

# Practitioner's introduction to measure theory

Himalaya Senapati  
email: [himalay@cmi.ac.in](mailto:himalay@cmi.ac.in)  
[logic-labs.in/mathfinance](http://logic-labs.in/mathfinance)

October 2, 2025

## Abstract

Here we provide a quick non-technical introduction to measure theory and give intuitive descriptions of concepts that will be useful in learning financial calculus.

## Contents

<b>1</b>	<b>Motivation</b>	<b>1</b>
<b>2</b>	<b>What is Measure theory?</b>	<b>2</b>
<b>3</b>	<b>Change of measures and Radon-Nikodym derivative</b>	<b>4</b>
<b>4</b>	<b>Cumulative and Probability density functions</b>	<b>5</b>
<b>5</b>	<b>Concluding first half of motivation</b>	<b>7</b>

## 1 Motivation

• Suppose  $X$  takes values uniformly randomly on the interval  $[0,1]$ . Then how to compute the following probabilities.

- $P(X = 0.5) = ?$

- $P(0.4 < X < 0.5) = ?$
- $P(X \in \mathbb{Q}) = ?$  Here  $\mathbb{Q}$  is the set of rational numbers.

- Applications in mathematical finance.

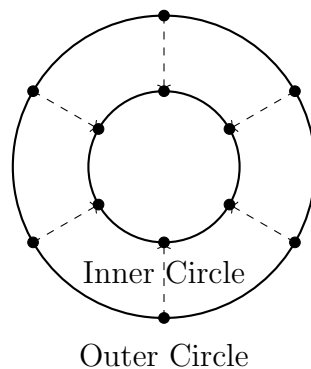
- Equivalent martingale measure / risk-neutral measure.
- T-Forward measure.
- Change of measure via Radon-Nikodym derivative, Change of numeraire.
- Girsanov's theorem.

## 2 What is Measure theory?

- As the name suggests, the theory is about measurements: **length, area, volume**. We want the defined measurement to agree with our intuitive understanding of the world:

- **additivity:**  $length(A \cup B) = length(A) + length(B)$  for disjoint sets  $A$  and  $B$ .
- $length(A \setminus B) = length(A) - length(B)$  for sets  $A \subset B$ .

- **Countable vs Uncountable additivity:** Let us look at the two concentric circles drawn below. As we can see, it is possible to have a one-to-one correspondence between the points in the inner circle to that of the outer circle (*bijection*). A natural question would be, why are the lengths of the two circles not equal. See also the video at this link: [Aristotle's Wheel Paradox](#).



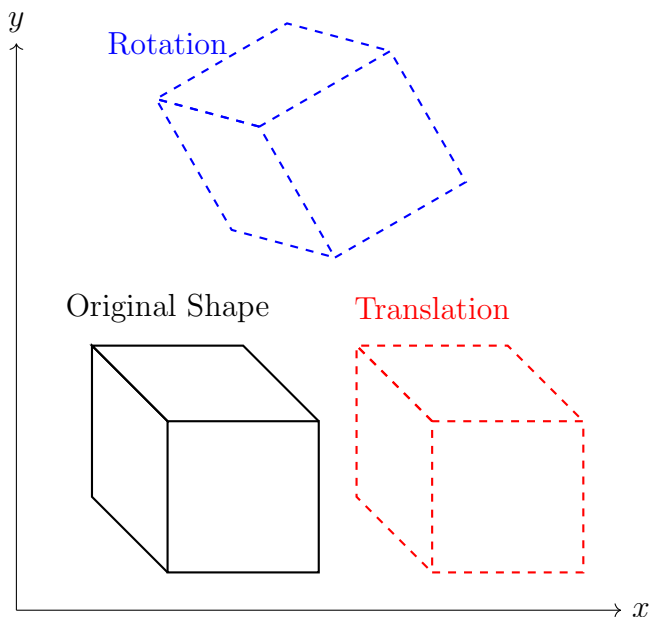
One point to note is that, the number of points on the circle is not *countable*<sup>1</sup>, it is *uncountable*<sup>2</sup>. This observations suggests that we cannot expect the measure of an uncountable union of pair-wise disjoint sets to be equal to the sum of the measure of the sets:

$$\text{length}(\cup_{r \in \mathbb{R}} X_r) \neq \sum_{r \in \mathbb{R}} \text{length}(X_r) \quad \text{for pair-wise disjoint sets } X_r \text{ in general.} \quad (1)$$

Thus, the additive property of measures should hold only up to countably infinite elements:

**countable additivity:**  $\text{length}(\cup_{i=1}^{\infty} X_i) = \sum_{i=1}^{\infty} \text{length}(X_i)$  for pair-wise disjoint sets  $X_i$ . (2)

• **Naive theory of measurement, define measure on all subsets:** First let us set up a naive theory of measurement on the three-dimensional closed unit cube  $\Omega = [0, 1] \times [0, 1] \times [0, 1]$ . Closed means the boundary points such as  $(0,0,0)$  are included. We would ‘naively’ expect that, we can define volume for any subset of  $\Omega$ . Let the function  $\lambda$  denote this volume measure: for any given subset  $X \subset \Omega$ ,  $\lambda(X)$  is the volume of  $X$ . As mentioned above, we expect it to retain the property  $\lambda(X \cup Y) = \lambda(X) + \lambda(Y)$  for disjoint sets  $X, Y \subset \Omega$ . Generalizing this by induction, for pairwise disjoint sets  $X_i \subset \Omega$ ,  $\lambda(\cup_{i=1}^k X_i) = \sum_{i=1}^k \lambda(X_i)$ . We would expect this notion of volume to be translation and rotation invariant: The translated-red and (translated+rotated)-blue cubes below should have the same volume as the original.



• **Inconsistency in this naive theory, Banach–Tarski paradox:** A strong version of the paradox roughly states the following: Given any two three-dimensional bounded subsets  $A$  and  $B$

<sup>1</sup> *Countable infinity:* A set is countably infinite if we can match the elements 1-for-1 with the natural numbers.

<sup>2</sup> *Uncountable infinity:* We cannot match the elements 1-for-1 with 1, 2, 3, ... We will always have some left over no matter what scheme we use. Look up [Cantor’s diagonal argument](#).

with non-empty interior, we can cut both of them into *finitely many* equal number of *congruent* pieces,  $A = \cup_{i=1}^k A_i$  and  $B = \cup_{i=1}^k B_i$  with  $A_i \cap A_j = B_i \cap B_j = \emptyset$  for all  $i \neq j$  such that  $A_i \equiv B_i$ . Here  $A_i \equiv B_i$  means congruence in the geometrical sense. Thus, going by our previous notion of our previous volume measure,

$$\lambda(A) = \lambda\left(\cup_{i=1}^k A_i\right) = \sum_{i=1}^k \lambda(A_i) = \sum_{i=1}^k \lambda(B_i) = \lambda\left(\cup_{i=1}^k B_i\right) = \lambda(B). \quad (3)$$

- In other words, if we try to define volume on all the subsets, while requiring it to respect certain intuitive properties (additivity, rotation & translation invariance), we end up proving that the volume of all subsets are equal, say  $\nu$ . Appealing again to additivity,  $\lambda(X \cup Y) = \lambda(X) + \lambda(Y) \Rightarrow \nu = 2\nu \Rightarrow \nu = 0$  or  $\infty$ . Thus, it turns out to be a trivial measure: volume of all subsets is equal to zero or infinity.

- **Measure space:** As we see above, it is not possible to define measure (say, volume) on all subsets of a set  $\Omega$  that we are interested in. Thus, we need to introduce  $\Sigma$  which is the set off all ‘measurable’ subsets of  $\Omega$ . Moreover, if we can define measure of disjoint sets  $X$  and  $Y$ , we should also be able to define measure of  $X \cup Y$ . Similarly, if both  $U$  and  $V$  are measurable (we can define their measure) with  $U \subset V$ , then we can also define measure of  $V \setminus U$ . In other words,

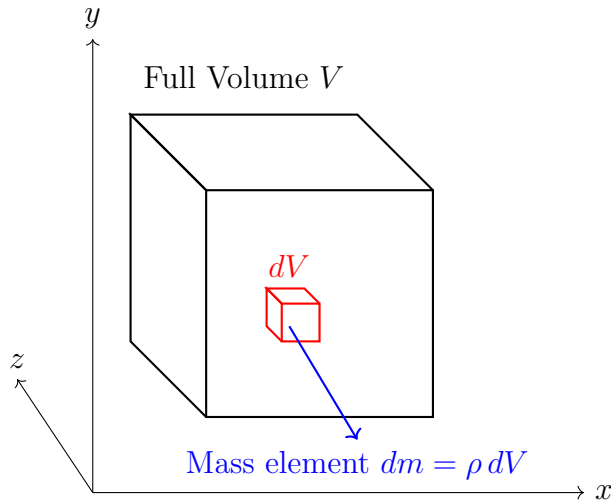
$$\begin{aligned} X, Y \in \Sigma \text{ with } X \cap Y = \emptyset &\Rightarrow X \cup Y \in \Sigma, \\ U, V \in \Sigma \text{ with } U \subset V &\Rightarrow V \setminus U \in \Sigma \end{aligned} \quad (4)$$

As  $\Sigma$  respects union/set-addition and intersection/set-subtraction, it is an algebra. It is called a  $\sigma$ -algebra. The pair  $(\Omega, \Sigma)$  is called a *measurable space* as ‘it can be measured’. Once we define the measure, say  $\mu$ , the tuple  $(\Omega, \Sigma, \mu)$  is called a *measure space* as the measure has already been defined.

- **Practitioner’s view:** As we see above, all notions from the *naive theory* are kept except for the fact that we can define measure, say volume, for all possible subsets. In practice, we do not need to worry about if a set is measurable or not. All sets that we will come across in industry settings will be measurable. We just need to respect certain intuitive properties such as *countable additivity* and we are good to go.

### 3 Change of measures and Radon-Nikodym derivative

- **Volume measure vs mass measure:** Given an existing volume measure  $dV = dx \cdot dy \cdot dz$ , we can always derive a new mass measure  $dm = \rho(x, y, z)dV$ .



It is intuitively clear that we can always find/there exists a density function  $\rho = \frac{dm}{dV}$ . However,  $\frac{dV}{dm}$  may not exist. It boils down to the property that  $dm > 0 \Rightarrow dV > 0$ . But  $dV > 0$  can still mean  $dm = 0$  if that volume element is ‘empty’ and there is no mass. Formally, it may be stated as below.

- **Radon-Nikodym theorem:** The theorem guarantees the existence of a Radon-Nikodym density/derivative  $\rho = \frac{d\mathbb{P}}{d\mathbb{Q}}$  when  $\mathbb{P} \ll \mathbb{Q}$ , i.e., if  $\mathbb{P}$  is absolutely continuous with respect to  $\mathbb{Q}$ .  $\rho$  is analogous to the density seen above, while  $\mathbb{P}$  and  $\mathbb{Q}$  are equivalent to mass and volume measures satisfying  $d\mathbb{P} > 0 \Rightarrow d\mathbb{Q} > 0$ .

- Radon-Nikodym theorem tells us when we can go from one measure to another. For example, we can always represent mass measure in terms of volume measure:  $dm = \rho dV$ , but not the other way. It finds application extensively in math finance, such as in *Girsanov's*.

#### 4 Cumulative and Probability density functions

- The cumulative distribution function of a real-valued<sup>3</sup> random variable  $X$  is defined as

$$F_X(r) = P(X \leq r). \tag{5}$$

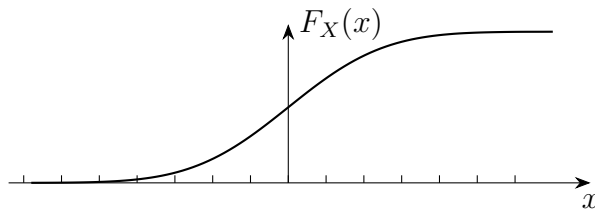
It is clear that  $F_X(-\infty) = 0$  and  $F_X(\infty) = 1$ . Moreover, it has to be a monotonically increasing function of  $x$ . That is, if  $x > y$ ,  $F_X(x) \geq F_X(y)$ . Moreover,

$$P(r_0 < X \leq r_1) = F_X(r_1) - F_X(r_0). \tag{6}$$

Do note the  $<$  and  $\leq$ , it becomes important when CDF becomes discontinuous. Below, we plot the CDF of a normal distribution with mean 0 and variance 1.

---

<sup>3</sup> $X$  takes values between  $-\infty$  and  $\infty$ . For example,  $X$  cannot be  $1 + 5i$ .



- The probability density function  $f(x)$  is used to specify the probability of the random variable falling within a particular range of values, as opposed to taking on any one value. It becomes important in the continuous case where the random variable can take a continuous range of values, say on the interval  $[0,1]$  as opposed to taking discrete values, say  $1, 2, 3, \dots$ .

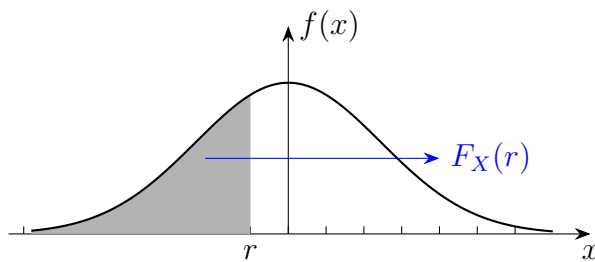
- Roughly speaking,  $P(r < X < r + dr) = f_X(r) dr$ . More formally,

$$P(r_0 < X < r_1) = \int_{r_0}^{r_1} f_X(r) dr.$$

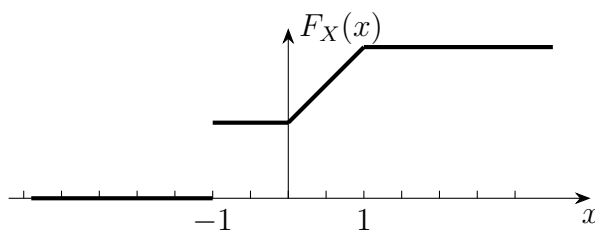
- If  $f_X$  is continuous at  $r$ ,

$$f_X(r) = \frac{d}{dx} F_X(r).$$

- Below, we plot the probability density function ( PDF )  $f(x)$  of a normal distribution while indicating the cumulative distribution function ( CDF ) as area under the curve.



- Note that PDF need not always exist (unless we appeal to Dirac-Delta functions). For example, if the random variable takes the value  $-1$  with probability half, and values uniformly randomly on the interval  $[0,1]$  with another probability half. As we see below, the CDF is not continuous and hence not differentiable at  $-1$ .



• **PDF under coordinate/variable transformation:** We will show an one-dimensional example. For more than one dimensions, the Jacobian becomes a useful tool to do this computation. Say, we have a random variable  $Y = 2X$ , and are asked to compute the PDF of  $Y$ . One way is to see that

$$\begin{aligned} F_Y(r) &= F_X\left(\frac{r}{2}\right) \\ \Rightarrow f_Y(r) &= \frac{d}{dr}F_Y(r) = \frac{d}{dr}F_X\left(\frac{r}{2}\right) = \frac{d(r/2)}{dr} \frac{d}{d(r/2)}F_X\left(\frac{r}{2}\right) = \frac{1}{2}f_X\left(\frac{r}{2}\right). \end{aligned} \quad (7)$$

We can also try to derive it from the first principles from the observation that  $P(r < X < r + dr) = P(2r < Y < 2r + dr)$ .

## 5 Concluding first half of motivation

• **P( $\mathbf{X} = 0.5$ ):** Recall from Sec. 1 that  $X$  takes values uniformly randomly on the interval  $[0,1]$ . As there are infinitely many numbers, and they have equal probability, with the probabilities summing up to 1,  $P(X = 0.5)$  should be equal to  $1/\infty = 0$ . A bit more formally,  $\{0.5\}$  is a measure zero set on the interval  $[0,1]$  as its length-measure is 0. Thus,  $P(X = 0.5)$  is equal to the PDF times the length-measure, which has to vanish.

• **P( $\mathbf{0.4} < \mathbf{X} < \mathbf{0.5}$ ):** The PDF of the uniform distribution over interval  $[0,1]$  is a constant given by  $f_X(r) = 1$  for  $0 \leq r \leq 1$ . Thus,  $P(0.4 < X < 0.5)$  is the integral of the constant function from .4 to .5, giving the value 0.1.

• **P( $\mathbf{X} \in \mathbb{Q}$ ):** As we see above,  $\{0.5\}$  is a measure zero set on the interval  $[0,1]$ . We also note that the set of rationals  $\mathbb{Q}$  is countable. Thus, the length-measure of  $\mathbb{Q}$  is the countable sum of 0's implying that  $\mathbb{Q}$  is also of zero length-measure on the interval  $[0,1]$ . This implies that  $P(X \in \mathbb{Q}) = 0$ .

• Note that, the set of numbers  $\{r : 0.4 < r < 0.5\}$  is uncountable. Thus, even though the sets  $\{0.45\}$  etc are of measure zero, it **does not** imply that, their union  $\{r : 0.4 < r < 0.5\}$  will also be of measure zero, as it is an uncountable union. There are both kinds: (a) **Cantor sets** are uncountable union of point sets and still of measure zero, (b) non-empty **Open intervals (a,b)** are uncountable union of point sets with length measure  $b - a$ .

## References

- [1] Marek Capiński and Ekkehard Kopp, *Measure, Integral and Probability*, Springer-Verlag, New York (1998).