Measure Theory

Felix Chen

Contents

§1 Integrals

§1.1 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 1.1.1 (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 1.1.2 (Integrals for simple functions)**.** Let $\{A_i\}$ be a partition of X, $a_i \ge 0$ 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 1.1.3

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 1.1.4 — $f :=\uparrow \lim_{n\to\infty} f_n$ 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 1.1.5 (Integrals for non-negative measurable functions)**.** Let f be a non-negative measurable function. We know that $\exists f_1, f_2, \ldots$ 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 1.1.6

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)

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

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

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

Hence the desired equality holds. □

Remark 1.1.7 — 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 1.1.8

Continuing on the properties of integrals:

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

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

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

Proof. Use the previous proposition. □

Definition 1.1.9 (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 $\int_X f \,\mathrm{d}\mu \ne \pm\infty$, 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 1.1.10 (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 1.1.11 (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 1.1.12

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 1.1.13

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 1.1.14

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 $f = 0, a.e.$.

§1.2 Properties of integrals

Theorem 1.2.1 (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 1.2.2

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 1.2.3

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 1.2.4 — 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 1.2.5 (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 1.2.6

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.

§1.3 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 1.3.1 (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 1.3.2

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 1.3.3 — 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 1.3.4

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 $\mathscr{F}$. If we don't require $f \geq 0$, ν will become a signed measure which we'll cover later.

Theorem 1.3.5 (Fauto's Lemma)

Let $\{f_n\}$ be non-negative measurable functions almost everywhere, then

$$\int_X \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n \, \mathrm{d}\mu.$$

Proof. Let $g_k := \inf_{n\geq k} f_n \uparrow g := \liminf_{n\to\infty} f_n$. By monotone convergence theorem,

$$\int_X g \, \mathrm{d}\mu = \lim_{k\to\infty} \int_X g_k \, \mathrm{d}\mu \leq \lim_{k\to\infty} \inf_{n\geq k} \int_X f_n \, \mathrm{d}\mu = \liminf_{n\to\infty} \int_X f_n \, \mathrm{d}\mu.$$

$\square$

Corollary 1.3.6

If there exists integrable g s.t. $f_n \geq g$, then $\int_X \liminf_{n\to\infty} f_n$ exists and

$$\int_X \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n \, \mathrm{d}\mu.$$

Theorem 1.3.7 (Lebesgue)

Let $f_n \to f, a.e.$ or $f_n \xrightarrow{\mu} f$, if there exists non-negative integrable function g s.t. $|f_n| \leq g, \forall n$, then

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

Proof. When $f_n \to f, a.e.$, by Fatou's lemma,

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

Since $|f_n| \le g$,

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

which gives the desired.

When $f_n \xrightarrow{\mu} f$, for all subsequence $\{n_k\}$, exists a subsequence $\{n'\}$ s.t. $f_{n'} \to f, a.e.$.
Thus $\int_X f_{n'}\,\mathrm{d}\mu \to \int_X f\,\mathrm{d}\mu$, hence $\int_X f_n\,\mathrm{d}\mu \to \int_X f\,\mathrm{d}\mu$.(Why?) □

Corollary 1.3.8

Let f_n be random variable on $(\Omega_n, \mathscr{F}_n, P_n)$, $f_n \xrightarrow{d} f$, then we have

$$\lim_{n\to\infty} \int_{X_n} f_n\,\mathrm{d}P_n = \int_X f\,\mathrm{d}P.$$

Proposition 1.3.9 (Transformation formula of integrals)

Let $g : (X, \mathscr{F}, \mu) \to (Y, \mathscr{S})$ be a measurable map. For all measurable f on $(Y, \mathscr{S})$, then

$$\int_Y f\,\mathrm{d}\mu \circ g^{-1} = \int_X f \circ g\,\mathrm{d}\mu$$

if one of them exists.

Proof. By the typical method, we only need to prove for indicator function f. □

Remark 1.3.10 — μ and $\mu \circ g^{-1}$ are the same measure in different spaces.

§1.4 Expectations

Let ξ be a r.v. on $(\Omega, \mathscr{F}, P)$,

Definition 1.4.1 (Expectations)**.** If $\int_\Omega \xi\,\mathrm{d}P$ exists, then we call it the **expectation** of ξ, denoted by $E(\xi)$ or $E\xi$.

Consider the distribution $\mu_\xi = P \circ \xi^{-1}$, $F_\xi(x) = P(\xi \le x)$.
Let $f = \mathrm{id} : \mathbb{R} \to \mathbb{R}$, then $E(\xi) = E(\mu_\xi)$:

$$\int_{\mathbb{R}} x\,\mathrm{d}F_\xi(x) = \int_{\mathbb{R}} f\,\mathrm{d}\mu_\xi = \int_{\mathbb{R}} f\,\mathrm{d}P \circ \xi^{-1} = \int_{\mathbb{R}} f \circ \xi\,\mathrm{d}P = \int_\Omega \xi\,\mathrm{d}P = E(\xi).$$

Let f be a measurable function on $(\mathbb{R}, \mathscr{B}_{\mathbb{R}})$, then $f(\xi)$ is a measurable function on $(\Omega, \mathscr{F})$, and

$$Ef(\xi) = \int_{\mathbb{R}} f\,\mathrm{d}F_\xi.$$

Let $\eta = f \circ \xi$, by the transformation formula,

$$\begin{aligned} Ef(\xi) &= \int_\Omega \eta(\omega)\,\mathrm{d}P(\omega) \\ &= \int_{\overline{\mathbb{R}}} y\,\mathrm{d}P\circ\eta^{-1}(y) = \int_{\overline{\mathbb{R}}} y\,\mathrm{d}\mu_\eta(y) = \int_{\overline{\mathbb{R}}} y\,\mathrm{d}\mu_\xi\circ f^{-1}(y) \\ &= \int_{\mathbb{R}} f(x)\,\mathrm{d}\mu_\xi(x) = \int_{\mathbb{R}} f\,\mathrm{d}F_\xi. \end{aligned}$$

Example 1.4.2

Possion distribution: $P(\xi = i) = \frac{\lambda^i}{i!}e^{-\lambda} =: p_i$. Its expectation is

$$\int_{\mathbb{R}} x\,\mathrm{d}\mu = \int_{\mathbb{N}} x\,\mathrm{d}\mu = \sum_{i=0}^{\infty} ip_i.$$

For continuous distribution, the density function p is actually a non-negative, integrable function, and $\int_{\mathbb{R}} p(x)\,\mathrm{d}x = 1$. So $\mu(B) = \int_B p(x)\,\mathrm{d}x$ is a probability measure.

Since $\mu_\xi\big|_{\mathscr{P}_{\mathbb{R}}} = \mu\big|_{\mathscr{P}_{\mathbb{R}}}$, $\mu_\xi = \mu$. By typical method, we can prove

$$Ef(\xi) = \int_{\mathbb{R}} f\,\mathrm{d}\mu_\xi = \int_{\mathbb{R}} f(x)p(x)\,\mathrm{d}x.$$

§1.5 L_p spaces

Definition 1.5.1 (L_p spaces)**.** Let $1 \le p < \infty$. Define

$$\|f\|_p := \left(\int_X |f|^p\right)^{\frac{1}{p}}, \quad L_p(X, \mathscr{F}, \mu) := \{f : \|f\|_p < \infty\}.$$

Sometimes we'll simplify the notation as $L_p(\mu), L_p(\mathscr{F})$ or just L_p.

- $f \in L_1$ iff f integrable, let $\|f\| := \|f\|_1$.
- $f \in L_p \iff f^p \in L_1 \implies f$ is finite a.e..

In fact, L_p is a normed vector space under the norm $\|\cdot\|_p$:

Lemma 1.5.2

Let $1 \le p < \infty$, let $C_p = 2^{p-1}$, then

$$|a + b|^p \le C_p(|a|^p + |b|^p), \quad a, b \in \mathbb{R}.$$

Proof. It's a single-variable inequality, it's obvious by taking the derivative. □

Thus by taking integral on both sides,

$$\int_X |f + g|^p\,\mathrm{d}\mu \le C_p\left(\int_X |f|^p\,\mathrm{d}\mu + \int_X |g|^p\,\mathrm{d}\mu\right).$$

So L_p space is a vector space.

Lemma 1.5.3 (Holder's inequality)
Let $1 < p, q < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$.

$$\|fg\| \leq \|f\|_p \|g\|_q, \quad \forall f \in L_p, g \text{ measurable.}$$

Proof. WLOG $\|f\|_p > 0$, $0 < \|g\|_q < \infty$. Let

$$a = \left(\frac{|f|}{\|f\|_p}\right)^p = \frac{|f|^p}{\int_X |f|^p \,\mathrm{d}\mu}, \quad b = \left(\frac{|g|}{\|g\|_q}\right)^q = \frac{|g|^q}{\int_X |g|^q \,\mathrm{d}\mu}.$$

By weighted AM-GM,

$$\int_X \frac{|fg|}{\|f\|_p \|g\|_q} \,\mathrm{d}\mu \leq \int_X \left(\frac{a}{p} + \frac{b}{q}\right) \mathrm{d}\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

The equality holds iff $a = b$, i.e. $\exists \alpha, \beta \geq 0$ not all zero s.t. $\alpha |f|^p = \beta |g|^q, a.e.$. □

Theorem 1.5.4 (Minkowski's inequality)
Let $1 \leq p < \infty$, then

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p, \quad \forall f, g \in L_p.$$

The equality holds iff $(1) : p = 1, fg \geq 0; (2) p > 1, \exists \alpha, \beta \geq 0, s.t. \alpha f = \beta g, a.e.$.

Proof. When $p = 1$, it follows by $|f + g| \leq |f| + |g|$.
When $p \geq 1$, let $q = \frac{p}{p-1}$, by Holder's inequality,

$$|f + g|^p \leq |f||f + g|^{p-1} + |g||f + g|^{p-1},$$

$$\implies \|f + g\|_p^p \leq (\|f\|_p + \|g\|_p) \cdot \||f + g|^{p-1}\|_q.$$

Note that

$$\||f + g|^{p-1}\|_q = \left(\int_X |f + g|^p \,\mathrm{d}\mu\right)^{\frac{1}{q}} = \|f + g\|_p^{\frac{p}{q}}.$$

Since $f + g \in L_p$, we can divide both sides by $\|f + g\|_p^{\frac{p}{q}}$ to get the result. □

In L_p space, we view two functions $f = g, a.e.$ as the same function, i.e. the original function space modding the equivalence relation out.