ and since $\mathcal{G}$ is a π-system, $\mathcal{G} \subseteq \mathcal{D}_G$, for all $G \in \mathcal{G}$. To see this, note that for $G \in \mathcal{G}$ and $\mathcal{D}_G$ is as defined above, for any $G' \in \mathcal{G}$, we have $G' \cap G \in \mathcal{G}$, since $\mathcal{G}$ is a π-system. But since $\mathcal{G} \subseteq \delta(\mathcal{G})$, $G' \in \mathcal{D}_G$. But $\mathcal{D}_G$ is a Dynkin system and consequently, by definition $\delta(\mathcal{G}) \subseteq \mathcal{D}_G$, $\forall G \in \mathcal{G}$.

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

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 $(\mathbb{X}, \sigma(\mathcal{G}))$ be a measurable space and $\mathcal{G}$ a collection of subsets of $\mathbb{X}$, such that:*

1. *$\mathcal{G}$ is a π-system,*
2. *there exists $\{G_j\}_{j\in\mathbb{N}} \subseteq \mathcal{G}$ with $G_1 \subseteq G_2 \subseteq \dots$ such that $\cup_{j=1}^{\infty} G_j := \lim_{j\to\infty} G_j = \mathbb{X}$ (the sequence $\{G_j\}_{j\in\mathbb{N}}$ is exhausting).*

Then, if μ and v are measures that coincide on $\mathcal{G}$ and are finite for all G_j, $\mu(A) = v(A)$, for all $A \in \sigma(\mathcal{G})$.

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

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

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

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

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

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

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

Since $\mathcal{G}$ is a π-system, by Theorem 2.3 $\delta(\mathcal{G}) = \sigma(\mathcal{G})$ and $\mathcal{G} \subseteq \mathcal{D}_j$ by definition of $\delta(\mathcal{G})$, hence $\sigma(\mathcal{G}) \subseteq \mathcal{D}_j$. But by construction $\mathcal{D}_j \subseteq \sigma(\mathcal{G})$ and we conclude that $\mathcal{D}_j = \sigma(\mathcal{G})$. So, for all $A \in \sigma(\mathcal{G})$ and $j = 1, 2, , \ldots,$

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

By continuity of measures from below and noting that $(A_1 \cap G_1) \subseteq (A \cap G_2) \subseteq \cdots$, letting $j \to \infty$ in (2.1) we have for all $A \in \sigma(\mathcal{G})$,

$$\mu(A) = \lim_{j\to\infty} \mu(A \cap G_j) = \lim_{j\to\infty} v(A \cap G_j) = v(A).$$

■

We take the following proven 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 extension is known as Carathéodory's Extension Theorem. Before stating this theorem we need the following definition.

Definition 2.2. *A nonempty collection $\mathcal{S} \subseteq 2^{\mathbb{X}}$ of subsets of $\mathbb{X}$ is called a semi-ring if:*

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* $\{S_j\}_{j=1}^m \in \mathcal{S}$ *that is pairwise disjoint such that* $B - A = \cup_{j=1}^m S_j$.

Remark 2.1. *1.* $\mathcal{S}$ *is a* π*-system in view of condition 2.*

2. Condition 3 says that the difference between two sets in a semi-ring can be expressed as a finite partition of sets in the semi-ring.

3. A ring is a nonempty collection $\mathcal{R} \subseteq 2^{\mathbb{X}}$ *such that* $A, B \in \mathcal{R} \implies A \cup B \in \mathcal{R}$ *and* $A - B \in \mathcal{R}$.

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

5. 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* $A_1 \subseteq A$, $A_1, A \in \mathcal{R}$. $A = A_1 \cup (A - A_1) \in \mathcal{R}$, *so every ring is a semi-ring.*

It follows from these remarks that we have the following hierarchy of collections of algebras $\mathcal{A}$, rings $\mathcal{R}$ and semi-rings $\mathcal{S}$, viz., $\mathcal{A} \subseteq \mathcal{R} \subseteq \mathcal{S}$.

Definition 2.3. *(Alternative) A semi-ring* $\mathcal{S} \subseteq 2^{\mathbb{X}}$ *is a collection of subsets of* $\mathbb{X}$ *such that*

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

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

3.' *For all* $A, A_1 \in \mathcal{S} \implies A_1 \subseteq A$ *then* A *has the representation* $A = \cup_{j=1}^m A_j$ *for* $\{A_j\}$ *pairwise disjoint.*

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 $\{S_j\}_{j\in\mathbb{N}} \in \mathcal{S}$ with $S_1 \subseteq S_2 \ldots$ such that $S_j \to \mathbb{X}$ as $j \to \infty$ with $\mu(S_j) < \infty$, 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 $A \subseteq \mathbb{X}$ define the collection of countable covers for A that are composed of sets in $\mathcal{S}$ by

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

Since a semi-ring $\mathcal{S}$ does not necessarily contain $\mathbb{X}$, it is possible that $C(A) = \emptyset$. 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) $A \subseteq B$ implies that every cover for B is also covers A, i.e., $C(B) \subseteq C(A)$. Therefore,

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

c) Let $A_n \subseteq \mathbb{X}$ 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

$$\sum_{k=1}^{\infty} \mu(S_{nk}) \le \mu^*(A_n) + \epsilon/2^n.$$

Now, $\underset{n\in\mathbb{N}}{\cup} A_n \subseteq \underset{n\in\mathbb{N}}{\cup} \underset{k\in\mathbb{N}}{\cup} S_{nk}$ and by the definition of infimum

$$\begin{aligned} \mu^*\left(\underset{n\in\mathbb{N}}{\cup} A_n\right) &\le \sum_{n\in\mathbb{N}} \sum_{k\in\mathbb{N}} \mu(S_{nk}) \\ &\le \sum_{n=1}^{\infty} (\mu^*(A_n) + \epsilon/2^n) = \sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon. \end{aligned}$$

Hence, $\mu^*(\underset{n\in\mathbb{N}}{\cup} A_n) \le \sum_{n\in\mathbb{N}} \mu^*(A_n)$.

Since μ^* satisfies properties a)-c) it is called an outer-measure.

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 $\mathcal{S}_U = \{S : S = \cup_{j=1}^m S_j, S_j \in \mathcal{S} \text{ and } m \in \mathbb{N}\}$ 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 the image of $\bar{\mu}(S)$ is independent of the disjoint finite union used in its representation. 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 $S_j = S_j \cap (\cup_{k=1}^n T_k) = \cup_{k=1}^n (T_k \cap S_j)$ and $S_j \cap T_k \in \mathcal{S}$, since a semi-ring is a π-system and since μ is a pre-measure (additive) on $\mathcal{S}$. $\mu(S_j) = \sum_{k=1}^n \mu(T_k \cap S_j)$. Given that T_k are disjoint, then

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

This shows that the set function $\bar{\mu}$ can be unambiguously defined on $\mathcal{S}_U$. 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 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. Also, since, $S_j - T_k \in \mathcal{S}$

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

where the unions are over disjoint sets. Lastly, we conclude that

$$A \cup B = (A - B) \cup (A \cap B) \cup (B - A) \in \mathcal{S}_U$$

where all sets in the union are disjoint.

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

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

where the collection $\{S_{n_{(k-1)}+1,\cdots,S_{n_k}}\}$ is disjoint. Let $U_l = \underset{j\in J_l}{\cup} S_j$ for disjoint index sets $J_1, \ldots, J_N$ such that $\cup_{l=1}^N J_l = \mathbb{N}$, and note that $U_l \in \mathcal{S}$. Now, $T = \underset{k\in\mathbb{N}}{\cup} T_k = \cup_{l=1}^N U_l$ and

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

Now, observe that for all $S \in \mathcal{S}$ and all $\{S_j\}_{j\in\mathbb{N}} \in C(S)$

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

Taking the infimum over $C(S)$, we have $\mu(S) \leq \mu^*(S)$. Now, taking $(S, \emptyset, \ldots) \in C(S)$ 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 $\mathcal{S} \subseteq \mathcal{A}^*$ where

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

Let $S, T \in \mathcal{S}$. Since $\mathcal{S}$ is a semi-ring and $T = (T \cap S) \cup (T \cap S^c) = (T \cap S) \cup (T - S) = (T \cap S) \cup \{\cup_{j=1}^m S_j\}$ with $\{S_j\}_{j=1}^m$ disjoint. 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 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 $Q \subseteq \mathbb{X}$ 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 $Q \subseteq \cup_{j\in\mathbb{N}} T_j$ 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 $Q \subseteq \mathbb{X}$, $Q \cap \mathbb{X} = Q$ and $Q \cap \mathbb{X}^c = \emptyset$. Since $\mu^*(\emptyset) = 0$ we have that $\mathbb{X} \in \mathcal{A}^*$.

2. For all $Q \subseteq \mathbb{X}$ 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 $Q \subseteq \mathbb{X}$

$$\mu^*(Q \cap (A \cup A')) + \mu^*(Q - (A \cup A'))$$

$$= \mu^*(Q \cap (A \cup (A' - A))) + \mu^*(Q - (A \cup A'))$$

$$= \mu^*((Q \cap A) \cup [Q \cap (A' - A)]) + \mu^*(Q - (A \cup A'))$$

$$\leq \mu^*(Q \cap A) + \mu^*(Q \cap (A' - A)) + \mu^*(Q - (A \cup A'))$$

using subadditivity of μ^*.

$$= \mu^*(Q \cap A) + \mu^*((Q - A) \cap A') + \mu^*((Q - A) - A')$$

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

using the defining expression for $\mathcal{A}^*$ twice, once for $Q - A$ and once for Q.

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 $A \cap A' = \emptyset$, then the equality $[\mu^*(Q \cap A) + \mu^*(Q - A) = \mu^*(Q)]$, becomes for $Q := (A \cup A') \cap P$, $P \subseteq \mathbb{X}$ and

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

By induction for a disjoint collection $A_j \in \mathcal{A}^*$,

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

In particular, 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

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

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

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

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, A^* is closed under finite intersections, so by Theorem 2.2, $\mathcal{A}^*$ is a σ-algebra.

Now, we show that μ^* is a measure on $\sigma(\mathcal{S})$. From above, $\mathcal{S} \subseteq A^*$, so $\sigma(\mathcal{S}) \subseteq \mathcal{A}^*$. From above, μ^* is a measure on $\mathcal{A}^*$ and on $\sigma(\mathcal{S})$, which extends μ on $\mathcal{S}$. By Theorem 2.4, any two extensions μ^* and v^* of μ coincide on $\sigma(\mathcal{S})$. ■

Remark 2.2. *If $E \in \mathcal{A}^*$ and $\mu^*(E) = 0$ then, if $B \subseteq E$, we have $B \in \mathcal{A}^*$ and $\mu^*(B) = 0$. To see this, we first show that $\mu^*(B) = 0$, assuming that B is μ^* measurable. For all $B, E \subseteq \mathbb{X}$, such that $B \subseteq E$, $\mu^*(B) \leq \mu^*(E)$. Since, $\mu^*(E) = 0$, it must be that $\mu^*(B) = 0$. Let $Q \subseteq \mathbb{X}$ (Q arbitrary), we want to show that $\mu^*(Q) = \mu^*(Q \cap B) + \mu^*(Q \cap B^c)$. Now $Q \cap B \subseteq B \implies \mu^*(Q \cap B) \leq \mu^*(B) = 0$, thus $\mu^*(Q \cap B) = 0$. $Q \cap B^c \subseteq Q \implies \mu^*(Q \cap B^c) \leq \mu^*(Q)$. Hence, $\mu^*(Q) \geq \mu^*(Q \cap B^c) + \mu^*(Q \cap B)$. By sub-additivity $\mu^*(Q) \leq \mu^*(Q \cap B^c) + \mu(Q \cap B)$ and thus $\mu^*(Q) = \mu^*(Q \cap B^c) + \mu^*(Q \cap B)$ and $B \in \mathcal{A}^*$. Every subset of a μ^* measurable set of measure zero is μ^*-measurable, and has measure zero.*

In what follows we let $R_n = (a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n]$ be a half-open rectangle in $\mathbb{R}^n$ with $(a_i, b_i] = \emptyset$ if $b_i \leq a_i$ and $a_i, b_i \in \mathbb{R}$. Let $\mathcal{I}^n$ be the collection of all such rectangles R^n.

Theorem 2.6. *$\mathcal{I}^n$ for $n \in \mathbb{N}$ is a semi-ring.*

Proof. Let $\mathcal{I}^1 = \{(a_i, b_i] : a_i \leq b_i \text{ where } a_i,\ b_i \in \mathbb{R}\}$ and consider the following cases:

1. If $b_i = a_i$, $(a_i, b_i] = \emptyset$

2. $(a_i, b_i] \cap (a_j, b_j] = \begin{cases} \emptyset & \in \mathcal{I}^1 \\ (a_j, b_i] & \in \mathcal{I}^1 \\ (a_i, b_j] & \in \mathcal{I}^1 \\ (a_i, b_i] & \in \mathcal{I}^1 \end{cases}$

3. If $(a_1, b_1] \subseteq (a_2, b_2]$ 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, $\mathcal{I}^1$ is a semi-ring.

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

$$(R_n \times R_1) \cap (I_n \times I_1) = (R_n \cap I_n) \times (R_1 \cap I_1)$$

where the righthand side of the equality is an element of $\mathcal{I}^{n+1}$.

Now,

$$(R_n \times R_1) - (I_n \times I_1) = (R_n \times R_1) \cap (I_n \times I_1)^c.$$

Note that,

$$\begin{aligned}(I_n \times I_1)^c &= \{(x, y) : x \notin I_n, y \notin I_1, \text{ or } x \in I_n \text{ and } y \notin I_1, \text{ or } x \notin I_n \text{ and } y \in I_1\} \\ &= (I_n^c \times I_1^c) \cup (I_n \times I_1^c) \cup (I_n^c \times I_1)\end{aligned}$$

where the components of the union are disjoint. Thus,

$$\begin{aligned}(R_n \times R_1) - (I_n \times I_1) &= [(R_n \times R_1) \cap (I_n^c \times I_1^c)] \cup [(R_n \times R_1) \cap (I_n \times I_1^c)] \\ &\cup [(R_n \times R_1) \cap (I_n^c \times I_1)] \\ &= [(R_n - I_n) \times (R_1 - I_1)] \cup [(R_n \cap I_n) \times (R_1 - I_1)] \\ &\cup [(R_n - I_n) \times (R_1 \cap I_1)].\end{aligned}$$

By the induction assumption, $R_n - I_n$ and $R_1 - I_1$ can be expressed as finite unions of disjoint rectangles, which completes the proof. ■

Definition 2.4. *Let $\lambda^n : \mathcal{I}^n \to [0, \infty)$ be defined as $\lambda^n(R_n) = \prod_{j=1}^{n}(b_j - a_j)$ whenever $b_j > a_j$ for $j = 1, \cdots, n$ and $\lambda^n(R_n) = 0$ if $b_j \leq a_j$ for some j.*

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

Proof. We need only verify conditions a), b) and c") from Theorem 1.6 and Remark 1.4. First, let's consider $n = 1$. a) If $b \leq a$ then $\lambda^1(\emptyset) = 0$. b) Let $I = (a, b]$ and choose $a < c < b$. Then, $I = (a, c] \cup (c, b]$.

$$\lambda^1((a, c]) + \lambda^1((c, b]) = (c - a) + (b - c) = b - a = \lambda^1(I).$$

c") Let $\{I_j = (a_j, b_j]\}_{j \in \mathbb{N}}$ be such that $(a_1, b_1] \supseteq (a_2, b_2] \supseteq \cdots$ and $I_j \downarrow \emptyset$ as $j \to \infty$ or equivalently, $\cap_{j=1}^{\infty}(a_j, b_j] = \emptyset$. Thus, we must have $\lim_{j\to\infty}(b_j - a_j) = 0$, that is $\lim_{j\to\infty} \lambda^1(I_j) = 0$.

Now, let's consider $n = 2$. a) If $b_i \leq a_i$ for $i = 1$ or $i = 2$ then $\lambda^2(\emptyset) = 0$. b) Let $I = (a_1, b_1] \times (a_2, b_2]$ and choose $a_2 < c_2 < b_2$. Then, $I = ((a_2, c_2] \times (a_1, b_1]) \cup ((c_2, b_2] \times (a_1, b_1])$ and the component sets of the union are disjoint.

$$\begin{aligned}\lambda^2((a_2, c_2] \times (a_1, b_1]) + \lambda^1((c_2, b_2] \times (a_1, b_1]) &= (c_2 - a_2)(b_1 - a_1) + (b_2 - c_2)(b_1 - a_1) \\ &= (b_1 - a_1)(b_2 - a_2) = \lambda^2(I).\end{aligned}$$

c) Let $\{I_j = (a_j, b_j] \times (\alpha_j, \beta_j]\}_{j \in \mathbb{N}}$ be such that $I_1 \supseteq I_2 \supseteq \cdots$ and $I_j \downarrow \emptyset$ as $j \to \infty$ or equivalently, $\cap_{j=1}^{\infty} I_j = \emptyset$. Thus, we must have either $\lim_{j\to\infty}(b_j - a_j) = 0$ or $\lim_{j\to\infty}(\beta_j - \alpha_j) = 0$. Thus,

$$\lambda^2(I_j) = (\beta_j - \alpha_j)(b_j - a_j) \to 0 \text{ as } j \to \infty.$$

The cases for $n = 3, 4, \cdots$ follow similarly. ■

Theorem 2.8. *There exists a unique extension of λ^n from $\mathcal{I}^n$ to a mesure 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 $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{I}^n)$. Since, $(-k, k]^n = \times_1^n(-k, k] \uparrow \mathbb{R}^n$ is an exhausting sequence of n-cubes (intervals) and since $\lambda^n((-k, k]^n) = (2k)^n < \infty$, all conditions of Carathéodory's Theorem are fulfilled. ■