

Chapter II. PROBABILITY BACKGROUND.

1. Measure

The language of option pricing involves that of probability, which in turn involves that of *measure theory*. This originated with Henri LEBESGUE (1875-1941), in his 1902 thesis, '*Intégrale, longueur, aire*'. We begin with the simplest case.

Length. The length $\mu(I)$ of an interval $I = (a, b), [a, b], [a, b)$ or $(a, b]$ should be $b - a$: $\mu(I) = b - a$. The length of the disjoint union $I = \bigcup_{r=1}^n I_r$ of intervals I_r should be the sum of their lengths:

$$\mu\left(\bigcup_{r=1}^n I_r\right) = \sum_{r=1}^n \mu(I_r) \quad (\text{finite additivity}).$$

Consider now an infinite sequence I_1, I_2, \dots (*ad infinitum*) of disjoint intervals. Letting $n \rightarrow \infty$ suggests that length should again be additive over disjoint intervals:

$$\mu\left(\bigcup_{r=1}^{\infty} I_r\right) = \sum_{r=1}^{\infty} \mu(I_r) \quad (\text{countable additivity}).$$

For I an interval, A a subset of length $\mu(A)$, the length of the complement $I \setminus A := I \cap A^c$ of A in I should be

$$\mu(I \setminus A) = \mu(I) - \mu(A) \quad (\text{complementation}).$$

If $A \subseteq B$ and B has length $\mu(B) = 0$, then A should have length 0 also:

$$A \subseteq B \ \& \ \mu(B) = 0 \ \Rightarrow \ \mu(A) = 0 \quad (\text{completeness}).$$

Let \mathcal{F} be the smallest class of sets $A \subset \mathbb{R}$ containing the intervals, closed under countable disjoint unions and complements, and complete (containing all subsets of sets of length 0 as sets of length 0). The above suggests – what Lebesgue showed – that length can be sensibly defined on the sets \mathcal{F} on the line, but on no others. There are others – but they are hard to construct (in technical language: the Axiom of Choice (AC), or some variant of it such

as Zorn's Lemma, is needed to demonstrate the existence of non-measurable sets – but all such proofs are highly non-constructive). So: some but not all subsets of the line have a length.¹ These are called the *Lebesgue-measurable sets*, and form the class \mathcal{F} described above; length, defined on \mathcal{F} is called *Lebesgue measure* μ (on the real line, \mathbb{R}).

Area. The area of a rectangle $R = (a_1, b_1) \times (a_2, b_2)$ – with or without any of its perimeter included – should be $\mu(R) = (b_1 - a_1) \times (b_2 - a_2)$. The area of a finite or countably infinite union of disjoint rectangles should be the sum of their areas:

$$\mu \left(\bigcup_{n=1}^{\infty} R_n \right) = \sum_{n=1}^{\infty} \mu(R_n) \quad (\text{countable additivity}).$$

If R is a rectangle and $A \subseteq R$ with area $\mu(A)$, the area of the complement $R \setminus A$ should be

$$\mu(R \setminus A) = \mu(R) - \mu(A) \quad (\text{complementation}).$$

If $B \subseteq A$ and A has area 0, B should have area 0:

$$A \subseteq B \ \& \ \mu(B) = 0 \ \Rightarrow \ \mu(A) = 0 \quad (\text{completeness}).$$

Let \mathcal{F} be the smallest class of sets, containing the rectangles, closed under finite or countably infinite unions, closed under complements, and complete (containing all subsets of sets of area 0 as sets of area 0). Lebesgue showed that area can be sensibly defined on the sets in \mathcal{F} and no others. The sets $A \in \mathcal{F}$ are called the *Lebesgue-measurable sets* in the plane \mathbb{R}^2 ; area, defined on \mathcal{F} , is called *Lebesgue measure* in the plane. So: some but not all sets in the plane have an area.

Volume. Similarly in three-dimensional space \mathbb{R}^3 , starting with the volume of a cuboid $C = (a_1, b_1) \times (a_2, b_2) \times (a_3, b_3)$ as

$$\mu(C) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot (b_3 - a_3).$$

Euclidean space. Similarly in k -dimensional Euclidean space \mathbb{R}^k . We start with

$$\mu \left(\prod_{i=1}^k (a_i, b_i) \right) = \prod_{i=1}^k (b_i - a_i),$$

¹There are alternatives to AC, under which *all sets are measurable*. So it is not a question of whether AC is *true* or not, but of what axioms of Set Theory we assume. Background: Model Theory in Mathematical Logic, etc.

and obtain the class \mathcal{F} of *Lebesgue-measurable* sets in \mathbb{R}^k , and *Lebesgue measure* μ in \mathbb{R}^k .

Probability.

The unit cube $[0, 1]^k$ in \mathbb{R}^k has Lebesgue measure 1. It can be used to model the *uniform distribution* (density $f(x) = 1$ if $x \in [0, 1]^k$, 0 otherwise), with probability = length/area/volume if $k = 1/2/3$.

Note. If a property holds everywhere except on a set of measure zero, we say it holds *almost everywhere* (a.e.) [French: *presque partout*, p.p.; German: *fast überall*, f.u.]. If it holds everywhere except on a set of probability zero, we say it holds *almost surely* (a.s.) [or, with probability one].

2 Integral.

1. *Indicators.* We start in dimension $k = 1$ for simplicity, and consider the simplest calculus formula $\int_a^b 1 \, dx = b - a$. We rewrite this as

$$I(f) := \int_{-\infty}^{\infty} f(x) \, dx = b - a \quad \text{if } f(x) = I_{[a,b]}(x),$$

the *indicator* function of $[a, b]$ (1 in $[a, b]$, 0 outside it), and similarly for the other three choices about end-points.

2. *Simple functions.* A function f is called *simple* if it is a finite linear combination of indicators: $f = \sum_{i=1}^n c_i f_i$ for constants c_i and indicator functions f_i of intervals I_i . One then extends the definition of the integral from indicator functions to simple functions by linearity:

$$I\left(\sum_{i=1}^n c_i f_i\right) := \sum_{i=1}^n c_i I(f_i)$$

for constants c_i and indicators f_i of intervals I_i .

3. *Non-negative measurable functions.* Call f a (*Lebesgue-*) *measurable function* if, for all c , the sets $\{x : f(x) \leq c\}$ is a Lebesgue-measurable set (§1). If f is a non-negative measurable function, we quote that it is possible to construct f as the increasing limit of a sequence of simple functions f_n :

$$f_n(x) \uparrow f(x) \quad \text{for all } x \in \mathbb{R} \quad (n \rightarrow \infty), \quad f_n \text{ simple.}$$

We then define the integral of f as

$$I(f) := \lim_{n \rightarrow \infty} I(f_n) \quad (\leq \infty)$$

(we quote that this does indeed define $I(f)$: the value does not depend on *which* approximating sequence (f_n) we use). Since f_n increases in n , so does $I(f_n)$ (the integral is *order-preserving*), so either $I(f_n)$ increases to a finite limit, or diverges to ∞ . In the first case, we say f is (*Lebesgue-*) *integrable* with (*Lebesgue-*) *integral* $I(f) = \lim I(f_n)$, or $\int f(x) dx = \lim \int f_n(x) dx$, or simply $\int f = \lim \int f_n$.

4. *Measurable functions.* If f is a measurable function that may change sign, we split it into its positive and negative parts, f_{\pm} :

$$\begin{aligned} f_+(x) &:= \max(f(x), 0), & f_-(x) &:= -\min(f(x), 0), \\ f(x) &= f_+(x) - f_-(x), & |f(x)| &= f_+(x) + f_-(x) \end{aligned}$$

If both f_+ and f_- are integrable, we say that f is too, and define

$$\int f := \int f_+ - \int f_-.$$

Then, in particular, $|f|$ is also integrable, and

$$\int |f| = \int f_+ + \int f_-.$$

Note. The Lebesgue integral is, by construction, an *absolute integral*: f is integrable iff $|f|$ is integrable. Thus, for instance, the well-known formula

$$\int_0^{\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}$$

has no meaning for Lebesgue integrals, since $\int_1^{\infty} \frac{|\sin x|}{x} dx$ diverges to $+\infty$ like $\int_1^{\infty} \frac{1}{x} dx$. It has to be replaced by the limit relation

$$\int_0^X \frac{\sin x}{x} dx \rightarrow \frac{\pi}{2} \quad (X \rightarrow \infty).$$

The class of (*Lebesgue-*) integrable functions f on \mathbb{R} is written $L(\mathbb{R})$ or (for reasons explained below) $L_1(\mathbb{R})$ – abbreviated to L_1 or L .

Higher dimensions. In \mathbb{R}^k , we start instead from k -dimensional boxes. If f is the indicator of a box $B = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_k, b_k]$, $\int f := \prod_{i=1}^k (b_i - a_i)$. We then extend to simple functions by linearity, to non-negative measurable functions by taking increasing limits, and to measurable functions by splitting into positive and negative parts.