AN APPLICATION OF PROBABILITY THEORY IN FINANCE: THE BLACK-SCHOLES FORMULA

EFFY FANG

ABSTRACT. In this paper, we start from the building blocks of probability theory, including σ-field and measurable functions, and then proceed to a formal introduction of probability theory. Later, we introduce stochastic processes, martingales and Wiener processes in particular. Lastly, we present an application - the Black-Scholes Formula, a model used to price options in financial markets.

CONTENTS

1. THE BUILDING BLOCKS

Consider the set Ω of all possible outcomes of a random process. When the set is small, for example, the outcomes of tossing a coin, we can analyze case by case. However, when the set gets larger, we would need to study the collections of subsets of Ω that may be taken as a suitable set of events. We define such a collection as a σ-field.

Definition 1.1. A σ-field on a set Ω is a collection $\mathcal{F}$ of subsets of Ω which obeys the following rules or axioms:

(a) $\Omega \in \mathcal{F}$
(b) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
(c) if $\{A_n\}_{n=1}^{\infty}$ is a sequence in $\mathcal{F}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

The pair $(\Omega, \mathcal{F})$ is called a measurable space.

Proposition 1.2. *Let $(\Omega, \mathcal{F})$ be a measurable space. Then*

(1) $\emptyset \in \mathcal{F}$
(2) if $\{A_n\}_{n=1}^{k}$ is a finite sequence of $\mathcal{F}$-measurable sets, then $\bigcup_{n=1}^{K} A_n \in \mathcal{F}$
(3) if $\{A_n\}_{n=1}^{\infty}$ is a sequence of $\mathcal{F}$-measurable sets, then $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$

Given Ω we let 2^Ω denote the set of all subsets of Ω and call it the power set of Ω.

The σ-field 2^Ω is the largest σ-field on Ω.

Proposition 1.3. *If $\{\mathcal{F}_\alpha\}_{\alpha \in \tau}$ is a collection of σ-fields on Ω, then $\bigcap_{\alpha \in \tau} \mathcal{F}$ is a σ-field on Ω.*

Proposition 1.4. *If $\mathcal{A}$ is a collection of subsets of Ω, then there exists a unique smallest σ-field on ω, containing $\mathcal{A}$, which is contained in every σ-field that contains $\mathcal{A}$. We denote this σ-field by $\mathcal{F}(\mathcal{A})$ and call it the σ-field generated by $\mathcal{A}$.*

Proposition 1.5. *if $\mathcal{A}, \mathcal{A}_1$ and $\mathcal{A}_2 \subseteq 2^\Omega$, then the following hold.*
- *(a) If $\mathcal{A}_1 \subseteq \mathcal{A}_2$, then $\mathcal{F}(\mathcal{A}_1) \subseteq \mathcal{F}(\mathcal{A}_2)$*
- *(b) If $\mathcal{A}$ is a σ-field, then $\mathcal{F}(\mathcal{A}) = \mathcal{A}$*
- *(c) $\mathcal{F}(\mathcal{F}(\mathcal{A})) = \mathcal{F}(\mathcal{A})$*
- *(d) If $\mathcal{A}_1 \subseteq \mathcal{F}(\mathcal{A}_2)$, then $\mathcal{F}(\mathcal{A}_1) \subseteq \mathcal{F}(\mathcal{A}_2)$*

Proof.
- (a) $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{F}(\mathcal{A}_2)$. Since $\mathcal{F}(\mathcal{A}_2)$ is a σ-field and $\mathcal{F}(\mathcal{A}_1)$ is the smallest σ-field containing $\mathcal{A}_1$, this implies $\mathcal{F}(\mathcal{A}_1) \subset \mathcal{F}(\mathcal{A}_2)$.
- (b) if $\mathcal{A}$ is a σ-field, it must be the smallest σ-field containing $\mathcal{A}$. Hence $\mathcal{A} = \mathcal{F}(\mathcal{A})$.
- (c) $\mathcal{F}(\mathcal{F}(\mathcal{A})) = \mathcal{F}(\mathcal{A}) = \mathcal{A}$
- (d) by (a) $\mathcal{F}(\mathcal{A}_1) \subset \mathcal{F}(\mathcal{F}(\mathcal{A}_2))$, and by (c) $\mathcal{F}(\mathcal{F}(\mathcal{A}_2)) = \mathcal{F}(\mathcal{A}_2)$.

$\square$

Definition 1.6. Let $(\Omega, \mathcal{F})$ be a measurable space.
- (a) A discrete filtration on $(\Omega, \mathcal{F})$ is an increasing sequence of σ-fields $(\mathcal{F}_n)_{n=1}^\infty$ such that $\mathcal{F}_1 \subset \mathcal{F}_2 \subset ... \subset \mathcal{F}_i \subset ... \subset \mathcal{F}$.
- (b) A continuous filtration on $(\Omega, \mathcal{F})$ is a set of σ-fields $(\mathcal{F}_t)_{t \in I}$, where I is an interval in $\mathbb{R}$, such that for all $t, s \in I, t < s$, we have $\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}$.

We call $\mathcal{F}_n$ (respectively $\mathcal{F}_t$) the history up to time n (respectively time t).

Example 1.7. Consider tossing a fair coin 3 times. The 3 tables below demonstrate the filtrations on outcome space after the first, second and the third toss. The outcome space can be filtered into two subsets by the outcomes of the first flip, and further filtered into four subsets by the outcomes of the first two flips, eight singleton sets by the outcomes of all three flips.

Flip 1	Flip 2	Flip 3	Flip 1	Flip 2	Flip 3	Flip 1	Flip 2	Flip 3
H	H	H	H	H	H	H	H	H
H	H	T	H	H	T	H	H	T
H	T	H	H	T	H	H	T	H
H	T	T	H	T	T	H	T	T
T	H	H	T	H	H	T	H	H
T	H	T	T	H	T	T	H	T
T	T	H	T	T	H	T	T	H
T	T	T	T	T	T	T	T	T

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3$$

Our next goal is to develop the tools to measure the likelihood of the events, collected into a σ-field.

Definition 1.8. The Borel field on $\mathbb{R}, \mathcal{B}(\mathbb{R})$, is the σ-field generated by the open intervals in $\mathbb{R}$. Subsets of $\mathbb{R}$ which belong to $\mathcal{B}(\mathbb{R})$ are Borel sets.

Proposition 1.9. *The Borel field is generated by the closed intervals.*

Proposition 1.10. *Let $X : \Omega \to \mathbb{R}$ denote a real-valued function. The collection of sets $X^{-1}(B)$, where B ranges over the Borel subsets of $\mathbb{R}$, is the σ-field generated by X and is denoted $\mathcal{F}_X$.*

Definition 1.11. A mapping $X : \Omega \to \mathbb{R}$, where $(\Omega, \mathcal{F})$ is a measurable space, is called $\mathcal{F}$-measurable if $X^{-1}(B) \in \mathcal{F}$ for every Borel subset $B \subset \mathbb{R}$.

Proposition 1.12. *A mapping $X : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable if and only if $\mathcal{F}_X \subset \mathcal{F}$.*

Proposition 1.13. *If the collection $\mathcal{A}$ of subsets of $\mathbb{R}$ generated the Borel field, then $X : \Omega \to \mathbb{R}$ is $\mathcal{F}$ measurable if and only if $X^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{A}$.*

Proposition 1.14. *If c is a real number and X and Y are $\mathcal{F}$ measurable functions defined on Ω, then $X + Y, X - Y, X \cdot Y$, and cX are $\mathcal{F}$ measurable. If $Y(\omega) \neq 0$ for all $\omega \in \Omega$, then $\frac{X}{Y}$ is also measurable.*

Example 1.15. The indicator function of a set A in Ω, $\mathbb{1}_A : \Omega \to \mathbb{R}$, is $\mathcal{F}$-measurable, where $\mathcal{F} = \{A, A^c, \emptyset, \Omega\}$ is measurable:

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

Proposition 1.16. *If $f : \mathbb{R} \to \mathbb{R}$ is continuous then f is Borel measurable.*

Proposition 1.17. *If the sequence $(X_n)_{n=1}^{\infty}$ of $\mathcal{F}$-measurable functions on Ω converges pointwise to X, then X is $\mathcal{F}$-measurable.*

2. Probability

We formally define probability space, expected value, continuity and integrability using σ-fields and measurable functions in this section.

Definition 2.1. A probability space is a triple $(\Omega, \mathcal{F}, P)$ where Ω is a set (the sample space), $\mathcal{F}$ is a σ-field on Ω and P, the probability measure, is a mapping from $\mathcal{F}$ into $[0, 1]$ such that
$$P(\Omega) = 1,$$
and if $(A_n)_{n=1}^{\infty}$ is any sequence of pairwise disjoint events in $\mathcal{F}$, then
$$P(\textstyle\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

Definition 2.2. If $(\Omega, \mathcal{F}, P)$ is a probability space, then $A \in \mathcal{F}$ and $B \in \mathcal{F}$ are independent events if
$$P(A \cap B) = P(A) \cdot P(B).$$

Definition 2.3. If $(\Omega, \mathcal{F}, P)$ is a probability space and $X : \Omega \to \mathbb{R}$ is measurable, we call X a random variable on $(\Omega, \mathcal{F}, P)$.

Definition 2.4. The random variables X and Y on the probability space $(\Omega, \mathcal{F}, P)$ are independent if the σ-field they generate, $\mathcal{F}_X$ and $\mathcal{F}_Y$, are independent.

Definition 2.5. A stochastic process X is a collection of random variables $(X_t)_{t \in T}$ on a probability space $(\Omega, \mathcal{F}, P)$, indexed by a subset T of the real numbers.

Definition 2.6. If $X = (X_t)_{t \in T}$ is a stochastic process on $(\Omega, \mathcal{F}, P)$ and $(\mathcal{F}_t)_{t \in T}$ is a filtration on $(\Omega, \mathcal{F}, P)$, then X is adapted to the filtration if X_t is $\mathcal{F}_t$ measurable for all $t \in T$.

In the next two sections, we will introduce two stochastic processes, martingales and Wiener Processes.

Definition 2.7. X is a simple random variable if and only if X is a finite linear combination of indicators. A simple random variable can only take finetly many values.

Definition 2.8. Let X denote a simple random variable on the probability space $(\Omega, \mathcal{F}, P)$. If X has range $(x_i)_{i=1}^{n}$ and $\omega_i \in X^{-1}(\{x_i\})$ for all i, then
$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i P_X(\{x_i\}) = \sum_{i=1}^{n} X(\omega_i) P_X(\{X(\omega_i)\}).$$

Definition 2.9. A random variable X on a probability space $(\Omega, \mathcal{F}, P)$ is integrable if its positive and negative parts, X^+ and X^- where

$$X^+(\omega) = \begin{cases} X(\omega) & \text{if } X(\omega) \geq 0 \\ 0 & \text{if } X(\omega) < 0 \end{cases}$$

$$X^-(\omega) = \begin{cases} -X(\omega) & \text{if } X(\omega) < 0 \\ 0 & \text{if } X(\omega) \geq 0 \end{cases}$$

are both integrable. If X is integrable we let

$$\mathbb{E}[X] := \mathbb{E}[X^+] - \mathbb{E}[X^-] = \int_\Omega X^+ dP - \int_\Omega X^- dP = \int_\Omega X dP.$$

Proposition 2.10. *If $(\Omega, \mathcal{F}, P)$ is a probability space with $\Omega = \{\omega_n : n \in \mathbb{N}\}$ and $\mathcal{F} = 2^\Omega$, then $X : \Omega \to \mathbb{R}$ is integrable if and only if*

$$\sum_{n=1}^\infty |X(\omega_n)| P(\{\omega_n\}) < \infty.$$

If X is integrable

$$\mathbb{E}[X] = \int_\Omega X dP = \sum_{n=1}^\infty X(\omega) P(\{\omega_n\}).$$

In probability theory, one says that an event happens *almost surely* if it happens with probability one. The concept is analogous to the concept of "almost everywhere" in measure theory. A property holds almost everywhere if the set for which the property does not hold has measure zero.

Proposition 2.11. *(Monotone Convergence Theorem) If $(X_n)_{n=1}^\infty$ is an increasing sequence of positive random variables on the probability space $(\Omega, \mathcal{F}, P)$, then there exists an integrable random variable X such that $X_n \to X$ almost surely as $n \to \infty$ if and only if $\lim_{n\to\infty} \int_\Omega X_n dP < \infty$. When the limit is finite we have*

$$\int_\Omega (\lim_{n\to\infty} X_n) dP = \int_\Omega X dP = \lim_{n\to\infty} \int_\Omega X_n dP$$

Proposition 2.12. *(Dominated Convergence Theorem) Let $(X_n)_{n=1}^\infty$ denote a sequence of random variables on the probability space $(\Omega, \mathcal{F}, P)$ and suppose $(X_n)_{n=1}^\infty$ converges almost surely to the random variable X. If there exists an integrable random variable Y such that for all n, $|X_n| \leq Y$ almost surely, then X and each X_n are integrable and*

$$\lim_{n\to\infty} \int_\Omega X_n dP = \int_\Omega X dP.$$

Definition 2.13. a probability density function (PDF), or density of a continuous random variable, is a function that describes the relative likelihood for this random variable to take on a given value.

Proposition 2.14. *If the random variable X on $(\Omega, \mathcal{F}, P)$ has density function f_X and g is a Borel measurable function such that $g(X)f_X$ is Riemann integrable, then $g(X)$ is an integrable random variable and*

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx.$$

Proposition 2.15. *(The Central Limit Theorem)*
    Let $(X_n)_{n=1}^{\infty}$ be a sequence of independent identically distributed random variables in $L^2(\Omega, \mathcal{F}, P)$ and for all n, let $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. If $\mathbb{E}[X_i] = \mu$ and $Var(X_i) = \sigma^2$ for all i, then

$$\lim_{n\to\infty} P[\tfrac{Y_n - \mu}{\sigma/\sqrt{n}} \leq x] = \lim_{n\to\infty} P[\tfrac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leq x] = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$$

for all $x \in \mathbb{R}$.

Proposition 2.16. *(The Radon-Nikodým Theorem) If P and Q are finite measures on the measurable space $(\Omega, \mathcal{F})$ and $Q(A) = 0$ whenever $A \in \mathcal{F}$ and $P(A) = 0$, then there exists a positive measurable function Y on Ω such that*

$$(2.17) \qquad\qquad Q(A) = \int_A Y dP$$

for all $A \in \mathcal{F}$. Moreover, any $\mathcal{F}$-measurable function Z on Ω satisfying Equation 2.17 for all $A \in \mathcal{G}$ is equal to Y almost everywhere.

Proposition 2.18. *If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$ and $\mathcal{G}$ is a σ-field on Ω such that $\mathcal{G} \subset \mathcal{F}$, then there exists a $\mathcal{G}$-measurable integrable random variable on $(\Omega, \mathcal{F}, P)$, $\mathbb{E}[X|\mathcal{G}]$, such that*

$$(2.19) \qquad\qquad \int_A \mathbb{E}[X|\mathcal{G}] dP = \int_A X dP$$

for all $A \subset \mathcal{G}$. Moreover, if Y is any $\mathcal{G}$-measurable integrable random variable satisfying

$$(2.20) \qquad\qquad \int_A Y dP = \int_A X dP$$

for all $A \subset \mathcal{G}$, then $Y = \mathbb{E}[X|\mathcal{G}]$ almost surely in $(\Omega, \mathcal{F}, P)$.

Proof. Without loss of generality, consider $X \geq 0$. Define a measure Q on $\mathcal{G}$ by

$$Q(A) = \int_A X dP$$

for all $A \in \mathcal{G}$. It is immediate that this is a measure, absolutely continuous relative to the measure P restricted to $\mathcal{G}$.

Then by *the Radon-Nikodým Theorem*, there exists a positive $\mathcal{G}$-measurable function function Z such that

$$\int_A X dP = \int_A Z dP$$

for all $A \in \mathcal{G}$.

For general random variable X, write $X = X^+ - X^-$, we can find random variables Z^+ and Z^- with respect to Z^+ and Z^-. Then such random variable satisfying Equation 2.19 exists and $\mathbb{E}[X|\mathcal{G}] = Z^+ - Z^-$.

If Y is any $\mathcal{G}$-measurable integrable random variable satisfying Equation 2.20, it satisfies 2.17, so $Y = Z = \mathbb{E}[X|\mathcal{G}]$ almost surely in $(\Omega, \mathcal{F}, P)$.

$\square$

If $A = \Omega$ in (2.19), then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \int_{\Omega} \mathbb{E}[X|\mathcal{G}]dP = \int_{\Omega} XdP = \mathbb{E}[X].$$

Definition 2.21. If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$ and $\mathcal{G}$ is a σ-filed on Ω such that $\mathcal{G} \subset \mathcal{F}$, then we call the $\mathcal{G}$-measurable random variable $\mathbb{E}[X|\mathcal{G}]$ satisfying

$$\int_A \mathbb{E}[X|\mathcal{G}]dP = \int_A XdP$$

for all $A \subset \mathcal{G}$, the *conditional expectation* of X given $\mathcal{G}$.

When Y is a random variable, we let $\mathbb{E}[X|Y] = \mathbb{E}[X|\mathcal{F}_Y]$ and call $\mathbb{E}[X|Y]$ the conditional expectation of X given Y.

3. MARTINGALES

A martingale is a stochastic process satisfying a condition that removes bias. There are two basic types of martingales: discrete and continuous. We focus on the discrete martingales here.

Definition 3.1. Let $(\mathcal{F}_n)_{n=1}^{\infty}$ be a filtration on the probability space $(\Omega, \mathcal{F}, P)$. A discrete martingale on $((\Omega, \mathcal{F}, P), (\mathcal{F}_n)_{n=1}^{\infty})$ is a sequence $(X_n)_{n=1}^{\infty}$ of integrable random variables on $(\Omega, \mathcal{F}, P)$, such that X_n is $\mathcal{F}_n$ measurable and

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$$

for all $n \geq 1$.

Proposition 3.2. *If $(X_n)_{n=1}^{\infty}$ is a martingale on $((\Omega, \mathcal{F}, P), (\mathcal{F}_n)_{n=1}^{\infty})$, then*
$$\mathbb{E}[X_n] = \mathbb{E}[X_m]$$
for all n and m.

Proof. $\mathbb{E}[X_n] = \int_{\Omega} X_n dP = \int_{\Omega} \mathbb{E}[X_{n+1}|\mathcal{F}_n]dP = \int_{\Omega} X_{n+1} dP = \mathbb{E}[X_{n+1}]$. Hence $\mathbb{E}[X_n] = \mathbb{E}[X_{n+1}] = \mathbb{E}[X_{n+2}] = ... = \mathbb{E}[X_m]$ for all n and m. $\square$

Example 3.3. Martingales are the mathematical fomulation of a sequence of fair games. Let X_n denote the winings per unit stake on the nth game in a sequence of fair games. Then $\mathbb{E}[X_n] = 0$ (since the games are fair) and $Y_n := \sum_{i=1}^{n} X_n$ are the winnings accumulated by the end of the nth game. Then $(Y_n)_{n=1}^{\infty}$ is a martingale.

Definition 3.4. If $(X_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P)$ adapted to the filtration $(\mathcal{F}_n)_{n=1}^{\infty}$, then $(X_n(\omega))_{n=1}^{\infty}$ is a sample path for each $\omega \in \Omega$.

We investigate if this stabilizes with time, that is, whether or not $\lim_{n \to \infty} X_n(\omega)$ exists.

Definition 3.5. A collection of integrable random variables $(X_\alpha)_{\alpha \in \tau}$ is L^1-bounded if

$$\sup_{\alpha \in \tau} \mathbb{E}[|X_\alpha|] < \infty.$$

Proposition 3.6. *If $(X_n)_{n=1}^{\infty}$ is a martingale on $(\Omega, \mathcal{F}, P)$ and $(X_\alpha)_{\alpha \in \tau}$ is L^1-bounded, then there exists an integrable random variable X on $(\Omega, \mathcal{F}, P)$ such that $\lim_{n \to \infty} X_n = X$ almost surely.*

Proof. Let $A := \{\omega \in \Omega : (X_n(\omega))_{n=1}^{\infty} \text{ converges}\}$. If $\omega \notin A$, then

$$\lim_{n \to \infty} \inf X_n(\omega) < \lim_{n \to \infty} \sup X_n(\omega)$$

and, since the rationals are dense in the reals, there exist rational numbers p and $q, p < q$, such that

$$\omega \in A_{p,q} := \{\omega \in \Omega : \lim_{n \to \infty} \inf X_n(\omega) < p < q < lim_{n \to \infty} \sup X_n(\omega)\}.$$

Hence, if $P(A_{p,q}) = 0$ for all $p, q \in \mathbb{Q}, p < q$, then

$$P(A^c) = P(\textstyle\bigcup_{p,q \in Q, p<q} A_{p,q}) \leq \textstyle\sum_{\{p,q \in Q, p<q\}} P(A_{p,q}) = 0$$

and the sequence $(X_n(\omega))_{n=1}^{\infty}$ converges almost surely. By the Monotone Convergence Theorem $\lim_{n \to \infty} X_n = X$ almost surely. $\qquad\square$

Definition 3.7. A set of integrable random variables $(X_i)_{i \in I}$ on the probability space $(\Omega, \mathcal{F}, P)$ is uniformly integrable if

$$\lim_{m \to \infty} \left(\sup_{i \in I} \int_{\{|X_i| \geq m\}} |X_i| dP \right) = 0$$

Definition 3.8. Let $(\mathcal{F}_t)_{t \in I}$ denote a filtration on $(\Omega, \mathcal{F}, P)$, indexed by an interval I of the real numbers, and let $(X_t)_{t \in I}$ denote a set of integrable random variables on $(\Omega, \mathcal{F}, P)$ adapted to the filtration; that is, X_t is $\mathcal{F}_t$ measurable for all $t \in I$. Then $(X_t)_{t \in I}$ is a continuous martingale if $\mathbb{E}[X_t | \mathcal{F}_s] = X_s$ for all $s, t \in I, s \leq t$.

4. Wiener Processes

Definition 4.1. Let $(W_t)_{t \geq 0}$ denote a collection of random variables on $(\Omega, \mathcal{F}, P)$ with the following properties:

- (a) $W_0 = 0$ almost surely;
- (b) W_t is $N(0,1)$ distributed for all $t \geq 0$ (Gaussian increments);
- (c) for any n and any $\{0 = t_0 < t_1 < ... < t_{n+1}\}$, $(W_{t_i} - W_{t_{i-1}})_{i=1}^{n}$ is a set of independent random variables (independent increments);
- (d) the probability distribution of $W_t - W_s$ depends only on $t - s$ for $0 \geq s \geq t$ (stationary increments).

A stochastic process satisfying the above properties is called Wiener process or Brownian motion.

Proposition 4.2. *If $(W_t)_{t \geq 0}$ is a Wiener process, then $(W_t)_{t \geq 0}$ and $(W_t^2 - t)_{t \geq 0}$ are martingales.*

Proof. By definition, $W_t - W_s$ is independent of $\mathcal{F}_s$ for all t and $s, 0 \leq s \leq t$, so is $(W_t - W_s)^2$. For $0 \leq s \leq t$

$$
\begin{aligned}
\mathbb{E}[W_t|\mathcal{F}_s] &= \mathbb{E}[W_t - W_s|\mathcal{F}_s] + \mathbb{E}[W_s|\mathcal{F}_s] \\
&= \mathbb{E}[W_t - W_s|\mathcal{F}_s] + W_s \\
&= \mathbb{E}[W_t - W_s] + W_s \\
&= \mathbb{E}[W_t] - \mathbb{E}[W_s] + W_s \\
&= W_s, \text{ since } \mathbb{E}[W_t] = \mathbb{E}[W_s] = 0.
\end{aligned}
$$

Hence $(W_t)_{t \geq 0}$ is a martingale.

For $0 \geq s \geq t, W_t^2 = (W_t - W_s)^2 + 2(W_t - W_s)W_s + W_s^2$ and

$$
\begin{aligned}
\mathbb{E}[W_t^2|\mathcal{F}_s] &= \mathbb{E}[(W_t - W_s)^2|\mathcal{F}_s] + 2\mathbb{E}[(W_t - W_s)W_s|\mathcal{F}_s] + \mathbb{E}[W_s^2|\mathcal{F}_s] \\
&= \mathbb{E}[(W_t - W_s)^2|\mathcal{F}_s] + 2W_s\mathbb{E}[(W_t - W_s)|\mathcal{F}_s] + W_s^2 \\
&= \mathbb{E}[(W_t - W_s)^2] + 2W_s\mathbb{E}[W_t - W_s] + W_s^2 \\
&= t - s + 2W_s \cdot 0 + W_s^2 \text{since } W_t - W_s \text{ is } N(0, t - s).
\end{aligned}
$$

Hence $\mathbb{E}[W_t^2 - t|\mathcal{F}_s] = W_s^2 + t - s - t = W_s^2 - s$. so $(W_t^2 - t)_{t \geq 0}$ is a martingale. □

Definition 4.3. If $(W_t)_{t \geq 0}$ is a Wiener process, fixing $\omega \in \Omega$, we get a function of time $X^\omega(t) = X(t, \omega)$, called a sample path of the process.

Proposition 4.4. *If $(W_t)_{t \geq 0}$ is a Wiener process, then there exists a probability space $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$ and a filtration $(\mathcal{F}_t)_{t \geq 0}$ on $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty, W)$, such that $(W_t)_{t \geq 0}$ is a stochastic process adapted to the filtration. Moreover, paths of the process are almost surely continuous and almost surely nowhere differentiable with respect to the measure W.*

The process $(\mu t + \sigma W_t)_{t \geq 0}$ is called Brownian motion with drift, while the process $(C \exp(\mu t + \sigma W_t))_{t \geq 0}$ is called a geometric or exponential Brownian motion.

5. The Black-Scholes Formula

In this section, we present an example of the application of stochastic process. We examine share price as the random variable, and derive the Black-Scholes fomula for pricing a call option.

We introduce some finance background first, for example, interest rate, options, and hedging.

We begin with the *interest rate*.

If an amount A is borrowed or saved for T years at rate r of *simple interest*, then the repayment due at time T is

$$
A + ArT = A(1 + rT).
$$

Interest can, of course, be compounded at various intervals of time. And the more frequent the compounding the greater the interest earned.

If the same amount is compounded at a total of nT intervals of times, the total repayment at time T will be $A(1 + rT)^{nT}$.

If we compound over smaller and smaller intervals, we obtain in the limit *continuously compounded interest.*

Proposition 5.1. *For any real number r*

$$\lim_{n \to \infty} (1 + \tfrac{r}{n})^n = e^r.$$

Proof. We know

$$\tfrac{d}{dx} \log(x) = \lim_{\Delta \to 0} \tfrac{\log(x + \Delta x) - \log(x)}{\Delta x} = \tfrac{1}{x}.$$

Let $x = 1$ and $\Delta x = r/n$, then $\Delta x \to 0$ as $n \to \infty$, then

$$\lim_{n \to \infty} \tfrac{\log(1 + \frac{r}{n}) - \log(1)}{\frac{r}{n}} = \lim_{n \to \infty} \tfrac{n}{r} \log(1 + \tfrac{r}{n}) = \tfrac{1}{r} \lim_{n \to \infty} \log(1 + \tfrac{r}{n})^n = 1.$$

Hence $\lim_{n \to \infty} \log(1 + \tfrac{r}{n})^n = r$, and as exp and log are inverse functions and both are continuous, this implies

$$\lim_{n \to \infty} (1 + \tfrac{r}{n})^n = \exp(\lim_{n \to \infty} \log(1 + \tfrac{r}{n})^n) = \exp(r) = e^r.$$

$\square$

The following two corollaries illustrate a basic functional relationship between time and money mathematically.

Corollary 5.2. *An amount A earning continuously compounded interest at a constant rate r per year is worth Ae^{rT} after T years.*

Corollary 5.3. *The discounted value of an amount A at a future time T, assuming a constant continuously compounded interest rate r, is given by Ae^{-rT}.*

Now we introduce the finance term *option.* An option is a contract which gives the buyer (the owner or holder) the right, but not the obligation, to buy or sell an underlying asset or instrument at a specified strike price on or before a specified date.

Since it is not an obligation, the investor can choose to not exercise the option. Hence the return is non-negative.

A *call option* is an option to buy a certain asset. An option to sell is called a *put option.*

If the option can only be exercised at any time prior to the maturity date, it is called a *European Option*; while if it can be exercised at any time prior to the maturity date, it is called an *American Option.* We only consider European options here.

We also introduce the term *arbitrage* and *hedge* in finance.

The simultaneous purchase and sale of an asset in order to profit from a difference in the price is called *arbitrage.* It is a trade that profits by exploiting price differences of identical or similar financial instruments, on different markets or in different forms.

If the market prices do not allow for profitable arbitrage, the prices are said to be *arbitrage-free.*

Hedging is making an investment to reduce the risk of adverse price movements in an asset. Normally, a hedge consists of taking an offsetting position in a related security, such as a futures contract

A *risk-neutral measure* is a probability measure such that each share price is exactly equal to the discounted expectation of the share price under this measure.

Now we present the *binomial model* for pricing options.

Proposition 5.4. *Suppose the interest rate is r, the share price of a certain stock is S at time 0 and that at a future time T it will either be Su or Sd where*

$$0 < d < 1 < e^{rT} < u.$$
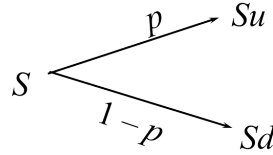
The risk neutral probability p that the share price will go up is

$$(5.5) \qquad p = \frac{e^{rT} - d}{u - d}$$

The arbitrage-free price for a call option, C_T, with strike price k, $Sd < k < Su$, and maturity date T is

$$(5.6) \qquad C_T = \frac{Su - k}{u - d} \cdot (1 - e^{-rT}d).$$

The arbitrage-free price for a put option, P_T, with strike price k, $Sd < k < Su$, and maturity date T is

$$(5.7) \qquad P_T = \frac{Sd - k}{u - d} \cdot (1 - e^{-rT}d).$$

Proposition 5.8. *No arbitrage opportunities for a call option exist if there exists a probability measure under which the discounted share price is a martingale.*

Proposition 5.9. *All claims on a call option can be hedged if there is at most one probability measure under which the discounted share price is a martingale.*

Proposition 5.10. *The seller's portfolio for hedging the call option consists of Δ shares and B bonds where*

$$(5.11) \qquad \Delta = \frac{Su - k}{Su - Sd} \ \text{ and } B = -de^{-rT}\left(\frac{Su - k}{u - d}\right)$$

Proposition 5.12. *If a stock has drift μ and volatility σ, then there exist two probability measures: W, the Wiener measure, and P_N, the risk neutral probability measure, on the measurable space $(\mathbb{R}^{[0,\infty)}, \mathcal{F}_\infty)$ such that the share price X_t has the following properties:*

(a) under $W, X_t = X_0 \exp(\mu t + \sigma W_t)$, and $(W_t)_{t \geq 0}$ is a Wiener process;

(b) under $P_N, e^{-rT} X_t = X_0 \exp(-\frac{\sigma^2}{2} t + \sigma \tilde{W}_t)$ and $(\tilde{W}_t)_{t \geq 0}$ is a Wiener process.

Corollary 5.13. *The discounted share price $(e^{-rT} X_t)_{t \geq 0}$ is a martingale with respect to the risk neutral probability measure P_n.*

Now we try to find the risk neutral probability.

We assume the stock has drift μ and strictly positive volatility σ. We partition the interval $[0, t]$ into n adjacent subintervals, each of length $\Delta t = t/n$. Then share price changes on each subinterval by a fraction $\exp(\mu \Delta t \pm \sigma \Delta x)$ where $(\Delta x)^2 = \Delta t$. We use taylor expension for $\exp(x)$ to second order. By proposition 5.4, the risk neutral probability p that the discounted share price rises by Δx over a typical subinterval $[s, s + \Delta t]$ so that a fair price is maintained or equivalently that the martingale property is satisfied, is given by

$$
\begin{aligned}
p &= \frac{e^{r(\Delta x)^2} - e^{\mu(\Delta x)^2 - \sigma \Delta x}}{e^{\mu(\Delta x)^2 + \sigma \Delta x} - e^{\mu(\Delta x)^2 - \sigma \Delta x}} \\
&= \frac{e^{(r-\mu)(\Delta x)^2} - e^{-\sigma(\Delta x)}}{e^{\sigma(\Delta x)} - e^{-\sigma(\Delta x)}} \\
&\approx \frac{(r-\mu)(\Delta x)^2 + \sigma \Delta x - \sigma^2 (\Delta x)^2/2}{\sigma \Delta x + \sigma^2 (\Delta x)^2/2 + \sigma \Delta x - \sigma^2 (\Delta x)^2/2} \\
&= \frac{\sigma + (r - \mu - \sigma^2/2)\Delta x}{2\sigma} \\
&= \frac{1}{2}(1 + \frac{(r - \mu - \frac{\sigma^2}{2})}{\sigma} \Delta x).
\end{aligned}
$$

Proposition 5.14. *(Black-Scholes Formula)*

Suppose the share price of a stock with volatility σ is X_0 today. For the buyer

$$
(5.15) \qquad X_0 N\left(\frac{\log(\frac{X_0}{k}) + (r + \frac{1}{2}\sigma^2)T}{\sigma \sqrt{T}}\right) - ke^{-rT} N\left(\frac{\log(\frac{X_0}{k}) + (r - \frac{1}{2}\sigma^2)T}{\sigma \sqrt{T}}\right)
$$

is a fair price for a call option with maturity date T and strike price k given that r is the risk-free interest rate.

Proof. By Proposition 5.12 and Corollary 5.13

$$
V_0 = \mathbb{E}_{P_N}[e^{-rT}(X_T - k)^+ | \mathcal{F}_0] = \mathbb{E}_{P_N}[e^{-rT}(X_T - k)^+]
$$

is the buyer's fair price for the option, and it suffices to show that this reduces to Equation 5.15. By Proposition 5.12,

$$
e^{-rT}(X_T - k)^+ = e^{-rT}(X_0 \cdot e^{(r - \frac{1}{2}\sigma^2)T + \sigma \sqrt{T} Y} - k)^+
$$

where Y is an $N(0, 1)$ distributed random variable. By proposition 2.14,

$$V_0 = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-rT}(X_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}x} - k)^+ e^{-\frac{1}{2}x^2} dx.$$

Since

$$X_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}x} - k \geq 0 \Leftrightarrow e^{\sigma\sqrt{T}x} \geq (\frac{k}{X_0})e^{-(r-\frac{1}{2}\sigma^2)T}$$

$$\Leftrightarrow x \geq \frac{1}{\sigma\sqrt{T}}(\log(\frac{k}{X_0}) - (r - \frac{1}{2}\sigma^2)T) =: T_1$$

using the substitution $y = x - \sigma\sqrt{T}$,

$$V_0 = \frac{X_0 e^{-\frac{1}{2}\sigma^2 T}}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{\sigma\sqrt{T}x-\frac{1}{2}x^2} dx - \frac{ke^{-rT}}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{-\frac{1}{2}x^2} dx$$

$$= \frac{X_0}{\sqrt{2\pi}} \int_{T_1}^{\infty} e^{-\frac{1}{2}(x-\sigma\sqrt{T})^2} dx - ke^{-rT}(1 - N(T_1))$$

$$= \frac{X_0}{\sqrt{2\pi}} \int_{T_1-\sigma\sqrt{T}}^{\infty} e^{-\frac{1}{2}y^2} - ke^{-rT}(1 - N(T_1))$$

$$= X_0(1 - N(T_1 - \sigma\sqrt{T})) - ke^{-rT}(1 - N(T_1)).$$

Since

$$T_1 - \sigma\sqrt{T} = \frac{1}{\sigma\sqrt{T}}(-\log(\frac{X_0}{k}) - (r + \frac{1}{2}\sigma^2)T)$$

we have

$$1 - N(T_1 - \sigma\sqrt{T}) = N(-T_1 + \sigma\sqrt{T}) = N(\frac{\log(\frac{X_0}{k})+(r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}).$$

Similarly

$$1 - N(T_1) = N(\frac{\log(\frac{X_0}{k})+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}).$$

Substituting these two formulas into the integral respresentation for V_0, we obtain the Black-Scholes formula. $\square$

Example 5.16. We use the Black-Scholes formula to price a call option with strike price \$26, maturity date 6 months, and interest rate 8% given that the stock has volatility 10%, that is $\sigma = 0.1$, and the share price is \$25 today. The price of the option is

$$25N(\frac{\log(\frac{25}{26})+(0.08+\frac{1}{2}(0.1)^2)\frac{1}{2}}{(0.1)\sqrt{0.5}}) - 26e^{-0.04}N(\frac{\log(\frac{25}{26})+(0.08-\frac{1}{2}(0.1)^2)\frac{1}{2}}{(0.1)\sqrt{0.5}})$$
$$= 25N(0.0495) - (24.98)N(-0.0212) = 0.51.$$

Proposition 5.17. *(Call-Put Parity)*
Suppose the share price of a stock with volatility σ is X_0 today. If C_T and P_T denote, respectively, fair prices for a call opoion and a put option with maturity date T, strike price k and risk-free interest rate r, then

(5.18) $$C_T - P_T = S - ke^{-rT}.$$

Proof. Since $C_T = \mathbb{E}[e^{-rT}(X_T-k)^+]$ and $P_T = \mathbb{E}[e^{-rT}(X_T-k)^-]$, Proposition 5.12 implies

$$\begin{aligned} C_T - P_T &= \mathbb{E}[e^{-rT}(X_T - k)^+] - \mathbb{E}[e^{-rT}(X_T - k)^-] \\ &= \mathbb{E}[e^{-rT}(X_T - k)] \\ &= X_0 - ke^{-rT}. \end{aligned}$$

$\square$

Using the *Call-Put Parity*, we can find the fomula to price a put option.
$$P_T = C_T - X_0 + ke^{-rT}$$
where
$$C_T = X_0 N\big(\frac{\log(\frac{X_0}{k})+(r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\big) - ke^{-rT} N\big(\frac{\log(\frac{X_0}{k})+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\big).$$

References

[1] Dineen, Sea. Probability Theory in Finance: A Mathematical Guide to the Black-Scholes Formula. Providence, R.I.: American Mathematical Society, 2005.
[2] Arbitrage Definition — Investopedia (Investopedia) http://www.investopedia.com/terms/a/arbitrage.asp
[3] Hedge Definition — Investopedia (Investopedia) http://www.investopedia.com/terms/h/hedge.asp