Measure Theory

Felix Chen

Contents

§1 Introduction

Teacher: Zhang Fuxi
Email: zhangfxi@math.pku.edu.cn
Homepage: http://www.math.pku.edu.cn/teachers/zhangfxi

§1.1 Starting from probablistics

Definition 1.1 (σ-algebra)**.** Let $\mathscr{F}$ be a family of subsets of a set Ω, if

- $\Omega \in \mathscr{F}$;
- If $A \in \mathscr{F}$, $A^c \in \mathscr{F}$;

- If $A_1, A_2, \cdots \in \mathscr{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathscr{F}$.(Countable union)

then we call $\mathscr{F}$ a σ-algebra.

Some intros about probablistics (left out because I haven't studied probablistics yet;)

§1.2 What is measure theory?

It's an abstract theory, different from probablistics and real analysis. In this course we study a general set X, focus on mathematical thinking and skills, from the simple to construct the complex.

Measure theory studies the intrinsic structure of mathematical objects, and the map between different measure spaces.

§2 Measure spaces and measurable maps

§2.1 Sets and set operations

Definition 2.1. A non-empty set X is our space(universal set), its elements (points) are denoted by lower case letters $x, y, \ldots$.

Some notations:

$$x \in A, x \notin A, x \in A^c, A \subset B, A \cup B, AB = A \cap B,$$

$$B \backslash A (B - A \text{ when } A \subset B), A \Delta B.$$

A family of sets $\{A_t, t \in T\}$.

$$\bigcup_{t \in T} A_t := \{x : \exists t \in T, s.t. x \in A_t\}, \quad \bigcap_{t \in T} A_t := \{x : x \in A_t, \forall t \in T\}.$$

Sometimes we write the union of disjoint sets as sums to emphasize the disjoint property.

Monotone sequence of sets:

$$A_n \uparrow : A_n \subset A_{n+1}, \forall n; \quad A_n \downarrow : A_n \supseteq A_{n+1}, \forall n.$$

Next we define the limits of sets:

Definition 2.2. For monotone sequences:

$$\lim_{n \to \infty} A_n := \bigcup_{n=1}^{\infty} A_n \text{ or } \bigcap_{n=1}^{\infty} A_n.$$

For general sequence of sets:

$$\limsup A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_n; \quad \liminf A_n := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_n.$$

A clearer intepretation of limsup and liminf:

limsup is the set of elements which occurs infinitely many times in A_n, and liminf is the elements which doesn't occur in only finitely many A_n's.

§2.2 Families of sets

Definition 2.3. A family of sets is denoted by script letters $\mathscr{A}, \mathscr{B}, \ldots$.

- A family is a **π-system** if $\mathscr{P} \neq \emptyset$ and it's closed under intersections, e.g. $\{(-\infty, a] : a \in \mathbb{R}\}$.
- **Semi-rings**: $\mathscr{Q}$ is a π-system, and for all $A \subset B$, then there exists finitely many pairwise disjoint sets $C_1, \ldots, C_n \in \mathscr{Q}$ s.t.

$$B \backslash A = \bigcup_{k=1}^{n} C_k = \sum_{k=1}^{n} C_k.$$

  e.g. $\mathscr{Q} = \{(a, b] : a, b \in \mathbb{R}\}$.

  Remark 2.4 — The condition $A \subset B$ can be removed.

- **Rings**: $\mathscr{R}$ is nonempty, and it's closed under union and substraction.
  e.g. $\mathscr{R} = \{\bigcup_{k=1}^{n}(a_k, b_k] : a_k, b_k \in \mathbb{R}\}$.
- **Algebras (fields)**: $\mathscr{A}$ is a π-system that contains X, and is closed under completion.

Proposition 2.5

Semi-rings are π-systems, rings are semi-rings, algebras are rings.

Proof. By definition we only need to check rings are π-systems: $A \cap B = A \backslash (A \backslash B)$.
For algebras, $A \cup B = (A^c \cap B^c)^c, A \backslash B = A \cap B^c$, so they are rings. □

Remark 2.6 — Rings are semi-rings with unions, Algebras are rings with universal set X.

Definition 2.7. Some other families that start from taking limits:

- **Monotone class**: If $A_1, \cdots \in \mathscr{U}$ and A_n monotone, then $\lim_{n\to\infty} A_n \in \mathscr{U}$.
- **λ-system**:

$$X \in \mathscr{L}; \quad A_1, A_2, \cdots \in \mathscr{L}, A_n \uparrow A \implies A \in \mathscr{L};$$
$$A, B \in \mathscr{L}, A \supseteq B \implies A \backslash B \in \mathscr{L}.$$

  notes: $A_n \in \mathscr{L} \iff B_n = A_n^c \in \mathscr{L}$.

- **σ-algebra**:

$$X \in \mathscr{F}; \quad A \in \mathscr{F} \implies A^c \in \mathscr{F};$$
$$A_1, A_2, \cdots \in \mathscr{F} \implies \bigcup_{n=1}^{\infty} A_n \in \mathscr{F}.$$

Proposition 2.8

σ-algebra = algebra & monotone class = λ-system & π-system.

Definition 2.9. **σ-rings**: $\mathscr{R}$ nonempty, $A, B \in \mathscr{R} \implies A \backslash B \in \mathscr{R}$;

$$A_1, A_2, \cdots \in \mathscr{R} \implies \bigcup_{n=1}^{\infty} A_n \in \mathscr{R}.$$

Note: We only need to verify the case when A_n's are disjoint.

Definition 2.10 (Measurable space)**.** Let $\mathscr{F}$ be a σ-algebra on a set X, we say $(X, \mathscr{F})$ is a **measurable space**.

Proposition 2.11

Let $(X, \mathscr{F})$ be a measurable space, A is a subset of X. Then $(A, A \cap \mathscr{F})$ is also a measurable space.

The smallest σ-algebra is $\{\emptyset, X\}$, the largest σ-algebra is the power set $\mathscr{T} = \mathcal{P}(X)$.

In some cases, $\mathscr{T}$ is too large, for example, when $X = \mathbb{R}$, we can't assign a "measure" to every subset that fits our common sense.

§2.3 Generation of σ-algebras

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

Definition 2.12 (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 2.13

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. $\square$

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 2.14

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 constitute a monotone class:

  > **Claim 2.15.** 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 constitute 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 constitute 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 2.16 (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 2.17.** 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 2.18 — 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 2.19
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 2.20 — 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. □

§2.4 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 2.21

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 2.22 (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 2.23

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. $\square$

Definition 2.24 (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 2.25 (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 2.26

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

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

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 2.27

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 2.28

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

Theorem 2.29

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

Theorem 2.31

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

Proof. If $f_1, f_2, \ldots$ 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 2.32 — 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 2.33 — The inferior or superior of **countable** many measurable functions are measurable as well.

Definition 2.34 (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 2.35

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 $f^+ - f^-$. □

Theorem 2.36

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 2.37 — 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 $h \in \mathcal{H}_3$, $\exists h_1, h_2, \ldots \uparrow h$.
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 2.38 — 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. $\square$

Remark 2.39 — This is the typical method we'll use frequently in measure theory: to start from simple functions and extend to general functions.

§3 Measure spaces

§3.1 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 $A_1, A_2, \cdots \in \mathscr{E}$ 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 3.1 (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 $\exists A_1, A_2, \cdots \in \mathscr{E}$ 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 $A_1, \ldots, A_n \in \mathscr{E}$, 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 3.2
Measure satisfies finite additivity and subtractivity.

Example 3.3 (Counting measure)
Let $\mu(A) = \#A$, $\forall A \in \mathscr{T}_X$. Then μ is a measure.

Example 3.4 (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 3.5 (Length)
Let $\mathscr{E} = \mathscr{Q}_{\mathbb{R}} = \{(a, b] : a, b \in \mathbb{R}\}$, $a \le b$, 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 3.6
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 $F(b_{n+1}) - F(a_1) < F(b) - F(a)$,
$\implies \sum_{i=1}^{n} \mu((a_i, b_i]) \le \mu((a, b])$.

For the reversed inequality, first we prove that for intervals

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

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.

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

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.

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

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

Definition 3.7 (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 3.8 (Discrete measure)
If X is countable, $p : X \to [0, \infty]$, $\mu(A) := \sum_{x\in A} p(x)$ is a measure.

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

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

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

- 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 3.9
The measure on a semi-ring has all the above properties.

Proof. In fact,

- 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,

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

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

Theorem 3.10

Let μ be a set function on a ring with finite additivity, then $1 \iff 2 \iff 3 \implies 4 \implies 5$.

- μ is countablely additive;
- μ is countablely subadditive;
- μ is lower continuous;
- μ is upper continuous;
- μ is continuous at $\emptyset$.

§3.2 Outer measure

Once we construct a measure on a semi-ring, we want to extend it to a σ-algebra. Since we can't directly do this, we shall relax some of our restrictions, say reduce countable additivity to subadditivity.

Definition 3.11 (Outer measure)**.** Let $\tau : \mathscr{T} \to [0, \infty]$ satisfying:

- $\tau(\emptyset) = 0$;
- If $A \subset B \subset X$, then $\tau(A) \leq \tau(B)$;
- (Countable subadditivity) $\forall A_1, A_2, \cdots \in \mathscr{T}$, we have

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

We call τ an **outer measure** on X.

It's easier to extend a measure on semi-ring to an outer measure:

Theorem 3.12

Let μ be a non-negative set function on a collection $\mathscr{E}$, where $\emptyset \in \mathscr{E}$ and $\mu(\emptyset) = 0$. Let

$$\tau(A) := \inf\left\{\sum_{n=1}^{\infty} \mu(B_n) : B_n \in \mathscr{E}, \bigcup_{n=1}^{\infty} B_n \supseteq A\right\}, \quad \forall A \in \mathscr{T}.$$

By convention, $\inf \emptyset = \infty$. (*μ need not be a measure!*)

Then τ is called the outer measure generated by μ.

Proof. Clearly $\tau(\emptyset) = 0$, and $\tau(A) \leq \tau(B)$ for $A \subset B$.

$$\bigcup_{n=1}^{\infty} B_n \supseteq B \implies \bigcup_{n=1}^{\infty} B_n \supseteq A.$$

For all $A_1, A_2, \cdots \in \mathscr{T}$, WLOG $\tau(A_n) < \infty$. Take $B_{n,k}$ s.t. $\bigcup_{k=1}^{\infty} B_{n,k} \supseteq A_n$, such that

$$\sum_{k=1}^{\infty} \mu(B_{n,k}) < \tau(A_n) + \frac{\varepsilon}{2^n}.$$

Therefore

$$\bigcup_{n=1}^{\infty}\bigcup_{k=1}^{\infty} B_{n,k} \supseteq A_n,$$

$$\tau\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mu(B_{n,k}) + \varepsilon \leq \sum_{n=1}^{\infty} \tau(A_n) + \varepsilon.$$

□

Example 3.13

Let $\mathscr{E} = \{X, \emptyset\}$, $\mu(X) = 1$, $\mu(\emptyset) = 0$. Then $\tau(A) = 1$, $\forall A \neq \emptyset$.

Example 3.14

Let $X = \{a, b, c\}$, $\mathscr{E} = \{\emptyset, \{a\}, \{a, b\}, \{c\}\}$. $\mu(A) = \#A$ for $A \in \mathscr{E}$.

Here something strange happens: $\tau(\{b\}) = 2$ instead of 1, and $\tau(\{b, c\}) = 3$ instead of 2.

In the above example, we found the set $\{b\}$ somehow behaves badly: if we divide $\{a, b\}$ to $\{a\} + \{b\}$, the outer measure is not the sum of two smaller measure.

Hence we want to get rid of this kind of inconsistency to get a proper measure:

Definition 3.15 (Measurable sets)**.** Let τ be an outer measure, if a set A satisfies *Caratheodory condition*:

$$\tau(D) = \tau(D \cap A) + \tau(D \cap A^c), \quad \forall D \in \mathscr{T},$$

we say A is **measurable**.

Remark 3.16 — Inorder to prove A measurable, we only need to check

$$\tau(D) \geq \tau(D \cap A) + \tau(D \cap A^c), \quad \forall D \in \mathscr{T}.$$

Let $\mathscr{F}_\tau$ be the collection of all the τ measurable sets,

Definition 3.17 (Complete measure space)**.** Let $(X, \mathscr{F}, \mu)$ be a measure space, if for all null set A, and $\forall B \subset A, B \in \mathscr{F} \implies \mu(B) = 0$, we say $(X, \mathscr{F}, \mu)$ is **complete**.

Theorem 3.18 (Caratheodory's theorem)

Let τ be an outer measure, then $\mathscr{F} := \mathscr{F}_\tau$ is a σ-algebra, and $(X, \mathscr{F}, \tau)$ is a complete measure space.

Proof. First we prove $\mathscr{F}$ is an algebra:

Note $\emptyset \in \mathscr{F}$, and $\mathscr{F}$ is closed under completements.

For measurable sets A_1, A_2,

$$\begin{aligned}
\tau(D) &= \tau(D \cap A_1) + \tau(D \cap A_1^c) \\
&= \tau(D \cap A_1 \cap A_2) + \tau(D \cap (A_1) \cap A_2^c) + \tau(D \cap A_1^c) \\
&= \tau(D \cap (A_1 \cap A_2)) + \tau(D \cap (A_1 \cap A_2)^c).
\end{aligned}$$

So $A_1 \cap A_2$ is measurable.

Secondly, we prove $\mathscr{F}$ is a σ-algebra.

Let $A_1, A_2, \cdots \in \mathscr{F}$,

$$B_n := A_n \backslash \bigcup_{i=1}^{n-1} A_i \in \mathscr{F},$$

Then B_i pairwise disjoint and $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$. Let $B_f = \bigcup_{i=1}^{\infty} B_i$.

It's sufficient to prove

$$\tau(D) \geq \tau(D \cap B_f) + \tau(D \cap B_f^c).$$

Let $D_n = \sum_{i=1}^{n} B_i \cap D$, $D_f = D \cap B_f$, $D_\infty = D \backslash D_f$.

Since B_i are measurable,

$$\tau(D) = \tau(D_n) + \tau(D \backslash D_n) \geq \tau(D_n) + \tau(D_\infty) = \sum_{i=1}^{n} \tau(D \cap B_i) + \tau(D_\infty).$$

Now we take $n \to \infty$,

$$\tau(D) \geq \sum_{i=1}^{\infty} \tau(D \cap B_i) + \tau(D_\infty) \geq \tau\left(D \cap \sum_{i=1}^{\infty} B_i\right) + \tau(D_\infty).$$

Where the last step follows from countable subadditivity.

This implies B_f measurable $\implies \mathscr{F}$ is a σ-algebra.

Next we prove $\tau|\mathscr{F}$ is a measure: Just let $D = \sum_{i=1}^{\infty} B_i$ in the previous equation.

Last we prove $(X, \mathscr{F}, \tau)$ is complete:

If $\tau(A) = 0$, $\tau(D) \geq \tau(D \cap A^c) = \tau(D \cap A) + \tau(D \cap A^c)$. Thus $A \in \mathscr{F}$. □

§3.3 Measure extension

Definition 3.19 (Measure extension)**.** Let μ, ν be measures on $\mathscr{E}$ and $\overline{\mathscr{E}}$, and $\mathscr{E} \subset \overline{\mathscr{E}}$. If

$$\nu(A) = \mu(A), \quad \forall A \in \mathscr{E},$$

we say ν is a extension of μ on $\overline{\mathscr{E}}$.

If we start from a measure μ on $\mathscr{E}$, ideally, μ can generate an outer measure τ, and we can take $\mathscr{F}_\tau$ to construct a mesaure space.

However, things could go wrong:

Example 3.20

Let $X = \{a, b, c\}, \mathscr{E} = \{\emptyset, \{a, b\}, \{b, c\}, X\}$ with

$$\mu(\emptyset) = 0, \mu(\{a, b\}) = 1, \mu(\{b, c\}) = 1, \mu(X) = 2.$$

Then μ is a measure on $\mathscr{E}$, and the outer measure

$$\tau(\emptyset) = 0.$$

Observe that $\mathscr{F}_\tau = \{\emptyset, X\}$, so in this case $\tau|_{\mathscr{F}}$ is the trivial measure.

Example 3.21

Let $X = \mathbb{R}$, $\mathscr{E} = \{(a,b] : a < b, a, b \in \mathbb{R}\}$. Let $\mu(\emptyset) = 0$, and $\mu(A) = \infty$ for $A \neq \emptyset$.

Then μ can be extend to the Borel σ-algebra on $\mathbb{R}$ with $\mu_\alpha = \sum_{q\in\mathbb{Q}} \alpha\delta_q$, $\forall \alpha \geq 0$. So the extension is not unique.

Therefore in order to get a "proper" extension, we must put some requirements on both the starting collection and the set function μ.

Proposition 3.22

Let $\mathscr{P}$ be a π system. If two measures μ, ν on $\sigma(\mathscr{P})$ satisfying

$$\mu|_{\mathscr{P}} = \nu|_{\mathscr{P}}, \quad \mu|_{\mathscr{P}} \text{ is } \sigma\text{-finite},$$

Then $\mu = \nu$.

Proof. Let $A_1, A_2, \cdots \in \mathscr{P}$ s.t. $X = \sum_{n=1}^{\infty} A_n$ and $\mu(A_n) < \infty$.

Fix n, let $B = A_n$, we want to prove that

$$\mu(B \cap A) = \nu(B \cap A), \quad \forall A \in \sigma(\mathscr{P}).$$

Let $B \in \mathscr{P}$ with $\mu(B) < \infty$,

$$\mathscr{L} := \{A \in \sigma(\mathscr{P}) : \mu(A \cap B) = \nu(A \cap B)\}.$$

We'll prove $\mathscr{L}$ is a λ system, so that $\mathscr{L} \supseteq \sigma(\mathscr{P})$.

Suppose $A_1, A_2 \in \mathscr{L}$ and $A_1 \supseteq A_2$, by $\mu(B) < \infty$,

$$\mu((A_1 - A_2)B) = \mu(A_1B) - \mu(A_2B) = \nu(A_1B - A_2B) = \nu((A_1 - A_2)B).$$

So $A_1 - A_2 \in \mathscr{L}$.

Let $A_1, A_2, \cdots \in \mathscr{L}$ and $A_n \uparrow A$, then

$$\mu(AB) = \lim_{n\to\infty} \mu(A_nB) = \lim_{n\to\infty} \nu(A_nB) = \nu(AB).$$

Which implies $A \in \mathscr{L}$.

Hence $\sigma(\mathscr{P}) \subset \mathscr{L}$, i.e.

$$\mu(A \cap A_n) = \nu(A \cap A_n), \quad \forall A \in \sigma(\mathscr{P}).$$

Therefore

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A \cap A_n) = \sum_{n=1}^{\infty} \nu(A \cap A_n) = \nu(A), \quad \forall A \in \sigma(\mathscr{P}).$$

□

Example 3.23

In probability, let $\mathscr{E}_1, \mathscr{E}_2$ be collections of sets. We say they're independent if

$$P(AB) = P(A)P(B), \quad \forall A \in \mathscr{E}_1, B \in \mathscr{E}_2.$$

By the previous theorem we can derive $\lambda(\mathscr{E}_1), \lambda(\mathscr{E}_2)$ are independent.

If $A_1, A_2, \ldots$ satisfy

$$P(A_{i_1} \cdots A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}),$$

we say they are independent.

Let $\{1, 2, \ldots\} = I + J$, then the σ-algebra generated by

$$\mathscr{E}_1 = \{A_\alpha \mid \alpha \in I\}, \quad \mathscr{E}_2 = \{A_\alpha \mid \alpha \in J\}$$

are independent.

Theorem 3.24 (Measure extension theorem)

Let μ be a measure on a semi-ring $\mathscr{Q}$, τ is the outer meaasure generated by μ. We have

$$\sigma(\mathscr{Q}) \in \mathscr{F}_\tau, \quad \tau|_{\mathscr{Q}} = \mu.$$

Remark 3.25 — Any measure on a semi-ring $\mathscr{Q}$ can extend to the $\sigma(\mathscr{Q})$, and if μ is σ-finite, the extension is unique.

Proof. For any $A \in \mathscr{Q}$, let $B_1 = A$, $B_n = \emptyset, n \geq 2$. Then $\tau(A) \leq \sum \mu(B_n) = \mu(A)$.

On the other hand, if $A_1, A_2, \cdots \in \mathscr{Q}$ s.t. $\bigcup_{n=1}^\infty A_n \supseteq A$, then

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

Thus $\tau(A) = \mu(A)$, where we used the fact that μ is countable subadditive.

Next we prove $A \in \mathscr{F}_\tau$. We need to show that

$$\tau(D) \geq \tau(D \cap A) + \tau(D \cap A^c).$$

WLOG $\tau(D) < \infty$. Take $B_1, B_2, \cdots \in \mathscr{Q}$ s.t.

$$\bigcup_{n=1}^\infty B_n \supseteq D, \quad \sum_{n=1}^\infty \mu(B_n) < \tau(D) + \varepsilon.$$

Denote $\hat{D} := B_n \in \mathscr{Q}$ for a fixed n. Suppose $\hat{D} \cap A^c = \hat{D} \backslash A = \sum_{i=1}^n C_i$.

$$\mu(\hat{D}) = \mu(\hat{D} \cap A) + \sum_{i=1}^n \mu(C_i) \geq \tau(\hat{D} \cap A) + \tau(\hat{D} \cap A^c).$$

Apply this inequality to each B_n,

$$\tau(D) + \varepsilon > \sum_{n=1}^\infty (\tau(B_n \cap A) + \tau(B_n \cap A^c)) \geq \tau(D \cap A) + \tau(D \cap A^c).$$

this implies $\tau(D) \geq \tau(D \cap A) + \tau(D \cap A^c) \implies A \in \mathscr{F}_\tau$.

At last by Caratheodory's theorem, τ is a measure on $\mathscr{F}_\tau \supseteq \sigma(\mathscr{Q})$. $\square$

Theorem 3.26 (Equi-measure hull)

Let τ be the outer measure generated by μ,

- $\forall A \in \mathscr{F}_\tau$, $\exists B \in \sigma(\mathscr{Q})$ s.t. $B \supseteq A$ and $\tau(A) = \tau(B)$;
- If μ is σ-finite, then $\tau(B \backslash A) = 0$.

Remark 3.27 — This theroem states that $\mathscr{F}_\tau$ is just $\sigma(\mathscr{Q})$ appended with null sets.

Proof. If $\tau(A) = \infty$, $B = X$ suffices.

By definition, there exists $B_n = \bigcup_{k=1}^{k_n} B_{n,k} \supseteq A$ s.t. $\tau(B_n) < \tau(A) + \frac{1}{n}$. Let $B = \bigcap_{n=1}^{\infty} B_n$, we must have $\tau(B) = \tau(A)$.

Now for the second part, let $X = \sum_{n=1}^{\infty} A_n$, $A_n \in \mathscr{Q}$, $\mu(A_n) < \infty$.

Since $A = \sum_{n=1}^{\infty} AA_n$, we have

$$AA_n \in \mathscr{F}_\tau, \quad \tau(AA_n) \leq \tau(A_n) = \mu(A_n) < \infty.$$

Let $B_n \in \sigma(\mathscr{Q})$ s.t. $B_n \supseteq AA_n$ and $\tau(B_n) = \tau(AA_n) < \infty$. Let $B := \bigcup_{n=1}^{\infty} B_n$ we have

$$\tau(B - A) = \tau\left(\bigcup_{n=1}^{\infty} (B_n - AA_n)\right) \leq \sum_{n=1}^{\infty} \tau(B_n - AA_n) = 0.$$

$\square$

Let $\mathscr{R}, \mathscr{A}, \mathscr{F}$ be the ring, algebra, σ-algebra generated by $\mathscr{Q}$, respectively. The outer measure τ restricts to a measure on each of these collections, denoted by μ_1, μ_2, μ_3. Each μ_i can generate an outer measure τ_i, but actually they're all the same as our original τ, since τ_i are "build up" from τ, intuitively τ_i cannot be any better than τ. (The proof says exactly the same thing, so I'll omit it)

Proposition 3.28

Let μ be a measure on an algebra $\mathscr{A}$. τ is the outer measure generated by μ, for all $A \in \sigma(\mathscr{A})$, if $\tau(A) < \infty$, then $\forall \varepsilon > 0$, $\exists B \in \mathscr{A}$ s.t. $\tau(A \Delta B) < \varepsilon$.

Remark 3.29 — In practice we often replace τ with a σ-finite measure μ on $\sigma(\mathscr{A})$. (Here σ-finite is on $\mathscr{A}$)

Proof. Choose $B_1, B_2, \cdots \in \mathscr{A}$ s.t.

$$\hat{B} := \bigcup_{n=1}^{\infty} B_n \supseteq A, \quad \sum_{n=1}^{\infty} \mu(B_n) < \tau(A) + \frac{\varepsilon}{2}.$$

Let N be a sufficiently large number, $B := \bigcup_{n=1}^{N} B_n \in \mathscr{A}$,

$$\tau(A \backslash B) \leq \tau\left(\bigcup_{n=N+1}^{\infty} B_n\right) \leq \sum_{n=N+1}^{\infty} \tau(B_n) \leq \frac{\varepsilon}{2}.$$

As $\tau(B \backslash A) \leq \tau(\hat{B} \backslash A) < \frac{\varepsilon}{2}$, $\tau(A \Delta B) < \varepsilon$. □

Example 3.30

Consider the Bernoulli test, recall $C_{i_1,\ldots,i_n}$ we defined earlier. A measure(probability) μ is defined on the semi-ring $\{C_{i_1,\ldots,i_n}\} \cup \{\emptyset, X\}$, then it can extend uniquely to the σ-algebra generated by it. This is how the probability of Bernoulli test comes from.

Let $(X, \mathscr{F}, P)$ be a probability space, $A_1, A_2, \cdots \in \mathscr{F}$. We define the **tail σ-algebra** $\mathscr{T}$:

$$\mathscr{G}_n := \sigma(\{A_{n+1}, A_{n+2}, \ldots\}), \quad \mathscr{T} := \bigcap_{n=1}^{\infty} \mathscr{G}_n.$$

Let $f_1, f_2, \ldots$ be random variable, the tail σ-algebra generated by them is defined similarly:

$$\mathscr{G}_n := \sigma(\{f_{n+1}, f_{n+2}, \ldots\}), \quad \mathscr{T} := \bigcap_{n=1}^{\infty} \mathscr{G}_n.$$

Theorem 3.31 (Kolmogorov's 0-1 law)

If $A_1, A_2, \cdots \in \mathscr{F}$ are independent, then $P(A) \in \{0, 1\}$, $\forall A \in \mathscr{T}$.

Proof. Let $\mathscr{F}_n := \sigma(\{A_1, \ldots, A_n\})$ and $\mathscr{G}_n$. They are clearly independent.

Note that $\mathscr{A} := \bigcup_{n=1}^{\infty} \mathscr{F}_n$ is an algebra.

Let $\mathscr{H} := \sigma(\mathscr{A}) \supseteq \mathscr{G}_n \supseteq \mathscr{T}$.

Hence $\forall A \in \mathscr{T} \subset \mathscr{H}$, $\forall \varepsilon > 0$, exists $B \in \mathscr{A}$ s.t. $P(A \Delta B) < \varepsilon$, so

$$P(A) - P(AB) \leq \varepsilon, \quad |P(A) - P(B)| \leq \varepsilon.$$

Since $B \in \mathscr{F}_n$ for some n, thus it is independent to A.

$$|P(A) - P(A)^2| \leq |P(A) - P(AB)| + |P(AB) - P(A)^2| \leq 2\varepsilon.$$

Let $\varepsilon \to 0$, we'll get $P(A) \in \{0, 1\}$. □

Remark 3.32 — When A_i's are replace by random variables, this theorem also holds.

Example 3.33

finite Markov chain

§3.4 The completion of measure spaces

Let $(X, \mathscr{F}, \mu)$ be a measure space, and

$$\widetilde{\mathscr{F}} := \{A \cup N : A \in \mathscr{F}, \exists B \in \mathscr{F} \ s.t. \ \mu(B) = 0, N \subset B\}.$$

Another way to define it is: $\widetilde{\mathscr{F}} := \{A \backslash N\}$, since

$$A \cup N = A + NA^c = (A \cup B) \backslash (BA^c \backslash N);$$

$$A \backslash N = A - NA = (A \backslash B) + (BA \backslash N).$$

In fact, we can do even more: $\widetilde{\mathscr{F}} := \{A \Delta N\}$.

Next we define the measure

$$\widetilde{\mu}(A \cup N) := \mu(A), \quad \forall A \cup N \in \widetilde{\mathscr{F}}$$

We need to check several things:

- $\widetilde{\mathscr{F}}$ is a σ-algebra.
- $\widetilde{\mu}$ is well-defined.
- $(X, \widetilde{\mathscr{F}}, \widetilde{\mu})$ is a complete measure space.

Remark 3.34 — The mesure $\widetilde{\mu}$ is the *minimal complete extension* of μ, i.e. if $(X, \mathscr{G}, \nu)$ is another complete extension, then

$$\nu(B) = \mu(B) = 0 \implies \forall N \subset B, N \in \mathscr{G}, \nu(N) = 0.$$

$$\mu(A) = \nu(A) \leq \nu(A \cup N) \leq \nu(A) + \nu(N) = \nu(A).$$

Thus $\mathscr{G} \supseteq \widetilde{\mathscr{F}}$ and $\nu(A) = \widetilde{\mu}(A)$ for $A \in \widetilde{\mathscr{F}}$.

Therefore we call $(X, \widetilde{\mathscr{F}}, \widetilde{\mu})$ the **completion** of $(X, \mathscr{F}, \mu)$.

Obviously $\emptyset \in \widetilde{\mathscr{F}}$; For $A \cup N \in \widetilde{\mathscr{F}}$, $(A \cup N)^c = A^c - A^c N \in \widetilde{\mathscr{F}}$.

$$\bigcup_{n=1}^{\infty} (A_n \cup N_n) = \bigcup_{n=1}^{\infty} A_n \cup \bigcup_{n=1}^{\infty} N_n, \quad N = \bigcup_{n=1}^{\infty} N_n \subset \bigcup_{n=1}^{\infty} B_n.$$

Thus $\widetilde{\mathscr{F}}$ is a σ-algebra.

For $\widetilde{\mu}$, if $A_1 \cup N_1 = A_2 \cup N_2$,

$$\mu(A_1) = \mu(A_1 \cup B_2) \geq \mu(A_2).$$

Last we prove the countable additivity of $\widetilde{\mu}$. It's easy to check, so left out.

For the completeness, if $C \subset A \cup N, \mu(A) = 0$, then $C \subset A \cup B$ which is null.

Combining with the previous results we have

Theorem 3.35

Let τ be the outer measure generated by μ, a σ-finite measure on a semi-ring $\mathscr{Q}$. We have $(X \mathscr{F}_\tau, \tau)$ is the completion of $(X, \sigma(\mathscr{Q}), \tau)$.

Proof. Let $\mathscr{F} = \sigma(\mathscr{Q})$, we'll prove that $\widetilde{\mathscr{F}} = \mathscr{F}_\tau$.

Since $(X, \mathscr{F}_\tau, \tau)$ is complete, we have $\mathscr{F}_\tau \supseteq \widetilde{\mathscr{F}}$.

For all $C \in \mathscr{F}_\tau$, it suffices to prove $C = A + N$ for some $A \in \mathscr{F}, N \subset B$ with B null.

Since $C^c \in \mathscr{F}_\tau$, $\exists B \in \mathscr{F}$ s.t.

$$B \supseteq C^c, \quad \tau(B \backslash C^c) = 0.$$

□

§3.5 Distributions

Let $F : \mathbb{R} \to \mathbb{R}$ be a non-decreasing, right continuous function (called a **quasi-distribution function**). Let ν be the measure on $\mathscr{Q}_\mathbb{R}$,

$$\nu : (a, b] \mapsto \max\{F(b) - F(a), 0\}.$$

Let τ be the outer measure generated by ν. We call the sets in $\mathscr{F}_\tau$ to be the Lebesgue-Stieljes measurable sets (L-S measurable), a measurable function

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

is L-S measurable, and $\tau\big|_{\mathscr{F}_\tau}$ is the L-S measure.

In fact finite L-S measures and the quasi-distribution functions are 1-1 correspondent (ignoring the difference of a constant), since $\mathscr{B}_\mathbb{R} = \sigma(\mathscr{Q}_\mathbb{R})$, $(\mathbb{R}, \mathscr{F}_\tau, \tau)$ is the completion of $(\mathbb{R}, \mathscr{B}_\mathbb{R}, \tau)$, and $\mu_F = \tau\big|_{\mathscr{B}_\mathbb{R}}$ is the unique extension of ν.

Conversely, given a measure μ on $(\mathbb{R}, \mathscr{B}_\mathbb{R})$, if $\mu((a, b]) < \infty$ for all $a < b$, then $\mu = \mu_F$, where

$$F = F_\mu : x \mapsto \mu((-\infty, x]), \quad x \in \mathbb{R}.$$

We say a probability measure on $(\mathbb{R}, \mathscr{B}_\mathbb{R})$ is a **distribution**. Let $F : \mathbb{R} \to \mathbb{R}$ be a quasi-distribution function, if F satisfies:

$$F(-\infty) := \lim_{x \to -\infty} F(x) = 0, \quad F(+\infty) := \lim_{x \to +\infty} F(x) = 1,$$

then we say F is a distribution function (d.f.).

From the previous example we know distribution and d.f. are one-to-one correspondent.

Theorem 3.36

Let $g : (X, \mathscr{F}) \to (Y, \mathscr{S})$, μ is a measure on $\mathscr{F}$. Let

$$\nu(B) := \mu(g^{-1}(B)) = \mu \circ g^{-1}(B), \quad \forall B \in \mathscr{S}.$$

Then ν is a measure on $\mathscr{S}$.

Proof. Trivial. Just check the definition one by one. □

Let $(\Omega, \mathscr{F}, P)$ be a probability space, $f : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B}_\mathscr{R})$. We say

$$P \circ f^{-1} : B \mapsto P(f \in B)$$

is the **distribution** of f, denoted by μ_f, i.e. $\mu_f(B) = P(f \in B)$ for Borel sets B.

If $\mu_f = \mu$, we say f obeys the distribution μ, denoted by $f \sim \mu$.

Let $F_f = F_{\mu_f}$ be the distribution function of f.

$$F_f := \mu_f((-\infty, x]) = P(f \le x), \quad x \in \mathbb{R}.$$

We can also say f obeys F_f, denoted by $f \sim F_f$.

If $F_f = F_g$, then we say f and g is **equal in distribution**, denoted by $f \overset{d}{=} g$.

Theorem 3.37
Any d.f. is the distribution function of some random variable.

Proof. Let $\Omega = \mathbb{R}, \mathscr{F} = \mathscr{B}_{\mathbb{R}}, P = \mu_F$, and $f = \mathrm{id}$. It's clear that the distribution function of f is precisely F. □

§3.6 The convergence of measurable functions

Let $(X, \mathscr{F}, \mu)$ be a measure space.

For any statement, if there exists null set N s.t. it holds for all $x \in N^c$, then we say this statement holds *almost everywhere*. (Often abbreviated as *a.e.*)

Definition 3.38. If a sequence of functions f_n satisfies

$$\mu\left(\lim_{n\to\infty} f_n \neq f\right) = 0,$$

(here f is finite a.e.) we say $\{f_n\}$ converges to f **almost everywhere**, denoted by $f_n \to f, a.e.$.

Definition 3.39. If $\forall \delta > 0$, $\exists A \in \mathscr{F}$ s.t. $\mu(A) < \delta$ and

$$\lim_{n\to\infty} \sup_{x\notin A} |f_n(x) - f(x)| = 0,$$

then we say $\{f_n\}$ converges to f **almost uniformly**, denoted by $f_n \to f, a.u.$.

If $f_n \to f, a.u.$, $\forall \varepsilon > 0$, $\exists m = m_k(\varepsilon)$ s.t. when $n \geq m$, $|f_n(x) - f(x)| < \varepsilon, \forall x \in C_k$, but we could have $\sup_k m_k(\varepsilon) = \infty$, thus $f_n \rightrightarrows f$ doesn't hold. e.g. $f_n(x) = x^n, f(x) = 0, x \in [0, 1), f(1) = 1$.

Proposition 3.40
$f_n \to f, a.u. \implies f_n \to f, a.e.$.

Proof. For all n, $\exists A_n$ s.t. $\mu(A_n) < \frac{1}{n}$, and $f_n \to f$ in A_n^c. Let $A := \bigcap_n A_n$.

Then $\{f_n \not\to f\} \cup \{|f| = \infty\} \subset A$, $\mu(A) = 0$, hence $f_n \to f, a.e.$. □

Proposition 3.41
$f_n \to f, a.e.$ iff $\forall \varepsilon > 0$,

$$\mu\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{|f_m - f| \geq \varepsilon\}\right) = 0.$$

Note: If $f(x) - g(x)$ is not defined, we regard it as $+\infty$.

Proof. Let $A_\varepsilon := \bigcap\bigcup\{|f_m - f| > \varepsilon\}$.

$$\left\{\lim_{n\to\infty} f_n \neq f\right\} \cup \{|f| = \infty\} = \bigcup_{k=1}^{\infty} A_{\frac{1}{k}} =\uparrow \lim_{k\to\infty} A_{\frac{1}{k}}.$$

□

Proposition 3.42

$f_n \to f, a.u.$ iff $\forall \varepsilon > 0$, we have

$$\downarrow \lim_{m\to\infty} \mu\left(\bigcup_{n=m}^{\infty}\{|f_n - f| \geq \varepsilon\}\right) = 0.$$

Proof. If $f_n \to f, a.u.$, $\forall \delta, \exists A \in \mathscr{F}$ s.t. $\mu(A) < \delta$ and $f_n \rightrightarrows f, x \in A^c$.

This means for any fixed ε, $\exists m$ s.t. when $n \geq m$, $x \notin A \implies |f_n(x) - f(x)| < \varepsilon$. Thus $A \supseteq \bigcup_{n=m}^{\infty} |f_n - f| \geq \varepsilon$.

Conversely, $\forall \delta > 0$, $\exists m_k$ s.t.

$$\mu\left(\bigcup_{n=m_k}^{\infty}\{|f_n - f| \geq \frac{1}{k}\}\right) < \frac{\delta}{2^k}.$$

Denote the above set by A_k, and $A = \bigcup_{k=1}^{\infty} A_k$, then $\mu(A) < \delta$, and $f_n(x) \rightrightarrows f(x)$ for $x \in A^c$. $\square$

Definition 3.43. If $\forall \varepsilon > 0$,

$$\lim_{n\to\infty} \mu(|f_n - f| \geq \varepsilon) = 0,$$

then we say $\{f_n\}$ converges to f **in measure**, denoted by $f_n \xrightarrow{\mu} f$.

Theorem 3.44

$$f_n \to f, a.u. \implies f_n \to f, a.e., \quad f_n \xrightarrow{\mu} f.$$

If $\mu(X) < \infty$, then

$$f_n \to f, a.u. \iff f_n \to f, a.e. \implies f_n \xrightarrow{\mu} f.$$

Theorem 3.45

$f_n \to f$ in measure iff for any subsequence of $\{f_n\}$, exists its subsequence $\{f_{n'}\}$ s.t.

$$f_{n'} \to f, a.u.$$

Proof. When $f_n \to f$ in measure, let $n_0 = 0$. Take $n_k > n_{k-1}$ inductively such that

$$\mu\left(\left\{|f_n - f| \geq \frac{1}{k}\right\}\right) \leq \frac{1}{2^k}, \quad \forall n \geq n_k.$$

Then $\forall \varepsilon > 0$, $\exists \frac{1}{m} < \varepsilon$, $\{|f_{n_k} - f| \geq \varepsilon\} \subset \{|f_{n_k} - f| \geq \frac{1}{k}\}$,

$$\mu\left(\bigcup_{k=m}^{\infty}\{|f_{n_k} - f| \geq \varepsilon\}\right) \leq \mu\left(\bigcup_{k=m}^{\infty}\left\{|f_{n_k} - f| \geq \frac{1}{k}\right\}\right) \leq \frac{1}{2^{m-1}} \to 0.$$

Conversely, we assume for contradiction that $\exists \varepsilon > 0$ s.t. $\mu(\{|f_n - f| \geq \varepsilon\}) \not\to 0$.
So $\exists \delta > 0$ and subsequence $\{n_k\}$ s.t. $\mu(\{|f_{n_k} - f| \geq \varepsilon\}) > \delta$.
Hence there doesn't exist a subsequence $\{f_{n'}\}$ of $\{f_{n_k}\}$ s.t. $f_{n'} \to f, a.u.$. $\square$

Example 3.46

Consider measure space $(\mathbb{R}, \mathscr{B}_{\mathbb{R}}, \lambda)$, the Lebesgue measure, $f_n = \mathbf{I}_{|x|>n}$, then

$$f_n \to 0, \forall x \implies f_n \to 0, a.e..$$

let $\varepsilon = 1$, it's clear that f_n doesn't converge to f in measure, hence not almost uniformly.

Example 3.47

Let $f_{k,i} = \mathbf{I}_{\frac{i-1}{k} < x \le \frac{i}{k}}$, $i = 1, \ldots, k$. It's clear that $f_{k,i} \to 0$ in measure, but not almost everywhere, and hence not almost uniformly.

§3.7 Probability space

Let $(\Omega, \mathscr{F}, P)$ be a probability space. Here almost everywhere is renamed to almost surely.

Let F be a real function, let $C(F)$ be the continuous points of F.

Let $F, F_1, F_2, \ldots$ be non-decreasing functions, if

$$\lim_{n\to\infty} F_n(x) = F(x), \quad \forall x \in C(F),$$

then we say $\{F_n\}$ converge to F weakly, $F_n \xrightarrow{w} F$.

Let $F, F_1, F_2, \ldots$ be distribution functions, $f_n \sim F_n$,

Definition 3.48. If $F_n \xrightarrow{w} F$, then we say $\{f_n\}$ converge to F in distribution, denoted by $f_n \xrightarrow{d} F$. For $f \sim F$, we can also write $f_n \xrightarrow{d} f$.

Theorem 3.49

$f_n \xrightarrow{P} f \implies f_n \xrightarrow{d} f$.

Proof.

$$\begin{aligned} P(h \le y) &\le P(h \le y, |h-g| < \varepsilon) + P(h \le y, |h-g| \ge \varepsilon) \\ &\le P(g \le y+\varepsilon) + P(|h-g| \ge \varepsilon). \end{aligned}$$

Let $F_n(x) = P_n(f \le x)$ Let $h = f_n$, $g = f$, $y = x$.

$$\limsup_{n\to\infty} F_n(x) \le F(x+\varepsilon), \quad \forall \varepsilon > 0.$$

Thus $\limsup_{n\to\infty} F_n(x) \le F(x)$. TODO ☐

Theorem 3.50 (Skorokhod)

If $f_n \xrightarrow{d} f$, then exists a probability space $(\widetilde{X}, \widetilde{\mathscr{F}}, \widetilde{P})$, with random variables $\{\tilde{f}_n\}$ and $\tilde{f}$, such that

$$\tilde{f}_n \overset{d}{=} f_n, \tilde{f} \overset{d}{=} f, \quad \tilde{f}_n \to \tilde{f}, a.s.$$

Proof. If $F_n \to F$ weakly, then $F_n^{\leftarrow} \to F^{\leftarrow}$ weakly. (Prove this yourself!)
Since $\mathbb{R}\backslash C(F_n^{\leftarrow})$ is countable, TODO □

If f is defined almost everywhere, we can extend it to $\tilde{f} = f \cdot \mathbf{I}_{N^c}$. So from now on when we talk about $f = g$, we mean $f = g, a.e.$.

§3.8 Review of first two sections

Here we list some concepts so that you can recall their definition and properties.
Collections of sets:

- π-system
- Semi-ring
- Ring, algebra
- σ-algebra
- Monotone class, λ-system

Measure:

- σ-finite
- Outer measure
- Caratheodory condition, measurable sets
- Measure extension, semi-ring $\to$ σ-algebra
- Complete measure space, completion
- For $\mathscr{F} = \sigma(\mathscr{A})$, $\forall F \in \mathscr{F}$, $\varepsilon > 0, \exists A \in \mathscr{A}$ s.t. $F = A \Delta N_\varepsilon$, $\mu(N_\varepsilon) \le \varepsilon$.

Functions:

- Measurable map
- $h \in \sigma(g) \implies h = f \circ g$ for some f.
- Typical method, simple non-negative functions $\to$ measurable functions
- Almost uniformly, almost everywhere, converge in measure

§4 Integrals

§4.1 Definition of Integrals

The idea of integration of f over μ is to compute the weighted sum of the values of f.
The definition of integrals is another example of typical method.

- For an indicator function $\mathbf{I}_A$, define $\int \mathbf{I}_A \, \mathrm{d}\mu = \mu(A)$.
- For simple function $f = \sum_{i=1}^n a_i \mathbf{I}_{A_i}$, just let $\int f \, \mathrm{d}\mu = \sum_{i=1}^n a_i \mu(A_i)$.
- For non-negative measurable function f, let $\int f \, \mathrm{d}\mu = \sup_{g \le f} \int g \, \mathrm{d}\mu$, where g is non-negative simple functions.

- For generic function f, write $f = f_+ - f_-$, define $\int f = \int f_+ - \int f_-$.

Definition 4.1 (Measurable partitions)**.** If a collection of sets $\{A_i\}$ satisfies

$$\mu(A_i \cap A_j) = 0, \quad \mu((\bigcup A_i)^c) = 0,$$

then we say $\{A_i\}$ is a **measurable partition** of X.

Definition 4.2 (Integrals for simple functions)**.** Let $\{A_i\}$ be a partition of X, $a_i \geq 0$ are reals. Let

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

define

$$\int_X f \,\mathrm{d}\mu := \sum_{i=1}^{n} a_i \mu(A_i).$$

Check it's well-defined: if $f = \sum_{j=1}^{m} b_j \mathbf{I}_{B_j}$, then

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_i m(A_i \cap B_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} b_j \mu(A_i \cap B_j).$$

Proposition 4.3

Let f, g be non-negative simple functions.

(1) $\int_X \mathbf{I}_A \,\mathrm{d}\mu = \mu(A), \quad \forall A \in \mathscr{F}$;

(2) $\int_X f \,\mathrm{d}\mu \geq 0$;

(3) $\int_X (af) \,\mathrm{d}\mu = a \int_X f \,\mathrm{d}\mu$;

(4) $\int_X (f + g) \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu + \int_X g \,\mathrm{d}\mu$;

(5) If $f \geq g$, then $\int_X f \,\mathrm{d}\mu \geq \int_X g \,\mathrm{d}\mu$.

(6) If $f_n \uparrow$ and $\lim_{n\to\infty} f_n \geq g$, then $\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu \geq \int_X g \,\mathrm{d}\mu$.

Remark 4.4 — $f :=\uparrow \lim_{n\to\infty} f_n$ need not be simple function. Even if f is simple, we don't know $\lim \int f_n \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu$ yet.

Proof of (4), (5). Since $\{A_i \cap B_j\}$ is a partition of X, on $A_i \cap B_j$,

$$f + g = a_i + b_j, \quad f = a_i, g = b_j.$$

□

Proof of (6). For all $\alpha \in (0, 1)$, let $A_n(\alpha) := \{f_n \geq \alpha g\} \uparrow X$. Then

$$f_n \mathbf{I}_{A_n(\alpha)} \geq \alpha g \mathbf{I}_{A_n(\alpha)}.$$

Thus if $g = \sum_{j=1}^m b_j \mathbf{I}_{B_j}$,

$$\int_X f_n \,\mathrm{d}\mu \geq \int_X f_n \mathbf{I}_{A_n(\alpha)} \,\mathrm{d}\mu \geq \alpha \int_X g \mathbf{I}_{A_n(\alpha)} \,\mathrm{d}\mu.$$

$$RHS = \alpha \sum_{j=1}^m b_j \mu(B_j \cap A_n(\alpha)) \uparrow \alpha \int_X g \,\mathrm{d}\mu.$$

Hence

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu \geq \alpha \int_X g \,\mathrm{d}\mu, \quad \forall \alpha < 1,$$

which completes the proof. □

Definition 4.5 (Integrals for non-negative measurable functions)**.** Let f be a non-negative measurable function. We know that $\exists f_1, f_2, \dots$ s.t. $f_n \uparrow f$. If we define the integral of f to be the limit of $\int f_n \,\mathrm{d}\mu$, we still need to prove this is well-defined. Therefore we use another definition:

$$\int_X f \,\mathrm{d}\mu := \sup\left\{\int_X g \,\mathrm{d}\mu : g \leq f \text{ is simple and non-negative}\right\}.$$

Proposition 4.6

Let f be a non-negative measurable function.

(1) If f is simple, then the two definition is the same.

(2) If $\{f_n\}$ is a series of simple non-negative functions, and $f_n \uparrow f$, then

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu.$$

(3)

$$\int_X f \,\mathrm{d}\mu = \lim_{n\to\infty} \left[\sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mu\left(\left\{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\right\}\right) + n\mu(\{f \geq n\})\right].$$

Proof of (2). By definition, $\int_X f_n \,\mathrm{d}\mu \leq \int_X f \,\mathrm{d}\mu$. Since for all simple function g, if $f_n \uparrow f \geq g$,

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu \geq \int_X g \,\mathrm{d}\mu.$$

Hence the desired equality holds. □

Remark 4.7 — The integral of f relies only on $\mu\big|_{\sigma(f)}$: if $f \in \mathscr{G} \subset \mathscr{F}$, then the integral of f is the same on $(X, \mathscr{G}, \mu\big|_{\mathscr{G}})$ and $(X, \mathscr{F}, \mu\big|_{\mathscr{F}})$.

Proposition 4.8

Continuing on the properties of integrals:

(1) $\int_X f \, \mathrm{d}\mu \geq 0$;

(2) $\int_X (af + g) \, \mathrm{d}\mu = a \int_X f \, \mathrm{d}\mu + \int_X g \, \mathrm{d}\mu$;

(3) If $f \geq g$, then $\int_X f \, \mathrm{d}\mu \geq \int_X g \, \mathrm{d}\mu$.

Proof. Use the previous proposition. □

Definition 4.9 (Integrals for generic functions)**.** Let f be a measurable function, and $f = f^+ - f^-$. If

$$\min\left\{\int_X f^+ \, \mathrm{d}\mu, \int_X f^- \, \mathrm{d}\mu\right\} < \infty,$$

we say the integral of f exists and define it to be

$$\int_X f \, \mathrm{d}\mu := \int_X f^+ \, \mathrm{d}\mu - \int_X f^- \, \mathrm{d}\mu.$$

If $\int_X f \, \mathrm{d}\mu \neq \pm\infty$, we say f is **integrable**.

For any $A \in \mathscr{F}$, $(A, \mathscr{F}_A, \mu_A)$ is a measure space. Define the integral of f on A to be

$$\int_A f \, \mathrm{d}\mu := \int_A f\big|_A \, \mathrm{d}\mu_A = \int_X f \mathbf{I}_A \, \mathrm{d}\mu.$$

where the latter equality holds since it holds for indicator functions.

Example 4.10 (The Lebesgue-Stieljes integral)

Let $(\mathbb{R}, \mathscr{B}_{\mathbb{R}}, \mu_F)$ be a measure space, where F is a quasi-distribution function. For a Borel function g,

$$\int_{\mathbb{R}} g \, \mathrm{d}F = \int_{\mathbb{R}} g(x) \, \mathrm{d}F(x) = \int_{\mathbb{R}} g(x) F(\mathrm{d}x) := \int_{\mathbb{R}} g \, \mathrm{d}\mu_F.$$

In particular, when $F(x) = x$, the integral is Lebesgue integral. Let λ be Lebesgue measure,

$$\int_{\mathbb{R}} g(x) \, \mathrm{d}x := \int_{\mathbb{R}} g \, \mathrm{d}\lambda.$$

If μ is a distribution, $F = F_\mu$, $g = \mathrm{id}$, we say

$$\int_{\mathbb{R}} x \, \mathrm{d}F(x) = \int_{\mathbb{R}} x \mu(\mathrm{d}x) = \int_{\mathbb{R}} \mathrm{id} \, \mathrm{d}\mu.$$

is the **expectation** of the distribution μ.

Example 4.11 (The integral on discrete measure)

Let $X = \{x_1, x_2, \dots\} = \{1, 2, \dots\}$, $\mu(\{x_i\}) = a_i$.

Let $I^+ = \{i : f(x_i) \geq 0\}, I^- = \{i : f(x_i) < 0\}$.

Let $I_n^+ = I^+ \cap \{1, \dots, n\}$, $f\mathbf{I}_{I_n^+}$ is a non-negative simple function and converges to f^+. Hence

$$\int_X f^+ \,\mathrm{d}\mu = \sum_{i \in I^+} f(x_i)a_i, \quad \int_X f^- \,\mathrm{d}\mu = -\sum_{i \in I^-} f(x_i)a_i.$$

$$\int_X f \,\mathrm{d}\mu = \sum_{i \in I} \sum_{i=1}^{\infty} f(x_i)a_i.$$

So f is integrable iff the series absolutely converges.

Theorem 4.12

Let f be a measurable function.

(1) If $\int_X f \,\mathrm{d}\mu$ exists, then $|\int_X f \,\mathrm{d}\mu| \leq \int_X |f| \,\mathrm{d}\mu$.

(2) f integrable $\iff$ $|f|$ integrable.

(3) If f is integrable, then $|f| < \infty, a.e.$.

Proof of (3). WLOG $f \geq 0$, then $f \geq f\mathbf{I}_{\{f=\infty\}}$.

$$\int_X f \,\mathrm{d}\mu \geq \int_X f\mathbf{I}_{\{f=\infty\}} \geq n\mu(\{f = \infty\}), \quad \forall n.$$

Thus $\mu(\{f = \infty\})$ must be 0. □

Theorem 4.13

Let f, g be measurable functions whose integral exists.

- $\int_A f \,\mathrm{d}\mu = 0$ for all null set A;
- If $f \geq g, a.e.$ then $\int_X f \,\mathrm{d}\mu \geq \int_X g \,\mathrm{d}\mu$.
- If $f = g, a.e.$, then their integrals exist simultaneously, $\int_X f \,\mathrm{d}\mu = \int_X g \,\mathrm{d}\mu$.

Proof. By definition, just check them one by one. □

Corollary 4.14

If $f = 0, a.e.$, then $\int_X f \,\mathrm{d}\mu = 0$; If $f \geq 0, a.e.$ and $\int_X f \,\mathrm{d}\mu = 0$, then $f = 0, a.e.$.

§4.2 Properties of integrals

Theorem 4.15 (Linearity of integrals)
Let f, g be functions whose integral exists.

- $\forall a \in \mathbb{R}$, the integral of af exists, and $\int_X (af)\,\mathrm{d}\mu = a\int_X f\,\mathrm{d}\mu$;
- If $\int_X f\,\mathrm{d}\mu + \int_X g\,\mathrm{d}\mu$ exists, then $f+g$ *a.e.* exists, its integral exists and

$$\int_X (f+g)\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu + \int_X g\,\mathrm{d}\mu.$$

Proof. The first one is trivial by definition.
As for the second,

1. First we prove $f+g$ *a.e.* exists. If $|f| < \infty, a.e.$, we're done.

   If $\mu(f=\infty) > 0$, then $\int_X f\,\mathrm{d}\mu = \infty$. This means $\int_X g\,\mathrm{d}\mu \neq -\infty$, so $\mu(g=-\infty)=0$. Thus $f+g$ *a.e.* exists. Similarly we can deal with the case $\mu(f=-\infty)>0$.

2. Next we prove the equality. $f+g = (f^+ + g^+) - (f^- + g^-)$. Let $\varphi = f^+ + g^+, \psi = f^- + g^-$. Our goal is

$$\int_X (\varphi - \psi)\,\mathrm{d}\mu = \int_X \varphi\,\mathrm{d}\mu - \int_X \psi\,\mathrm{d}\mu.$$

   Since $f+g$ *a.e.* exists, so $\varphi - \psi$ exists almost everywhere. If $\int_X \varphi\,\mathrm{d}\mu = \int_X \psi\,\mathrm{d}\mu = \infty$, then the integral of f, g must be $+\infty$ and $-\infty$, which contradicts with our condition. So both sides of above equation exist.

   Since $\max\{\varphi, \psi\} = \psi + (\varphi-\psi)^+ = \varphi + (\varphi - \psi)^-$, by the linearity of non-negative integrals,

$$\int_X \psi\,\mathrm{d}\mu + \int_X (\varphi-\psi)^+\,\mathrm{d}\mu = \int_X \varphi\,\mathrm{d}\mu + \int_X (\varphi-\psi)^-\,\mathrm{d}\mu.$$

   which rearranges to the desired equality.

Note: we need to verify that we didn't add infinity to the equation in the last step. □

Proposition 4.16
Let f, g be integrable functions, If $\int_A f\,\mathrm{d}\mu \geq \int_A g\,\mathrm{d}\mu, \forall A \in \mathscr{F}$, then $f \geq g, a.e.$.

Proof. Let $B = \{f < g\}$, then $(g-f)\mathbf{I}_B \geq 0$,

$$\int_B (g-f)\,\mathrm{d}\mu = \int_B (g-f)\mathbf{I}_B\,\mathrm{d}\mu \geq 0.$$

By the linearity of integrals we get $(g-f)\mathbf{I}_B = 0, a.e.$, i.e. $\mu(B) = 0$. □

Proposition 4.17
If μ is σ-finite, the integral of f, g exists, the conclusion of previous proposition also holds.

Proof. Let $X = \sum_n X_n$, $\mu(X_n) < \infty$. By looking at X_n, we may assume $\mu(X) < \infty$.

Since $\{f < g\} = \{-\infty \neq f < g\} + \{f = -\infty < g\}$.

Let $B_{M,n} = \{|f| \leq M, f + \frac{1}{n} < g\}$. By condition,

$$\int_{B_{M,n}} f \,\mathrm{d}\mu \geq \int_{B_{M,n}} g \,\mathrm{d}\mu \geq \int_{B_{M,n}} f \,\mathrm{d}\mu + \frac{1}{n}\mu(B_{M,n}).$$

Since $\int_{B_{M,n}} f \,\mathrm{d}\mu \leq M\mu(X)$ is finite, we get $\mu(B_{M,n}) = 0$. This implies $\{-\infty \neq f < g\} = \bigcup B_{M,n}$ is null.

Let $C_M = \{g > -M\}$, similarly,

$$-\infty \cdot \mu(C_M) = \int_{C_M} f \,\mathrm{d}\mu \geq \int_{C_M} g \,\mathrm{d}\mu = -M\mu(C_M).$$

Hence $\mu(C_M) = 0$, $\{-\infty = f < g\} = \bigcup C_M$ is null. □

Remark 4.18 — When $\geq$ is replaced by $=$, the conclusion holds as well. This proposition tells us that the integrals of f totally determines f. (In calculus, taking the derivative of integrals gives original functions)

Theorem 4.19 (Absolute continuity of integrals)
Let f be an integrable function, $\forall \varepsilon > 0$, $\exists \delta > 0$, such that $\forall A \in \mathscr{F}$,

$$\mu(A) < \delta \implies \int_A |f| \,\mathrm{d}\mu < \varepsilon.$$

Proof. Take non-negative simple functions $g_n \uparrow |f|$. Since $\int |f| \,\mathrm{d}\mu < \infty$, $\exists N$ s.t.

$$\int_X (|f| - g_N) \,\mathrm{d}\mu = \int_X |f| \,\mathrm{d}\mu - \int_X g_N \,\mathrm{d}\mu < \frac{\varepsilon}{2}.$$

Let $M = \max_{x \in X} g_N(x)$, $\delta = \frac{\varepsilon}{2M}$, so

$$\int_A |f| \,\mathrm{d}\mu < \frac{\varepsilon}{2} + \int_A g_N \,\mathrm{d}\mu = \frac{\varepsilon}{2} + M\mu(A) < \varepsilon.$$

□

Example 4.20
Fundamental theorem of Calculus, Lebesgue version: Let g be a measurable function, then g is absolutely continuous iff $\exists f : [a, b] \to \mathbb{R}$ Lebesgue integrable, s.t.

$$g(x) - g(a) = \int_a^x f(z) \,\mathrm{d}z.$$

The absolute continuity can be implied by the absolute continuity of integrals.

§4.3 Convergence theorems

Levi, Fatou, Lebesgue.

In this section we mainly discuss the commutativity of integrals and limits, i.e. if $f_n \to f$, we care when does the following holds:

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu.$$

Theorem 4.21 (Monotone convergence theorem, Levi's theorem)

Let $f_n \uparrow f, a.e.$ be non-negative functions, then

$$\int_X f_n \,\mathrm{d}\mu \uparrow \int_X f \,\mathrm{d}\mu.$$

Proof. By removing countable null sets, we may assume $0 \le f_n(x) \uparrow f$.

Take non-negative simple functions $f_{n,k} \uparrow f_n$. Let $g_k = \max_{1\le n\le k} f_{n,k}$ be simple functions.

$$g_k = \max_{1\le n\le k} f_{n,k} \le \max_{1\le n\le k+1} f_{n,k+1} = g_{k+1}.$$

So $g_k \uparrow$, say $g_k \to g$ for some function g. Clearly $g \le f$ as $g_k \le f_k$, $\forall k$.

Note as $k \to \infty$, $g_k \ge f_{n,k} \implies g \ge f_n, \forall n$. so $g = f$.

By definition of integrals,

$$\int_X f \,\mathrm{d}\mu = \lim_{k\to\infty} \int_X g_n \,\mathrm{d}\mu,$$

and

$$\int_X g_n \,\mathrm{d}\mu \le \int_X f_n \,\mathrm{d}\mu \le \int_X f \,\mathrm{d}\mu.$$

So the conclusion follows. □

Corollary 4.22

Let f_n be functions whose integrals exist, if

$$f_n \uparrow f, a.e. \quad \int_X f_1^- \,\mathrm{d}\mu < \infty, \quad \text{or} \quad f_n \downarrow f, a.e. \quad \int_X f_1^+ \,\mathrm{d}\mu < \infty,$$

then the integral of f exists, and $\int_X f_n \,\mathrm{d}\mu \to \int_X f \,\mathrm{d}\mu$.

Remark 4.23 — Counter example when $\int_X f_1^+ \,\mathrm{d}\mu = \infty$: let $X = \mathbb{R}$,

$$f_n = \mathbf{I}_{[n,\infty)} \downarrow f = 0, \quad \int_X f_n \,\mathrm{d}\mu = \infty, \quad \int_X f \,\mathrm{d}\mu = 0.$$

Corollary 4.24

If the integral of f exists, then for any measure partition $\{A_n\}$,

$$\int_X f \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_{A_n} f \,\mathrm{d}\mu.$$

If $f \geq 0$, then $\nu : A \mapsto \int_A f \,\mathrm{d}\mu$ is a measure on $\mathscr{F}$.