# Stochastics II: Martingales

This first (mathematical) treatment will be introductory and of a non-axiomatic approach.

We begin by defining some basic probability concepts in terms of the terminology used.

Martingales are defined and we look at examples of stochastic processes and application to option pricing.

# Some basic probability terminology and definitions:

A set $\Omega$ of all possible outcomes of some given experiment is called the *sample space*.

A particular outcome $\omega \in \Omega$ is called a *sample point*, or *sample path* for a stochastic process.

An *event* $\Psi$ is a set of outcomes, i.e. $\Psi \subset \Omega$.

$\boxed{\text{Example 1}}$ Experiment:

A dice is rolled and the number appearing on top is observed. The sample space consists of the 6 possible numbers:

$$\Omega = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$$

If the number 4 appears then $\omega = 4$ is a sample point, clearly $4 \in \Omega$.

Let $\Psi_1$, $\Psi_2$, $\Psi_3$ = events that an even, odd, prime number occurs respectively.

So

$$\Psi_1 = \{2, 4, 6\}, \Psi_2 = \{1, 3, 5\}, \Psi_3 = \{2, 3, 5\}$$

$\Psi_1 \cup \Psi_3 = \{2, 3, 4, 5, 6\}$ – event that an even or prime number occurs.

$\Psi_2 \cap \Psi_3 = \{3, 5\}$ – event that odd or prime number occurs.

$\Psi_3^c = \{1, 4, 6\}$ – event that prime number does not occur (complement of event).

$\boxed{\text{Example 2}}$ Experiment:

Toss a coin twice and observe the sequence of heads (H) and tails (T) that appears. Sample space

$$\Omega = \{HH, TT, HT, TH\}$$

Let $\Psi_1$ be event that at least one head appears, and $\Psi_2$ be event that both tosses are the same:

$$\Psi_1 = \{HH, HT, TH\}, \ \Psi_2 = \{HH, TT\}$$

$\Psi_1 \cap \Psi_2 = \{HH\}$

Events are subsets of $\Omega$, but not all subsets of $\Omega$ are events.

Now consider a stock price $S_t : (t = 1, \ldots, T)$.
$\Omega$ is the set of all possible values over time, i.e.

$$\Omega = \left\{\omega : \omega = \left(S_1, S_2, \ldots\ldots S_T\right)\right\}.$$

Assume $S \uparrow$ by factor $u$ and $S \downarrow$ by factor $v$.
Hence problem is equivalent to knowledge of price movements at each time, i.e.

$$\Omega = \left\{\omega : \omega = \left(a_1, a_2, \ldots\ldots a_T\right)\right\} \text{ where } a_t = u \text{ or } v$$

**Modelling Uncertainty (Future Prices):**

List all possible future prices - *possible states of the world.*

Unknown future - one of many possible outcomes - *true state of the world.*

Increasing time →more and more information becomes available.

An *event* is a subset of $\Omega$ which can be given a probability of occurring.

If $A \subseteq \Omega$ is an event, then $\mathbb{P}(A)$ is the actual history being one of the possible histories contained in $A$ :

**1.** $A$ could be the event that the interest rate reaches a certain level

**2.** Path dependency -

$$A = \left\{\omega \in \Omega : S_0 = 100,\ 110 \le S_{t_0} \le 130\right\}$$

i.e. event that process $S_t$ has value 100 at time 0, and $S_t \in \left[110,\ 130\right]$ at time $t_0$.

# Family of events

$\mathcal{F}$ is set of all events, so if $A \in \mathcal{F}$, then $A \subseteq \Omega$, and we can define the probability of $A$ occurring can be defined.

If $B \subseteq \Omega$ but $B \notin \mathcal{F}$, then $B$ is not an event and cannot be assigned a probability.

# Filtration

For each $t$ consider the process $S_t$.

As time increases we learn more and more about which $\omega \in \Omega$ is the actual history.

Allows steady accumulation of information, enabling us to distinguish between increasing number of events, as $\exists$ more and more sets $A \subseteq \Omega$ to which probabilities can be assigned.

Suppose at time $t$ we can distinguish events in a set - if

$$t_1 < t_2 < T_{\max} \text{ then } \mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2} \subseteq \mathcal{F}_T \equiv \mathcal{F}.$$

The set $\{\mathcal{F}_t\}_{t \in [0,\, T]}$ is called *filtration*.

A filtration allows us to alter probabilities we assign to future events, conditioned by what we have observed to have happened already.

Returning to $\{S_t\}_{t \in [0,\, T]}$ – if the value of $S_t$ is included in the information sets $\mathcal{F}_t$ at each $t \geq 0$, then we say $\{S_t : t \in [0,\, T]\}$ is *adapted* to $\mathcal{F}_t$, i.e. the value $S_t$ will be known given the information set $\mathcal{F}_t$.

$\mathbb{Q}$ is called a *Probability Measure* on the sample space $\Omega$.

For $A \in \mathcal{F}$, $\mathbb{Q}(A) = \mathbb{P}(A)$, i.e. the probability of $A$ containing the actual history.

$\mathbb{Q}(A) = 0$ : $A$ is a set of measure zero (i.e. with zero probability of occurring).

## Properties of $\mathbb{Q}$

**1.** $\mathbb{Q}(\Omega) = 1$ : $\mathbb{P}$ of actual history contained within the set of all possible histories is 1.

**2.** $\mathbb{Q}(\Phi) = 0$ : $\mathbb{P}$ of no history occurring is zero.

In addition ...

If $A_1, A_2$ are events and $A_1 \cap A_2 = \Phi$ then

$$\mathbb{Q}(A_1 \cup A_2) = \mathbb{Q}(A_1) + \mathbb{Q}(A_2).$$

If $A_1, A_2$ are independent then $\mathbb{P}$ of either events occurring is the sum of the probabilities of each occurring.

A **probability space** is a triple $\left(\Omega, \mathcal{F}, \mathbb{Q}\right)$

$\mathcal{F}$ : is the family of events to which probabilities may be assigned. It is a collection of subsets of $\Omega$, called events.

$\mathbb{Q}$ : is a **probability measure** which assigns probabilities to events (more later).

# Special Expectations:

Given any PDF $p(x)$ of $X$.

Mean $\mu = \mathbb{E}[X] = \int_{\mathbb{R}} x\, p(x)dx$.

Variance $\sigma^2 = \mathbb{V}[X] = \mathbf{E}\left[(X-\mu)^2\right] = \int_{\mathbb{R}} x^2 p(x)dx - \mu^2$ ($2^{\text{nd}}$ moment about the mean).

The $n^{\text{th}}$ moment about zero is defined as

$$\begin{aligned}\mu_n &= \mathbb{E}\left[X^n\right]\\ &= \int_{\mathbb{R}} x^n p(x)dx.\end{aligned}$$

In general, for any function $h$

$$\mathbb{E}[h(X)] = \int_{\mathbb{R}} h(x)p(x)dx.$$

where $X$ is a RV following the distribution given by $p(x)$.

Moments about the mean are given by

$$\mathbb{E}[(X-\mu)^n]; \;\; n = 2, 3, \ldots$$

The special case $n = 2$ gives the variance $\sigma^2$.

If $X$, $Y$ are RV's, then their covariance is defined as:

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\end{aligned}$$

The *moment generating function* of $X$, denoted $M_X(u)$ is given by

$$M_X(u) = \mathbb{E}[e^{uX}] = \int_{\mathbb{R}} e^{ux} p(x)dx$$

provided the expectation exists. We can expand as a power series to obtain

$$M_X(u) = \sum_{n=0}^{\infty} \frac{u^n \mathbb{E}(X^n)}{n!}$$

so the $n^{\text{th}}$ moment is the coefficient of $u^n/n!$, or the $n^{\text{th}}$ derivative evaluated at zero.

# Conditional Expectations

Given a probability space we can calculate useful quantities, e.g. the expected interest rate at time $T$, given that we know all that has happened up to some time $t < T$.

We write this expectation (formally) as

$$\mathbb{E}\left[ r_T \mid \mathcal{F}_t \right]$$

or simply as $\mathbb{E}[r_T]$.

This is the conditional expectation, conditional upon our knowledge at time $t$ as represented by the family of events $\mathcal{F}_t$. We are essentially generating forecasts of a process $r_T$ using information available as of time $t < T$.

# Measures

A probability distribution is sometimes called a *probability measure* for a RV $X$.

We can think of a probability measure $\mathbb{P}$ on the real line as assigning a weight $\mathbb{P}(dx) = f_X(x)dx$ to a closed interval $\left[x, x+dx\right]$.

When we have more general probability spaces the measure assigns a weight $\mathbb{P}(dw)$ to an elementary event $dw$. We then have an expectation

$$\mathbb{E}_{\mathbb{P}}\left[f(X)\right] = \int_{\Omega} f(X(w))\mathbb{P}(dw).$$

If $\{X_t : t \in [0,T]\}$ is Brownian motion, then the probability space $(\Omega, \mathbb{P})$ is the set of all paths, and $\mathbb{P}$ the probability of each path.

Two probability measures $\mathbb{P}$ and $\mathbb{Q}$ are *equivalent* if they operate over the same sample space $\Omega$ and agree on all possibilities. If $\Psi \subset \Omega$ is any event, then

$$\mathbb{P}(\Psi) \neq 0 \Leftrightarrow \mathbb{Q}(\Psi) \neq 0$$

i.e. if $\Psi$ is possible under $\mathbb{P}$ then it is possible under $\mathbb{Q}$. Similarly if $\Psi$ is impossible under $P$ then it is also impossible under $\mathbb{Q}$.

# Martingales

For each time $t$ , $S(t)$ and $r(t)$ represent the stock price and interest rate in turn.

As time progresses, we have increasingly more information about these prices and rates. This allows us to accumulate information which enables us to distinguish more and more events.

The flow of information associated with a stochastic process is called a ***filtration*** . We will write $\mathcal{F}_t$ for the history up to time $t$, i.e. what we know at time $t$.

A discrete time stochastic process

$$\{M_n : n = 0, \ldots\ldots, N\}$$

is a ***martingale*** if $\mathbf{E}\,|M_n| < \infty$ and

$$\mathbf{E}(M_{n+1}|\mathcal{F}_n) = M_n.$$

If we take expectations of both sides we have

$$\mathbf{E}(M_{n+1}) = E(M_n) \quad \forall n$$

and hence martingales have constant mean.

Note:

**1.** Brownian Motion $X(t)$ is a martingale.

**2.** One reason why martingales are nice is that they make it very simple to calculate $\mathbf{E}[X(t)]$.

A process which has

$$\mathbf{E}(M_{n+1}|\mathcal{F}_n) \geq M_n$$

is called a *submartingale* and if it has

$$\mathbf{E}(M_{n+1}|\mathcal{F}_n) \leq M_n$$

it is called a *supermartingale*.

A process whose value at time $t$ can be dependent on the history $\mathcal{F}_t$, but not the future (i.e. a time $T > t$) is called *adapted* to the filtration $\mathcal{F}$.

The martingale property is different from the Markov property which says that the past and future behaviour of the process are independent conditional on the present.

In continuous time we have the definition :

A stochastic process

$$\{M_t : t \in \mathbb{R}_+\}$$

is a **martingale** if $\mathbb{E}\,|M_t| < \infty$ and

$$\mathbb{E}(M_t|\mathcal{F}_s) = M_s \;\; \forall\; 0 \leq s \leq t.$$

There are similar definitions of sub and super martingale.

A martingale can be thought of as a zero-drift stochastic process.

A stochastic variable $S$ is a martingale if it evolves according to the process

$$dS = \sigma dX$$

where $dX$ is the usual increment in a Wiener process and $\sigma$ itself may be stochastic.

From the definition

$$\mathbb{E}[S_T] = S_t \;\; \forall \text{ times } t \leq T,$$

i.e. a martingale has the property that its expected value at any time in the future is equal to its value today.

Consider a small time interval $dt \to 0$, then $dS \sim N(\mu = 0)$. This implies that over many small time intervals the mean change in $S$ must also be zero. Hence the expected value of $S$ in the future must be its current value.

An asset process

$$dS = \mu S dt + \sigma S dX$$

is not a martingale unless $\mu = 0$ because otherwise the drift and its expected value will move farther away from its value today as time goes on.

A *supermartingale* is a stochastic process with the property that

$$\mathbb{E}[S_T] \leq S_t,$$

and a *submartingale* one with the property that

$$\mathbb{E}[S_T] \geq S_t \text{ }, \forall \text{ times } \ 0 \leq t \leq T.$$

# Examples

## 1. Exponential Martingale

Consider the process

$$dS = \mu dt + \sigma dX; \;\; S(0) = 0$$

Find a function $f(t)$ such that

$$\exp\left(S(t) + f(t)\right)$$

is a martingale.

First set $\Psi = \exp\left(S_t + f(t)\right) \rightarrow \Psi_t = \Psi f'; \;\; \Psi_S = \Psi_{SS} = \Psi$

Then Itô on $\Psi$

$$
\begin{aligned}
d\Psi &= \frac{\partial \Psi}{\partial t} dt + \frac{\partial \Psi}{\partial S} dS + \tfrac{1}{2} \frac{\partial^2 \Psi}{\partial S^2} dS^2 \\
&= d\Psi = \left(f'(t) + \mu + \tfrac{1}{2}\sigma^2\right)\Psi dt + (\sigma\Psi) dX.
\end{aligned}
$$

For this to be a martingale, i.e. a zero-drift stochastic process, we require

$$f'(t) + \mu + \tfrac{1}{2}\sigma^2 = 0 \Rightarrow \frac{df}{dt} = -(\mu + \tfrac{1}{2}\sigma^2)$$

Hence

$$f(t) = -(\mu + \tfrac{1}{2}\sigma^2)t.$$

Therefore $e^{\left(S_t - \left(\mu + \frac{1}{2}\sigma^2\right)t\right)}$ is a martingale.

# 2. Itô Integral

In the stochastics lecture we defined the Itô integral. We saw how to use the definition to integrate from first principles. Now consider

$$I(T) = \int_0^T X(t)dX(t) = \lim_{N\to\infty} \sum_{n=0}^{N-1} X_n(X_{n+1} - X_n)$$

$$= \tfrac{1}{2}\left(X^2(T) - T\right)$$

Taking expectations gives

$$\begin{aligned}\mathbb{E}[I(T)] &= \tfrac{1}{2}\mathbb{E}\big[X^2(T) - T\big] \\ &= \tfrac{1}{2}\mathbb{E}\big[X^2(T)\big] - \tfrac{T}{2}\end{aligned}$$

We know

$$\mathbb{V}[X(T)] = \mathbb{E}\big[X^2(T)\big] - \underbrace{\mathbb{E}^2[X(T)]}_{=0} = T$$

$$\therefore\ \mathbb{E}\big[X^2(T)\big] = T$$

$$\Rightarrow \mathbb{E}[I(T)] = 0$$

$\Rightarrow I(T)$ is a martingale.

The Itô integral is also consistent with Itô's lemma, as

$$
\begin{aligned}
d\left(X^2\right) &= 2XdX + dt \\
\int_0^T d\left(X^2\right) &= 2\int_0^T XdX + \int_0^T dt \\
\int_0^T XdX &= \tfrac{1}{2}\left(X^2(T) - T\right)
\end{aligned}
$$

# 3. Stratonovich Integral

Consider the special integral

$$S(T) = \int_0^T X(t)dX(t) = \lim_{N\to\infty} \sum_{n=0}^{N-1} \tfrac{1}{2}(X_{n+1} + X_n)(X_{n+1} - X_n)$$

$$= \tfrac{1}{2}X^2(T)$$

called the *Stratanovich Integral.* Taking expectations we have

$$\mathbb{E}[S(T)] = \tfrac{1}{2}\mathbb{E}\left[X^2(T)\right] = \tfrac{1}{2}\mathbb{V}[X(T)]$$

$$= \tfrac{T}{2} \quad \therefore \ S(T) \text{ is not a martingale.}$$

Now consider another example. If

$$G_t = X(t)t^2 - 2\int_0^t sX_s ds$$

then $G_t$ is a martingale.

$$\begin{aligned}
\mathbb{E}[G_{t+s}|\mathcal{F}_t] &= \mathbb{E}\Big[X_{t+s}(t+s)^2 - 2\int_0^{t+s} uX_u du \Big| \mathcal{F}_t\Big] \\
&= X_t(t+s)^2 - 2\int_0^t sX_s ds - 2\mathbb{E}\Big[\int_t^{t+s} uX_u du\Big] \\
&= X_t t^2 - 2\int_0^t sX_s ds + s(2t+s)X_t - 2\mathbb{E}\Big[\int_t^{t+s} uX_u du\Big]
\end{aligned}$$

$G(t)$ is a martingale because as the last two terms cancel.

The integral $\int_0^{t+s} uX_u du$

$$= \int_0^t sX_s ds + \int_t^{t+s} uX_u du$$

and the result is a consequence of

$$\begin{aligned} 2\mathbb{E}\left[\int_t^{t+s} uX_u du\right] &= 2\left[\int_t^{t+s} u\mathbf{E}[X_u]du\right] \\ &= 2X_t\left[\int_t^{t+s} udu\right] \\ &= X_t\left[(t+s)^2 - t^2\right] \\ &= s(2t+s)X_t \end{aligned}$$

$\Rightarrow$

$$\mathbb{E}[G_{t+s}|\mathcal{F}_t] = X_t t^2 - 2\int_0^t sX_s ds$$
$$= G_t$$

# Change of measure: The Girsanov Theorem

For a fixed time $t$ consider $\phi_t \sim N(0,1)$ and density

$$f(\phi_t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\phi_t^2}.$$

Let $\mathbb{P}$ denote the probability measure such that

$$d\mathbb{P}(\phi_t) = f(\phi_t)d\phi_t.$$

Now introduce a function $\xi = \xi(\phi_t)$ where

$$\xi(\phi_t) = e^{\phi_t \mu - \frac{1}{2}\mu^2}.$$

Then the product of $\xi(\phi_t)$ and $d\mathbb{P}(\phi_t)$ gives

$$d\mathbb{P}(\phi_t)\xi(\phi_t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\phi_t^2+\mu\phi_t-\frac{1}{2}\mu^2}d\phi_t\ ,$$

which is a new probability.

Simplifying gives

$$d\widetilde{\mathbb{P}}(\phi_t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\phi_t-\mu)^2}d\phi$$

with $d\widetilde{P}(\phi_t)$ denoting a new probability measure The density here is now for a RV of mean $\mu$ and variance $1$.

The new measure is defined by

$$d\widetilde{\mathbb{P}}(\phi_t) = d\mathbb{P}(\phi_t)\xi(\phi_t)$$

We see that $d\mathbb{P}(\phi_t) \times \xi(\phi_t)$ and then working in $\widetilde{\mathbb{P}}$, changes the mean of $\phi_t$ from 0 to $\mu$. In this particular example the shape of probability measure is also preserved.

The important point is that $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are different measures because they have different means, and assign different weights to intervals on the $\phi$ axis.

We know

$$\mathbb{E}_{\mathbb{P}}[\phi_t] = 0 \ \& \ \mathbb{E}_{\mathbb{P}}[\phi_t^2] = 1$$

whereas

$$\mathbf{E}_{\widetilde{\mathbb{P}}}[\phi_t] = \mu.$$

So $\exists$ a function $\xi(\phi_t)$ with the property that multiplication with a probability measure gives a new probability. The resulting RV is again normal but of different mean.

Given $\mu$ and $\sigma$ the (unique) transformation (which leaves the variance unchanged) of measure is a reversible process, i.e. if

$$d\widetilde{\mathbb{P}}(\phi_t) = \xi(\phi_t)d\mathbb{P}(\phi_t)$$

then

$$d\mathbb{P}(\phi_t) = \xi^{-1}(\phi_t)d\widetilde{\mathbb{P}}(\phi_t).$$

We can rewrite the above as a derivative term, i.e.

$$\frac{d\widetilde{\mathbb{P}}(\phi_t)}{d\mathbb{P}(\phi_t)} = \xi(\phi_t),$$

the change of measure $\widetilde{\mathbb{P}}$ with respect to change of measure $\mathbb{P}$ is given by the function $\xi(\phi_t)$.

The term $\xi(\phi_t)$ is then regarded as being the density of this derivative term.

A Derivative of this type has a special name and is called a ***Radon-Nikodym Derivative.***

If this derivative exists, then the resulting density $\xi(\phi_t)$ allows us to transform the mean of $\phi_t$ whilst leaving the variance unchanged.

**Note:** If $\xi(\phi_t) = e^{-\phi_t\mu+\frac{1}{2}\mu^2}$ then the function subtracts a mean from the RV whereas the inverse function $\xi^{-1}(\phi_t) = e^{\phi_t\mu-\frac{1}{2}\mu^2}$ adds a mean $\mu$ to $\phi_t$ (which originally had a zero mean).

There are two simple methods for changing measure.

**Subtraction of means:** Given a RV $\Phi \sim N(\mu,\, \sigma^2)$, define a new RV $\widetilde{\Phi}$ by transforming $\Phi$, according to

$$\widetilde{\Phi} = \frac{\Phi - \mu}{\sigma} \sim N(0,\, \sigma^2)$$

Then $\widetilde{\Phi}$ will have zero mean.

**Using equivalent measures:** Given a RV $\Phi$ with probability $\mathbb{P}$, i.e. $\Phi \sim \mathbb{P} = N(\mu,\, \sigma^2)$, transform the probabilities $d\mathbb{P}$ through multiplication by $\xi(\Phi)$ to obtain a new probability $\widetilde{\mathbb{P}}$ such that $Z \sim \widetilde{\mathbb{P}}(0,\, \sigma^2)$

# Change of Drift: Using Girsanov

We frequently need to change the drift of a stochastic process, (e.g. to remove the effects of risk premia) which evolves according to a SDE. This can be accomplished again by use of this theorem:

Let

$$dG = A(G,t)dt + B(G,t)dX$$

Then

$$dG = [A(G,t) - a(G,t)B(G,t)]dt + B(G,t)d\widetilde{X}$$

where

$$d\widetilde{X} = a(G,t)dt + dX$$

Now $d\widetilde{X}$ is an increment in Brownian motion under the transformed probability measure

$$d\,\widetilde{\mathbb{P}} = \xi(t).\,d\mathbb{P}$$

where

$$\xi(t) = \exp\left(-\tfrac{1}{2}\int_0^t a^2(G(\tau))d\tau - \int_0^t a(G(\tau))dX\right)$$

Girsanov's Theorem "distorts" the probabilities of the original Brownian Motion.

# Example:

Consider the process followed by GBM

$$dS = \mu S dt + \sigma S dX.$$

If we now wish to consider this in a risk-neutral world

$$dS = rSdt + \sigma S d\widetilde{X}$$

we must use a change of measure

$$\xi(t) = \exp\left(-\frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 t - \left(\frac{\mu - r}{\sigma}\right)X\right)$$

$$= \frac{d\widetilde{X}}{dX}.$$

# Some Finance Terminology:

*Self-financing* means a strategy which never needs to be "topped-up" with extra cash, nor can it allow withdrawals.

*Replicating strategy* is a self-financing portfolio strategy which hedges a claim precisely.

*Arbitrage-free* market is one where there is no possibility of making riskless profits. The existence of an arbitrage opportunity would be a self-financing trading strategy which started with zero value and terminated at a future time $T$ with a positive value.

*Complete market* is one where any possible derivative can be hedged by trading with a self-financing portfolio.

From the definition of a martingale, it is quite clear that, it implies no-arbitrage. That is, its future expectation, conditional upon the history is just its current value - not higher nor lower.

# Numeraires

A *numeraire* is a particular asset that can be used to price relative to.

Suppose $p_t$ is the value at time $t$ of the numeraire, then for an asset with value $\upsilon_t$ its value $\widetilde{\upsilon}_t$ relative to the numeraire is

$$\widetilde{\upsilon}_t = \frac{\upsilon_t}{p_t}$$

A numeraire must have strictly positive value and be self-financing.

There are many assets (or portfolios) that could be chosen as a numeraire, and this is done solely for the purpose of convenience.

A numeraire may have volatility and can be any tradeable instrument. The choice does not affect the eventual prices in any way. Two examples of numeraire are:

**1.** Money market account

$$p_t = \exp\left(\int_0^t r_\tau d\tau\right),$$

which is the value at time $t$ of 1 unit invested at time 0 in the short rate and continuously compounded.

**2.** A discount bond maturing at a convenient date $T$

$$p_t = Z_t(T), \;\; t < T.$$

# Risk Neutral Measure

In the standard Black-Scholes world, we can price derivatives for all times. The Feynman Kac equation allows the Black-Scholes solution to be expressed in probabilistic terms by taking the risk-neutral pricing formula

$$V\left(S,\, t\right) = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}\left[f\left(S_T\right)\middle|\mathcal{F}_t\right].$$

$r$ is the risk-free rate and $f\left(S_T\right)$ the payoff function. Then the option price $V$ is the present worth of the expected value ($\mathbb{E}_{\mathbb{Q}}[\cdot]$) of the payoff with respect to the risk-neutral probability density $\mathbb{Q}$.

It then follows in this Black-Scholes framework

$$S_T = S_t \exp[(r - \tfrac{1}{2}\sigma^2)(T - t) + \sigma(X_T - X_t)]$$

where the notation $F_t \equiv F(t)$, i.e. $F$ is a function of variable $t$.

A surprising fact about the Black-Scholes equation is the vanishing of the drift rate $\mu$, due to the $\Delta$–hedging which eliminates all the risk.

The important point to note is that the stochastic process evolves according to the risk-neutral SDE

$$dS = rSdt + \sigma SdX.$$

# Option Pricing

For the pricing of European Calls we know

$$\Psi(S_T) = \max\left[S_T - E,\ 0\right]$$

and the solution of the Black-Scholes equation gives the following pricing scheme:

$$C\left(S,t;\ T,E\right) = SN(d_1) - Ee^{-r(T-t)}N(d_2)$$

where $N(x)$ is the standard Normal cumulative distribution function

$$N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}\phi^2}\,d\phi.$$

$$d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = \frac{\log(S/E) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

where $S = S(t) \geq 0$ is the spot price, $t \leq T$ is the time, $E > 0$ is the strike, $T > 0$ the expiry date, $r \geq 0$ the interest rate and $\sigma$ is the volatility of $S$.

To obtain the solution we use the Kolmogorov solution

$$\begin{aligned} C\left(S,t;\ T,E\right) &= e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}[\Psi(S_T)|\mathcal{F}_t] \\ &= e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}\left[\max\left[S-E,\ 0\right]\right] \end{aligned}$$

So

$$\begin{aligned} C\left(S,t;\ T,E\right) &= e^{-r(T-t)}\int_0^\infty \max\left[S-E,\ 0\right]p(S)dS \\ &= e^{-r(T-t)}\int_E^\infty (S-E)p(S)dS \qquad \because \int_0^E \to 0 \\ &= e^{-r(T-t)}\int_E^\infty Sp(S)dS - Ee^{-r(T-t)}\int_E^\infty p(S)dS \end{aligned}$$

Let us consider the second term initially, this is the same as

$$Ee^{-r(T-t)}\mathbb{P}[S > E].$$

Now

$$\mathbb{P}[S > E] = \mathbb{P}[\log S > \log E]$$

and we know

$$S(T) = S(t)\exp\left[(r - \tfrac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{(T-t)}\,\phi\right]$$

therefore the probability $\mathbb{P}[\log S_T > \log E]$ becomes

$$\log S + \left[(r - \tfrac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{(T-t)}\,\phi\right] > \log E$$

rearranging to give

$$\phi > \frac{-\log\frac{S}{E} - (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}} \rightarrow$$

$$\phi > -\left\{\underbrace{\frac{\log\frac{S}{E} + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}}}_{=d_2}\right\}$$

Hence

$$\mathbb{P}[S > E] = \mathbb{P}\left[\phi > -\left\{\underbrace{\frac{\log\frac{S}{E} + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}}}_{=d_2}\right\}\right]$$

$$\mathbb{P}[S > E] = 1 - N(-d_2) = N(d_2)$$

So finally we have

$$Ee^{-r(T-t)}N(d_2).$$

Solving for the first integral is slighty more complex

$$e^{-r(T-t)}\int_E^\infty S\,p(S)dS = e^{-r(T-t)}\int_{\log E}^\infty e^Y p(Y)dY,$$

where $Y$ is normally distributed with $Y = \log S$. This allows us to use the (lognormal) transition probability density function

$$\frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}}\int_{\log E}^\infty \exp\left\{Y - \frac{\left(Y-\left(r-\frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2(T-t)}\right\}dY$$

$$= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}}\int_{\log E}^\infty \exp\left\{Y - \frac{(Y-\mu_Y)^2}{2\sigma_Y^2}\right\}dY$$

If we expand and manipulate the exponent we have

$$\frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}} \int_{\log E}^{\infty} \exp\left\{-\frac{1}{2\sigma_Y^2}\left((Y-(\mu_Y+\sigma_Y^2))^2 - 2\mu_Y\sigma_Y^2 - \sigma_Y^4\right)\right\} dY$$

$$= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}} e^{\mu_Y+\frac{1}{2}\sigma_Y^2} \int_{\log E}^{\infty} \exp\left\{-\frac{1}{2\sigma_Y^2}\left((Y-(\mu_Y+\sigma_Y^2))^2\right)\right\} dY$$

Now $e^{\mu_Y+\frac{1}{2}\sigma_Y^2}$ simplifies nicely because

$$\mu_Y + \frac{1}{2}\sigma_Y^2 = \log S(t) + (r - \frac{1}{2}\sigma^2)(T-t) + \frac{1}{2}\sigma^2(T-t)$$

$$= \log S(t) + r(T-t) \;\therefore$$

$$e^{-r(T-t)} e^{\mu_Y+\frac{1}{2}\sigma_Y^2} = S(t)$$

Introduce another substitution

$$
\begin{aligned}
U &= Y - (\mu_Y + \sigma_Y^2) \\
&= Y - \log S(t) - (r + \tfrac{1}{2}\sigma^2)(T - t)
\end{aligned}
$$

to give

$$
\frac{S(t)}{\sqrt{2\pi\sigma_Y^2}} \int_{\log E}^{\infty} \exp\left\{-\tfrac{1}{2\sigma_Y^2}\left((Y - (\mu_Y + \sigma_Y^2))^2\right)\right\} dY
$$

$$
= \frac{S(t)}{\sqrt{2\pi\sigma_Y^2}} \int_{\log E - \log S(t) - \left(r + \frac{1}{2}\sigma^2\right)(T-t)}^{\infty} \exp\left\{-\tfrac{U^2}{2\sigma_Y^2}\right\} dU
$$

and one final transformation

$$
Z = \frac{U}{\sigma_Y}
$$

to yield

$$\frac{S(t)}{\sqrt{2\pi}} \int_{\frac{\log(E/S(t))-\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}}^{\infty} \exp\{-\tfrac{1}{2}Z^2\}dZ$$

$$= S(t)\mathbb{P}\left[\phi > \frac{\log(E/S(t))-\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right]$$

$$= S(t)\mathbb{P}\left[\phi > -\left(\underbrace{\frac{\log(S(t)/E)+\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}}_{=d_1}\right)\right]$$

$$= S(t)\mathbb{P}[\phi > -d_1] = S(t)(1 - N(-d_1))$$
$$= S(t)N(d_1)$$

Hence we finally get the Call Option price as

$$C\left(S, t;\ T, E\right) = S(t)N(d_1) - Ee^{-r(T-t)}N(d_2)$$