Measure Theory

Felix Chen

Contents

§0.1 Generation of σ-algebras

Let $\mathscr{E}$ be a nonempty collection of sets.

Definition 0.1 (Generate rings)**.** We say $\mathscr{G}$ is the ring (algebra, etc.) generated by $\mathscr{E}$, if

- $\mathscr{G} \supseteq \mathscr{E}$;
- For any ring $\mathscr{G}'$, $\mathscr{G}' \supseteq \mathscr{E} \implies \mathscr{G}' \supseteq \mathscr{G}$

Proposition 0.2

The ring (or whatever) generated by $\mathscr{E}$ always exists.

Proof. Let $\mathbf{A}$ be the set consisting of the rings containing $\mathscr{E}$, then $\bigcap_{\mathscr{G}\in\mathbf{A}} \mathscr{G}$ is the desired ring. □

Denote $r(\mathscr{E}), m(\mathscr{E}), p(\mathscr{E}), l(\mathscr{E}), \sigma(\mathscr{E})$ the ring/monotone class/π-system/λ-system/σ-algebra generated by $\mathscr{E}$.

Theorem 0.3

Let $\mathscr{A}$ be an algebra, then $\sigma(\mathscr{A}) = m(\mathscr{A})$.

Proof. Clearly $\sigma(\mathscr{A}) \supseteq m(\mathscr{A})$.

On the other hand, we only need to prove $m(\mathscr{A})$ is a σ-algebra.

Since $\mathscr{A}$ is an algebra, so $X \in \mathscr{A} \subset m(\mathscr{A})$.

For the completion:

Let $\mathscr{G} := \{A : A^c \in m(\mathscr{A})\}$, we want to prove $\mathscr{G} \supseteq m(\mathscr{A})$.

Clearly $\mathscr{A} \subset \mathscr{G}$; If $A_1, A_2, \cdots \in \mathscr{G}$, $A_n \uparrow A$, then

$$A_n^c \in m(\mathscr{A}) \implies A^c =\downarrow \lim_n A_n^c \in m(\mathscr{A}).$$

Similarly if $A_n \downarrow A$, we can also deduce $A^c \in m(\mathscr{A})$.

So $\mathscr{G}$ is a monotone class containing $\mathscr{A}$, hence it must contain $m(\mathscr{A}) \implies \forall A \in m(\mathscr{A})$, $A^c \in m(\mathscr{A})$.

For the intersection:

- $\forall A \in \mathscr{A}, B \in m(\mathscr{A}), AB \in m(\mathscr{A})$: If $B \in \mathscr{A}$, this clearly holds;

  Moreover, such B's constitude a monotone class:

  > **Claim 0.4.** Let $\mathscr{M}$ be a monotone class, then $\forall C \in \mathscr{M}$, $\mathscr{G}_C = \{D : CD \in \mathscr{M}\}$ is a monotone class.

  If $D_1, D_2, \cdots \to D$ satisfy $C \cap D_i \in m(\mathscr{A})$, then $D \cap C = \lim_n D_i \cap C \in \mathscr{M}$.

  Therefore such B's constitude a monotone class $\mathscr{G}_A$ containing $\mathscr{A} \implies \mathscr{G}_A \supseteq m(\mathscr{A})$.

- All the A's which satisfies the first condition constitude a monotone class:

  Let $\mathscr{G}_B = \{A : AB \in m(\mathscr{A})\}$, then $\mathscr{G} = \bigcup_{B \in m(\mathscr{A})} \mathscr{G}_B$ is a monotone class containing $\mathscr{A}$.

  Hence $\mathscr{G} \supseteq m(\mathscr{A}) \implies \forall A \in m(\mathscr{A})$, $\forall B \in m(\mathscr{A})$, we have $AB \in m(\mathscr{A})$.

□

Theorem 0.5 (λ-π theorem)

Let $\mathscr{P}$ be a π-system, then $\sigma(\mathscr{P}) = l(\mathscr{P})$.

Proof. Obviously $\sigma(\mathscr{P}) \supseteq l(\mathscr{P})$.

We only need to check that $l(\mathscr{P})$ is a π-system, i.e. closed under intersection.

> **Claim 0.6.** If $\mathscr{L}$ is a λ-system, then $\forall C \in \mathscr{L}$, $\mathscr{G}_C$ is a λ-system, where
> $$\mathscr{G}_C := \{D : CD \in \mathscr{L}\}.$$

Proof of the claim. First of all, $X \in \mathscr{G}_C$ as $CX = C \in \mathscr{G}_C$.

Second, if $D_1, D_2 \in \mathscr{G}_C$,

$$CD_1, CD_2 \in \mathscr{L} \implies C(D_1 - D_2) = CD_1 - CD_2 \in \mathscr{L} \implies D_1 - D_2 \in \mathscr{G}_C.$$

Lastly, if $D_n \in \mathscr{G}_C$, $D_n \to D$,

$$CD_n \in \mathscr{L} \implies CD = \lim_n CD_n \in \mathscr{L} \implies D \in \mathscr{G}_C$$

□

The rest is similar to the previous theorem:

- $\forall A \in \mathscr{P}, B \in l(\mathscr{P}), AB \in l(\mathscr{P})$: If $B \in \mathscr{P}$ this clearly holds;

  By the claim, $\mathscr{G}_A = \{B : AB \in l(\mathscr{P})\}$ is a λ-system, so $\mathscr{G}_A \supseteq l(\mathscr{P})$.

- For $B \in l(\mathscr{P})$, let
  $$\mathscr{G}_B = \{A : AB \in l(\mathscr{P})\}.$$
  By our claim, $\mathscr{G}_B$'s are λ-systems. So $\mathscr{G} = \bigcap_{B \in l(\mathscr{P})} \mathscr{G}_B$ is a λ-system.

  Moreover $\mathscr{G} \supseteq \mathscr{P}$ (This is proved above), so $\mathscr{G} \supseteq l(\mathscr{P})$.

  This means $\forall A, B \in l(\mathscr{P}), AB \in l(\mathscr{P})$.

□

Remark 0.7 — These two proofs are very similar. Note how we make use of the conditions.

Let X be a topological space, $\mathscr{O}$ is the collection of all the open sets.

Let $\mathscr{B}_X := \sigma(\mathscr{O})$ be the **Borel σ-algebra** on the space X, $B \in \mathscr{B}_X$ are called **Borel sets**, and $(X, \mathscr{B}_X)$ is called the **topological measurable space**.

Theorem 0.8

Let $\mathscr{Q}$ be a semi-ring, then

$$r(\mathscr{Q}) = \mathscr{G} := \bigcup_{n=1}^{\infty} \left\{ \sum_{k=1}^{n} A_k : A_1, \ldots, A_n \in \mathscr{Q} \text{ and pairwise disjoint} \right\}.$$

Remark 0.9 — If $\mathscr{R}$ is a ring, then $\mathscr{A} = a(\mathscr{R}) = \mathscr{R} \cup \{A^c : A \in \mathscr{R}\}$ can also be written out explicitly, while $\sigma(\mathscr{A})$ usually cannot be expressed explicitly.

Proof. Since $r(\mathscr{Q})$ is closed under finite unions, so $r(\mathscr{Q}) \supseteq \mathscr{G}$.

Reversely, $\mathscr{G}$ is nonempty. In fact we only need to prove it's closed under subtraction, as

$$A \cap B = A \backslash (A \backslash B) \in \mathscr{G}, A \cup B = (A \backslash B) \cup (B \backslash A) \cup (A \cap B) \in \mathscr{G}.$$

Suppose $A = \sum_{i=1}^{n} A_i, B = \sum_{j=1}^{m} B_j$.

Then $A_i \backslash B_1$ can be split to several disjoint sets C_k in $\mathscr{Q}$. Continue this process, each C_k can be split again into smaller set. When all of the B_j's are removed, we end up with many tiny sets which are in $\mathscr{Q}$ and pairwise disjoint. (This process can be formalized using induction)

Therefore $A \backslash B \in \mathscr{G}$, the conclusion follows. $\square$

§0.2 Measurable maps and measurable functions

For a map $f : X \to Y$, we say the **preimage** of $B \subset Y$ is $f^{-1}(B) := \{x : f(x) \in B\}$.

Some properties of preimage:

$$f^{-1}(\emptyset) = \emptyset, \quad f^{-1}(Y) = X;$$
$$B_1 \subset B_2 \implies f^{-1}(B_1) \subset f^{-1}(B_2), \quad (f^{-1}(B))^c = f^{-1}(B^c);$$
$$f^{-1}\left(\bigcup_{t \in T} A_t\right) = \bigcup_{t \in T} f^{-1}(A_t), \quad f^{-1}\left(\bigcap_{t \in T} A_t\right) = \bigcap_{t \in T} f^{-1}(A_t).$$

Proposition 0.10

Let $\mathscr{T}$ be a σ-algebra on Y, then $f^{-1}(\mathscr{T})$ is also a σ-algebra on X.

Furthermore, for $\mathscr{E}$ on Y,

$$\sigma(f^{-1}(\mathscr{E})) = f^{-1}(\sigma(\mathscr{E})).$$

Proof. $f^{-1}(\mathscr{E}) \subset f^{-1}(\sigma(\mathscr{E})) \implies f^{-1}(\sigma(E)) \supseteq \sigma(f^{-1}(\mathscr{E}))$.

Again, let

$$\mathscr{G} := \{B \subset Y : f^{-1}(B) \in \sigma(f^{-1}(\mathscr{E}))\}.$$

We need to prove $\mathscr{G}$ is a σ-algebra. This can be checked easily by previous properties, so I leave them out. Hence $\mathscr{G} \supseteq \mathscr{E} \implies \mathscr{G} \supseteq \sigma(\mathscr{E}) \implies f^{-1}(\sigma(\mathscr{E})) \subset \sigma(f^{-1}(\mathscr{E}))$. $\square$

Definition 0.11 (Measurable maps)**.** Let $(X,\mathscr{F})$ and $(Y,\mathscr{S})$, and $f: X \to Y$ a map. We say f is **measurable** if $f^{-1}(\mathscr{S}) \subset \mathscr{F}$, i.e. the preimage of measurable sets are also measurable, denoted by

$$f:(X,\mathscr{F}) \to (Y,\mathscr{S}) \quad \text{or} \quad (X,\mathscr{F}) \xrightarrow{f} (Y,\mathscr{S}) \quad \text{or} \quad f \in \mathscr{F}.$$

Clearly the composition of measurable maps is measurable as well.

A map f is measurable is equivalent to $\sigma(f) \subset \mathscr{F}$, where

$$\sigma(f) := f^{-1}(\mathscr{S})$$

is the smallest σ-algebra which makes f measurable, called the generate σ-algebra of f.

Theorem 0.12

Let $\mathscr{E}$ be a nonempty collection on Y, then

$$f:(X,\mathscr{F}) \to (Y,\sigma(\mathscr{E})) \iff f^{-1}(\mathscr{E}) \subset \mathscr{F}.$$

Proof. Trivial. □

Definition 0.13 (Generalize real numbers)**.** Let $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. Similarly we can assign an order to $\overline{\mathbb{R}}$.

For the calculations, we assign 0 to $0 \cdot \pm\infty$, and $\infty - \infty$, $\frac{\infty}{\infty}$ is undefined.

For all $a \in \overline{\mathbb{R}}$, define $a^+ = \max\{a, 0\}$, $a^- = \max\{-a, 0\}$, so $a = a^+ - a^-$.
Define the Borel σ-algebra on $\overline{\mathbb{R}}$:

$$\mathscr{B}_{\overline{\mathbb{R}}} := \sigma(\mathscr{B}_{\mathbb{R}} \cup \{+\infty, -\infty\}).$$

For any set A, $A \in \mathscr{B}_{\overline{\mathbb{R}}} \iff A = B \cup C$, where $B \in \mathscr{B}_{\mathbb{R}}$, $C \subset \{+\infty, -\infty\}$.

Definition 0.14 (Measurable functions)**.** We say a function f is **measurable** if

$$f:(X,\mathscr{F}) \to (\overline{\mathbb{R}}, \mathscr{B}_{\overline{\mathbb{R}}}).$$

A **random variable (r.v.)** is a measurable map to $(\mathbb{R}, \mathscr{B}_{\mathbb{R}})$.

Measurable functions are in fact random variables that can take $\pm\infty$ as its value.

Theorem 0.15

Let $(X,\mathscr{F})$ be a measurable space, $f: X \to \overline{\mathbb{R}}$ if and only if

$$\{f \le a\} \in \mathscr{F}, \quad \forall a \in \mathbb{R}.$$

Proof. Just note that these sets can generate $\mathscr{B}_{\overline{\mathbb{R}}}$.
Let $\mathscr{E} = \{[-\infty, a] : \forall a \in \mathbb{R}\}$. Then

$$f \text{ measurable} \iff \sigma(f) = f^{-1}\mathscr{B}_{\overline{\mathbb{R}}} = f^{-1}\sigma(\mathscr{E}) \subset \mathscr{F} \iff \sigma(f^{-1}\mathscr{E}) \subset \mathscr{F}.$$

□

Example 0.16

The contant functions are measurable; the indicator functions of a measurable set are measurable $\implies$ *step functions* are measurable.

We say a function f is **Borel function** if it's a measurable function from Borel measurable space to itself.

Corollary 0.17

If f, g are measurable functions, then $\{f = a\}, \{f > g\}, \dots$ are measurable sets.

Theorem 0.18

The arithmetic of measurable funtions are also measurable functions (if they are well-defined).

Proof. Here we only proof $f + g$ is measurable for f, g measurable. For all $a \in \mathbb{R}$, decompose $\{f + g < a\}$ to $A_1 \cup A_2 \cup A_3$:

$$A_1 := \{f = -\infty, g < +\infty\} \cup \{g = -\infty, f < +\infty\} \in \mathscr{F};$$

$$A_2 := \{f = +\infty, g > -\infty\} \cup \{g = +\infty, f > -\infty\} \in \mathscr{F};$$

$$A_3 := \{f < a - g\} \cap \{f, g \in \mathbb{R}\} = \left(\bigcup_{r \in \mathbb{Q}} (\{f < r\} \cap \{g < a - r\})\right) \cap \{f, g \in \mathbb{R}\} \in \mathscr{F}.$$

□

Remark 0.19 — All the measurable functions (or random variables) constitude a vector space.

Theorem 0.20

The limit inferior and limit superior of measurable functions are measurable.

Proof. If $f_1, f_2, \dots$ are measurable, then $\inf f_n$ is measurable:

$$\left\{\inf_{n \geq 1} f_n \geq a\right\} = \bigcap_{n=1}^{\infty} \{f_n \geq a\}.$$

Remark 0.21 — In particular, f measurable $\implies f^+, f^-$ measurable.

Hence

$$\liminf_{n \to \infty} f_n = \lim_{N \to \infty} \inf_{n \geq N} f_n = \sup_{N \geq 1} \inf_{n \geq N} f_n,$$

which is clearly measurable. □

Remark 0.22 — The inferior or superior of **countable** many measurable functions are measurable as well.

Definition 0.23 (Simple functions)**.** Let $(X, \mathscr{F})$ be a measurable space. A **measurable partition** of X is a collection of subsets $\{A_1, \ldots, A_n\}$ with $\sum_{i=1}^n A_i = X$, and $A_i \in \mathscr{F}$.

A **simple function** is defined as

$$f = \sum_{i=1}^n a_i \mathbf{I}_{A_i}.$$

where $\{A_1, \ldots, A_n\}$ is a measurable partition of X, and $a_i \in \mathbb{R}$.

It's clear that simple functions are measurable.

Theorem 0.24

Let f be a measurable function, there exists simple functions $f_1, \ldots$ s.t. $f_n \to f$.

- If $f \geq 0$, we have $0 \leq f_n \leq f$;
- If f is bounded, we have $f_n \rightrightarrows f$.

Proof. This is a standard truncation. For $f \geq 0$, let

$$f_n = \sum_{k=0}^{n2^n - 1} \frac{k}{2^n} \mathbf{I}_{\{k \leq 2^n f \leq k+1\}} + n \mathbf{I}_{f \geq n}.$$

It's clear that $f_n \geq 0$, $f_n \uparrow$, and $f_n(x) \to f(x)$:

$$0 \leq f(x) - f_n(x) \leq \frac{1}{2^n}, \quad f(x) < n;$$

$$n = f_n(x) \leq f(x), \quad f(x) \geq n.$$

Therefore if f is bounded, when $n > \max f(x)$ we have $|f_n(x) - f(x)| < \frac{1}{2^n}$ for all $x \in X$.

For general measurable functions, just decompose f to $f^+ - f^-$. $\square$

Theorem 0.25

Let $g : (X, \mathscr{F}) \to (Y, \mathscr{S})$. Let h be a map $X \to \mathbb{R}$.

Then $h : (X, g^{-1}\mathscr{S})$ iff $h = f \circ g$, where $f : (Y, \mathscr{S}) \to (\mathbb{R}, \mathscr{B}_{\mathbb{R}})$.

Remark 0.26 — For $\overline{\mathbb{R}}$ or $[a, b]$, this theorem also holds.

Proof. There's a typical method for proving something related to measurable functions:

We'll prove the statement for $h \in \mathcal{H}_i$ in order:

- $\mathcal{H}_1$: indicator functions $h = \mathbf{I}_A$, $\forall A \in g^{-1}\mathscr{S}$;
- $\mathcal{H}_2$: non-negative simple functions;
- $\mathcal{H}_3$: non-negative measurable functions;

- $\mathcal{H}_4$: measurable functions.

When $h \in \mathcal{H}_1$, suppose $h = \mathbf{I}_A$, then

$$A = g^{-1}B, B \in \mathscr{S} \implies f = \mathbf{I}_B \text{ suffices.}$$

When $h = \sum_{i=1}^n a_i \mathbf{I}_{A_i} \in \mathcal{H}_2$, since $A_i \in g^{-1}\mathscr{S}$,

$$\exists B_i \in \mathscr{S} \quad s.t. \quad A_i = \{h = a_i\} = g^{-1}B_i.$$

Thus $f = \sum_{i=1}^n a_i \mathbf{I}_{B_i}$ is the desired function.

When $h \in \mathcal{H}_3$, $\exists h_1, h_2, \dots \uparrow h$.

Assume $h_n = f_n \circ g$, let

$$f(y) := \begin{cases} \lim\limits_{n\to\infty} f_n(y), & \text{if it exists;} \\ 0, & \text{otherwise.} \end{cases}$$

Remark 0.27 — Here we still need to prove f is measurable.

Hence for any $x \in X$,

$$h(x) = \lim_{n\to\infty} h_n(x) = \lim_{n\to\infty} f_n(g(x)) = f(g(x)),$$

as f_n's limit must exist at $y = g(x)$.

So for general h, let $h = h^+ - h^-$ and we're done. NOTE: We need to assert that $\infty - \infty$ doesn't occur. □

Remark 0.28 — This is the typical method we'll use frequently in measure theory: to start from simple functions and extend to general functions.

§1 Measure spaces

§1.1 The definition of measure and its properties

The concept of "measure" is frequently used in our everyday life: length, area, weight and even propbability. They all share a similarly: the measure of a whole is equal to the sum of the measure of each part.

In the language of mathematics, let $\mathscr{E}$ be a collection of sets, and there's a function $\mu : \mathscr{E} \to [0, \infty]$ which stands for the measure.

countable additivity: Let $A_1, A_2, \dots \in \mathscr{E}$ be pairwise disjoint sets, and $\sum_{i=1}^\infty A_i \in \mathscr{E}$, then

$$\mu\left(\sum_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu(A_i).$$

Definition 1.1 (Measure)**.** Suppose $\emptyset \in \mathscr{E}$, if a non-negative function

$$\mu : \mathscr{E} \to [0, \infty]$$

satisfies countable additivity, and $\mu(\emptyset) = 0$, then we say μ is a **measure** on $\mathscr{E}$.

If $\mu(A) < \infty$ for all $A \in \mathscr{E}$, we say μ is finite. (In practice we'll just simplify this to $\mu(X) < \infty$)
If $\exists A_1, A_2, \dots \in \mathscr{E}$ are pairwise disjoint sets, s.t.

$$X = \sum_{n=1}^{\infty} A_n, \quad \mu(A_n) < \infty, \forall n.$$

Then we say μ is σ-finite.

There's a weaker version of countable additivity, that is **finite additivity**: If $A_1, \dots, A_n \in \mathscr{E}$, pairwise disjoint, and $\sum A_i \in \mathscr{E}$,

$$\mu\left(\sum_{i=1}^{n} A_i\right) = \sum_{1=i}^{n} \mu(A_i),$$

then we say μ is finite additive.

Subtractivity: $\mu(B - A) = \mu(B) - \mu(A)$, where $A, B, B - A \in \mathscr{E}$, and $\mu(A) < \infty$.

Proposition 1.2

Measure satisfies finite additivity and subtractivity.

Example 1.3 (Counting measure)

Let $\mu(A) = \#A$, $\forall A \in \mathscr{T}_X$. Then μ is a measure.

Example 1.4 (Point measure)

Let $(X, \mathscr{F})$ be a measurable space, define $\delta_x(A) = \mathbf{I}_A(x)$. Then we can define a measure

$$\mu(A) = \sum_{i=1}^{n} p_i \delta_{x_i}(A)$$

Example 1.5 (Length)

Let $\mathscr{E} = \mathscr{Q}_{\mathbb{R}} = \{(a, b] : a, b \in \mathbb{R}\}$, $a \leq b$, then $\mu((a, b]) = b - a$ gives a measure.

Another classical example is the so-called "coin space":
Let $X = \{x = (x_1, x_2, \dots) : x_i \in 0, 1, \forall n\}$.

$$C_{i_1, \dots, i_n} := \{x : x_1 = i_1, \dots, x_n = i_n\},$$

Let

$$\mathscr{Q} = \{\emptyset, X\} \cup \{C_{i_1, \dots, i_n} : n \in \mathbb{N}, i_1, \dots, i_n \in \{0, 1\}\}$$

be a semi-ring. Then $\mu(C_{i_1, \dots, i_n}) = \frac{1}{2^n}$ gives a measure.

We need to check the countable additivity, but actually this can be realized as a compact space and the C's are open sets, so in fact we only need to check finite additivity. (Or we can prove this explicitly)

Another more complex example: finite markov chain.

Proposition 1.6

Let $X = \mathbb{R}$, $\mathscr{E} = \mathscr{R}_{\mathbb{R}}$. $F : \mathbb{R} \to \mathbb{R}$ is non-decreasing, right continuous, then $\mu((a,b]) = F(b) - F(a)$ gives a measure on $\mathscr{E}$.

Proof. First $\mu(\emptyset) = 0$, suppose

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

Since every partial sum has measure at most $F(b_{n+1}) - F(a_1) < F(b) - F(a)$,
$\implies \sum_{i=1}^{n} \mu((a_i, b_i]) \leq \mu((a, b])$.

For the reversed inequality, first we prove that for intervals

$$\bigcup_{i=1}^{n}(c_i, d_i] \supseteq (a, b] \implies \sum_{i=1}^{n} \mu((c_i, d_i]) \geq \mu((a, b]).$$

This can be easily proved by induction, WLOG $b_{n+1} = \max_i b_i$.

Our idea is to extend each $(a_i, b_i]$ a little bit to apply above inequality.

For all $\varepsilon > 0$, take $\delta_i > 0$ s.t.

$$\tilde{b}_i := b_i + \delta_i, \quad F(\tilde{b}_i) - F(b_i) \leq \frac{\varepsilon}{2}.$$

Hence for all $\delta > 0$, $\bigcup_{i=1}^{\infty}(a_i, \tilde{b}_i) \supseteq [a + \delta, b]$, by compactness exists a finite open cover.

$$F(b) - F(a + \delta) \leq \sum_{i=1}^{n} \left(F(\tilde{b}_i) - F(a_i)\right) \leq \varepsilon + \sum_{i=1}^{\infty}(F(b) - F(a)).$$

Let $\varepsilon, \delta \to 0$ to conclude. $\square$

Definition 1.7 (Measure space)**.** A triple $(X, \mathscr{F}, \mu)$ is called a **measure space**, if $(X, \mathscr{F})$ is a measurable space and μ is a measure on $\mathscr{F}$.

If $N \in \mathscr{F}$ s.t. $\mu(N) = 0$, we say N is a **null set**.

A probability space is a measure space $(X, \mathscr{F}, P)$ with $P(X) = 1$.

Example 1.8 (Discrete measure)

If X is countable, $p : X \to [0, \infty]$, $\mu(A) := \sum_{x \in A} p(x)$.

There are other important properties which we think a sensible measure would have:

- Monotonicity: If $A, B \in \mathscr{E}$, $A \subset B$, then $\mu(A) \leq \mu(B)$.
- Countable subadditivity: $A_1, A_2, \cdots \in \mathscr{E}$,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

- Lower continuity: $A_1, A_2, \cdots \in \mathscr{E}$ and $A_n \uparrow A \in \mathscr{E}$.

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

- Similarly there's upper continuity (which requires $\mu(A_1) < \infty$).

Theorem 1.9

The measure on a semi-ring has all the above properties.

Proof. We'll prove that:

- Finite additivity $\implies$ monotonicity, subtractivity;
- Countable additivity $\implies$ subadditivity, upper and lower continuity.

Here we only prove the subadditivity, since others are trivial.
Let $A_1, A_2, \cdots \in \mathscr{Q}$, and $\bigcup_{i=1}^{\infty} A_i \in \mathscr{Q}$.

$$B_n := A_n \backslash \bigcup_{i=1}^{n-1} A_i \in r(\mathscr{Q}) \implies B_n = \sum_{k=1}^{k_n} C_{n,k}, \quad C_{n,k} \in \mathscr{Q}.$$

$$A_n \backslash B_n \in r(\mathscr{Q}) \implies A_n \backslash B_n = \sum_{l=1}^{l_n} D_{n,l}, \quad D_{n,l} \in \mathscr{Q}.$$

Thus by countable additivity,

$$\begin{aligned} \mu\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} \left(\sum_{k=1}^{k_n} \mu(C_{n,k})\right) \\ &\leq \sum_{n=1}^{\infty} \left(\sum_{k=1}^{k_n} \mu(C_{n,k}) + \sum_{l=1}^{l_n} \mu(D_{n,l})\right) = \sum_{n=1}^{\infty} \mu(A_n). \end{aligned}$$

Using similar technique we can deduce the upper and lower continuity. □