

2 Set Functions

Notation Let $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, +\infty\}$ and $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\} \cup \{+\infty\}$. Here $+\infty$ and $-\infty$ are symbols satisfying obvious conditions:

$$\begin{aligned} & \text{for any real number } x \in \mathbb{R} : -\infty < x < +\infty, \\ & (\pm\infty) + (\pm\infty) = x + (\pm\infty) = (\pm\infty) + x = \pm\infty, \\ & (\pm\infty)(\pm\infty) = +\infty \quad \text{and} \quad (\pm\infty)(\mp\infty) = -\infty, \\ & (\pm\infty)x = x(\pm\infty) = \begin{cases} \pm\infty & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ \mp\infty & \text{if } x < 0, \end{cases} \\ & \frac{x}{+\infty} = \frac{x}{-\infty} = 0. \end{aligned}$$

We see that $(+\infty) + (-\infty)$ is not defined. Also $\frac{1}{0}$ is still not defined.

Example 6 Let I^* be the collection of all intervals of types (a, b) , $(a, b]$, $[a, b)$ and $[a, b] \subseteq \mathbb{R}^*$. Possible set functions are

- (i) $\ell : I^* \rightarrow \mathbb{R}^+$, $\ell(a, b) = b - a$,
- (ii) If $f : \mathbb{R} \rightarrow \mathbb{R}$ is a non-negative Riemann integrable function then define $F : I^* \rightarrow \mathbb{R}^+$ by $F((a, b)) = \int_a^b f(t)dt$.

***Definition** Let \mathcal{A} be a collection of subsets of a non-empty set X . We say that the extended real-valued set function $\phi : \mathcal{A} \rightarrow \mathbb{R}^*$ is

non-negative if $\phi(A) \geq 0$ for all $A \in \mathcal{A}$,

non-positive if $\phi(A) \leq 0$ for all $A \in \mathcal{A}$,

finite at A if $|\phi(A)| < \infty$,

finite if $|\phi(A)| < \infty$ for all $A \in \mathcal{A}$,

σ -finite if for all $A \in \mathcal{A}$ there exist a countable collection $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$ such that $A \subseteq \bigcup_{n \geq 1} A_n$ and $|\phi(A_n)| < \infty$ for all $n \geq 1$.

2.1 (Finitely) additive functions

Definition $\phi : \mathcal{A} \rightarrow \mathbb{R}^*$ is *finitely additive* if

- (i) $\phi(\emptyset) = 0$,
- (ii) if $\{A_n\}_{1 \leq n \leq N} \subseteq \mathcal{A}$ is a finite collection of disjoint sets **and** $\bigcup_{n=1}^N A_n \in \mathcal{A}$ then

$$\phi\left(\bigcup_{n=1}^N A_n\right) = \sum_{n=1}^N \phi(A_n).$$

Example 7 Let $X = (0, 1]$, and $\mathcal{A} = \{(a, b] : 0 \leq a < b \leq 1\}$. Define

$$\mu((a, b]) = \begin{cases} b - a & \text{if } a \neq 0 \\ +\infty & \text{if } a = 0. \end{cases}$$

This is an additive function. Verification is left to the student.

2.2 σ -additive set functions.

Definition $\mu : \mathcal{A} \rightarrow \mathbb{R}^*$ is σ -additive if

- (i) $\mu(\emptyset) = 0$,
- (ii) If $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$ is a countable collection of disjoint sets **and** $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Notes (A) Since all but a finite number of the sets A_n in (ii) may be empty the definition for σ -additive contains that for additive, that is, σ -additive implies additive.

(B) If we look at Example 7, take $A_n = (\frac{1}{n+1}, \frac{1}{n}]$ for each $n \geq 1$. Then

$$\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1$$

while

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu((0, 1]) = +\infty.$$

Hence being additive does not imply being σ -additive.

(C) If \mathcal{A} is a σ -ring then the definition is simplified since necessarily we will have $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

(D) If \mathcal{A} is a σ -field then μ cannot take both values $+\infty$ and $-\infty$.

***Proof of (D)** Assume otherwise so there exist $F, G \in \mathcal{A}$ such that $\mu(F) = +\infty$ and $\mu(G) = -\infty$.

Since \mathcal{A} is a σ -field we have $X \in \mathcal{A}$ and so $\mu(X)$ should be defined. But we can decompose X in two ways. In the first $X = F \cup F^c$ and so $\mu(X) = \mu(F) + \mu(F^c)$. We cannot have $\mu(F^c) = -\infty$ since $(+\infty) + (-\infty)$ is not defined. So whatever the value for $\mu(F^c)$ we have $\mu(X) = +\infty$. Yet we also have $X = G \cup G^c$ and so $\mu(X) = \mu(G) + \mu(G^c)$. This time we cannot have $\mu(G^c) = +\infty$ and so $\mu(X) = -\infty$. This is the required contradiction. ■

Definition A map $F : \mathbb{R} \rightarrow \mathbb{R}$ is a *distribution function* if F is monotonically increasing, so if $x < y$ then $F(x) \leq F(y)$, and right continuous, so $\lim_{x \rightarrow x_0+} F(x) = F(x_0)$. (*It might well be that $\lim_{x \rightarrow x_0-} F(x) \neq F(x_0)$.)

A limited collection of such F are given by $F(x) = \int_0^x f(t)dt$, where $f \geq 0$. (c.f. example 6(ii)).

Notation If $F : \mathbb{R} \rightarrow \mathbb{R}$ is a distribution function then define the set function $\mu_F : \mathcal{P} \rightarrow \mathbb{R}^+$ by

$$\mu_F((a, b]) = F(b) - F(a).$$

Note that if $b \geq a$ then $F(b) \geq F(a)$ since F is increasing, and so $\mu_F((a, b]) \geq 0$. That is, μ_F is a non-negative set function.

The next result is important.

Theorem 2.1

The set function μ_F is σ -additive on \mathcal{P} .

Proof is in two steps.

(A) μ_F is additive on \mathcal{P} .

Let $(a_i, b_i] \in \mathcal{P}, 1 \leq i \leq n$ be disjoint sets for which $\bigcup_{i=1}^n (a_i, b_i] \in \mathcal{P}$. Then $\bigcup_{i=1}^n (a_i, b_i] = (a, b]$ for some a and b . Since we have a **finite** collection of intervals we can assume they are ordered such that

$$a = a_1 < b_1 = a_2 < b_2 = a_3 < \dots < b_n = b.$$

(This is not necessarily possible for an infinite collection.) Then, using $-F(a_j) + F(b_{j-1}) = 0$ for $2 \leq j \leq n$, we find

$$\begin{aligned}
\mu_F \left(\bigcup_{i=1}^n (a_i, b_i] \right) &= \mu_F(a, b] = F(b) - F(a) \\
&= F(b_n) + (-F(a_n) + F(b_{n-1})) \\
&\quad + (-F(a_{n-1}) + F(b_{n-2})) + \dots \\
&\quad \dots + (-F(a_2) + F(b_1)) - F(a_1) \\
&= (F(b_n) - F(a_n)) + (F(b_{n-1}) - F(a_{n-1})) \\
&\quad + (F(b_{n-2}) - F(a_{n-2})) + \dots \\
&\quad \dots - F(a_2) + (F(b_1) - F(a_1)) \\
&= \sum_{i=1}^n \mu_F((a_i, b_i]).
\end{aligned}$$

(B) μ_F is σ -additive on \mathcal{P} .

Let $\{(a_i, b_i]\}_{i \geq 1}$ be a countable collection of disjoint intervals such that $\bigcup_{i=1}^{\infty} (a_i, b_i] \in \mathcal{P}$. Then $\bigcup_{i=1}^{\infty} (a_i, b_i] = (a, b]$ for some a and b . We wish to prove

$$\mu_F(a, b] = \sum_{i=1}^{\infty} \mu_F(a_i, b_i].$$

The idea of the proof is to use the fact that μ_F is additive on \mathcal{P} . This means that we have to replace the infinite union by a finite one. This can be done easily by throwing away all but a finite number of the intervals in the union. Unsurprisingly, this introduces an error and we can only prove an inequality. To get the inequality in the other direction we have to work harder and use another result that allows us to replace an infinite union by a finite one. Such a result is the Heine-Borel Theorem, Theorem 1.3. But that result deals with open intervals covering a closed interval while we have open-closed intervals covering an open-closed interval. We will have to “alter” our intervals so they are of a form to which we can apply the Heine-Borel Theorem.

(B1) Let $N \geq 1$ be given. Then $(a, b] \supseteq \bigcup_{i=1}^N (a_i, b_i]$. Since this is a finite union relabel so that $a_1 < a_2 < \dots < a_N$. We then decompose $(a, b]$ as

$$(a, b] = (a, a_1] \cup \bigcup_{i=1}^N (a_i, b_i] \cup \bigcup_{i=1}^{N-1} (b_i, a_{i+1}] \cup (b_N, b].$$

Since μ_F is additive on \mathcal{P} , by part (A), we have

$$\mu_F(a, b] = \mu_F(a, a_1] + \sum_{i=1}^N \mu_F(a_i, b_i] + \sum_{i=1}^{N-1} \mu_F(b_i, a_{i+1}] + \mu_F(b_N, b].$$

On throwing away a number of these terms, we get

$$\mu_F(a, b] \geq \sum_{i=1}^N \mu_F(a_i, b_i].$$

Thus we have a sequence of partial sums bounded above. This sequence is increasing because, as we increase N we are adding non-negative terms, $\mu_F(a_i, b_i]$, to the partial sums. An increasing sequence bounded above converges. Hence the infinite series converges and its sum satisfies

$$\sum_{i=1}^{\infty} \mu_F(a_i, b_i] \leq \mu_F(a, b].$$

(B2) Let $\varepsilon > 0$ be given such that $a + \varepsilon < b$. Recall that the distribution function F is right continuous. This means, that at a point b_i , if we move to the right the function does not change its value by much. That is, we can find $b'_i > b_i$ for which

$$F(b_i) \leq F(b'_i) < F(b_i) + \frac{\varepsilon}{2^i}. \quad (1)$$

(Note the factor of $1/2^i$. This is a “Trick”. Later we will be adding together infinitely many of these terms and without this factor such a sum would diverge, however small ε .) Consider

$$[a + \varepsilon, b] \subseteq (a, b] = \bigcup_{i=1}^{\infty} (a_i, b_i] \subseteq \bigcup_{i=1}^{\infty} (a_i, b'_i).$$

So we now have a closed interval covered by a countable collection of open intervals. This is exactly the situation in which we can use the Heine-Borel Theorem. Thus we can conclude that we have a finite sub-cover of $[a + \varepsilon, b]$. Let M be the largest i occurring in this subcover. Then

$$[a + \varepsilon, b] \subseteq \bigcup_{i=1}^M (a_i, b'_i). \quad (2)$$

Though we have a finite union the intervals are not in \mathcal{P} and they are not disjoint. Nonetheless, since we have a finite union we can relabel such that

$a \leq a_1 < a + \varepsilon$ and $a_1 < a_2 < a_3 \dots < a_M < b$. Also, since the open intervals in (2) must overlap we have $b'_i > a_{i+1}$ for all $1 \leq i \leq M - 1$. Using the sequence $\{a_i\}$ we can decompose $(a_1, b]$ as

$$(a_1, b] = \bigcup_{i=1}^{M-1} (a_i, a_{i+1}] \cup (a_M, b].$$

So, since μ_F is additive, and this is a finite disjoint union, we get

$$\mu_F(a_1, b] = \sum_{i=1}^{M-1} \mu_F(a_i, a_{i+1}] + \mu_F(a_M, b].$$

But

$$\begin{aligned} \mu_F(a_i, a_{i+1}] &= F(a_{i+1}) - F(a_i) && \text{by definition} \\ &\leq F(b'_i) - F(a_i) && \text{since } a_{i+1} < b'_i \\ &< F(b_i) - F(a_i) + \frac{\varepsilon}{2^i} && \text{by definition of } b'_i \\ &= \mu_F(a_i, b_i] + \frac{\varepsilon}{2^i}. \end{aligned}$$

Similarly $\mu_F(a_M, b] \leq \mu_F(a_M, b'_M] < \mu_F(a_M, b_M] + \frac{\varepsilon}{2^M}$. Hence

$$\begin{aligned} \mu_F(a_1, b] &< \sum_{i=1}^{M-1} \left(\mu_F(a_i, b_i] + \frac{\varepsilon}{2^i} \right) + \mu_F(a_M, b_M] + \frac{\varepsilon}{2^M} \\ &< \sum_{i=1}^{\infty} \mu_F(a_i, b_i] + \varepsilon. \end{aligned} \tag{3}$$

We now see the importance of weighting the ε in (1) by $1/2^i$. From above we have that $a \leq a_1 < a + \varepsilon$. So as $\varepsilon \rightarrow 0$ we have that $a_1 \rightarrow a$ and, again since F is right continuous, $F(a_1) \rightarrow F(a)$. In particular this means that as $\varepsilon \rightarrow 0$ we have $\mu_F(a_1, b] \rightarrow \mu_F(a, b]$. So letting $\varepsilon \rightarrow 0$ in (3) we obtain

$$\mu_F(a, b] \leq \sum_{i=1}^{\infty} \mu_F(a_i, b_i].$$

Finally, combining parts (A) and (B) we get our result. ■

2.3 Extending a σ -additive function

Theorem 2.1 has given us an important example of a σ -additive set function defined on a semi-ring of subsets of \mathbb{R} . Our aim is to enlarge the collection

of sets on which the set function is defined and this is what we achieve in the next few sections. We do this in general but then apply the general results, in corollaries and examples, to the specific case of μ_F . To highlight the difference between the general results and specific, I will mark the specific results with a †.

Assume that \mathcal{C} is a semi-ring.

Theorem 2.2 *If $\mu : \mathcal{C} \rightarrow \mathbb{R}^+$ is a non-negative σ -additive set function on the semi-ring \mathcal{C} , then there is a unique σ -additive set function ν on $\mathcal{R}(\mathcal{C})$ such that $\nu = \mu$ on \mathcal{C} . (So we say ν extends μ .) Further ν is non-negative.*

Proof

Let $A \in \mathcal{R}(\mathcal{C})$ so, by Theorem 1.7, $A = \bigcup_{k=1}^n E_k$ for some disjoint sets $A_k \in \mathcal{C}$. Define

$$\nu(A) = \sum_{k=1}^n \mu(E_k).$$

We need first show that this is well defined. So assume that we have another decomposition of A as $A = \bigcup_{j=1}^m F_j$, a disjoint union of $F_j \in \mathcal{C}$. Put $H_{kj} = E_k \cap F_j$. Since \mathcal{C} is a semi-ring we have that $H_{kj} \in \mathcal{C}$ for all k and j .

Trivially, $E_k \subseteq A = \bigcup_{j=1}^m F_j$ but this means that

$$E_k = E_k \cap A = E_k \cap \left(\bigcup_{j=1}^m F_j \right) = \bigcup_{j=1}^m H_{kj}.$$

Similarly $F_j = \bigcup_{k=1}^n H_{kj}$. Then

$$\begin{aligned} \sum_{k=1}^n \mu(E_k) &= \sum_{k=1}^n \sum_{j=1}^m \mu(H_{kj}) && \text{since } \mu \text{ is additive on } \mathcal{C}, \\ &= \sum_{j=1}^m \sum_{k=1}^n \mu(H_{kj}) && \text{interchanging sums,} \\ &= \sum_{j=1}^m \mu(F_j) && \text{since } \mu \text{ is additive on } \mathcal{C}. \end{aligned}$$

So the result is independent of the decomposition and ν is well-defined.

To show it is σ -additive take any collection $\{E_k\}_{k \geq 1} \subseteq \mathcal{R}$ of distinct sets for which $\bigcup_{k \geq 1} E_k \in \mathcal{R}$. Let $E = \bigcup_{k \geq 1} E_k$. Then by Theorem 1.7 $E \in \mathcal{R}$ means that $E = \bigcup_{r=1}^n A_r$ for disjoint sets $A_r \in \mathcal{R}$. Similarly, by Theorem 1.7 again, $E_k \in \mathcal{R}$ means that each $E_k = \bigcup_{i=1}^{m_k} B_{ki}$ for disjoint sets $B_{ki} \in \mathcal{R}$.

Put $C_{rki} = A_r \cap B_{ki}$, a collection of disjoint sets from \mathcal{C} (having used the fact that \mathcal{C} is a semi-ring.) Note that

$$A_r = \bigcup_{k \geq 1} \bigcup_{i=1}^{m_k} C_{rki} \quad \text{and} \quad B_{ki} = \bigcup_{r=1}^n C_{rki}.$$

Then

$$\begin{aligned} v \left(\bigcup_{k \geq 1} E_k \right) &= v \left(\bigcup_{r=1}^n A_r \right) \\ &= \sum_{r=1}^n \mu(A_r) && \text{by definition of } v, \\ &= \sum_{r=1}^n \mu \left(\bigcup_{k \geq 1} \bigcup_{i=1}^{m_k} C_{rki} \right) \\ &= \sum_{r=1}^n \sum_{k=1}^{\infty} \sum_{i=1}^{m_k} \mu(C_{rki}) && \text{since } \mu \text{ is } \sigma\text{-additive on } \mathcal{C}, \\ &= \sum_{k=1}^{\infty} \sum_{i=1}^{m_k} \sum_{r=1}^n \mu(C_{rki}) && \text{interchange allowed since } \mu \geq 0, \\ &= \sum_{k=1}^{\infty} \sum_{i=1}^{m_k} \mu(B_{ki}) && \text{since } \mu \text{ is additive on } \mathcal{C}, \\ &= \sum_{k=1}^{\infty} v(E_k) && \text{by definition of } v. \end{aligned}$$

Hence v is σ -additive on \mathcal{R} .

Let τ be any other extension of μ from \mathcal{C} to \mathcal{R} . Then for $A \in \mathcal{R}$ decomposed as $A = \bigcup_{i=1}^n E_i$, with $E_i \in \mathcal{C}$, we have

$$\begin{aligned} \tau(A) &= \sum_{i=1}^n \tau(E_i) && \text{since } \tau \text{ is additive on } \mathcal{R}, \\ &= \sum_{i=1}^n \mu(E_i) && \text{since } \tau = \mu \text{ on } \mathcal{C}, \\ &= v(A) && \text{by definition of } v. \end{aligned}$$

Hence v is unique. ■

Corollary 2.3† *The μ_F on \mathcal{P} of Theorem 2.1 can be extended to a σ -additive function on \mathcal{E} that we still denote by μ_F .*

Proof By Corollary 1.8 \mathcal{E} is the ring generated by \mathcal{P} . ■

2.4 Measure and Outer Measure.

Definition Any non-negative set function $\mu : \mathcal{C} \rightarrow \mathbb{R}^+$ which is σ -additive is a *measure* on \mathcal{C} .

Note Assume that μ is a measure on a ring \mathcal{R} and $E, F \in \mathcal{R}$ with $E \subseteq F$. Then $F = E \cup (F \setminus E)$, a disjoint union with $F \setminus E \in \mathcal{R}$. So $\mu(F) = \mu(E) + \mu(F \setminus E) \geq \mu(E)$ since $\mu \geq 0$. Hence $E \subseteq F$ implies $\mu(E) \leq \mu(F)$. We say that μ is *monotonic*.

Definition If \mathcal{C} is the collection of all subsets of X then $\lambda : \mathcal{C} \rightarrow \mathbb{R}^+$ is an *outer measure on X* if

- (i) $\lambda(\phi) = 0$,
- (ii) (monotone) If $E \subseteq F$ then $\lambda(E) \leq \lambda(F)$,
- (iii) (Countable subadditivity) For any countable collection of subsets $\{A_n\}$ we have

$$\lambda \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \lambda(A_n).$$

***Note** that in (iii) we do not require the A_n to be disjoint. We will see later that, for some outer measures, even if the A_n are disjoint we don't necessarily have equality.

The main result here is

Theorem 2.4

Let \mathcal{R} be a ring of subsets of X such that $X = \bigcup_{i=1}^{\infty} E_i$ for some $E_i \in \mathcal{R}$. Let μ be a measure on \mathcal{R} . For $A \subseteq X$ define

$$\mu^*(A) = \inf \sum_{j=1}^{\infty} \mu(A_j) \tag{4}$$

where the inf is over all covers of $A \subseteq \bigcup_{j=1}^{\infty} A_j$ by sets $A_j \in \mathcal{R}$. Then μ^* is an outer measure on X .

Proof

We need show that the conditions of the definition of an outer measure are satisfied.

Since μ is a measure we have that $\mu(A_j) \geq 0$ for all j and so $\mu^*(A) \geq 0$ for all A , that is $\mu^* : \mathcal{C} \rightarrow \mathbb{R}^+$ where \mathcal{C} is the collection of all subsets of X .

- (i) The empty set ϕ is in \mathcal{R} so it is covered by itself and thus $\mu^*(\phi) \leq \mu(\phi) = 0$. Since we have seen that $\mu^* \geq 0$ we can conclude that $\mu^*(\phi) = 0$.

(ii) If $E \subseteq F$ then every cover of F is a cover of E . So the set of values of $\sum_{j=1}^{\infty} \mu(A_j)$ for covers of F is contained in the set of values for covers of E . In general, if $A \subseteq B \subseteq \mathbb{R}$ then $\inf B \leq \inf A$. In our case this means that $\mu^*(E) \leq \mu^*(F)$.

(iii) Let any countable collection of sets, $\{A_n\}$, be given. If $\mu^*(A_n) = +\infty$ for some n then $\sum_{j=1}^{\infty} \mu^*(A_j) = +\infty$ since $\mu^* \geq 0$. The result is then trivial.

Assume that $\mu^*(A_n) < +\infty$ for all n . Let $\varepsilon > 0$ be given. Then by the definition of infimum or equivalently, the least upper bound, we can find a cover of $A_n \subseteq \bigcup_{i=1}^{\infty} G_{ni}$, with $G_{ni} \in \mathcal{R}$ such that

$$\mu^*(A_n) \leq \sum_{i=1}^{\infty} \mu(G_{ni}) < \mu^*(A_n) + \frac{\varepsilon}{2^n}. \quad (5)$$

(Note how we have used the same “trick” as in the proof of Theorem 2.1 of weighting the errors in (5) by $1/2^n$.) But then

$$\begin{aligned} \sum_{n=1}^{\infty} \mu^*(A_n) &\geq \sum_{n=1}^{\infty} \left(\sum_{i=1}^{\infty} \mu(G_{ni}) - \frac{\varepsilon}{2^n} \right) \\ &\geq \mu^* \left(\bigcup_{n=1}^{\infty} A_n \right) - \varepsilon \end{aligned} \quad (6)$$

having used the definition of μ^* and the fact that $\bigcup_{n=1}^{\infty} \bigcup_{i=1}^{\infty} G_{ni}$ is a cover for $\bigcup_{n=1}^{\infty} A_n$. Since (6) is true for all $\varepsilon > 0$ we deduce countable subadditivity. ■

Corollary 2.5† *From μ_F on \mathcal{E} we can define μ_F^* the Lebesgue-Stieltjes outer measure on \mathbb{R} .*