# CHAPTER 1: MEASURE THEORY

In this chapter we define the notion of measure $\mu$ on a space $\Omega$, construct integrals on this space, and establish their basic properties under limits. The measure $\mu(E)$ will be defined for appropriate subsets $E$ of $\Omega$, called "measurable sets", and generalizes the familiar notions of length, area and volume. As such, it is a non-negative number which may well be infinite. In general, not all subsets of $\Omega$ can be taken as measurable. To require otherwise may sometimes be inconsistent with geometric properties of $\Omega$, as in the case of the real line $\mathbf{R}$ with the translation-invariant Euclidean measure (see Exercise $A.1$ in Appendix $A$) – and sometimes not even desirable, as in the probabilistic notion of conditional expectation (section 2.8). Thus, a family $\mathcal{F}$ of measurable subsets of $\Omega$ must be given, and must have some specific properties consistent with the definition of measure, namely those of a $\sigma$-*algebra*. The properties of $\sigma$-algebras and of measures are self-evident enough if we consider only *finite* unions of measurable sets. It is their strengthening to *countable* unions which accounts for the power and flexibility of the theory, but also for the difficulties encountered in the construction of measures.

This strengthening is necessary for the development of a *theory of integration* with powerful convergence properties, which allow the interchange of integrals and limits under fairly general conditions. Such a theory was developed in the beginning of the twentieth century by H. Lebesgue, who built on the more "classical" theory that culminated with the Riemann integral in the middle of the nineteenth century.

The subject has its roots in the *method of exhaustion*, invented in antiquity by Eudoxos ($4^{th}$ century BC) and greatly developed by Archimedes ($3^{rd}$ century BC) for the purpose of calculating areas and volumes of geometric figures. This method is the precursor of the modern concept of *limit:* according to it, a convex region is approximated by inscribed (or circumscribed) polygons, whose areas are relatively easy to calculate, and then the number of vertices of the polygons is increased until the area of the region has been "exhausted". The much later work of Newton and Leibniz in the late $17^{th}$ century made this method a systematic and powerful tool for such calculations. By the twentieth century the main applications of this theory had shifted from geometry and elementary mechanics to differential equations, convergence of orthogonal expansions, and probability theory.

## 1.1. MEASURES AND INTEGRALS

Let $\Omega$ be a given non-empty set. A family $\mathcal{F}$ of subsets of $\Omega$ is said to be an *algebra*, if it contains the empty set $\emptyset$ and is closed under complements and finite unions; that is, if

(i) $E \in \mathcal{F}$ implies $E^c \equiv \Omega \setminus E \in \mathcal{F}$;

(ii) for any sets $E_1, \cdots, E_N$ in $\mathcal{F}$ and $N \in \mathbf{N}$, we have $\cup_{n=1}^N E_n \in \mathcal{F}$.

An algebra $\mathcal{F}$ is called *$\sigma$-algebra*, if it is closed under countable unions; that is, if

(iii) for any sequence $\{E_n\}_{n\in\mathbf{N}} \subseteq \mathcal{F}$ we have $\cup_{n\in\mathbf{N}} E_n \in \mathcal{F}$.

Given a nonempty set $\Omega$ its *power set* $\mathcal{P}(\Omega)$, the set of all its subsets, is the largest $\sigma$-algebra with which it can be endowed; and the *trivial* $\sigma$-algebra $\{\emptyset, \Omega\}$ the smallest.

With $\Omega = \mathbf{R}$, the collection $\mathcal{F} = \{A \subset \mathbf{R} : A \text{ or } A^c \text{ is countable}\}$ is easily checked to be a $\sigma$-algebra. On the other hand, with $\Omega = (0,1]$ the collection $\mathcal{A}$ that contains the empty set and all finite disjoint unions of intervals of the form $(a,b]$ with $0 \le a < b \le 1$ is an algebra, but not a $\sigma$-algebra.

A pair $(\Omega, \mathcal{F})$ with $\mathcal{F}$ a *$\sigma$-algebra* of subsets of $\Omega$ will be called a *measurable space.* Then a set function $\mu : \mathcal{F} \to [0,\infty]$ is called a **measure**, if $\mu(\emptyset) = 0$ and if $\mu$ is *countably additive* in the sense that

$$\mu(\cup_{n\in\mathbf{N}} E_n) = \sum_{n\in\mathbf{N}} \mu(E_n) \tag{1.1}$$

holds for every sequence $\{E_n\}_{n\in\mathbf{N}}$ of pairwise disjoint subsets in $\mathcal{F}$. The triple $(\Omega, \mathcal{F}, \mu)$ is said to be a **measure space**, and the subsets $E$ in $\mathcal{F}$ are said to be *measurable.*

A set function $\mu : \mathcal{P}(\Omega) \to [0,\infty]$ with $\mu(E) \le \mu(F)$ for $E \subseteq F \subseteq \Omega$ is called *finite*, if $\mu(\Omega) < \infty$; it is called *$\sigma$-finite*, if there exists a sequence $\{E_n\}_{n\in\mathbf{N}} \subseteq \mathcal{F}$ such that $\Omega = \cup_{n\in\mathbf{N}} E_n$ and $\mu(E_n) < \infty$, $\forall\, n \in \mathbf{N}$. A measure $\mu$ on $(\Omega, \mathcal{F})$ is called a *probability*, if $\mu(\Omega) = 1$; then $(\Omega, \mathcal{F}, \mu)$ is called *probability space.*

Occasionally we shall abuse notation and refer to $\Omega$ itself as the "measure space", unless we consider simultaneously several different $\sigma$-algebras $\mathcal{F}$ and/or measures $\mu$. It should be noted that we have the *monotonicity property* $\mu(E) \le \mu(F)$ for any sets $E \subseteq F$ in $\mathcal{F}$, and that the *countable subadditivity property*

$$\mu(\cup_{n\in\mathbf{N}} E_n) \le \sum_{n\in\mathbf{N}} \mu(E_n) \qquad \text{holds for } \textit{any} \text{ sequence } \{E_n\}_{n\in\mathbf{N}} \subseteq \mathcal{F}. \tag{1.1$'$}$$

Given a measurable space $(\Omega, \mathcal{F})$, a function $f : \Omega \to \mathbf{R}$ will be said to be **measurable** if $f^{-1}((a,\infty))$ belongs to the $\sigma$-algebra $\mathcal{F}$ of measurable sets, for every $a \in \mathbf{R}$. Taking complements, countable unions and intersections we deduce that $f^{-1}(I) \in \mathcal{F}$ *holds for any interval $I$.* Useful here are the observations $[a,\infty) = \cap_{n\in\mathbf{N}} (a - 1/n, \infty)$,

$$[a,b] = \bigcap_{n\in\mathbf{N}} (a - 1/n, b + 1/n), \qquad (a,b) = \bigcup_{n\in\mathbf{N}} [a + 1/n, b - 1/n]$$

(which show that any $\sigma-$algebra that contains the open intervals also contains the closed ones, and vice-versa), as well as

$$f^{-1}\left(\cup_{\alpha\in A} E_\alpha\right) = \cup_{\alpha\in A}\, f^{-1}\left(E_\alpha\right)\,, \qquad f^{-1}\left(\cap_{\alpha\in A} E_\alpha\right) = \cap_{\alpha\in A}\, f^{-1}\left(E_\alpha\right)\,.$$

These latter are valid for any function $f : \Omega \to X$ and any collection $\{E_\alpha\}_{\alpha\in A}$ of nonempty subsets of the (arbitrary) space $X$.

When $f : \Omega \to \mathbf{R}$ takes only a finite number of values, we say that $f$ is a **simple** function; we shall denote by $\mathcal{S}$ the class of all simple functions, and by $\mathcal{S}_+$ its subclass of non-negative simple functions. We shall also consider $\mathcal{S}_+^*$, the class of functions $f : \Omega \to [0,\infty]$ that take only finite number of values, possibly $+\infty$. The integral of such a function $f \in \mathcal{S}_+^*$ is defined as

$$I(f) \equiv \int_\Omega f\, d\mu := \sum_{k=1}^{K} y_k \cdot \mu\big(f^{-1}\{y_k\}\big), \tag{1.2}$$

where $f(\Omega) := \{y_1, \cdots, y_K\} \subseteq [0,\infty]$ is the range of $f$. The right-hand side of (1.2) is well-defined even when $y_k$ or $\mu(f^{-1}\{y_k\})$ is infinite, provided we set $0\cdot\infty = 0$. Since $f$ is non-negative, no ambiguous expression of the form $\infty - \infty$ can appear. We have also

$$I(cf) = c\, I(f) \qquad \text{for any real constant } c \geq 0 \text{ and any } f \in \mathcal{S}_+\,. \tag{1.2$'$}$$

- We define the **Lebesgue integral** of any measurable function $f : \Omega \to [0,\infty)$ as

$$I(f) \equiv \int_\Omega f\, d\mu \;:=\; \sup_{\substack{g\in\mathcal{S}\\ 0\le g\le f}} \int_\Omega g\, d\mu\,. \tag{1.3}$$

Here the supremum on the right-hand side is taken over all non-negative simple functions $g \in \mathcal{S}_+$ satisfying $g \leq f$ *pointwise*, that is, $g(\omega) \leq f(\omega)$ for every $\omega \in \Omega$. (The Eudoxian method of "exhaustion" once again!) We have clearly the *monotonicity property*

$$I(f_1) \;\leq\; I(f_2)\,, \qquad \text{if} \quad 0 \leq f_1 \leq f_2 \tag{1.3$'$}$$

for this integral, as well as the property

$$I(cf) = c\, I(f) \quad \text{for any } c \in [0,\infty) \text{ and any measurable } f : \mathbf{R} \to [0,\infty)\,. \tag{1.2$''$}$$

Such a function $f$ is said to be **integrable**, if the supremum of (1.3) is finite; to wit, if $I(f) < \infty$.

• Now if $f : \Omega \to \mathbf{R}$ is any real-valued, measurable function, it can be seen easily (cf. Exercise 1.9) that its positive and negative parts $f^{\pm}$, as well as its absolute value $|f|$, namely

$$f^+ = \max(f, 0), \quad f^- = \max(-f, 0), \quad |f| = f^+ + f^- \,, \tag{1.4}$$

are all measurable. The real-valued function $f$ is then said to be *integrable*, if $|f|$ is integrable; in this case $f^{\pm}$ are also clearly integrable, and we can set

$$I(f) \;\equiv\; \int_\Omega f\,d\mu \;:=\; \int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu \;=\; I(f^+) - I(f^-). \tag{1.5}$$

It is seen from (1.2)″ that we have

$$I(cf) = c\,I(f) \qquad \text{for any real constant } \; c\,, \; \text{ and any integrable } \; f : \Omega \to \mathbf{R}\,. \qquad (1.2)'''$$

Here are some simple examples of measure spaces.

**(a):** The *counting measure* on a countable space $\Omega = \{\omega_1, \cdots, \omega_N\}$ with $1 \le N \le \infty$, with $\mu(E)$ the number of elements in $E$ (which can be infinite); here all subsets $E$ are measurable.

**(b):** The *normalized counting measure*, defined on finite spaces $\Omega = \{\omega_1, \cdots, \omega_N\}$; this measure $\mu$ is the counting measure divided by $N$ and satisfies $\mu(\Omega) = 1$.

**(c):** The measure space for $n$ *coin-tosses*; this is just the normalized counting measure on a space $\Omega$ with $N = 2^n$ elements, conveniently represented as $n-$tuples $\omega = (\omega_1, \cdots, \omega_n)$ with $\omega_i \in \{0, 1\}$. Each $\omega_i$ can be interpreted as the outcome of the $i$-th coin toss, with $\omega_i = 1$ when the outcome is "heads" and $\omega_i = 0$ when the outcome is "tails".

**(d):** The *Dirac measure* $\delta_c$ at the point $c \in \mathbf{R}^d$, on the Euclidean space $\Omega = \mathbf{R}^d$ with all subsets of $\Omega$ measurable, and with $\delta_c(E)$ equal to either 1 or 0, depending on whether the set $E$ contains the point $c$ or not.

To illustrate the subsequent discussion, we shall also assume for the moment that

**(e):** there exists a unique measure $\lambda$ on the real line $\mathbf{R}$, which is defined on a $\sigma$-algebra that contains all intervals in $\mathbf{R}$ and coincides with the notion of length when restricted to intervals, namely

$$\lambda(E) = b - a, \quad \text{when} \quad E = (a, b]. \tag{1.6}$$

This measure $\lambda$ is called the **Lebesgue measure** on the real line. We have the following generalization.

**(f):** For every increasing, right-continuous function $F : \mathbf{R} \to \mathbf{R}$ there exists a unique measure $\mu_F$ defined on a $\sigma$-algebra containing all intervals in $\mathbf{R}$, so that

$$\mu_F(E) = F(b) - F(a) \quad \text{when} \quad E = (a, b]. \tag{1.7}$$

This measure is called the **Lebesgue-Stieltjes measure** associated with $F$. It reduces to the Lebesgue measure when $F(x) = x$. We shall construct this measure in Section 1.4, and study it further in Appendix $A$.

We postpone to §1.4.C a discussion of measures on finite-dimensional Euclidean spaces, and to §1.6.A the more delicate issue of constructing measures on infinite-dimensional spaces.

**1.1 Remark:** The intersection of $\sigma$-algebras is a $\sigma$-algebra; thus, for any given family $\mathcal{G}$ of subsets of $\Omega$, there is a smallest $\sigma$-algebra, denoted by $\sigma(\mathcal{G})$, that contains $\mathcal{G}$. We say then that the $\sigma$-algebra $\sigma(\mathcal{G})$ *is generated by the family* $\mathcal{G}$.

**1.1 Definition:** If $\Omega$ is a *metric space*, with $\mathcal{O}$ denoting the collection of its open sets, we denote by $\mathcal{B}(\Omega) := \sigma(\mathcal{O})$ the collection of **Borel subsets;** this is the smallest $\sigma$-algebra that contains the open sets of $\Omega$. A function $f : \Omega \to \mathbf{R}$ is then called **Borel measurable,** if $f^{-1}((a,\infty)) \in \mathcal{B}(\Omega)$ holds for every $a \in \mathbf{R}$.

Since any open set in $\mathbf{R}$ is a countable union of open intervals, *the class of Borel measurable functions includes all continuous functions;* that is, all functions $f : \Omega \to \mathbf{R}$ such that $f^{-1}(U) \in \mathcal{O}$ holds for every open set $U$ of $\mathbf{R}$.

**1.2 Remark:** The pointwise limit of a sequence of continuous functions can easily fail to be continuous. This is a big nuisance in classical analysis and in the theory of the Riemann integral: taking the limit under the Riemann-integral sign is always a headache. By contrast, as we shall see in the next section, the pointwise limit of a sequence of measurable functions is always measurable, and the Lebesgue integral behaves very well as far as interchange of limits and integrals is concerned.

**1.3 Remark:** In the case of the real line $\Omega \equiv \mathbf{R}$, the Borel $\sigma$-algebra is also generated by intervals, for instance of the type $(a, b]$ with $-\infty < a < b < \infty$. *There exist subsets of the real line that are not Borel* (see Proposition $A$.2, Appendix $A$); such sets are necessarily uncountable.

In the preceding examples (a)-(d), all functions $f : \Omega \to \mathbf{R}$ are measurable. When $\Omega$ is finite and we consider the normalized counting measure, the integral of $f$ is just the average of the values of $f$. For a countable set $\Omega = \{\omega_n\}_{n\in\mathbf{N}}$ it is an easy exercise to verify that

$$\int_\Omega f \, d\mu = \sum_{n\in\mathbf{N}} f(\omega_n) \cdot \mu(\{\omega_n\}) \tag{1.8}$$

when $f$ is non-negative, or when $f$ is integrable – which in this case implies simply that the series on the right-hand side is absolutely convergent.

We shall use the notation $\int_E f \, d\mu \equiv I(f\chi_E)$ for any $E \in \mathcal{F}$ and any integrable function $f : \Omega \to \mathbf{R}$. It will be seen in §1.4.B that *if* $f : \mathbf{R} \to \mathbf{R}$ *is a continuous function* *and $E$ is any bounded interval, then the function $f\chi_E$ is integrable* (in the above sense) with respect to Lebesgue measure $\lambda$ on the real line. In this case the integral of $f$ coincides with the usual Riemann integral of $f$ over the set $E$.

**1.1 Exercise:** Let $(\Omega, \mathcal{F}, \mu)$ be a finite measure space.

(i) : Show that $E \in \mathcal{F},\ F \in \mathcal{F},\ \mu(E \triangle F) = 0$ imply $\mu(E) = \mu(F)$.

(ii) : We write $E \sim F$ if $\mu(E \triangle F) = 0$ for $E \in \mathcal{F},\ F \in \mathcal{F}$. Show that "$\sim$" defines an equivalence relation on $\mathcal{F}$.

(iii) : Define $\rho(E, F) := \mu(E \triangle F)$ for $E \in \mathcal{F},\ F \in \mathcal{F}$. Establish the triangle inequality $\rho(D, F) \le \rho(D, E) + \rho(E, F)$ for any sets $E$, $F$, $D$ in $\mathcal{F}$, and argue that $\rho$ is a metric on the space of equivalence classes induced by "$\sim$".

**1.2 Exercise:** Consider a nonempty set $\Omega$ with all its subsets measurable, and for a given function $f : \Omega \to [0, \infty]$ define

$$\mu(E) \equiv \sum_{\omega \in E} f(\omega) := \sup_{\substack{F \subseteq E \\ card(F) < \infty}} \left( \sum_{\omega \in F} f(\omega) \right), \qquad A := \{\omega \in \Omega \,|\, f(\omega) > 0\}.$$

(i) : If $A$ is uncountable, then $\mu(A) = \mu(\Omega) = \infty$.

(ii) : If $A$ is countable and $g : \mathbf{N} \to A$ is any bijection, then $\mu(E) = \sum_{g(n) \in E} f(g(n))$.

(iii) : The set-function $\mu$ is a measure.

(iv) : The measure $\mu$ is $\sigma$-finite, *iff:* $A$ is countable, and $f(\omega) < \infty,\ \forall\, \omega \in \Omega$.

**1.3 Exercise:** Let the set-function $\mu : \mathcal{F} \to [0, \infty]$ be a *finitely-additive measure*, i.e., suppose that $\mu(\emptyset) = 0$ and that $\mu\left(\cup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} \mu(E_j)$ holds for any finite collection of pairwise disjoint measurable sets $E_1, \cdots, E_n$. Then

(i) $\mu$ is a measure, *iff*: it is continuous from below (that is, $\mu\left(\cup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$ holds for any increasing sequence $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$);

(ii) $\mu$ is a measure, *iff*: it is continuous from above at $\emptyset$ (that is, $\lim_{n\to\infty} \mu(A_n) = 0$ holds for any decreasing sequence $\{A_n\}_{n\in\mathbf{N}} \subseteq \mathcal{F}$ with $\cap_{n\in\mathbf{N}} A_n = \emptyset$), and $\mu(A_1) < \infty$.

**1.4 Exercise:** For any $E \subseteq \mathbf{N}$, denote by $\gamma_n(E)$ the number of integers in $E$ that do not exceed $n$. If $\mathcal{D}$ is the class of subsets $E$ of $\mathbf{N}$ for which the limit $\nu(E) = \lim_{n\to\infty} (\gamma_n(E)/n)$ exists, then show that $\nu$ is finitely additive, but not countably additive, on $\mathcal{D}$.

**1.5 EXERCISE: Regular Measure.** Suppose that $\Omega$ is a metric space, and that $\mathcal{F}$ is a $\sigma$-algebra that contains the Borel sets: $\mathcal{B}(\Omega) \subseteq \mathcal{F}$. A measure $\mu : \mathcal{F} \to [0, \infty]$ is called *regular*, if for every $E \in \mathcal{F}$ and $\varepsilon > 0$ there exist an open set $G$ and a closed set $F$ such that $F \subseteq E \subseteq G$ and $\mu(G \setminus F) < \varepsilon$.

(i) If $\mu$ as above is regular, then an arbitrary set $D \in \mathcal{F}$ is "very near a Borel set", in the sense that we have

$$B_1 \subseteq D \subseteq B_2, \quad \text{for some } \; B_1,\, B_2 \; \text{ in } \; \mathcal{B}(\Omega) \; \text{ with } \; \mu(B_2 \setminus B_1) = 0$$

as well as

$$D = E \cup F, \quad \text{for some } \; E \in \mathcal{B}(\Omega),\; F \subseteq B \in \mathcal{B}(\Omega) \; \text{ with } \; \mu(B) = 0.$$

(ii) *If a measure $\mu$ on $\mathcal{B}(\Omega)$ is finite on bounded Borel sets, then it is regular.*
(*Hint:* Start with the finite case $\mu(\Omega) < \infty$. Consider the collection $\mathcal{E}$ of subsets $E$ of $\Omega$, such that for each $\varepsilon$ there exist a closed set $F_\varepsilon$ and an open set $G_\varepsilon$, with $F_\varepsilon \subseteq E \subseteq G_\varepsilon$, $\mu\left(G_\varepsilon \setminus F_\varepsilon\right) < \varepsilon$. Argue that $\mathcal{E}$ is a $\sigma$-algebra that contains the closed sets, hence $\mathcal{B}(\Omega) \subseteq \mathcal{E}$. Then try to remove the assumption $\mu(\Omega) < \infty$.)

**1.6 Exercise:** An increasing function $f : \mathbf{R} \to \mathbf{R}$ is Borel-measurable.

**1.7 Exercise:** The pointwise supremum of an *uncountable* family of measurable functions $f_\alpha : \Omega \to \mathbf{R}$, $\alpha \in A$ can fail to be measurable.

**1.8 Exercise:** Let the function $f : \mathbf{R} \times \mathbf{R}^d \to \mathbf{R}$ be such that $f(x, \cdot)$ is $\mathcal{B}(\mathbf{R}^d)-$measurable for every $x \in \mathbf{R}$, and $f(\,\cdot\,, y)$ is continuous for every $y \in \mathbf{R}^d$.
(i) Show that $f$ is then $\mathcal{B}(\mathbf{R} \times \mathbf{R}^d)-$measurable.
(ii) Argue, by induction, that every function on $\mathbf{R}^d$ which is continuous in each of its variables separately, is $\mathcal{B}(\mathbf{R}^d)-$measurable.
(*Hint:* Consider first the case of $g$ simple; then approximate.)

**1.9 EXERCISE: Preservation of measurability under simple operations.** Let $f$, $g$ be real-valued, measurable functions of $(\Omega, \mathcal{F})$. If $c$ is a real number, show that the functions

$$c\,f\,, \qquad f^2\,, \qquad f + g\,, \qquad f\,g\,, \qquad |f|\,, \qquad f^{\pm}$$

are also measurable.