Basic Real Analysis

January 29, 2002

1 Notations and Terminologies

A **space** Ω is an arbitrary, *non-empty* set, whose elements are denoted generically by ω. The class of all subsets of Ω is denoted by $\mathcal{P}(\Omega)$. For any subset $E \subseteq \Omega$, its **complement** is defined as $E^c = \Omega \backslash E = \{\omega;\ \omega \notin E\}$.

Definition: A non-empty subset $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ is said to be an **algebra** if it is closed under *finite* unions and complements; that is

1. $E^c \in \mathcal{A}$ whenever $E \in \mathcal{A}$.
2. $\cup_{j=1}^n E_j \in \mathcal{A}$ whenever $E_j \in \mathcal{A}$, $1 \leq j \leq n$, for all n.

Moreover, $\mathcal{A}$ is said to be an **σ-algebra** if, in addition, $\mathcal{A}$ is closed under *countable* unions. That is,

3. $\cup_{j=1}^\infty E_j \in \mathcal{A}$ whenever $E_j \in \mathcal{A}$, $j \geq 1$.

Exercise: Show that an algebra is an σ-algebra if it is closed under countable disjoint unions.

Exercise: For any non-empty subset of $\mathcal{P}(\Omega)$, say $\mathcal{D}$, define $\sigma(\mathcal{D})$ as the intersection of all σ-algebras that contain $\mathcal{D}$. Show that $\sigma(\mathcal{D})$ is the *smallest* σ-algebra that contains $\mathcal{D}$. We say $\sigma(\mathcal{D})$ is the **smallest σ-algebra generated by** $\mathcal{D}$.

Definition: Let (Ω, ρ) be a metric space, and $\mathcal{O}$ be the collection of all open sets. Then $\mathcal{B} = \mathcal{B}(\Omega) \doteq \sigma(\mathcal{O})$ is said to be the **Borel σ-algebra** of Ω.

Exercise: Show that the Borel σ-algebra $\mathcal{B}(\mathsf{R})$ of the real line is also generated by each of the following family:

$$\mathcal{E}_1 = \{(a,b);\ a < b\}\,,\quad \mathcal{E}_2 = \{(a,b];\ a < b\}\,,\quad \mathcal{E}_3 = \{[a,b);\ a < b\}\,,\quad \mathcal{E}_4 = \{[a,b];\ a < b\}$$

Definition: A class of subsets $\mathcal{H} \subseteq \mathcal{P}(\Omega)$ is said to be an **elementary family** if

(1). $\emptyset \in \mathcal{H}$;

(2). $E \cap F \in \mathcal{H}$ whenever $E \in \mathcal{H},\ F \in \mathcal{H}$;

(3). for every $E \in \mathcal{H}$, its complement $E^c = \cup_{j=1}^n F_j$, where $(F_j)_1^n \subseteq \mathcal{H}$ are disjoint.

Example: Let $\mathcal{H}$ = collection of intervals of form $(a, b]$, $(a, \infty), \emptyset$ where $-\infty \leq a < b < \infty$. It is easy to show that $\mathcal{H}$ is an elementary family.

Exercise: Suppose $\mathcal{H}$ is an elementary family. Show that the class

$$\mathcal{E} \doteq \left\{ \cup_{j=1}^{n} E_j; \quad (E_j)_1^n \subseteq \mathcal{H} \text{ are disjoint}, \ n \geq 1 \right\}$$

is an algebra.

Definition: A nonempty class $\mathcal{D} \subseteq \mathcal{P}(\Omega)$ is said to be a π**-class** if $E \cap F \in \mathcal{D}$ whenever both $E, F \in \mathcal{D}$, whereas it is said to be a λ**-class** if

(1). $\Omega \in \mathcal{D}$;

(2). $A \backslash B \in \mathcal{D}$ if both $A, B \in \mathcal{D}$ and $A \supseteq B$;

(3). for every increasing sequence $\{A_n; \ n = 1, 2, \cdots\} \subseteq \mathcal{D}$, the limit $\lim A_n \doteq \cup A_n \in \mathcal{D}$.

Exercise: If $\mathcal{D}$ is both a π-class and a λ-class, then $\mathcal{D}$ is a σ-algebra.

Dynkin system theorem: If a λ-class $\mathcal{A}$ contains a π-class $\mathcal{D}$, then $\mathcal{A}$ also contains $\sigma(\mathcal{D})$, the smallest σ-algebra generated by $\mathcal{D}$.

Proof. Suppose $\mathcal{E}$ is the *minimal* λ-class that contains $\mathcal{D}$. Indeed, $\mathcal{E}$ is the intersection of all the λ-classes that contain $\mathcal{D}$. By definition, $\mathcal{A} \supseteq \mathcal{E}$. It suffices to show that $\mathcal{E}$ is an σ-algebra, or $\mathcal{E}$ is a π-class. To this end, define

$$\mathcal{E}_1 \doteq \{A; \ A \subseteq \Omega, \ A \cap D \in \mathcal{E}, \quad \text{for all } D \in \mathcal{D}\}.$$

It is easy to see that $\mathcal{E}_1$ is a λ-class that contains $\mathcal{D}$ (why?). Therefore $\mathcal{E}_1 \supseteq \mathcal{E}$, or $A \cap D \in \mathcal{E}$ whenever $A \in \mathcal{E}, \ D \in \mathcal{D}$. Now define

$$\mathcal{E}_2 \doteq \{E; \ E \subseteq \Omega, \ A \cap E \in \mathcal{E}, \quad \text{for all } A \in \mathcal{E}\}.$$

It follows that $\mathcal{E}_2$ is a λ-class that contains $\mathcal{D}$, in particular, $\mathcal{E}_2 \supseteq \mathcal{E}$. This impliese that $A \cap E \in \mathcal{E}$ whenever $A, E \in \mathcal{E}$. This completes the proof. □

2 Measurable space and Measure

If $\mathcal{F}$ is an σ-algebra of Ω, then the pair $(\Omega, \mathcal{F})$ is said to be a **measurable space**.

Definition: A set function $\mu: \ \mathcal{F} \to [0, \infty]$ is said to be a **measure** on the measurable space $(\Omega, \mathcal{F})$ if

(1). $\mu(\emptyset) = 0$;

(2). it is *countably additive*, that is,

$$\mu\left(\cup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$$

for any sequence $\{E_n; \ n = 1, 2, \cdots\} \subseteq \mathcal{F}$ of which E_n's are disjoint.

The triple $(\Omega, \mathcal{F}, \mu)$ is said to be a **measure space**. In particular, if $\mu(\Omega) = 1$, then the triple is said to be a **probability space**.

Remark: A measure μ is said to be a **finite** if $\mu(\Omega)$ is finite, and said to be σ-**finite** if there exists a sequence of subsets $\{A_n;\ n = 1, 2, \cdots\}$ such that $\Omega = \cup A_n$ and $\mu(A_n)$ is finite for all n.

Below are several example of measure spaces.

Example: Suppose $\Omega = \{\omega_1,\ \omega_2, \cdots,\ \omega_N\}$. The *counting measure* on Ω is defines as $\mu(E) = \text{card}(E)$ for all $E \subseteq \Omega$.

Example: The *Lebesgue-measure* λ on space $(\mathsf{R}, \mathcal{B}(\mathsf{R}))$ is a measure such that

$$\lambda\big\{(a,b)\big\} = \lambda\big\{(a,b]\big\} = \lambda\big\{[a,b)\big\} = \lambda\big\{[a,b]\big\} = b - a$$

The Lebesgue measure $\lambda^{(n)}$ on space $(\mathsf{R}^n, \mathcal{B}(\mathsf{R}^n))$ is such that

$$\lambda^{(n)}\left(\times_{j=1}^n I_j\right) = \prod_{j=1}^n \lambda(I_j)$$

where I_j are intervals on R.

Example: Suppose $(f_n;\ \ n = 0, 1, \cdots)$ is a sequence of non-negative numbers such that $\sum_{n=1}^\infty f_n = 1$. The measure

$$\mu\big(\{n\}\big) = f_n, \qquad \text{or} \qquad \mu(A) = \sum_{n \in A} f_n$$

defines a probability measure on natural numbers N_0. In particular, if $f_0 = p,\ f_1 = 1 - p$, the measure is said to be the *Bernoulli measure*. A *Binomial measure* is such that

$$f_k = \binom{n}{k} p^k (1-p)^{n-k}; \qquad k = 0, 1, 2, \cdots, n$$

If $p = p_n = \frac{\lambda}{n}$, then

$$f_k^{(n)} = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \quad \to \quad e^{-\lambda}\frac{\lambda^k}{k!} := u_k, \quad \text{as } n \to \infty; \quad k = 0, 1, 2, \cdots$$

The measure defines by $\{u_k;\ k = 0, 1, 2, \cdots\}$ is a *Poisson measure*.

Example: At the other extreme, for a function $f:\ \mathsf{R} \to [0, \infty)$ with $\int_{\mathsf{R}} f(x)\,dx = 1$, one can define a probability measure on $\big(\mathsf{R}, \mathcal{B}(\mathsf{R})\big)$ by

$$\mu(E) = \int_E f(x)dx, \qquad \forall\ E \in \mathcal{B}(R).$$

The function f is said to be the *density function*. Below are some examples.

(1). Exponential density function $\text{Exp}(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ; \quad x \geq 0 \\ 0 & ; \quad x < 0 \end{cases}$$

(2). Normal density function $N(\mu, \sigma)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \forall\, x \in \mathsf{R}$$

Theorem: For any measure space $(\Omega, \mathcal{F}, \mu)$, we have

(1). $\mu(E) \le \mu(F)$ whenever both $E, F \in \mathcal{F}$ and $E \subseteq F$.

(2). $\mu\left(\cup_{n=1}^{\infty} E_n\right) \le \sum_{n=1}^{\infty} \mu(E_n)$ for any sequence $\{E_n;\ n = 1, 2, \cdots\} \subseteq \mathcal{F}$.

(3). for any increasing sequence $E_1 \subseteq E_2 \subseteq \cdots$, $\mu\left(\cup_{n=1}^{\infty} E_n\right) = \lim_n \mu(E_n)$.

(4). for any decreasing sequence $E_1 \supseteq E_2 \supseteq \cdots$, $\mu\left(\cap_{n=1}^{\infty} E_n\right) = \lim_n \mu(E_n)$, provided that $\mu(E_k) < \infty$ for some k. In particulr, if μ is a probability measure, this condition automatically holds.

Remark: The condition in (4) is necessary. Indeed, let μ be the Lebesgue-measure on the real line and $E_n = (n, \infty)$. We have $\mu(E_n) \equiv \infty$ and $\cap_n E_n = \emptyset$.

Proof. The assertions (1), (2) are trivial. To show (3), define $F_n \doteq E_n \backslash E_{n-1}$ with convention $E_0 = \emptyset$. It is easy to see that (F_n) are disjoint and $\cup_{n=1}^{\infty} E_n = \cup_{n=1}^{\infty} F_n$. Therefore,

$$\mu\left(\cup_{n=1}^{\infty} E_n\right) = \mu\left(\cup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mu(F_n) = \sum_{n=1}^{\infty} \left[\mu(E_n) - \mu(E_{n-1})\right] = \lim_n \mu(E_n).$$

As for (4), assume $k = 1$ without loss of generality. Let $F_j \doteq E_1 \backslash E_j$. It follows from (3) that

$$\mu(E_1) - \lim_j \mu(E_j) = \lim_j \mu(F_j) = \mu\left(\cup_{j=1}^{\infty} F_j\right) = \mu\left(E_1 \backslash \cap_{j=1}^{\infty} E_j\right) = \mu(E_1) - \mu\left(\cap_{j=1}^{\infty} E_j\right).$$

The assertion follows since $\mu(E_1)$ is finite □

For a measure space $(\Omega, \mathcal{F}, \mu)$, a set $E \in \mathcal{F}$ is said to be a **null set** if $\mu(E) = 0$. We say a statement holds μ**-a.s.** (almost surely) or μ**-a.e.** (almost everywhere) if it fails only on a null set.

If $E \in \mathcal{F}$ is a null set, then $\mu(F) = 0$ whenever $F \subseteq E$ and $F \in \mathcal{F}$. However, it need not be the case that an arbitrary subset of any null set is a member of $\mathcal{F}$. If this is the case, the measure μ is called **complete**; i.e., if $F \subseteq E$ where $E \in \mathcal{F}$ and $\mu(E) = 0$, then $F \in \mathcal{F}$.

Completeness is a very useful property, as it simplifies technical arguments. It can always be achieved by a **completion** of the σ-algebra.

Theorem: Let $(\Omega, \mathcal{F}, \mu)$ be an arbitrary measure space, and

$$\mathcal{N} \doteq \{F \subseteq \Omega;\ \ F \subseteq E \ \text{ for some } E \in \mathcal{F},\ \mu(E) = 0\}$$

be the class of all *negligible sets*. Define

$$\bar{\mathcal{F}} \doteq \{F \cup N;\ \ F \in \mathcal{F},\ N \in \mathcal{N}\}$$

and a set function on $\bar{\mathcal{F}}$ by $\bar{\mu}(F \cup N) = \mu(F)$. Then the triple $(\Omega, \bar{\mathcal{F}}, \bar{\mu})$ is a complete measure space. It is called the **completion** of $(\Omega, \mathcal{F}, \mu)$. Indeed, it is the unique extension of μ to a measure on $\bar{\mathcal{F}}$.

Proof. We first show that $\bar{\mathcal{F}}$ is a σ-algebra. Note that $\mathcal{N}$ is closed under countable union since the countable union of null sets is still a null set. It follows readily that $\bar{\mathcal{F}}$ is closed under countable union. $\bar{\mathcal{F}}$ is also closed under complement: take $F \cup N \in \bar{\mathcal{F}}$ with $N \subseteq E \in \mathcal{F}$ and $\mu(E) = 0$. Assume $F \cap E = \emptyset$ (otherwise, just replace the pair N, E by $N \backslash F, E \backslash F$). It is not difficult to see that $(F \cup N)^c = (F \cup E)^c \cup (E \backslash N)$. But $(F \cup E)^c \in \mathcal{F}$ and $E \backslash N \in \mathcal{N}$, in other words, $F \cup N \in \bar{\mathcal{F}}$.

Secondly, we show that $\bar{\mu}$ defines a measure. Note $\bar{\mu}$ is well-defined: if $A = F_1 \cup N_1 = F_2 \cup N_2 \in \bar{\mathcal{F}}$, and $N_i \subseteq E_i \in \mathcal{F}$, then $\mu(F_1) \leq \mu(F_2) + \mu(E_2) = \mu(F_2)$. Similarly, $\mu(F_2) \leq \mu(F_1)$. Hence $\bar{\mu}(A) = \mu(F_1) = \mu(F_2)$ is well-defined. Since $\bar{\mu}(\emptyset) = 0$ is a trivial fact, it remains to show that $\bar{\mu}$ is countably additive. Let $\{F_j \cup N_j; \quad j = 1, 2, \cdots\} \subseteq \bar{\mathcal{F}}$ be disjoint sets. We have

$$\bar{\mu}\left\{\cup_{j=1}^{\infty}(F_j \cup N_j)\right\} = \bar{\mu}\left\{(\cup_{j=1}^{\infty} F_j) \cup (\cup_{j=1}^{\infty} N_j)\right\} = \mu\left\{\cup_{j=1}^{\infty} F_j\right\} = \sum_{j=1}^{\infty} \mu(F_j) = \sum_{j=1}^{\infty} \bar{\mu}\{F_j \cup N_j\};$$

here we have used the fact that $\cup_j N_j \in \mathcal{N}$.

Finally, the completeness of $\bar{\mu}$ is trivial. It remains to show that $\bar{\mu}$ is the unique extension to $\bar{\mathcal{F}}$: suppose ν is also an extension. Let $F \cup N \in \bar{\mathcal{F}}$, and $N \subseteq E \in \mathcal{F}$. It follows that

$$\mu(F) = \nu(F) \leq \nu(F \cup N) \leq \nu(F \cup E) \leq \nu(F) + \nu(E) = \mu(F) \quad \Rightarrow \quad \nu(F \cup N) = \mu(F) = \bar{\mu}(F \cup N).$$

This completes the proof. □

3 Measure extension, Lebsegue-Stieltjes measure

Consider a right-continuous, non-decreasing function $F: \mathsf{R} \to \mathsf{R}$. The question is: does there exist a measure on space $(\mathsf{R}, \mathcal{B}(\mathsf{R}))$ such that

$$\mu_F\{(a, b]\} = F(b) - F(a), \qquad \forall \quad -\infty < a < b < \infty.$$

The answer is affirmative. Indeed, such measure is also unique. It is called the **Lebesgue-Stieltjes measure** associated with F. In particular, when $F(x) = x$, it is called the **Lebesgue measure** (more precisely, the complete extension of μ_F).

How to construct such a measure? In general, a typical path to construct an "interesting" measure is as follows: (1) construct in a "natural" and "straightforward" way a "proto-measure" on an elementary family; for example, all the intervals of form $(a, b]$, (a, ∞), $\emptyset$. (2) extend the proto-measure to the algebra generate by the finite disjoint union of members in $\mathcal{H}$. (3) extend further to obtain a measure μ on $\mathcal{F} = \sigma(\mathcal{H})$, the σ-algebra generated by $\mathcal{H}$. (4) complete the measure space $(\Omega, \mathcal{F}, \mu)$. Here step (2) is usually some meticulous technical work, while steps (3),(4) are achieved with the help of outer measure and Karathéodory theorem.

3.1 Outer measure, Karathéodory theorem

The outer measure is a bridge connecting the "measure" on an algebra to a true measure on an σ-algebra.

Definition: A set function $\mu^*: \mathcal{P}(\Omega) \to [0, \infty]$ is called an **outer measure** if

(1). $\mu^*(\emptyset) = 0$;

(2). $\mu^*(E) \le \mu^*(F)$ whenever $E \subseteq F$;

(3.) $\mu^*\left(\cup_{j=1}^{\infty} E_j\right) \le \sum_{j=1}^{\infty} \mu^*(E_j)$ for any sequence $\{E_j;\ \ j = 1, 2, \cdots\} \subseteq \mathcal{P}(\Omega)$.

A set $E \in \mathcal{P}(\Omega)$ is said to be μ^***-measurable**, if

$$\mu^*(F) = \mu^*(F \cap E) + \mu^*(F \cap E^c), \qquad \forall\ F \in \mathcal{P}(\Omega).$$

Karathédory Theorem: Let μ^* be an outer measure on space Ω, and $\mathcal{M} \subseteq \mathcal{P}(\Omega)$ be the class of all μ^*-measurable sets. Then

(1). $\mathcal{M}$ is a σ-algebra.

(2). $\mu \doteq \mu^*|_{\mathcal{M}}$, the restriction of μ^* on $\mathcal{M}$, is a complete measure on $(\Omega, \mathcal{M})$.

Proof. $\mathcal{M}$ is clearly closed under complement. Now take $A, B \in \mathcal{M}$. For every $E \in \mathcal{P}(\Omega)$, by definition,

$$\begin{aligned}
\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\
&= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c) \\
&\ge \mu^*\big(E \cap (A \cup B)\big) + \mu^*\left(E \cap (A \cup B)^c\right) \qquad \text{(why?)}
\end{aligned}$$

However, the above "$\ge$" can be replaced by "$\le$" by the definition of outer measure. Hence it is indeed an equality, which implies $A \cup B \in \mathcal{M}$, or $\mathcal{M}$ is an algebra. In order to show $\mathcal{M}$ is an σ-algebra, it suffices now to prove that $\mathcal{M}$ is closed under countable disjoint unions. Take a sequence of disjoint sets $\{A_j;\ j = 1, 2, \cdots\} \subseteq \mathcal{M}$ and an arbitrary $E \in \mathcal{P}(\Omega)$. For any $n \ge 1$, we have

$$\begin{aligned}
\mu^*\left(E \cap (\cup_{j=1}^{n} A_j)\right) &= \mu^*\left(E \cap (\cup_{j=1}^{n} A_j) \cap A_n\right) + \mu^*\left(E \cap (\cup_{j=1}^{n} A_j) \cap A_n^c\right) \\
&= \mu^*(E \cap A_n) + \mu^*\left(E \cap (\cup_{j=1}^{n-1} A_j)\right) \\
&= \cdots = \sum_{j=1}^{n} \mu^*(E \cap A_j).
\end{aligned}$$

It follows that

$$\mu^*(E) = \mu^*\left(E \cap (\cup_{j=1}^{n} A_j)\right) + \mu^*\left(E \cap (\cup_{j=1}^{n} A_j)^c\right) \ge \sum_{j=1}^{n} \mu^*(E \cap A_j) + \mu^*\left(E \cap (\cup_{j=1}^{\infty} A_j)^c\right)$$

for all n. Letting $n \to \infty$, we obtain

$$\mu^*(E) \ge \sum_{j=1}^{\infty} \mu^*(E \cap A_j) + \mu^*\left(E \cap (\cup_{j=1}^{\infty} A_j)^c\right) \ge \mu^*\left(E \cap (\cup_{j=1}^{\infty} A_j)\right) + \mu^*\left(E \cap (\cup_{j=1}^{\infty} A_j)^c\right) \ge \mu^*(E).$$

This not only proves that $\mathcal{M}$ is closed under countable disjoint union, it also shows that

$$\mu^*\left(\cup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu^*(A_j)$$

by taking $E = \cup_{j=1}^{\infty} A_j$. In other words, μ^* is countably additive, and hence a measure on $(\Omega, \mathcal{M})$.

It remains to show that μ^* is complete on $(\Omega, \mathcal{M})$. Take $A \in \mathcal{P}(\Omega)$ such that $\mu^*(A) = 0$. For any $E \in \mathcal{P}(\Omega)$, we have $\mu^*(E \cap A) = 0$ and

$$\mu^*(E) \geq \mu^*(E \cap A^c) = \mu^*(E \cap A^c) + \mu^*(E \cap A) \geq \mu^*(E).$$

Therefore, all the inequalities are indeed equalities, which implies $A \in \mathcal{M}$, or $(\Omega, \mathcal{M}, \mu^*)$ is a complete measure space. □

3.2 Extend a "measure" from an algebra to a σ-algebra

The Karathéodory provides a birdge for this extension. Suppose $\mathcal{E}$ is an algebra.

Defintion: A set function $\nu : \mathcal{E} \to [0, \infty]$ is said to be a **pre-measure**, if

(1). $\nu(\emptyset) = 0$.

(2). $\nu\left(\cup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j)$ for any disjoint sequence $\{E_j\} \subseteq \mathcal{E}$ such that $\cup_{j=1}^{\infty} E_j \in \mathcal{E}$.

Any pre-measure ν on $\mathcal{E}$ will define an outer measure μ^* on Ω. Indeed, define

$$\mu^*(A) \doteq \inf_{A \subseteq \cup_j E_j,\ E_j \in \mathcal{E}} \sum_{j=1}^{\infty} \nu(E_j), \qquad \forall\ A \in \mathcal{P}(\Omega).$$

Proposition: The above defined μ^* is an outer measure on Ω, while $\mu^*|_{\mathcal{E}} = \nu$.

Proof. Clearly $\mu^*(\emptyset) = 0$ and $\mu^*(E) \leq \mu^*(F)$ whenever $E \subseteq F$. Now take an arbitrary sequence $\{A_1, A_2, \cdots\} \subseteq \mathcal{P}(\Omega)$. For any $\epsilon > 0$, there exist $\{E_{jk};\ k = 1, 2, \cdots\} \subseteq \mathcal{E}$ such that

$$\mu^*(A_j) \geq \sum_{k=1}^{\infty} \nu(E_{jk}) - \frac{\epsilon}{2^j}, \qquad \forall\ j \geq 1.$$

It follows that

$$\mu^*\left(\cup_{j=1}^{\infty} A_j\right) \leq \sum_{j,k} \nu(E_{jk}) \leq \sum_{j=1}^{\infty} \mu^*(A_j) + \sum_{j=1}^{\infty} \frac{\epsilon}{2^j} = \sum_{j=1}^{\infty} \mu^*(A_j) + \epsilon.$$

Since ϵ is arbitrary, we have $\mu^*\left(\cup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mu^*(A_j)$. Or, μ^* is an outer measure.

Note $\mu^*(A) \leq \nu(A)$ for all $A \in \mathcal{E}$ by defintion. To show the equality actually holds, it suffices to show that

$$\nu(A) \leq \sum_{j=1}^{\infty} \nu(E_j), \qquad \forall\ \ E_j \in \mathcal{E},\ A \subseteq \cup_{j=1}^{\infty} E_j.$$

Let $F_j = E_j \cap A$, we have $F_j \in \mathcal{E}$, $\cup_j F_j = A$, and $\nu(F_j) \leq \nu(E_j)$. However, this implies that

$$\nu(A) = \nu(F_1) + \nu(F_2 \backslash F_1) + \cdots + \nu\left(F_n \backslash (\cup_{j=1}^{n-1} F_j)\right) + \cdots \leq \sum_j \nu(F_j) \leq \sum_j \nu(E_j).$$

The first equality follows from the assumption that ν is a premeasure. □

Theorem (Hahn): Let $\mathcal{M}$ denote all the μ^*-measurable sets. We have

(1). $\mathcal{F} \doteq \sigma(\mathcal{E}) \subseteq \mathcal{M}$;

(2). $\mu \doteq \mu^*|_{\mathcal{F}}$ is a measure on space $(\Omega, \mathcal{F})$ that satisfies $\mu|_{\mathcal{E}} = \nu$;

(3). if ν is σ-finite, then μ is the unique extension of ν to $\mathcal{F}$;

(4). if ν is σ-finite, then the measure space $(\Omega, \mathcal{M}, \mu^*)$ is the completion of the measure space $(\Omega, \mathcal{F}, \mu)$.

Proof: (1). Since $\mathcal{M}$ is a σ-algebra. It suffices to show that $\mathcal{E} \in \mathcal{M}$. That is, we want to show

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c), \qquad \forall\ A \in \mathcal{P}(\Omega),\ E \in \mathcal{E}.$$

Given any $\epsilon > 0$, by definition,

$$\mu^*(A) + \epsilon \geq \sum_{j=1}^{\infty} \nu(E_j), \quad \text{ for some } \ E_j \in \mathcal{E}, \ \ A \subseteq \cup_j E_j.$$

But

$$\sum_{j=1}^{\infty} \nu(E_j) = \sum_{j=1}^{\infty} [\nu(E_j \cap E) + \nu(E_j \cap E^c)] \geq \mu^*(A \cap E) + \mu^*(A \cap E^c).$$

Since ϵ is arbitrary, we have

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c) \quad \Rightarrow \quad \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c).$$

(2) now follows from Karathéodory theorem and the preceding proposition. We will use Dynkin system theorem to prove (3): suppose ρ is another extension to $\mathcal{F}$, and define

$$\mathcal{G} \doteq \{F; \ \ F \in \mathcal{F}, \ \ \mu(F) = \rho(F)\}.$$

We first assume that $\mu(\Omega) = \nu(\Omega)$ to be finite. It is not difficult to see that $\mathcal{G}$ is a λ-class that contain the π-class $\mathcal{E}$, whence $\mathcal{G} \supseteq \sigma(\mathcal{E}) = \mathcal{F}$, or $\mathcal{G} = \mathcal{F}$. In general, there exists $\Omega_n \in \mathcal{E}$ such that $\Omega_n \uparrow \Omega$ with $\nu(\Omega_n)$ being finite. Define

$$\mathcal{E}_n \doteq \{\Omega_n \cap E; \ \ E \in \mathcal{E}\}; \qquad \mathcal{F}_n \doteq \{\Omega_n \cap F; \ \ F \in \mathcal{F}\}.$$

Then $\mathcal{E}_n$ is an algebra in space Ω_n and $\mathcal{F}_n = \sigma_n(\mathcal{E}_n)$ (see exercise below); here σ_n means the smallest σ-algebra generated relative to Ω_n. It follows that $\mu|_{\mathcal{F}_n} = \rho|_{\mathcal{F}_n}$ by the above discussion. That is, $\mu(A \cap \Omega_n) = \rho(A \cap \Omega_n)$ for all n and all $A \in \mathcal{F}$. Letting $n \to \infty$, we have $\mu(A) = \rho(A)$ for all $A \in \mathcal{F}$. This completes the proof of (3). The proof of (4) is left as an exercise. □

Exercise: For every nonempty $\mathcal{D} \subseteq \mathcal{P}(\Omega)$ and every nonempty set $F \subseteq \Omega$, $\sigma(\mathcal{D}) \cap F = \sigma_F(F \cap \mathcal{D})$. Here σ_F means the smallest σ-algebra is generated relative to set F; that is

$$\sigma_F(\mathcal{D} \cap F) = \text{smallest } \sigma\text{-algebra in } \mathcal{P}(F) \text{ that contains } \mathcal{D} \cap F.$$

By convention $\mathcal{D} \cap F = \{D \cap F; \ \ D \in \mathcal{D}\}$.

Proof. First of all, it is not difficult to verify that $\sigma(\mathcal{D}) \cap F$ is a σ-algebra that contains $F \cap \mathcal{D}$, which implies that $\sigma(\mathcal{D}) \cap F \supseteq \sigma_F(F \cap \mathcal{D})$. To show the reverse, define

$$\mathcal{G} \doteq \{B; \quad B \cap F \in \sigma_F(F \cap \mathcal{D})\}$$

Clearly, $\mathcal{G}$ is a σ-algebra and $\mathcal{D} \subseteq \mathcal{G}$, which implies that $\mathcal{G} \supseteq \sigma(\mathcal{D})$, or $\sigma(\mathcal{D}) \cap F \subseteq \sigma_F(F \cap \mathcal{D})$. □

Exercise: Suppose ν is σ-finite.

(1). For every $A \subseteq \Omega$, there exists $F \in \mathcal{F}$ such that $A \subseteq F$ and $\mu^*(A) = \mu^*(F) = \mu(F)$.

(2). If $E \in \mathcal{M}$, then $E = B \cup N$, where $B \in \mathcal{F}$ and $N \subseteq D$, $D \in \mathcal{F}$ with $\mu^*(D) = \mu(D) = 0$.

(3). $(\Omega, \mathcal{M}, \mu^*)$ is the completion of $(\Omega, \mathcal{F}, \mu)$. Or $\mathcal{M} = \bar{\mathcal{F}}$, $\mu^* = \bar{\mu}$.

Proof: (1). For any $A \subseteq \Omega$ and $n \in \mathsf{N}$, there exists $\{F_{n,k};\ k = 1, 2, \cdots\} \subseteq \mathcal{E}$ such that

$$A \subseteq \cup_k F_{n,k} := F_n \quad \text{and} \quad \mu^*(F_n) \le \sum_k \mu^*(F_{n,k}) \le \mu^*(A) + \frac{1}{n}.$$

Define $F \doteq \cap_n F_n$. We have $F \in \mathcal{F}$, $A \subseteq F$ and

$$\mu^*(A) \le \mu^*(F) \le \mu^*(F_n) \le \mu^*(A) + \frac{1}{n}, \qquad \forall\ n \ge 1.$$

Hence $\mu^*(A) = \mu^*(F)$.

(2). We first assume that $\mu^*(E)$ is finite. It follows from (1) that there exists a $F \in \mathcal{F}$ such that $E \subseteq F$ and $\mu^*(E) = \mu^*(F)$. Since $F \in \mathcal{F} \subseteq \mathcal{M}$, we have $\mu^*(F \backslash E) = 0$. By (1) again, there exists a set $D \in \mathcal{F}$ such that $D \supseteq F \backslash E$ and $\mu^*(D) = 0$. Now define

$$B \doteq F \backslash D, \qquad N \doteq E \backslash B \subseteq D.$$

In general, we can write $E = \cup_j E_j$ where $E_j \in \mathcal{M}$ and $\mu^*(E_j) < \infty$. Let (B_j, N_j) as constructed above for E_j. Just define $B = \cup_j B_j$, $N = \cup_j N_j$ and $D = \cup_j D_j$.

(3) now follows directly from (1), (2) and Karathéodory theorem. □

3.3 Lebegue-Stieltjes measure

For every finite measure μ on $(\mathsf{R}, \mathcal{B}(\mathsf{R}))$, one can define a function $F(x) \doteq \mu\big((-\infty, x]\big)$.

Exercise: Show that F is non-negative, non-decreasing and right-continuous.

Can we "reverse" this process? In other words, given a right-continuous, non-decreasing function F, can we associate with it a measure μ_F on $(\mathsf{R}, \mathcal{B}(\mathsf{R}))$ such that

$$\mu_F\big\{(a, b]\big\} = F(b) - F(a), \qquad \forall \quad -\infty < a < b < \infty.$$

The answer is affirmative and the measure is indeed unique.

The construction of this measure closely follows the steps described before.

Summary: *Step 1.* The collection of intervals of form $(a, b]$, (a, ∞), $\emptyset$ consist of an elementary family. One can define a "proto-measure" ρ by

$$\rho\big((a,b]\big] = F(b) - F(a), \qquad \rho\big((a,\infty)\big] = F(\infty) - F(a).$$

Step 2. Let $\mathcal{E}$ = finite disjoint unions of intervals from $\mathcal{H}$. Then $\mathcal{E}$ is an algebra. The proto-measure ρ can be naturely extended to $\mathcal{E}$ by

$$\nu(\emptyset) = 0, \qquad \nu\left(\cup_{j=1}^{n} I_j\right) = \sum_{j=1}^{n} \rho(I_j), \quad \text{where } I_j \in \mathcal{H} \text{ are disjoint intervals.}$$

It can be shown (though very technical) that ν is indeed a σ-finite pre-measure (that is, countably additive).

Step 3. From the pre-measure ν, we can define an outer measure μ^* on R. It follows from above discussion that $\mu \doteq \mu^*|_{\sigma(\mathcal{E})=\mathcal{B}(\mathsf{R})}$ is a measure and is the uniqueness of ν to a measure on $\mathcal{F}$.

Step 4. The completion of $(\mathsf{R}, \mathcal{B}(\mathsf{R}), \mu)$ is $(\mathsf{R}, \mathcal{M}, \mu^*|_{\mathcal{M}})$. Here $\mathcal{M}$ is the class of all μ^*-measurable sets.

Theorem: For any given non-decreasing, right-continuous function $F: \mathsf{R} \to \mathsf{R}$, there is a unique measure μ_F on $(\mathsf{R}, \mathcal{B}(\mathsf{R}))$ such that

$$\mu_F\big((a,b]\big] = F(b) - F(a), \qquad \forall\ -\infty < a < b < \infty.$$

Its completion, denoted by $\bar{\mu}_F$ is called the **Lebesgue-Stieltjes measure** induced by F, and it is σ-finite. When $F(x) \equiv x$, the measure is called the **Lebesgue measure**.

Proof. Following the steps described above, it suffices to show that the set function ν is σ-finite (exercise) and countably additive on the algebra $\mathcal{E}$. We first show that ν is well-defined. Suppose $\{I_i\}_1^n$ and ${K_j}_1^m$ are two sequence of disjoint intervals such that $\cup_i I_i = \cup_j K_j$. We have

$$\sum_i \nu(I_i) = \sum_i \nu\left\{\cup_j (I_i \cap K_j)\right\} = \sum_i \sum_j \nu(I_i \cap K_j) = \sum_j \nu\left\{\cup_i (I_i \cap K_j)\right\} = \sum_j \nu(K_j).$$

Here we are using the obvious fact that

$$F(b) - F(a) = \nu\big\{(a,b]\big\} = \sum_{i=1}^{n} \nu\big\{(a_i, b_i]\big\}, \qquad \text{if } (a_i, b_i] \text{ are disjoint and } (a,b] = \cup_i (a_i, b_i].$$

In order to prove the countably additivity of ν on algebra $\mathcal{E}$, it suffices to show that if $\cup_{i=1}^{n} I_i = \cup_{j=1}^{\infty} K_j$, where $\{I_i\}_1^n$ and $\{K_j\}_1^\infty$ are both disjoint sequence of intervals, then

$$\nu\big\{\cup_i I_i\big\} = \sum_{j=1}^{\infty} \nu(K_j).$$

All we need to show the case where $n = 1$, since this will imply that, for every n,

$$\nu\big\{\cup_i I_i\big\} = \sum_{i=1}^{n} \nu(I_i) = \sum_i \nu\left\{\cup_{j=1}^{\infty} (I_i \cap K_j)\right\} = \sum_{i=1}^{n} \sum_{j=1}^{\infty} \nu(I_i \cap K_j) = \sum_{j=1}^{\infty} \nu\left\{\cup_i (I_i \cap K_j)\right\} = \sum_{j=1}^{\infty} \nu(K_j).$$

Now assume that $I = (a, b] = \cup_{j=1}^{\infty} K_j = \cup_{j=1}^{\infty}(a_j, b_j]$, where $K_j = (a_j, b_j]$ are disjoint intervals.

For every $n \geq 0$, since $\cup_{j=1}^{n}(a_j, b_j] \subseteq (a, b]$, it is not difficult to see that

$$\sum_{j=1}^{n} \nu\big\{(a_j, b_j]\big\} = \sum_{j=1}^{n} \big[F(b_j) - F(a_j)\big] \leq F(b) - F(a) = \nu\big\{(a, b]\big\} \quad \Rightarrow \quad \sum_{j=1}^{\infty} \nu(K_j) \leq \nu(I).$$

For the reverse inequality, pick an arbitrary $0 < \epsilon < b - a$. It follows that

$$[a + \epsilon, b] \subseteq (a, b] = \cup_{j=1}^{\infty}(a_j, b_j] \subseteq \cup_{j=1}^{\infty}(a_j, b_j')$$

where $b_j' > b_j$ such that (by right-continuity)

$$\nu\big\{(a_j, b_j]\big\} = F(b_j) - F(a_j) \geq F(b_j') - F(a_j) - \frac{\epsilon}{2^j}.$$

It follows from Heine-Broel theorem that there is a finite cover of the compact interval $[a + \epsilon, b]$; that is, with a possible relabelling,

$$(a + \epsilon, b] \subseteq [a + \epsilon, b] \subseteq \cup_{j=1}^{N}(a_j, b_j') \subseteq \cup_{j=1}^{N}(a_j, b_j'] \qquad \text{for some } N.$$

It is not difficult to show that (exercise)

$$\nu\big\{(a + \epsilon, b]\big\} \leq \sum_{j=1}^{N}(a_j, b_j'] \leq \sum_{j=1}^{\infty}(a_j, b_j'] \leq \sum_{j=1}^{\infty} \nu(a_j, b_j] + \epsilon.$$

Letting $\epsilon \to 0$, we obtain from the right-continuity of F that

$$\nu\big\{(a, b]\big\} = F(b) - F(a) = \lim_{\epsilon}\big[F(b) - F(a + \epsilon)\big] = \lim_{\epsilon} \nu\big\{(a + \epsilon, b]\big\} \leq \sum_{j=1}^{\infty} \nu(a_j, b_j].$$

This completes the proof. □

4 Product space, product σ-algebra

Let $\{(X_\alpha, \mathcal{F}_\alpha);\ \alpha \in I\}$ be a collection of measurable space, here I is the index set. Let $X = \prod_{\alpha \in I} X_i$ denote the (Cartesian) **product space**. The α-th **projection** π_α is defined so that

$$\pi_\alpha(\omega) = \omega_\alpha, \qquad \text{where } \omega = (\omega_\alpha)_{\lambda \in I} \text{ is an arbitrary element in } X.$$

Remark: More precisely, X is defines as the collection of all maps ω on I such that $\omega(\alpha) \in X_\alpha$. In the interesting case where α is uncountable, the non-empty-ness of X follows from the so-called *Axiom of choice.*

The question is how to define a useful σ-algebra on X. Define the *cylinder sets*

$$\mathcal{C} \doteq \big\{\pi_\alpha^{-1}(E_\alpha);\quad E_\alpha \in \mathcal{F}_\alpha,\quad \alpha \in I\big\}$$

Definition: The σ-algebra generated by $\mathcal{C}$ is called the **product σ-algebra** on X, and is denoted by $\mathcal{F} = \otimes_{\alpha \in I} \mathcal{F}_\alpha$.

Exercise: Suppose $\mathcal{F}_\alpha = \sigma(\mathcal{E}_\alpha)$, $\forall\, \alpha \in I$. Then $\mathcal{F} = \sigma(\mathcal{C}')$ where $\mathcal{C}'$ is defined similarly as $\mathcal{C}$, except that $\mathcal{F}_\alpha$ is replaced by $\mathcal{E}_\alpha$.

Proposition: In the case where I is *countable*, $\mathcal{F} = \sigma(\mathcal{R})$ where $\mathcal{R}$ is the collection of "rectangles":

$$\mathcal{R} = \left\{ \prod_{\alpha\in I} E_\alpha; \quad E_\alpha \in \mathcal{F}_\alpha, \quad \forall\, \alpha \in I \right\}.$$

Moreover, if $\mathcal{F}_\alpha = \sigma(\mathcal{E}_\alpha)$, then $\mathcal{F} = \sigma(\mathcal{R}')$ where $\mathcal{R}'$ is similarly defined as $\mathcal{R}$, except that $\mathcal{F}_\alpha$ is replaced by $\mathcal{E}_\alpha$.

The proof of the proposition is left as an exercise.

Proposition: Suppose now $\{X_1, \cdots, X_n\}$ are metric spaces, and $X = \prod_{i=1}^n X_i$ is the product space with the *product measure*. Then $\mathcal{B}(X) \supseteq \otimes_{i=1}^n \mathcal{B}(X_i)$ and the equality holds if $X_1, \cdots, X_n$ are seperable. In particular, $\mathcal{B}(\mathsf{R}^n) = \mathcal{B}(\mathsf{R}) \otimes \cdots \otimes \mathcal{B}(\mathsf{R})$ n-times.

Proof. We shall give a proof for the case where $n = 2$. The proof for general n is the same, but with heavier notation. Assume that the product metric is $\rho(x, y) = \max_{i=1,2} \rho_i(x_i, y_i)$ where $x = (x_1, x_2)$, $y = (y_1, y_2)$. Note the specfic choice of ρ is not important, as long as it generates the same topology; for example, $\rho(x, y) \doteq \sqrt{\rho_1^2(x_1, y_1) + \rho_2^2(x_2, y_2)}$ is an equivalent measure. We have

$$\mathcal{B}(X_1) \otimes \mathcal{B}(X_2) = \sigma(\mathcal{C}'), \qquad \text{where } \mathcal{C}' = \left\{\pi_i^{-1}(A_i); \quad A_i \subseteq X_i \text{ is open set}, \quad i = 1, 2\right\}.$$

Since $\pi_i^{-1}(A_i)$ is open in X, $\mathcal{C}' \subseteq \mathcal{B}(X)$, which implies that $\mathcal{B}(X) \supseteq \mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$.

Now assume that $X_{1,2}$ are separable, and $D_1 = \{x_n\}_{n\in\mathsf{N}}$, $D_2 = \{y_n\}_{n\in\mathsf{N}}$ are countable dense set in X_1, X_2 respectively. It is not difficult to see that $D = D_1 \times D_2$ is a countable dense set in X. Any open set in X now can be expressed as a countable union of the open sets of form

$$\{B(x; r); \quad x \in D, \ r \text{ is a positive rational number}\}, \quad \text{here } B \text{ stands for the open ball in } X.$$

But $B(x, r) = B_1(x_i, r) \times B_2(y_j, r)$ by the definition of metric ρ, whenever $x = (x_i, y_j)$. Here B_i stands for the open ball in space X_i. It follows that $B(x, r) \in \mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$. This implies

$$\mathcal{B}(X) = \sigma\{B(x; r); \ x \in D, \ r > 0 \text{ is rational}\} \subseteq \mathcal{B}(X_1) \otimes \mathcal{B}(X_2).$$

We complete proof. □

5 Measurable functions, approximation by simple functions

Consider measurable spaces $(X, \mathcal{F})$ and $(Y, \mathcal{G})$. A map $f : X \to Y$ is said to be **$\mathcal{F}/\mathcal{G}$-measurable** if $f^{-1}(E) \in \mathcal{F}$ whenever $E \in \mathcal{G}$. We simply say f is **measurable** if no confusion is incurred. *Unless specified, whenever $Y = \mathsf{R}$, we take $\mathcal{G} = \mathcal{B}(\mathsf{R})$.*

Example: Suppose $\{E_1, \cdots, E_n\} \subseteq \mathcal{F}$ is a collection of sets, and $\{\alpha_1, \cdots, \alpha_n\}$ is a collection of real numbers. Then the **simple function**

$$f(x) = \sum_{j=1}^n \alpha_j 1_{E_j}(x), \qquad \forall\, x \in X$$

is a measurable function. Here $1_E(x)$ is the **indicator function** such that

$$1_E(x) = \left\{ \begin{array}{ccc} 1 & ; & x \in E \\ 0 & ; & x \notin E \end{array} \right\}.$$

Lemma: If $\mathcal{G} = \sigma(\mathcal{E})$, then f is measurable if and only if $f^{-1}(E) \in \mathcal{F}$ whenever $E \in \mathcal{E}$.

Proof. Let $\mathcal{A} \doteq \{E \subseteq Y;\ \ f^{-1}(E) \in \mathcal{F}\}$. It is not difficult to verify that $\mathcal{A}$ is a σ-algebra (exercise). Since $\mathcal{E} \in \mathcal{A}$, then $\mathcal{G} = \sigma(\mathcal{E}) \subseteq \mathcal{A}$. This completes the proof. □

Remark: Suppose X and Y are two metric spaces. Then any continuous function $f : X \to Y$ is $\mathcal{B}(X)/\mathcal{B}(Y)$-measurable.

Exercise: The composition of measruable functions is measurable.

Proposition: If $\{f_n\}_{n\in\mathsf{N}}$ is a sequence of measurable functions, then

$$\begin{array}{c} h_m(x) \doteq \sup_{n\geq m} f_n(x), \quad g_m(x) = \inf_{n\geq m} f_m(x) \qquad (m \in \mathsf{N}) \\ k(x) \doteq \limsup_{n\to\infty} f_n(x), \quad l(x) \doteq \liminf_{n\to\infty} f_n(x) \end{array}$$

are all measurable functions.

Remark: In the preceding proposition, the various functions could very possibly take value on the *extended real line* $\bar{\mathsf{R}} = [-\infty, \infty]$. The Borela σ-algebra on the $\bar{\mathsf{R}}$ is defined as

$$\mathcal{B}(\bar{\mathsf{R}}) \doteq \left\{E \subseteq \bar{\mathsf{R}};\ \ E \cap \mathsf{R} \in \mathcal{B}(\mathsf{R})\right\}.$$

This Borel σ-algebra is generated by the intervals of form $\{(a, \infty];\ \ a \in \mathsf{R}\}$ or $\{[-\infty, a);\ \ a \in \mathsf{R}\}$. Now a function $f :\ X \to \bar{\mathsf{R}}$ is called $\mathcal{F}$-measurable if it is $\mathcal{F}/\mathcal{B}(\bar{\mathsf{R}})$-measurable. The preceding proposition remains valid even if $\{f_n\}_{n\in\mathsf{N}}$ takes value in $\bar{\mathsf{R}}$.

Suppose now $Y = \prod_{\alpha\in I} Y_i$ is a product space of measurable spaces $\{(Y_\alpha, \mathcal{G}_\alpha);\ \alpha \in I\}$. We associate Y with the product σ-algebra $\mathcal{G} = \otimes_\alpha \mathcal{G}_\alpha$. We have the following result, whose proof is left as an exercise.

Proposition: A mapping $f : X \to Y$ is $\mathcal{F}/\mathcal{G}$-measurable if and only if $\pi_\alpha \circ f : X \to Y_\alpha$ is $\mathcal{F}/\mathcal{G}_\alpha$-measurable, for all $\alpha \in I$.

Remark: It follows from the proposition that, if $\{f_i : X \to \mathsf{R};\ i = 1, \cdots, n\}$ are measurable functions, and $\phi : \mathsf{R}^n \to \mathsf{R}$ is a measurable function, then $\phi \circ (f_1, f_2, \cdots, f_n) : X \to \mathsf{R}$ is a measurable function. In particulr, if we take $\phi(x_1, \cdots, x_n) = x_1 + \cdots + x_n$, which is continuous and hence measurable, we conclude that $f_1 + \cdots + f_n : X \to \mathsf{R}$ is measurable. Similarly, $f_1 \cdot \cdots \cdot f_n$, $\max_i f_i$ and $\min_i f_i$ are all measurable functions.

5.1 Approximation by simple functions

Consider a measurable space $(X, \mathcal{F})$. A simple function is of form $\sum_{j=1}^{n} \alpha_j 1_{E_j}$ where $\{\alpha_j\}$ are real numbers, and $\{E_j\} \subseteq \mathcal{F}$. The summation or product of simple functions is still a simple function.

Proposition:

1. If $f : X \to [0, \infty]$ is a measurable function, there is an increasing sequence of non-negative, simple function $\{\phi_n\}_{n \in \mathsf{N}}$ such that $f(\omega) = \lim_n \phi_n(\omega)$ for all $\omega \in X$; this convergence is uniform on any sets where f is bounded.
2. If $f : X \to \bar{\mathsf{R}}$ is a measurable function, there is a sequence $\{\phi_n\}_{n \in \mathsf{N}}$ of simple functions such that $0 \le |\phi_1| \le |\phi_2| \le \cdots \le |f|$, $\lim_n \phi_n(\omega) = f(\omega)$ for all $\omega \in \Omega$; this convergence is uniform on any sets where $|f|$ is bounded.

Proof. The proof is simply by construction.

1. Fix n, define

$$E_n^{(k)} \doteq \left\{ \omega \in X;\ \frac{k}{2^n} < f(\omega) \le \frac{k+1}{2^n} \right\}, \quad k = 0, \cdots, 4^n - 1; \quad \text{and } F_n \doteq \{\omega \in X;\ f(\omega) > 2^n\}.$$

The simple function

$$\phi \doteq \sum_{k=0}^{4^n - 1} \frac{k}{2^n} 1_{E_n^{(k)}} + 2^n 1_{F_n}.$$

Clearly $\phi_n \le \phi_{n+1}$ and $\phi_n \to f$. Moreover, $0 \le f - \phi_n \le 2^{-n}$ on $\{f \le 2^n\}$.

2. Write $f = f^+ - f^-$, where $f^\pm \doteq \max(\pm f, 0)$ are measurable functions. From part 1, there exists $\{g_n^\pm\}_{n \in \mathsf{N}}$ which increases pointwise to $f^\pm$. Let $\phi_n \doteq g_n^+ - g_n^-$. Note $|\phi_n| = g_n^+ + g_n^-$, the claim follows readily. □

6 Integration

Consider a measure space $(X, \mathcal{F}, \mu)$. For any measurable function $f : X \to \bar{\mathsf{R}}$, how can we define the integration $I(f) = \int_\Omega f(x) d\mu(x)$? We should start with the case where f is non-negative.

6.1 Definition of the integration

For a non-negative *simple function* of canonical form

$$\phi(\omega) = \sum_{j=1}^{n} \alpha_j 1_{E_j}(\omega); \quad \text{here } \{E_j\} \text{ are disjoint and } X = \cup_{j=1}^{n} E_j,\ \alpha_j \in [0, \infty],$$

its integration can be naturally defined as

$$I(\phi) = \int_\Omega \phi(\omega)\, d\mu(\omega) \doteq \sum_{j=1}^{n} \alpha_j \mu(E_j), \qquad \text{with convention } \infty \cdot 0 = 0.$$

Now for a general non-negative measurable function $f : X \to [0, \infty]$, the integral is defined as

$$I(f) = \int_\Omega f(\omega)\, d\mu(\omega) \doteq \sup_{\{\phi;\ 0 \le \phi \le f,\ \phi \text{ is simple}\}} I(\phi).$$

A general measurable function $f : X \to \bar{\mathsf{R}}$ can be written as $f = f^+ - f^-$ where $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$. Clearly, $f^\pm$ are both measurable. We say $I(f)$ is **well-defined** if $I(f^+) - I(f^-)$ is well defined (that is, $I(f^\pm)$ are not simultaneously ∞), and in this case define

$$I(f) = I(f^+) - I(f^-).$$

We say $I(f)$ is **integrable** if $I(f^\pm)$ are both finite.

We have the following elementary properties for the integration.

Proposition: Suppose f, g are measurable functions. The following relationships hold whenever the associated integrations are meaningful.

1. $I(cf) = c \cdot I(f)$, for an arbitrary constant c.
2. $I(f) \le I(g)$ if $f \le g$.
3. If $f \ge 0$, the mapping $E \mapsto \int_E f(\omega)\, d\mu(\omega) \doteq I(f1_E)$, for all $E \in \mathcal{F}$, defines a measure on the space $(X, \mathcal{F})$.

Proof. Part 1 is trivial. For part 2, it suffices to observe that $f^+ \le g^+$ and $f^- \ge g^-$. It remains to show part 3. It suffices to show the countable additivity, that is,

$$I(f1_E) = \sum_{n=1}^\infty I(f1_{E_n}), \qquad \text{where } \{E_n\}_{n \in \mathsf{N}} \text{ are disjoint and } \cup_{n=1}^\infty E_n = E.$$

Without loss of generality, we assume $E = X$. We first show that this is true when f is simple, say $f = \sum_{j=1}^k \alpha_j 1_{F_j}$. It follows that

$$I(f1_{E_n}) = I\left(\sum_{j=1}^k \alpha_j 1_{E_n \cap F_j}\right) = \sum_{j=1}^k \alpha_j \mu(E_n \cap F_j),$$

which implies that

$$\sum_{n=1}^\infty I(f1_{E_n}) = \sum_{n=1}^\infty \sum_{j=1}^k \alpha_j \mu(E_n \cap F_j) = \sum_{j=1}^k \sum_{n=1}^\infty \alpha_j \mu(E_n \cap F_k) = \sum_{j=1}^k \alpha_j \mu(F_k) = I(f).$$

Now assume that f is a general non-negative measurable function. For an arbitrary $0 \le \phi \le f$, we have $\phi 1_{E_n} \le f1_{E_n}$ for all n, and

$$\sum_{n=1}^\infty I(f1_{E_n}) \ge \sum_{n=1}^\infty I(\phi 1_{E_n}) = I(\phi) \quad \Rightarrow \quad \sum_{n=1}^\infty I(f1_{E_n}) \ge I(f).$$

As for the reverse inequality, let $0 \le \phi_n \le f1_{E_n}$ be simple functions. For any N,

$$\phi_1 + \phi_2 + \cdots + \phi_N \le f1_{E_1} + f1_{E_2} + \cdots + f1_{E_N} \le f$$

is a simple function. By definition,

$$I(f) \geq I(\phi_1 + \phi_2 + \cdots + \phi_N) = I(\phi_1) + I(\phi_2) + \cdots + I(\phi_N);$$

see the Remark below. It follows that

$$I(f) \geq \sum_{n=1}^{N} I(f 1_{E_n}) \qquad \Rightarrow \qquad I(f) \geq \sum_{n=1}^{\infty} I(f 1_{E_n}).$$

This completes the proof. □

Remark: It is not difficult to verify that, if ϕ_1 and ϕ_2 are two non-negative simple functions, then

$$I(\phi_1 + \phi_2) = I(\phi_1) + I(\phi_2).$$

This is also true for general integrable functions, whose proof will be deferred. As for simple function case. Take

$$\phi_1 = \sum_{j=1}^{n} \alpha_j 1_{E_j}, \qquad \phi_2 = \sum_{k=1}^{m} \beta_k 1_{F_k}$$

It follows that

$$\phi_1 + \phi_2 = \sum_j \lambda_j \sum_k 1_{E_j \cap F_k} + \sum_k \beta_k \sum_j 1_{F_k \cap E_j} = \sum_{j,k} (\alpha_j + \beta_k) 1_{E_j \cap F_k},$$

and

$$I(\phi_1 + \phi_2) = \sum_{j,k} (\lambda_j + \beta_k) \mu(E_j \cap F_k) = \sum_j \alpha_j \mu(E_j) + \sum_k \beta_k \mu(F_k) = I(\phi_1) + I(\phi_2).$$

6.2 Three main theorems

Three very important results in integration theory are Monotone Convergence Theorem (MCT), Dominated Convergence Theorem (DCT) and Fatou Lemma.

Monotone Convergence Theorem: Suppose $\{f_n\}_{n \in \mathsf{N}}$ is a sequence of *non-negative* mesurable functions such that $f_1 \leq f_2 \leq \cdots \leq f_n \leq \cdots$ and $f = \lim_n f$. Then $I(f) = \lim_n I(f_n)$.

Fatou Lemma: For any sequence $\{f_n\}_{n \in \mathsf{N}}$ of *non-negative* measure functions, we have

$$I\left(\liminf_n f_n\right) \leq \liminf_n I(f_n).$$

Dominated Convergence Theorem: Suppose $\{f_n\}_{n \in \mathsf{N}}$ is a sequence of measurable functions such that $\lim_n f_n = f$. If there exists a measurable function g such that $I(g)$ is finite (i.e. g is integrable) and $|f_n| \leq g$ for all n, then

$$I(f) = \lim_n I(f_n).$$

Proof of MCT. It follows from the assumption that $I(f_n) \le I(f_{n+1}) \le \cdots \le I(f)$, which implies that $I(f) \ge \lim_n I(f_n)$. It remains to show the reverse inequality. To this end, fix an arbitrary constant $\epsilon \in (0,1)$. Define sets

$$E_n \doteq \{\omega \in X; \quad f_n(\omega) \ge \epsilon \cdot f(\omega)\}.$$

By assumption, we have $E_1 \subseteq E_2 \subseteq \cdots$ and $\cup_{n=1}^{\infty} E_n = X$ (why?). It follows that

$$I(f_n) \ge I(f_n 1_{E_n}) \ge I(\epsilon f 1_{E_n}) = \epsilon I(f 1_{E_n}).$$

Note that the mapping $E \mapsto I(f1_E)$ defines a measure on $(X, \mathcal{F})$, we establish

$$\lim_{n\to\infty} I(f 1_{E_n}) = I(f 1_X) = I(f).$$

Therefore, $\lim_n I(f_n) \ge \epsilon I(f)$. Letting $\epsilon \to 1$, we have $\lim_n I(f_n) \ge I(f)$. □

Proof of Fatou Lemma. Define $g_n(\omega) = \inf_{k\ge n} f_k(\omega)$, which is clearly an increasing sequence of non-negative functions. We have

$$\liminf_n f_n = \sup_n \left(\inf_{k\ge n} f_k\right) = \sup_n g_n = \lim_n g_n.$$

Clearly, $I(g_n) \le \inf_{k\ge n} I(f_k)$, which implies that

$$I\left(\liminf_n f_n\right) = I(\lim_n g_n) = \lim_n I(g_n) \le \lim_n \left\{\inf_{k\ge n} I(f_k)\right\} = \liminf_n I(f_n);$$

here the second equality follows from MCT. □

Proof of DCT. Since $I(f_n) = I(f_n^+) - I(f_n^-)$ and $f_n^{\pm} \le g$, $\lim_n f_n^{\pm} = f^{\pm}$, it is without loss generality if we assume $\{f_n\}_{n\in\mathbb{N}}$ are non-negative. In order to proceed, we first prove that $I(h_1 + h_2) = I(h_1) + I(h_2)$ if $h_{1,2}$ are two arbitrary non-negative measurable functions. Indeed, let $\{\phi_n^{(i)}\}$ be a sequence of approximating simple functions for h_i such that $\lim_n \phi_n^{(i)} = h_i$. It follows from MCT that

$$I(h_1 + h_2) = \lim_n I\left(\phi_n^{(1)} + \phi_n^{(2)}\right) = \lim_n I(\phi_n^{(1)}) + \lim_n I(\phi_n^{(2)}) = I(h_1) + I(h_2).$$

Here we use MCT in the last step. By assumption $g \pm f_n \ge 0$, which implies $I(g \pm f_n) = I(g) \pm I(f_n)$ (why?). It follows from Fatou Lemma that

$$\begin{aligned} I(f+g) &= I\left(\liminf_n\{f_n + g\}\right) \le \liminf_n I(f_n + g) \le \liminf_n \{I(f_n) + I(g)\} = I(g) + \liminf_n I(f_n); \\ I(g-f) &= I\left(\liminf_n\{g - f_n\}\right) \le \liminf_n I(g - f_n) \le \liminf_n \{I(g) - I(f_n))\} = I(g) - \limsup_n I(f_n). \end{aligned}$$

These two inequalities imply

$$\limsup_n I(f_n) \le I(f) \le \liminf_n I(f_n) \qquad \Rightarrow \qquad \lim_n I(f_n) = I(f).$$

This completes the proof. □

6.3 Properties of integrals

In this section, we present a collection of useful properties of integral.

Proposition: For any measurable functions f, g, we have

$$I(f+g) = I(f) + I(g)$$

as long as $I(f), I(g)$ are well-defined, and $I(f)+I(g)$ is meaningful (that is, not $\infty - \infty$).

Proof. We have already seen from the proof of DCT that the claim is true if f and g are both non-negative. In general, let $h = f+g$, we have

$$h^+ - h^- = f^+ - f^- + g^+ - g^- \quad \Rightarrow \quad h^+ + f^- + g^- = h^- + f^+ + g^+$$

which implies that

$$I(h^+) + I(f^+) + I(g^+) = I(h^+ + f^- + g^-) = I(h^- + f^+ + g^+) = I(h^-) + I(f^+) + I(g^+)$$

or,

$$I(h) = I(h^+) - I(h^-) = I(f^+) - I(f^-) + I(g^+) - I(g^-) = I(f) + I(g).$$

This completes the proof. □

Exercise: A function f is integrable if and only if $|f|$ is integrable. Furthermore, $I(|f|) \geq |I(f)|$.

Proposition: Suppose $\{f_n\}_{n \in \mathsf{N}}$ is a sequence of measurable functions. Then

$$I\left(\sum_{n=1}^{\infty} f_n\right) = \sum_{n=1}^{\infty} I(f_n)$$

provided either one of the following conditions holds:

(1). f_n is non-negative for all n;

(2). $\sum_{n=1}^{\infty} |f_n|$ is integrable; that is, $\sum_{n=1}^{\infty} I(|f_n|)$ is finite.

Proof. (1) is just an application of MCT; (2). As $I(\sum_n |f_n|) = \sum_n I(|f_n|) < \infty$, it is not difficult to see that $\mu(\sum_n |f_n| = \infty) = 0$; in other words, $\sum_n |f_n|$ converge almost surely, which implies $\sum_n f_n$ also converge almost surely. The result follows from DCT. □

Proposition: Suppose $f \geq 0$ is measurable. Then $I(f) = 0$ if and only if $f = 0$ almost surely.

Proof. "$\Leftarrow$" is clear. "$\Rightarrow$": Define $E_n \doteq \left\{\omega;\ f(\omega) \geq \frac{1}{n}\right\}$. It follows that $E_n \subseteq E_{n+1}$ and $\cup_{n=1}^{\infty} E_n = \{\omega;\ f(\omega) > 0\}$. Moreover,

$$0 = I(f) \geq I\left(f 1_{E_n}\right) \geq I\left(\frac{1}{n} 1_{E_n}\right) = \frac{1}{n}\mu(E_n) \quad \Rightarrow \quad \mu(E_n) = 0; \quad \forall\ n.$$

Then $\mu(E) = \lim_n \mu(E_n) = 0$. □

Exercise: Consider two integrable functions f, g. Then $I(f1_E) = I(g1_E)$ for all $E \in \mathcal{F}$ if and only $f = g$ almost surely.

Below is a collection of exercises related to integration.

Exercise: Suppose μ, ν are two measures on measurable space $(X, \mathcal{F})$. If $\mu(E) \leq \nu(E)$ for all $E \in \mathcal{F}$, then $\int_X f\, d\mu \leq \int_X f\, d\nu$ for all non-negative measurable functions f.

Exercise: Consider a measure space $(X, \mathcal{F}, \mu)$ and a non-negative measurable function f. The set function $\nu : E \mapsto \int_E f\, d\mu$ defines a measure $(X, \mathcal{F})$. Show that for any non-negative measurable function g,

$$\int_X g\, d\nu = \int_X gf\, d\mu.$$

Exercise: For any integrable function f and any constant $\epsilon > 0$, we can find a simple function $\phi = \sum_{j=1}^n \alpha_j 1_{E_j}$ such that ϕ is integrable and $I(|f - \phi|) \leq \epsilon$.

Exercise: Let $[a, b]$ be a given bounded interval on R, and $f : [a, b] \times X \to \mathsf{R}$ be a given measurable function such that $f(t, \cdot)$ is integrable for all $t \in [a, b]$. Define $F(t) = \int_X f(t, \omega)\, d\mu(\omega)$, $a \leq t \leq b$. If the partial derivative $\frac{\partial f}{\partial t}$ exists and satisfies $\left|\frac{\partial f}{\partial t}(t, \omega)\right| \leq h(\omega)$ for some integrable function h. Then F is differentiable and

$$F'(t) = \int_X \frac{\partial f}{\partial t}(t, \omega)\, d\mu(\omega).$$

6.4 Product Measure, Fubuni-Tonelli theorem

Suppose we have two measure space $(X_i, \mathcal{F}_i, \mu_i)$, $i = 1, 2$. We have already defined the product σ-algebra $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$ on the product space $X = X_1 \times X_2$. We wish to construct a **product measure** $\mu = \mu_1 \otimes \mu_2$ such that

$$\mu(E_1 \times E_2) = \mu_1(E_1) \cdot \mu_2(E_2), \qquad \forall\ E_i \in \mathcal{F}_i, \quad i = 1, 2.$$

One way to go about this is to follow essentially the path that we used to construct the Lebesgue-Stieltjes measure on real line: starting by constructing $\pi(E_1 \times E_2) = \mu_1(E_1) \cdot \mu_2(E_2)$ on rectangles $E = E_1 \times E_2$, extended in an obvious way to a pro-measure on the algebra generated by the finite disjoint union of such rectangles, and then extend this pre-measure to a measure μ on $\mathcal{F}$ by Hahn-Karathéodory theorem.

But here we prefer to a more direct approach with the help of the integration theory at hand. This approach also yields an *integral representation* of the product measure $\mu = \mu_1 \otimes \mu_2$. The following elementary example is used for illustration.

Area of unit disk: Let $X_1 = X_2 = \mathsf{R}$, and $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{B}(\mathsf{R})$. The unit disk

$$E = \left\{(\omega_1, \omega_2);\ \ \omega_1^2 + \omega_2^2 \leq 1\right\}$$

is in $\mathcal{F} = \mathcal{B}(\mathsf{R}^2)$. For any $\omega_1 \in \mathsf{R}$, the "section"

$$E_{\omega_1} \doteq \{\omega_2;\ \ (\omega_1, \omega_2) \in E\} = \left\{ \begin{array}{ccc} \{\omega_2;\ |\omega_2| \leq \sqrt{1 - \omega_1^2}\} & ; & \text{if } |\omega_1| \leq 1 \\ \emptyset & ; & \text{if } |\omega_1| > 1 \end{array} \right\}.$$

It follows that
$$\mu_2(E_{\omega_1}) = \left\{ \begin{array}{ccc} 2\sqrt{1-\omega_1^2} & ; & \text{if } |\omega_1| \le 1 \\ 0 & ; & \text{if } |\omega_1| > 1 \end{array} \right\}.$$

The area of the unit disk is
$$\mu(E) = \text{area of } E = \int_{-\infty}^{\infty} \mu(E_{\omega_1})\, d\mu_1(\omega_1) = \int_{-1}^{1} 2\sqrt{1-\omega_1^2}\, d\omega_1 = \pi.$$

From now on, we are going to use the following notation. For any set $E \subseteq X$, define the "sections"
$$E_{\omega_1} \doteq \{\omega_2;\ (\omega_1, \omega_2) \in E\}, \quad E_{\omega_2} \doteq \{\omega_1;\ (\omega_1, \omega_2) \in E\}.$$
For any function $f : X_1 \times X_2 \to \mathsf{R}$, define the "section functions"
$$f_{\omega_1} :\ X_2 \to \mathsf{R} \ \text{ with } \ \omega_2 \mapsto f(\omega_1, \omega_2); \quad f_{\omega_2} :\ X_1 \to \mathsf{R} \ \text{ with } \ \omega_1 \mapsto f(\omega_1, \omega_2)$$

Lemma: For every $E \in \mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$, and every $\mathcal{F}$-measurable function $f : X \to \mathsf{R}$. we have

1. $E_{\omega_1} \in \mathcal{F}_2$ for all $\omega_1 \in X_1$; and $E_{\omega_2} \in \mathcal{F}_1$ for all $\omega_2 \in X_2$.
2. f_{ω_1} is $\mathcal{F}_2$-measurable; and f_{ω_2} is $\mathcal{F}_1$-measurable.

Proof. Let $\mathcal{E} = \{E \in X;\ E_{\omega_1} \in \mathcal{F}_2 \text{ for all } \omega_1 \in X_1 \text{ and } E_{\omega_2} \in \mathcal{F}_1 \text{ for all } \omega_2 \in X_2\}$. It is not difficult to verify that $\mathcal{E}$ is a σ-algebra. But $\mathcal{E}$ contains all the rectangles of form $E_1 \times E_2$, where $E_i \in \mathcal{F}_i$. It follows that $\mathcal{E} \supseteq \sigma\{E_1 \times E_2;\ E_i \in \mathcal{F}_i,\ i = 1, 2\} = \mathcal{F}$. Part 2 follows easily from part 1. □

Product Measure Theorem: Let $(X_i, \mathcal{F}_i, \mu_i)$, $i = 1, 2$ be σ-finite measure spaces. Let $X = X_1 \times X_2$, $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$. For any $E \in \mathcal{F}$, the functions
$$\omega_1 \mapsto \mu_2(E_{\omega_1}) \ \text{ is } \mathcal{F}_1\text{-measurable}; \quad \omega_2 \mapsto \mu_1(E_{\omega_2}) \ \text{ is } \mathcal{F}_2\text{-measurable}.$$

Furthermore, the set function
$$\mu(E) \doteq \int_{X_1} \mu_2(E_{\omega_1})\, d\mu_1(\omega_1), \qquad \forall\ E \in \mathcal{F}$$
defines a σ-finite measure on $(X, \mathcal{F})$. This measure satisfies
$$\mu(E) = \int_{X_1} \mu_2(E_{\omega_1})\, d\mu_1(\omega_1) = \int_{X_2} \mu_1(E_{\omega_2})\, d\mu_2(\omega_2)$$
and is the unique measure on $(X, \mathcal{F})$ such that $\mu(E_1 \times E_2) = \mu_1(E_1) \cdot \mu_2(E_2)$ for all $E_i \in \mathcal{F}_i$, $i = 1, 2$. Measure μ is called the **product measure** and denoted by $\mu_1 \otimes \mu_2$.

Exercise: For any $E \in \mathcal{F}$, $\mu(E) = 0$ if and only if $\mu_2(E_{\omega_1}) = 0$ for μ_1-a.e. $\omega_1 \in X_1$ if and only if $\mu_1(E_{\omega_2}) = 0$ for μ_2-a.e. $\omega_2 \in X_2$.

Tonelli Theorem: For any non-negative measurable function $f : X \to [0, \infty]$, we have

$$\begin{aligned} \int\int_{X_1 \times X_2} f(\omega_1, \omega_2)\, d(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) &= \int_{X_1} \left(\int_{X_2} f(\omega_1, \omega_2)\, d\mu_2(\omega_2) \right) d\mu_1(\omega_1) \\ &= \int_{X_2} \left(\int_{X_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) \right) d\mu_2(\omega_2) \end{aligned}$$

Fubini Theorem: For any *integrable* function $f : X \to \mathsf{R}$, f_{ω_1} is μ_2-integrable for μ_1-a.e. $\omega_1 \in X_1$, and f_{ω_2} is μ_1-integrable for μ_2-a.e. $\omega_2 \in X_2$. Furthermore, the equalities

$$\begin{aligned} \int\int_{X_1 \times X_2} f(\omega_1, \omega_2)\, d(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) &= \int_{X_1} \left(\int_{X_2} f(\omega_1, \omega_2)\, d\mu_2(\omega_2) \right) d\mu_1(\omega_1) \\ &= \int_{X_2} \left(\int_{X_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1) \right) d\mu_2(\omega_2) \end{aligned}$$

still hold.

Remark: Fubini-Tonelli theorems are frequently used "in tandem", when on wishes to invert the order of a double integrable: one first verifies $\int |f|\, d(\mu_1 \otimes \mu_2) < \infty$ using Tonelli theorem, and then apply Fubini theorem to conclude $\int f\, d(\mu_1 \otimes \mu_2) = \int\int f\, d\mu_2\, d\mu_1 = \int\int f\, d\mu_1\, d\mu_2$.

Proof of product measure theorem. We first assume the $\mu_1(X_1)$ and $\mu_2(X_2)$ are both finite. In order to prove the measurability of function $\omega_1 \mapsto \mu_2(E_{\omega_1})$ for all set $E \in \mathcal{F}$, define

$$\mathcal{E} \doteq \{E \subseteq X;\ \ \omega_1 \mapsto \mu_2(E_{\omega_1}) \text{ is } \mathcal{F}_1\text{-measurable}\}.$$

It is not difficult to see that

$$\mathcal{E} \supseteq \mathcal{R} \doteq \{A \times B;\ \ A \in \mathcal{F}_1,\ B \in \mathcal{F}_2\};$$

here $\mathcal{R}$ is the collection of all the rectangles. Indeed, if $E = A \times B$, $\mu_2(E_{\omega_1}) = \mu_2(B) \cdot 1_A$ is clearly $\mathcal{F}_1$-measurable. Since $\mathcal{R}$ is a π-class, it remains to show that $\mathcal{E}$ is also a λ-class. Obviously, $\emptyset \in \mathcal{E}$. Take $E, F \in \mathcal{E}$ and $E \subseteq F$, we have $\mu_2\{(F \backslash E)_{\omega_1}\} = \mu_2(F_{\omega_1}) - \mu_2(E_{\omega_1})$, which is also $\mathcal{F}_1$-measurable. Moreover, for an increasing sequence $\{E^{(n)}\}_n \subseteq \mathcal{E}$, and $E = \cup_n E^{(n)}$, we have

$$\mu_2(E_{\omega_1}) = \mu_2\left\{\cup_n E^{(n)}_{\omega_1}\right\} = \lim_n \mu_2\left\{E^{(n)}_{\omega_1}\right\}.$$

But limit of measurable functions is still measurable, we have $E \in \mathcal{E}$. This concludes that $\mathcal{E}$ is a λ-class, hence $\mathcal{E} \supseteq \sigma(\mathcal{R}) = \mathcal{F}_1 \otimes \mathcal{F}_2$.

Next we show that μ defines a measure. Indeed, for a disjoint sequence $\{E^{(n)}\}$ and $E = \cup_n E^{(n)}$, we have

$$\mu(E) = \int_{X_1} \mu_2(E_{\omega_1})\, d\mu_1(\omega_1) = \int_{X_1} \sum_n \mu_2(E^{(n)}_{\omega_1})\, d\mu_1(\omega_1) = \sum_n \int_{X_1} \mu_2(E^{(n)}_{\omega_1})\, d\mu_1(\omega_1) = \sum_n \mu(E^{(n)}).$$

Clearly, $\mu(E_1 \times E_2) = \mu_1(E_1) \cdot \mu_2(E_2)$. Uniqueness follows from the fact that $\mathcal{R}$ is an elementary family (hence the collection of its finite disjoint unions form an algebra), and the uniqueness part of Hahn extension theorem.

Now repeat the above steps to see that $\tilde{\mu}(E) \doteq \int_{X_2} \mu_1(E_{\omega_2})\, d\mu_2(\omega_2)$ also defines a measure on $(X, \mathcal{F})$ such that $\tilde{\mu}(E_1 \times E_2) = \mu_1(E_1) \cdot \mu_2(E_2)$. It follows from uniqueness that $\mu = \tilde{\mu}$, and the equalities follow.

In general, if $X_{1,2}$ are only σ-finite, we can express X as the limit of an increasing sequence of rectangles in $\mathcal{R}$. It suffices to establish the theorem in each of these rectangles, which we have already done. □

Proof of Tonelli theorem. The equalities obviously hold when $f = 1_E$, $\forall\, E \in \mathcal{F}$, thanks to the product measure thorem. Hence the equalities also hold for simple functions. For a general non-negative measurable function f, there exist a sequence of simple functions $\{f_n\}$ such that $f_n \uparrow f$ pointwise. The equalities follow by applying MCT twice. □

Proof of Fubini theorem. Write $f = f^+ - f^-$. It follows that $\int_X f^\pm \, d\mu < \infty$, which proves the first part of the theorem, and the equalities follow by applying Tonelli theorem to $f^\pm$ separately. □

7 Fundamental inequalities

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. For every $0 < p < \infty$, define

$$\mathsf{L}^p := \mathsf{L}^p(\Omega, \mathcal{F}, \mu) \doteq \{f : \Omega \to \mathsf{R}; \quad f \text{ is measurable and } I(|f|^p) < \infty\}$$

Furthermore, for any $f \in \mathsf{L}^p$, its **L^p-norm** is defined as $\|f\|_p \doteq \{I(|f|^p)\}^{\frac{1}{p}}$. Note, L^1 is just the space of integrable functions, and $f \in \mathsf{L}^p$ if and only if $|f|^p \in \mathsf{L}^1$.

Chebychev inequality: For any $0 < p < \infty$ and measurable function f, we have

$$\mu\left(\{\omega;\ |f(\omega)| \geq a\}\right) \leq \frac{I(|f|^p)}{a^p}; \qquad \forall\ a > 0.$$

The proof is left as an exercise.

Hölder inequality: For $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q$$

for any measurable functions. If $f \in \mathsf{L}^p$ and $q \in \mathsf{L}^q$, then the inequality implies $fg \in \mathsf{L}^1$. Furthermore, in this case, the equality holds if and only if $|f|^p$ and $|g|^q$ are linearly dependent (that is, there exists constants $c_{1,2}$ such that $c_1|f|^p + c_2|g|^q = 0$ almost surely).

Proof. Note that the function $\log(x)$ is a strictly concave function on R^+. It follows that $\log[\lambda x + (1-\lambda)y] \geq \lambda \log(x) + (1-\lambda)\log(y)$ for all $x, y \geq 0$ and $0 < \lambda < 1$, and the equality holds if and only if $x = y$. This inequality is equivalent to

$$x^\lambda y^{1-\lambda} \leq \lambda x + (1-\lambda) y, \qquad \forall\ x, y \geq 0,\ 0 < \lambda < 1.$$

If either $\|f\|_p \in \{0, \infty\}$ or $\|g\|_q \in \{0, \infty\}$, the inequality automatically holds. Otherwise, let

$$x \doteq \left(\frac{|f|}{\|f\|_p}\right)^p, \quad y \doteq \left(\frac{|g|}{\|g\|_q}\right)^q, \quad \lambda = \frac{1}{p}.$$

It follows that

$$\frac{|f|}{\|f\|_p} \cdot \frac{|g|}{\|g\|_q} \le \frac{1}{p}\left(\frac{|f|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g|}{\|g\|_q}\right)^q,$$

which implies that

$$\frac{\int_\Omega |fg|\, d\mu}{\|f\|_p \|g\|_q} \le \frac{1}{p} + \frac{1}{q} = 1.$$

This completes the proof. □

Minkowski inequality: For every $1 \le p < \infty$, and $f, g \in \mathsf{L}^p$, we have

$$\|f+g\|_p \le \|f\|_p + \|g\|_p.$$

Proof. The case where $\|f+g\|_p = 0$ or $p = 1$ is trivial. Assume now $p > 1$ and $\|f+g\|_p \neq 0$. First, observe that $f + g \in \mathsf{L}^p$ since $|f+g|^p \le c_p(|f|^p + |g|^p)$ for some constant c_p independent of f, g. Clearly

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

which implies, with $(p-1) \cdot q = p$ where $\frac{1}{p} + \frac{1}{q} = 1$,

$$I(|f+g|^p) \le I(|f| \cdot |f+g|^{p-1}) + I(|g| \cdot |f+g|^{p-1}) \le \|f\|_p \cdot \{I(|f+g|^p)\}^{\frac{1}{q}} + \|g\|_p \cdot \{I(|f+g|^p)\}^{\frac{1}{q}}$$

which implies

$$I(|f+g|^p)^{1-\frac{1}{q}} \le (\|f\|_p + \|g\|_p).$$

The Minkowski inequality follows readily. □

Jensen inequality: Suppose $(\Omega, \mathcal{F}, \mu)$ is a *probability space*; that is $\mu(\Omega) = 1$. For any convex function $\phi : \mathsf{R} \to \mathsf{R}$ and any integrable function $h : \Omega \to \mathsf{R}$, we have

$$\phi\big(I(h)\big) \le I\big(\phi(h)\big), \quad \text{or} \quad \phi\left(\int h\, d\mu\right) \le \int \phi\big(h(x)\big)\, d\mu(x)$$

Proof. Let $x_0 = I(h)$. Since ϕ is a convex function, there exists a number B such that

$$\phi(x) - \phi(x_0) \ge B(x - x_0); \qquad \forall\, x \in \mathsf{R}.$$

see the remark below. It follows that

$$I\big(\phi(h)\big) - \phi(x_0) \cdot \mu(\Omega) \ge B\left(\int_\Omega h\, d\mu - x_0 \cdot \mu(\Omega)\right) = B\big[I(h) - I(h)\big] = 0.$$

In other words, $I\big(\phi(h)\big) \ge \phi\big(I(h)\big)$. □

Remark: Every (proper) convex function f on R is continuous. Its left- and right- derivatives

$$D^+ f(x) \doteq \lim_{h \to 0+} \frac{f(x+h) - f(x)}{h}; \qquad D^- f(x) \doteq \lim_{h \to 0-} \frac{f(x+h) - f(x)}{h}$$

exist for every $x \in \mathsf{R}$. Furthermore, $D^\pm f$ are both non-decreasing functions, and $D^- f(x) \le D^+ f(x) \le D^- f(y)$ for all $x < y$. For every $B \in [D^- f(x_0), D^+ f(x_0)]$, the inequality

$$f(x) - f(x_0) \ge B(x - x_0), \qquad \forall\, x \in \mathsf{R}.$$

Exercise: Suppose $(\Omega, \mathcal{F}, \mu)$ is a finite measure space. Show that $\mathsf{L}^r \subseteq \mathsf{L}^p$ for all $0 < p < r < \infty$.

Exercise (Young's inequality): If $f : \mathsf{R}^n \to \mathsf{R} \in \mathsf{L}^1$, $g : \mathsf{R}^n \to \mathsf{R} \in \mathsf{L}^p$ for some $1 \le p < \infty$, then the convolution

$$(f * g)(x) \doteq \int_{\mathsf{R}^n} f(x-y)g(y)\,dy$$

satisfies

$$\|f * g\|_p \le \|f\|_1 \cdot \|g\|_p.$$

8 Modes of convergence

Definition: Consider a measure space $(\Omega, \mathcal{F}, \mu)$ and a sequence of measurable functions $\{f_n\}$. Let f be a measurable function. We say

1. $f_n \to f$ **almost surely** (a.s., a.e.) if $\lim_n f_n(\omega) = f(\omega)$ for μ-a.e. ω. Notation: $f_n \xrightarrow{a.s.} f$.
2. $f_n \to f$ **in measure** if $\mu\{|f_n - f| \ge \epsilon\} \to 0$ for every fixed $\epsilon > 0$. Notation: $f_n \xrightarrow{\mu} f$. In particular, if $\mu = \mathsf{P}$ is a probability measure, we say $f_n \to f$ **in probability**, which is denoted by $f_n \xrightarrow{P} f$.
3. $f_n \to f$ **in** L^p for some $p > 0$, if $\|f_n - f\|_p \to 0$. Notation: $f_n \xrightarrow{\mathsf{L}^p} f$.

Exercise: Suppose $f_n \xrightarrow{\mathsf{L}^p} f$ for some $p > 0$. Then $f_n \xrightarrow{\mu} f$.

Exercise: Suppose $f_n \xrightarrow{\mu} f$ and $f_n \xrightarrow{\mu} g$. Then $f = g$ almost surely.

Remark: In general, $f_n \xrightarrow{a.s.} f$ does not imply $f_n \xrightarrow{\mu} f$. For example $f_n = 1_{[n,n+1)}$. Then $f_n \to 0$ almost surely. But f_n does not converge to 0 in measure. However, this is true if μ is a finite measure.

Proposition: Suppose $\mu(\Omega) < \infty$. If $f_n \xrightarrow{a.s.} f$, then $f_n \xrightarrow{\mu} f$. Indeed, $f_n \xrightarrow{a.s.} f$ if and only if $\lim_n \mu\{|f_m - f| \ge \epsilon \text{ for some } m \ge n\} = 0$ for every $\epsilon > 0$.

Proof. We have

$$\{\omega;\ \lim_n f_n(\omega) = f(\omega)\} = \cap_{k=1}^{\infty} \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} \left\{\omega;\ |f_m(\omega) - f(\omega)| \le \frac{1}{k}\right\}.$$

It follows that

$$\{\omega;\ \lim_n f_n(\omega) \ne f(\omega)\} = \cup_{k=1}^{\infty} \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} \left\{\omega;\ |f_m(\omega) - f(\omega)| > \frac{1}{k}\right\} := \cup_{k=1}^{\infty} B_k := \cup_{k=1} \left(\cap_{n=1}^{\infty} B_{n,k}\right).$$

Since $\mu(\Omega) < \infty$ and $B_{n,k} \downarrow B_k$ as $n \to \infty$, we have $\mu(B_k) = \lim_n \mu(B_{n,k})$.

"$\Leftarrow$": by assumption, $\mu(B_{n,k}) \to 0$ as $n \to \infty$, hence $\mu(B_k) = 0$ for all k, which implies that

$$\mu\left\{\omega;\ \lim_n f_n(\omega) \ne f(\omega)\right\} = 0, \quad \text{or} \quad f_n \xrightarrow{a.s.} f.$$

"$\Rightarrow$": by assumption, $\mu\left(\cup_k B_k\right) = 0$, therefore $\mu(B_k) = 0$. □

Remark: In particular, when $\mu = \mathsf{P}$ is a probability measure, then $f_n \xrightarrow{a.s.} f$ if and only if $\mathsf{P}\{|f_n - f| \geq \epsilon, \text{ i.o.}\} = 0$ for all ϵ. Here "i.o." means "infinitely often".

Remark: In general, $f_n \xrightarrow{\mu} f$ does not imply $f_n \xrightarrow{a.s.} f$. For example, let $\Omega = (0,1]$, $\mathcal{F} = \mathcal{B}(\Omega)$ and μ be the Lebesgue measure. For every n, write $n = 2^k + j$ for some $j = 0, 1, \cdots, 2^k - 1$, and $k \in \mathsf{N}$. Take

$$f_n(\omega) = f_{2^k+j}(\omega) \doteq 1_{\left(j2^{-k},(j+1)2^{-k}\right]}(\omega).$$

Clearly, $I(f_n) = 2^{-k} \to 0$; in particular, $f_n \xrightarrow{\mu} 0$. But for every $\omega \in (0,1)$,

$$f_n(\omega) = \left\{ \begin{array}{l} 1, \text{ for infinitely many times} \\ 0, \text{ for infinitely many times} \end{array} \right\}.$$

Hence $\mu(\lim_n f_n = 0) = 0$.

Proposition: Suppose $f_n \xrightarrow{\mu} f$. There exists a subsequence $\{f_{n_k}\}_{k \in \mathsf{N}}$ such that $f_{n_k} \xrightarrow{a.s.} f$.

Proof. By assumption, we can find a subsequence $\{f_{n_k}\} := \{g_k\}$ such that

$$\mu(E_k) \leq \frac{1}{2^k} \quad \text{where} \quad E_k \doteq \left\{\omega; \quad |g_k(\omega) - f(\omega)| > \frac{1}{k}\right\}.$$

Now the set

$$\{\omega; \ \lim_k g_k(\omega) \neq f(\omega)\} = \cup_{k=1}^{\infty} \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} \left\{\omega; \ |g_m(\omega) - f(\omega)| > \frac{1}{k}\right\}.$$

However, for every $m \geq n \geq k$,

$$\mu\left\{\omega; \ |g_m(\omega) - f(\omega)| > \frac{1}{k}\right\} \leq \mu\left\{\omega; \ |g_m(\omega) - f(\omega)| > \frac{1}{m}\right\} \leq 2^{-m}.$$

Therefore, for $n \geq k$,

$$\mu\left\{\cup_{m=n}^{\infty}\left\{\omega; \ |g_m(\omega) - f(\omega)| > \frac{1}{k}\right\}\right\} \leq \sum_{m=n}^{\infty} 2^{-m} = 2^{-(n-1)},$$

which in turn implies

$$\mu\left\{\cap_{n=1}^{\infty} \cup_{m=n}^{\infty}\left\{\omega; \ |g_m(\omega) - f(\omega)| > \frac{1}{k}\right\}\right\} = 0.$$

This shows $g_k \xrightarrow{a.s.} f$. □

Below is a collection of exercises.

Exercise: Suppose $f_n \geq 0$ and $f_n \xrightarrow{\mu} f$. Show that

$$I(f) \leq \liminf_n I(f_n)$$

Exercise: If $f_n \xrightarrow{\mu} f$ and $g_n \xrightarrow{\mu} g$, then

1. $f_n + g_n \xrightarrow{\mu} f + g$
2. $f_n g_n \xrightarrow{\mu} fg$ if $\mu(\Omega) < \infty$; but not necessarily otherwise.
3. $\phi(f_n) \xrightarrow{\mu} \phi(f)$ if $\phi : \mathsf{R} \to \mathsf{R}$ is uniformly continuous.

Exercise: Suppose $\mu(\Omega) < \infty$. Then

$$f_n \xrightarrow{\mu} f \quad \text{if and only if} \quad \rho\left(f_n, f\right) \to 0;$$

where

$$\rho(f, g) \doteq I\left(\frac{|f-g|}{1+|f-g|}\right)$$

is a metric on the space of (almost surely equivalence class) measurable functions $f : \Omega \to \mathsf{R}$.

Exercise: If $|f_n| \leq g \in \mathsf{L}^1$ and $f_n \xrightarrow{\mu} f$. Then $\lim_n I(f_n) = I(f)$.