

1 Sigma Algebra

A measure is the generalized volume or length of a subset. For instance, measure on the real line provides a generalized notion of length for subsets of the real numbers. The normal length measure of a closed interval on the real line $[a, b]$ is $b - a$.

We will consider a abstract measure theory. For a set X , let the set of its subsets (powerset) be denoted as $P(X)$.

Definition 1.1 (Measurable subsets). *For a set X , let set $\mathcal{A} \subseteq P(X)$ contain some subsets of X . The set \mathcal{A} is a σ -algebra if it fulfills*

1. *The empty and whole set must be in \mathcal{A} , where $\{\}, X \in \mathcal{A}$.*
2. *The set complement of any set in \mathcal{A} must also be in \mathcal{A} , where*

$$A \in \mathcal{A} \implies A^c = X \setminus A \in \mathcal{A}$$

3. *The union of a countably sized number of sets in \mathcal{A} is also in \mathcal{A} , where for a countable selection of A_i*

$$i \in \mathbb{N}, A_i \in \mathcal{A} \implies \bigcup_{i=0}^{\infty} A_i \in \mathcal{A}$$

For such a σ -algebra \mathcal{A} on the set X , the subsets of X contained in \mathcal{A} (or just members of the σ -algebra) are called measurable subsets.

Measurable sets are closed in the σ -algebra under complements and countable unions.

Some examples of σ -algebras on the set X are

- $\mathcal{A} = \{\emptyset, X\}$, this is the smallest σ -algebra
- $\mathcal{A} = P(X)$, this is the largest σ -algebra

2 Borel Sigma Algebras

Theorem 2.1 (Intersection of σ -algebras). *For an arbitrary amount of sigma algebras $i \in I, \mathcal{A}_i$ on X , the intersection of them is also a sigma algebra on X ,*

$$\bigcap_{i \in I} \mathcal{A}_i$$

Proof. Consider the sigma algebras \mathcal{A}_i on X and their intersections $\mathcal{B} = \bigcap \mathcal{A}_i$, we need to show that \mathcal{B} is also a sigma algebra.

For 1, note that because both the empty set and full set are in all sigma algebras, then they are also in \mathcal{B}

$$\forall i \in I, \emptyset, X \in \mathcal{A}_i \implies \emptyset, X \in \mathcal{B}$$

For 2, if a set $B \in \mathcal{B}$, then B is in all sigma algebras as well as B^c , so B^c is in \mathcal{B}

$$B \in \mathcal{B} \implies \forall i \in I, B \in \mathcal{A}_i \implies B^c \in \mathcal{A}_i \implies B^c \in \mathcal{B}$$

For 3, for some countable number of elements $B_k \in \mathcal{B}$, each is also in all sigma algebras $B_k \in \mathcal{A}_i$, hence the union of B_k is in all sigma algebras, and thus also in \mathcal{B}

$$k \in \mathbb{N}, B_k \in \mathcal{B} \implies \forall i \in I, B_k \in \mathcal{A}_i \implies \forall i \in I, \bigcup_{k=0}^{\infty} B_k \in \mathcal{A}_i \implies \bigcup_{k=0}^{\infty} B_k \in \mathcal{B}$$

□

Definition 2.1 (Smallest σ -algebra). *For any set $M \in P(X)$ that is a set of subsets in X , the smallest σ -algebra that contains the elements of M is the intersection of all σ -algebra that contains the elements of M .*

$$M \subseteq \bigcap_{M \subseteq \mathcal{A}} \mathcal{A} = \sigma(M)$$

and \mathcal{A} are all σ -algebras.

The smallest σ -algebra of M is denoted by $\sigma(M)$, or called the σ -algebra generated by M .

To practically construct the smallest σ -algebra of a set M , we progressively apply the definitions of σ -algebras and append new subsets into the set.

Definition 2.2 (Borel Sigma Algebra). *For the set X which is a metric space (consists of a set and a metric, like \mathbb{R}^n and Euclidean distance), we define the Borel Sigma algebra to be the sigma algebra generated by all the open sets in X .*

An open set S in X is a subset where for every element s in S , there exists a distance $\epsilon > 0$ such that $\forall x, d(x, s) < \epsilon \implies x \in S$. That is, for each element in S there exists a small disc around which is also contained in S .

The Borel Sigma algebra is in essence

$$B(X) = \sigma(M)$$

where M contains all the open sets in X .

The Borel σ -algebra on space X is importantly not the power set of the space, $P(X)$. We require such structure to define a measure function with appropriate features.

3 Measure

Definition 3.1 (Measurable Space). A measurable space consists of an ordered-tuple (X, \mathcal{A}) where X is a set and \mathcal{A} is a σ -algebra on the set X .

Definition 3.2 (Measure). For a measurable space (X, \mathcal{A}) , a mapping μ from the σ -algebra to the zero and positive real numbers $\mu: \mathcal{A} \rightarrow [0, \infty]$ is called a measure if it satisfies:

1. For the empty set $\emptyset \in \mathcal{A}$ from the σ -algebra \mathcal{A} , its volume/measure should be zero

$$\mu(\emptyset) = 0$$

2. The σ -additive rule. For any countable number of mutually disjoint measurable sets in our sigma algebra, the volume of their union is the sum of their individual volumes. So given

$$i \in \mathbb{N}, A_i \in \mathcal{A}, i \neq j \implies A_i \cap A_j = \emptyset$$

then we have

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

Note that the union of the measurable sets has a measure because of their closure under the σ -algebra.

The co-domain of the measure includes both the positive reals and a symbol infinity, where $[0, \infty] = [0, \infty) \cup \{\infty\}$. This space has some special rules:

1. $\forall x \in [0, \infty], x + \infty = \infty$
2. $\forall x \in (0, \infty], x \times \infty = \infty$
3. $0 \times \infty = 0$ for some cases only in measure theory.

Definition 3.3 (Measure Space). A measure space is a ordered set (X, \mathcal{A}, μ) where X is a set, \mathcal{A} is a σ -algebra on X , and μ is a measure on the measurable space.

3.1 Examples of Measures

The following measures works on all sets X and any of their their σ -algebras \mathcal{A} .

The counting measure is defined by the mapping

$$\mu(A) = \begin{cases} |A| & A \text{ is finite sized} \\ \infty & \text{otherwise} \end{cases}$$

where $A \in \mathcal{A}$ is a member of the σ -algebra. The measure axioms for the finite cases is obvious to show. The counting measure maps subsets to their sizes as generalized volumes.

The Dirac measure on a point $p \in X$ maps subsets containing p to 1 and subsets not containing it to 0, it is defined by

$$\delta_p(A) = \begin{cases} 1 & p \in A \\ 0 & p \notin A \end{cases}$$

To search for a measure that generalizes volume in Euclidean space $X = \mathbb{R}^n$, we require that our measure adhere to some basis axioms:

1. The measure of a unit volume/cube (set of points from $[0, 1]$ in all dimensions) is always 1

$$\mu([0, 1]^n) = 1$$

2. The measure is invariant under translation by a vector

$$\forall x \in \mathbb{R}^n, \forall A \in \mathcal{A}, \mu(x + A) = \mu(A)$$

One measure that works is the Lebesgue measure, but it only works on σ -algebras that are not simply the powerset — it operates on the Borel σ -algebras on a metric space.

4 Lebesgue Measures On the Real Powerset Proof

Definition 4.1 (Equivalence class). *An equivalence class within the set X defined by the equivalence relation $x \sim y$ on $a \in X$ is*

$$[a] = \{x \in X : x \sim a\} \subseteq X$$

Definition 4.2 (Axiom of Choice). *For an indexed set of subsets (of set X), $i \in I, S_i \subseteq X$, there exists a set A containing a single element from each set*

$$i \in I, a_i = f(S_i), A = \{a_0, a_1, \dots\}$$

where $f: P(X) \rightarrow X$ is a choice function on the subsets S_i .

Theorem 4.1 (Monotonic Measures). *Consider a measure μ on a measurable space (X, \mathcal{A}) . If for some $A, B \in \mathcal{A}$, $A \subseteq B$, then*

$$\mu(A) \leq \mu(B)$$

Proof. For the given sets $A, B \in \mathcal{A}$ with $A \subseteq B$, let $C = B \setminus A$ so that $A \cup C = B$. Then

$$\mu(B) = \mu(A \cup C) = \mu(A) + \mu(C) \geq \mu(A)$$

for $\mu(C) \geq 0$. □

Lemma 4.2 (Powerset Zero Measure). *For all measures μ on the measurable space $(\mathbb{R}, P(\mathbb{R}))$ that satisfies*

1. $\mu((0, 1]) < \infty$
2. $\forall x \in \mathbb{R}, A \in P(\mathbb{R}), \mu(x + A) = \mu(A)$

the measure must be the zero measure

$$\forall A \in P(\mathbb{R}), \mu(A) = 0$$

Proof. Within the interval $I = (0, 1]$, consider the equivalence relation defined by

$$\forall x, y \in I, x \sim y \iff x - y \in \mathbb{Q}$$

We can partition the interval I with some infinite equivalence relations that may or may not be countable

$$I = \bigcup_i [a_i], a_i \in I$$

and notice that each equivalence class is disjoint with all others.

Using the Axioms of Choice, we create a set A by picking one unique element from each equivalence class

$$A = \{a_0, a_1, \dots\} \subseteq I$$

namely, the set A has the property that

1. For any equivalence class, there is an element in A that is in that class.

$$\forall[a_i], \exists a \in A, a \in [a_i]$$

2. That the element in (1) is unique in A .

$$\forall[a_i], \forall a, b \in A, a, b \in [a_i] \implies a = b$$

Consider the enumerated rational translations of such set A , where $r_n, n \in \mathbb{N}$ is an enumeration of rational numbers within $(-1, 1]$

$$A_n = A + r_n$$

We note that these translated sets are disjoint, namely

$$n \neq m \implies A_n \cap A_m = \emptyset$$

this can be proved by contraposition

$$\begin{aligned} A_n \cap A_m \neq \emptyset &\implies \exists x, x \in A_n \wedge x \in A_m \\ &\implies \exists a_n, a_m \in A, x = r_n + a_n = r_m + a_m \\ &\implies a_n - a_m = r_m - r_n \in \mathbb{Q} \\ &\implies a_n \sim a_m \\ &\implies a_n = a_m \\ &\implies r_m = r_n \implies n = m \end{aligned}$$

Then notice that both

$$(0, 1] \subseteq \bigcup_n A_n \subseteq (-1, 2]$$

as

$$\begin{aligned} x \in (0, 1] &\implies \exists a \in A, a \in [x] \\ &\implies \exists r \in \mathbb{Q}, a + r = x \\ &\implies \exists n, r_n = r, x \in A_n \\ &\implies x \in \bigcup_n A_n \\ x \in \bigcup_n A_n &\implies \exists n, x \in A_n \\ &\implies x \in r_n + A \\ r_n \in (-1, 1], A \subseteq (0, 1] &\implies -1 < x \leq 2 \\ &\implies x \in (-1, 2] \end{aligned}$$

Because measures are monotonic,

$$\mu((0, 1]) \leq \mu\left(\bigcup_n A_n\right) \leq \mu((-1, 2])$$

Suppose that $\mu((0, 1]) = c < \infty$

$$\mu((-1, 2]) = \mu((-1, 0] \cup (0, 1] \cup (1, 2]) = 3c$$

Then because $\mu(A_n) = \mu(r_n + A) = \mu(A)$ we have

$$c \leq \sum_n \mu(A) \leq 3c$$

As c is finite, the middle summation must also be finite. For n is countably infinite, the only way this identity holds is if $\mu(A) = 0$ as anything else results in infinity.

Hence $c = 0$, with

$$\mu(\mathbb{R}) = \mu\left(\bigcup_{z \in \mathbb{Z}} (0, 1] + z\right) = \sum_z \mu((0, 1]) = 0$$

As any subset of \mathbb{R} must also have zero measure (otherwise the measure for \mathbb{R} is non-zero)

$$\forall A \in P(\mathbb{R}) \implies A \subseteq \mathbb{R} \implies \mu(A) = 0$$

□

Theorem 4.3 (Non Lebesgue Measures). *For the measurable space $(X, P(X))$ where the set is the real number line, $X = \mathbb{R}$, there does not exist a Lebesgue measure — a measure that follows geometric intuition $\mu: P(X) \rightarrow [0, \infty]$ where*

1. $\mu([0, 1]) = 1$
2. $\forall x \in X, A \in P(X), \mu(x + A) = \mu(A)$

Proof. By the previous lemma, if such measure μ exists,

$$1 = \mu([0, 1]) = \mu(\{0\} \cup (0, 1]) \implies \mu((0, 1]) = 1 - \mu(\{0\}) < \infty$$

Therefore $\mu = 0$, which is a contradiction to the fact that $\mu([0, 1]) = 1 \neq 0$. Hence a measure μ can't exist. □

5 Measurable Maps

Definition 5.1 (Measurable Maps). *Given two measurable spaces $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$, a mapping $f: \Omega_1 \rightarrow \Omega_2$ is a measurable map when the pre-image of measurable sets are also measurable sets*

$$\forall A \in \mathcal{A}_2, f^{-1}(A) \in \mathcal{A}_1$$

This definition of measurable maps is important in defining an integral of a function f . If the function is measurable, then the pre-image of any measurable set of its range is also measurable on its domain, so we can take the measure of the pre-image and multiply it against the height (summing for every height) to compute the integral.

Theorem 5.1 (Composition of Measurable Mappings). *For the measurable spaces $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$, $(\Omega_3, \mathcal{A}_3)$ and the mappings $f: \Omega_1 \rightarrow \Omega_2$ and $g: \Omega_2 \rightarrow \Omega_3$. If both f and g are measurable with respect to the σ -algebras, then their composition $g \circ f: \Omega_1 \rightarrow \Omega_3$ is also measurable with the same σ -algebras.*

Proof. Consider the measurable set $S \in \mathcal{A}_3$. Using the identity of pre-images

$$(g \circ f)^{-1}(S) = f^{-1}(g^{-1}(S))$$

As g is measurable, $g^{-1}(S) \in \mathcal{A}_2$ is also a measurable set; As f is measurable, $f^{-1}(g^{-1}(S)) \in \mathcal{A}_1$ is also a measurable set. Hence the pre-image of S is a measurable set and the composition is also measurable. \square

5.1 Examples of Measurable Maps

Consider the measurable spaces (Ω, \mathcal{A}) and $(\mathbb{R}, B(\mathbb{R}))$, the characteristic map on set $A \subseteq \Omega$ is defined to be

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

The characteristic map is measurable when A is measurable $A \in \mathcal{A}$, for let $B = B(\mathbb{R})$:

1. if B contains neither 0 or 1, then its pre-image is the empty set, which is in \mathcal{A}
2. if B contains only 0 or 1, then its pre-image is either A or A^c , which are both in \mathcal{A}
3. if B contains both 0 and 1, then its pre-image is Ω which is in \mathcal{A}

The map is measurable because all pre-images of measurable sets in the co-domain is also a measurable set in the domain with respect to \mathcal{A} .

Theorem 5.2 (Measurable Real Number Functions). *For the measurable spaces (Ω, \mathcal{A}) and $(\mathbb{R}, B(\mathbb{R}))$. If the functions $f, g: \Omega \rightarrow \mathbb{R}$ are both measurable, then*

$$f + g, f - g, f \cdot g, |f|$$

are all measurable functions.

Proof. This proof depends on the properties of the Borel σ -algebra on the real numbers, which is really difficult. Take this theorem as given please. \square

Continuous functions in the real numbers are always measurable.

6 Lebesgue Integral

We will first define the Lebesgue integral on step/simple functions.

Consider a measure space consist of (X, \mathcal{A}, μ) , where X is a set, \mathcal{A} is a σ -algebra on X , and $\mu: \mathcal{A} \rightarrow [0, \infty]$ is a measure. We aim to define a notion of integration for a measurable map $f: X \rightarrow \mathbb{R}$ with respect to \mathcal{A} and $B(\mathbb{R})$.

Definition 6.1 (Characteristic Function Integrals). *For any characteristic function $\chi_A: X \rightarrow \mathbb{R}$ where $A \in \mathcal{A}$ is a measurable set, we define its integral to be the measure of the set*

$$I(\chi_A) = \mu(A)$$

Definition 6.2 (Simple Functions). *Simple functions (step functions) are linear combinations of characteristics functions. They are defined by*

$$f(x) = \sum_{i=1}^n c_i \chi_{A_i}(x)$$

for some $n \in \mathbb{N}$, $c_i \in \mathbb{R}$, and measurable sets $A_i \in \mathcal{A}$.

A function is simple if we can find measurable sets and constants such that their linear combination with the characteristic functions forms the function.

Because scalar multiples of measurable maps are measurable, and sums of measurable maps are measurable, simple functions are also measurable.

Notice that while the linear combination representation of simple functions are not unique, the resultant functions are. We should choose a suitable representation of simple functions for the task.

Definition 6.3 (Set of positive simple functions). *Define the set of positive simple functions S^+ by the set of simple functions with positive (or zero) coefficients*

$$S^+ = \{f: X \rightarrow \mathbb{R}: f \text{ is simple, } f \geq 0\}$$

This resembles a half vector space where addition and positive scalar multiplication are closed.

Theorem 6.1 (Simple Function Integrals). *For a simple function $f \in S^+$, choose any representation of the function*

$$f(x) = \sum_{i=1}^n c_i \chi_{A_i}(x), c_i \geq 0$$

The Lebesgue integral on the simple function f with respect to the measure μ is given by the linear combination

$$I(f) = \sum_{i=1}^n c_i \mu(A_i) \in [0, \infty]$$

The Lebesgue Integral is well-defined, where the value of the integral does not depend on the representation of the simple function.

The alternate syntax of the Lebesgue Integral is

$$\int_X f(x) d\mu = \int_X f d\mu = I(f) = \sum c_i \mu(A_i)$$

The properties of the Lebesgue integral are

1. That it is linear:

$$\forall a, b \geq 0, I(af + bg) = aI(f) + bI(g)$$

for positive simple functions $f, g \in S^+$.

2. Monotonicity, where for positive simple functions $f, g \in S^+$

$$\forall x \in X, f(x) \leq g(x) \implies I(f) \leq I(g)$$

Both property are trivially verified by the definition of Lebesgue Integral on simple functions.

To approximate a generic measurable positive function $f: X \rightarrow [0, \infty)$, we partition the real axis into finite sized heights $\{c_1, c_2, \dots\}$, then define the simple function $h \in S^+$ by

$$h(x) = \sum_{i=1}^{\infty} c_i \chi_{A_i}(x)$$

where A_i is the pre-image of the function f within each height gap

$$A_i = f^{-1}([c_i, c_{i+1}))$$

Notably, because we use the lower-bound of the discrete heights in our simple function for every pre-image domain, the simple function $h \leq f$. In theory, we can refine the heights so that our approximate simple function approaches the generic measurable positive function.

Definition 6.4 (Supremum and Infimum). *For a set $X \subseteq U$, define the supremum of the set to be the lowest upper-bound. That is*

$$\sup X = \min\{M \in U : \forall x \in X, x \leq M\}$$

Similarly, the infimum of the set is the highest lower-bound,

$$\inf X = \max\{M \in U : \forall x \in X, x \geq M\}$$

On the real numbers with standard orderings, the supremum and infimum of a set always exists.

Definition 6.5 (Lebesgue Integral). *For a positive measurable map $f: X \rightarrow \mathbb{R}, f \geq 0$ in a measure space (X, \mathcal{A}, μ) and $(\mathbb{R}, B(\mathbb{R}), \lambda)$, define the Lebesgue Integral on f with respect to the measure μ to be*

$$\int_X f d\mu = I(f) = \sup\{I(h) : h \in S^+, h \leq f\} \in [0, \infty]$$

Namely, the integral is the smallest upper-bound of all integrals of simple functions smaller than the measurable map.

A mapping f is called μ -integrable if it is finite $\int_X f d\mu < \infty$.

Definition 6.6 (Almost Everywhere). *The statement $p(x)$ almost everywhere where $p(x)$ is a predicate on a measure space (X, \mathcal{A}, μ) implies that the measure of the subset on X where the predicate don't hold is zero:*

$$\mu(\{x \in X : \neg p(x)\}) = 0$$

If $p(x)$ holds for all $x \in X$, then it must be $p(x)$ almost everywhere (but not the contrary).

For two positive measurable maps $f, g: X \rightarrow [0, \infty)$, properties of the Lebesgue Integral include

1. The equality,

$$(f = g) \text{ with } \mu \text{ almost everywhere} \implies \int_X f d\mu = \int_X g d\mu$$

2. Monotonicity,

$$f \leq g \text{ with } \mu \text{ almost everywhere} \implies \int_X f d\mu \leq \int_X g d\mu$$

3. The identity

$$f = 0 \text{ with } \mu \text{ almost everywhere} \iff \int_X f d\mu = 0$$

Lemma 6.2 (Measure Equivalent Simple functions). *For any simple function $h \in S^+$ on a measure space (X, \mathcal{A}, μ) , consider the set of functions $h_a(x)$ defined by*

$$h_a(x) = \begin{cases} h(x) & x \in \tilde{X} \\ a & x \in \tilde{X}^c \end{cases}$$

where $a \in \mathbb{R}^+$ is a positive real number and $\tilde{X} \in \mathcal{A}$ is a measurable set with its complement having zero measure $\mu(\tilde{X}^c) = 0$.

Then

$$\int_X h d\mu = \int_X h_a d\mu$$

Proof. First notice that all simple functions have finite domains, hence have a canonical linear representation

$$h(x) = \sum_{c \in h(X)} c \chi_{\{x \in X : h(x)=c\}}(x)$$

Therefore the canonical representation of $h_a(x)$ is

$$\begin{aligned}
h_a(x) &= a\chi_{\{x \in X: h_a(x)=a\}}(x) + \sum_{c \in h_a(X) \setminus \{a\}} c\chi_{\{x \in X: h_a(x)=c\}}(x) \\
&= a\chi_{\tilde{X}^c}(x) + a\chi_{\{x \in \tilde{X}: h_a(x)=a\}}(x) + \sum_{c \in h(\tilde{X}) \setminus \{a\}} c\chi_{\{x \in \tilde{X}: h_a(x)=c\}}(x) \\
&= a\chi_{\tilde{X}^c}(x) + \sum_{c \in h(\tilde{X})} c\chi_{\{x \in \tilde{X}: h_a(x)=c\}}(x) \\
&= a\chi_{\tilde{X}^c}(x) + \sum_{c \in h(\tilde{X})} c\chi_{\{x \in \tilde{X}: h(x)=c\}}(x) \\
&= a\chi_{\tilde{X}^c}(x) + \sum_{c \in h(X)} c\chi_{\{x \in \tilde{X}: h(x)=c\}}(x)
\end{aligned}$$

the last simplification is possible for if $c \in h(\tilde{X}^c)$ and $c \notin h(\tilde{X})$, the set $\{x \in \tilde{X}: h(x) = c\} = \emptyset$.

Hence,

$$\begin{aligned}
\int_X h d\mu &= \sum_{c \in h(X)} c\mu(\{x \in X: h(x) = c\}) \\
&= \sum_{c \in h(X)} c \left(\mu(\{x \in \tilde{X}: h(x) = c\}) + \mu(\{x \in \tilde{X}^c: h(x) = c\}) \right) \\
&= a\mu(\tilde{X}^c) + \sum_{c \in h(X)} c\mu(\{x \in \tilde{X}: h(x) = c\}) \\
&= \int_X h_a d\mu
\end{aligned}$$

for both

$$0 = \mu(\tilde{X}^c) = a\mu(\tilde{X}^c) = \mu(\{x \in \tilde{X}^c: h(x) = c\}) = 0$$

□

To prove the first (and second) monotonicity properties, let

$$\tilde{X} = \{x \in X: f(x) \leq g(x)\} \implies \mu(\tilde{X}^c) = 0$$

then we have

$$\begin{aligned}
\int_X f d\mu &= \sup\{I(h): h \in S^+, \forall x \in X, h(x) \leq f(x)\} \\
&= \sup\{I(h_a): h \in S^+, \forall x \in X, h(x) \leq f(x)\} \\
&= \sup\{I(h_a): h \in S^+, \forall x \in \tilde{X}, h(x) \leq f(x)\} \\
&\leq \sup\{I(h_a): h \in S^+, \forall x \in \tilde{X}, h(x) \leq g(x)\} \\
&= \sup\{I(h): h \in S^+, \forall x \in X, h(x) \leq g(x)\} \quad (\text{mirroring the first 3 statements}) \\
&= \int_X g d\mu
\end{aligned}$$

The third property in the forward direction is obvious from (1), the reverse is harder.

7 Monotone Convergence Theorem

The monotone convergence theorem describes how the Lebesgue integral interacts with infinite series. That if a series of increasing functions converges to another function, and the limit of the increasing functions' Lebesgue integrals also converges to the Lebesgue integral of the other function.

Theorem 7.1 (Monotone Convergence Theorem). *For a measure space (X, \mathcal{A}, μ) and a countable series of measurable functions*

$$n \in \mathbb{N}, f_n(x): X \rightarrow [0, \infty)$$

where

$$f_1 \leq f_2 \leq f_3 \leq \dots \quad \mu \text{ almost everywhere}$$

and there is a limit function $f(x): X \rightarrow [0, \infty)$ where

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad \mu \text{ almost everywhere}$$

Then the function $f(x)$ is measurable and

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = \int_X f d\mu$$

Proof.

□