Understanding the Risk Neutral Measure

February 2024

1 Probability Theory

1.1 Sample Space

- All possible outcomes from a random experiment, which will form a set called the *Sample Space* denoted by Ω
  E.g $\Omega = [H, T]$ Is the Sample Space for flipping a coin

  A random number between 0,1 corresponds to $\Omega = (0, 1)$ which is the entire segment between 0 and 1

1.1.1 Subsets of Ω

- All subsets of the sample space Ω form a set denoted 2^{Ω}

- $|\Omega|$ denotes the cardinal of Ω which is the number of elements
- This means the number of elements in a subset, 2^{Ω} of Ω is given by $2^{|\Omega|} = n$. Then 2^{Ω} has 2^n elements

1.2 Events and Probability

The set of parts 2^{Ω} satisfies the following properties

- It contains the empty set $\emptyset$
- If it contains set A, then it also contains its complement $\bar{A} = \frac{\Omega}{A}$
- It is closed with regard to unions, meaning a union of sets belonging to 2^{Ω} also belongs to 2^{Ω}

Any subset F of 2^{Ω} that satisfies the previous three properties is called a σ-field. The sets belonging to F are called events. *This means the complement of event, or union of events is also an event*

The chance of occurrence of an event is measured by a probability function $P : F \to [0, 1]$ which satisfies the following two properties,

- $P(\Omega) = 1$
- For any mutually disjoint events $\in$ F,

$$P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$$

The triplet (Ω, F, P) is called a probability space

2 Probability Measure

When we have a probability space formed by (Ω, F, P), the set function P is the probability Measure when it follows the requirements below

- $P : F \to [0, 1]$
- $P(\Omega) = 1$
- $P(\emptyset) = 0$

- P satisfies countable additivity property, that is

$$P(\cup_{i\in N}E_i) = \Sigma_{i\in N}P(E_i)$$

where E_i are events

- The conditional probability for Probability Measure P is defined as

$$P(B|A) = \frac{P(A\cap B)}{P(A)}$$

2.1 Equivalent Probability Measure

Let Q and P be Probability Measures on (Ω, F). The two measures Q and P are equivalent if and only if for all $A \in F$

$$P(A) = 1 \qquad\qquad Q(A) = 1 \tag{1}$$

From this, its follows that Q and P can assign completely different probabilities on the same event. **However, all events that are impossible under Q are impossible under P and vice versa.** This is also the same for certain $(P(A) = 1)$ events. Also if event A has strictly postie P-probability, then it also has strictly positive Q-probability

2.2 Measurable Function

A measurable function is a function between the underlying sets of two measurable spaces that preserves the structure of the spaces: the preimage of any measurable set is measurable.

Let (Ω, X) and (T, Y) be measurable spaces. A function $f : \Omega \to T$ is said to be measurable if for every $y \in Y$ the pre image of y under f is in X

$$f^{-1}(y) := [x \in \Omega | f(x) \in y \in Y] \in X$$

The pre image / inverse image is the set of all elements of the domain that map to the elements of the codomain / range.

In probability theory, a measurable function on a probability space is known as a random variable.

3 Random Variables

The σ-field provides F the knowledge about which events are possible on the considered probability space (Ω, F, P). A random variable X is a function which is defined as $X : \Omega \to \mathbb{R}$

Given any two numbers $a, b \in \mathbb{R}$, then all the states for which X takes values between a and b forms a set that is an event $\in F$

$$\{\omega \in \Omega : a < X(\omega) < b\} \in F$$

4 Stochastic Process

A stochastic process on the probability space (Ω, F, P) is a family of random variables X_t parameterized by $t \in T$, where $T \subset \mathbb{R}$. If T is an interval, then X_t is a stochastic process in continuous time.

4.1 Filtration

Consider that all the information accumulated until time t is contained by the σ-field $\mathbb{F}_t$. This means $\mathbb{F}_t$ contains the information containing events that have already occurred until time t, and which did not. Since information is growing in time, we have

$$\mathbb{F}_s \subset \mathbb{F}_t \subset \mathbb{F}$$

for any $s, t \in T$ with $s \leq t$. The family $\mathbb{F}_t$ (The family containing all information already occurred until time t) is called a filtration.

This means that the information at time t determines the value of the random variable X_t

5 Adaption to a Filtration: Adapted Process

A stochastic process is adapted (also referred to as a non-anticipating or non-anticipative process) if information about the value of the process at a given time is available at that same time.

The process X is said to be adapted to the filtration $\mathbb{F}$ if the random variable $X_i : \omega \rightarrow \mathbb{R}$ is a $(\mathbb{F}, \mathbb{R})$ measurable function

6 Martingale

A stochastic process X_t, $t \in T$, is called a martingale with respect to the filtration $\mathbb{F}_t$ if

- X_t is integrable for each $t \in T$
- X_t is adapted to the filtration $\mathbb{F}_t$
- $X_s = E[X_t|\mathbb{F}_S$ For all $s < t$

7 Integration in Probability Measure

Let Ω_i be a partition of Ω, then each Ω_i is an event and its associated probability is $P(\Omega_i$
Now consider the *characteristic function* of a set $A \subset \Omega$ which is defined by,

$$\chi_A = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

A *Simple Function* is a sum of characteristic functions $f = \Sigma_i^n c_i \chi_{\Omega_i}$ where $c_i \in \mathbb{R}$
This means $f(\omega) = c_k$ for $\omega \in \Omega_k$
The integral of the simple function f is defined by

$$\int_\Omega f dP = \Sigma_i^n c_i P(\Omega_i)$$

If $X : \Omega \rightarrow \mathbb{R}$ is a random variable, then from measure theory it is known that there is a sequence of simple functions $(f_n)_{n\geq 1}$ satisfying $lim_{n\rightarrow\infty} f_n(\omega) = X(\omega)$.

Furthermore, if $X \geq 0$ then we may assume that $f_n \leq f_{n+1}$. Then we can we can define

$$\int_\Omega X dP = lim_{n\rightarrow\infty} \int_\Omega f_n dP$$

Two widely used properties are as follows

- Linearity: For any two Random Variables X and Y and a,b $\in \mathbb{R}$

$$\int_\Omega (aX + bY) dP = a \int_\Omega X dP + b \int_\Omega Y dP$$

- Positivity: If X ≤ 0 then,

$$\int_\Omega X dP \leq 0$$

8 Expectation

A random variable $X : \Omega \rightarrow \mathbb{R}$ is called integrable if

$$\int_\Omega |X(\omega) dP(\omega) = \int_\mathbb{R} |x| p(x) dx < \infty$$

The *Expectation* of an integrable random variable X is defined by

$$E[X] = \int_\Omega X(\omega) dP(\omega) = \int_\mathbb{R} x p(x) dx$$

In general, for any measurable function g, we have

$$E[g(X)] = \int_{\Omega} g(X(\omega))dP(\omega) = \int_{\mathbb{R}} g(x)p(x)dx$$

The expectation operator is linear,

$$E[cX] = xE[X]$$
$$E[X+Y] = E[X] + E[Y]$$

If two random variables are independent and integrable,

$$E[XY] = E[X]E[Y]$$

9 Radon - Nikodym theorem and Derivative

The Radon–Nikodym theorem is a result in measure theory that expresses the relationship between two measures defined on the same measurable space, (Ω, F).

One way to derive a new measure from one already given is to assign a density to each point of the measurable space, then integrate over the measurable subset of interest, which is expressed as

$$v(A) = \int_A f d\mu$$

Where v is the new measure being defined for any measurable subset A and the function f is the density at a given point. The integral is with respect to an existing measure μ

The function f is the Radon Nikodym Derivative denoted by $\frac{dv}{d\mu}$. The probability density function of a random variable is the Radon Nikodym Derivative of the induced measure with respect to some base measure (usually the Lebesgue measure for continuous random variables)

9.1 Formal Definition for Probability Density Function

A random variable X with values in a measurable space (Ω, F) has as probability distribution measure $X^{-1} \circ P$ on (Ω, F). The density of X with respect to a reference measure μ on (Ω, F) is the Radon Nikodym derivative,

$$f = \frac{dX^{-1} \circ P}{d\mu} \qquad f d\mu = dX^{-1} \circ P \tag{2}$$

The measure $X^{-1} \circ P$ is defined on (Ω, F) via

$$X : \Omega \to \mathbb{R} \qquad X^{-1} : \mathbb{R} \to \Omega \qquad P : F \in \Omega \to [0,1] \tag{3}$$

$$X^{-1} \circ P = P(X^{-1})$$
$$X^{-1} \circ P : \mathbb{R} \to [0,1] \in \mathbb{R}$$

The term $X^{-1} \circ P = P_X$ is the distribution of the random variable X on $\mathbb{R}$

Specifically, μ is the Lebesgue measure on $\mathbb{R}$ for continuous random variables, where the Lebesgue measure on $\mathbb{R}$ is defined as

$$\mu : M \subset \mathbb{R} \to [0, \infty]$$

The subsets $M \subset \mathbb{R}$ are Lebesgue measurable sets

9.2 Deriving Expectation Definition: Measure Theory

To see where the Expectation function comes from. we will make use of a Change of Variables in a Measure Theory Sense.

Before we start, it is good practice to remember Function Composition. Suppose we have two functions f and g where,

$$f : X \to Y \qquad\qquad g : Y \to Z \tag{4}$$

Then the composition, $g \circ f = h$ is defined as,

$$X \to Z$$

Suppose we have a probability space (Ω, F, P), and a random variable $X : \Omega \to \mathbb{R}$, then the Expectation is defined by, as an abstract integral,

$$E[X] = \int_\omega X dP$$

A random variable is a function which maps elements from ω to $\mathbb{R}$. Then define a new function g which maps $\mathbb{R}$ to $\mathbb{R}$. Then together when can define a new Random Variable by $g \circ X = g(X)$ which will again map from Ω to $\mathbb{R}$. This is where the Change of Variables will be used.
Beginning with, since $g(X)$ is just another random variable and A is a subset of Ω where ω are events $\in A \in \Omega$,

$$\int_A g(X(\omega))dP(\omega)$$

Then for simplicity we will rename $X(\omega)$ as x since x is the range of the function $X(\omega)$, or defend by the inverse $X^{-1}(x) = \omega$

$$\int_A g(X(\omega)dP(\omega) = \int_{X(A)} g(x)d(P \circ X^{-1})(x) = \int_{X(A)} g(x)dP_X(x)$$

This is because $P \circ X^{-1} = P_X(x)$ which is saying the Probability Measure composed with the inverse of the Random Variable,

$$X^{-1}(x) : x \to \omega \qquad\qquad P : A \to [0, 1] \tag{5}$$

Thus $P \circ X^{-1}$ is mapping from x to $[0, 1]$ which is saying we are now assigning x a probability (Remember x is some numerical value assigned by the random variable X to an event). Then since we are only interested in the continuous case,

$$\int_{X(A)} g(x)dP_X(x) = \int_{X(A)} g(x)f_X(x)dx$$

where the $f_X(x)$ comes from is from using the Radon - Nikodym derivative, $fd\mu = d(X^{-1} \circ P)(x)$

Remember that μ is the Lebesgue Measure, $\mu : M \subset \mathbb{R} \to [0, \infty]$, so we can just claim in this case $d\mu$ is dx

9.3 Change of Measure in an Expectation

Let $P, Q : F \to \mathbb{R}$ be two probability measures on Ω, such that there is an integrable random variables $f : \Omega \to \mathbb{R}$, such that $dQ = fdP$. This means,

$$Q(A) = \int_A dQ = \int_A f(\omega)dP(\omega)$$

Then we can define the following,

$$E^Q[X] = \int_\Omega X(\omega)dQ(\omega) = \int_\Omega X(\omega)f(\omega)dP(\omega) = E^P[fX]$$

10 Girsanov's Theorem (For Brownian Motion)

The Girsanov theorem states how a stochastic process change with the change of measure. To be more precise, it relates a Wiener measure P to a different measure Q on the space of continuous paths by giving an explicit formula for the likelihood ratios, which is the Radon-Nikodym derivative, between them.

Essentially the Radon - Nikodym Theorem tells how a new probability measure is defined, while Girsanov's Theorem describes the distribution of the stochastic process under this new measure

Remember that f is the desnity process Q (the new measure) with respect to the original measure P

11 Risk Neutral Measure

A Risk Neutral Measure is a probability measure such that each price is exactly equal to the discounted expectation of the said price under this measure

Suppose at a future time T a derivative pays H_T units, where H_T is random variable on the probability space (Ω, F, Q) describing the market.
Further suppose that the discount factor from now until time T is $DF(0,T)$. Then today's fair value of the derivative is,

$$H_0 = DF(0,T)E^Q[H_T]$$

where any Martingale Measure Q that solves the equation is a risk-neutral measure

11.1 Change of Measure from Risk Neutral

The above can be re-stated in terms of an alternative measure Q as

$$H_0 = DF(0,T)E^P(\frac{dQ}{dP}H_T)$$

If in a financial market there is just one risk neutral measure, then there is a unique arbitrage-free price for each asset in the market.

If there are more risk neutral measures, then in an interval of unique prices no arbitrage is possible.

If no equivalent martingale measures exist, arbitrage opportunities do.

12 Fundamental Theorem of asset pricing

- The First Fundamental Theorem of Asset Pricing: A discrete market on a discrete probability space (Ω, F, P) is arbitrage-free if, and only if, there exists at least one risk neutral probability measure that is equivalent to the original probability measure, P.
- The Second Fundamental Theorem of Asset Pricing: An arbitrage-free market

  $$(S, B)$$

  consisting of a collection of stocks S and a risk-free bond B is complete if and only if there exists a unique risk-neutral measure that is equivalent to P and has numeraire B.

13 Risk Neutral Measure to Real World Measure

In the following, P is the risk neutral measure, while Q is the real world measure! (or any appropriate alternative measure)
Using the change of Measure in an Expectation when $dQ = fdP$,

$$E^Q[X] = \int_\Omega X(\omega)dQ(\omega) = \int_\Omega X(\omega)f(\omega)dP(\omega) = E^P[fX]$$

and the theory of Risk Neutral measure,

$$H_0 = DF(0,T)E^P[H_T]$$

And using the Radon-Nikodym Theorem, with Q as the risk neutral (reference) measure,

$$f = \frac{dQ}{dP}$$

then we get

$$H_0 = DF(0,T)E^Q[\frac{dQ}{dP}H_T] = DF(0,T)E^P[fH_T]$$

By the Radon - Nikodym theorem.

If you continue with the Radon - Nikodym theorem you will require a function f to price todays fair value of the derivative in the real world measure, which will usually be extremely hard or near impossible.

When we say f here, we specifically mean it is the risk-neutral probability density function, i,e the formal definition for a probability density function

One example is of course the Black Scholes model, which for pricing in the real world measure will require market risk premium. Though this of course requires an investors expected return

14 Breeden-Litzenberger formula for risk-neutral densities