Stochastic Finance

Complementary Notes for Textbook (SCFA)

Instructor: Jaehyuk Choi

2017-18 Module 3 (Spring 2018)

Contents

1 Random Walks and First Step Analysis **5**

1.1 First Step Analysis 5

1.2 Time and Infinity 6

1.3 Tossing an Unfair Coin 7

1.4 Numerical Calculation and Intuition 7

1.5 First Steps with Generating Functions 8

1.6 Exercises 8

2 First Martingale Steps **9**

2.1 Classic Examples 9

2.2 New Martingales from Old 11

2.3 Revisiting the Old Ruins 12

2.4 Submartingales 12

2.5 Doob's Inequalities 12

2.6 Martingale Convergence 12

2.7 Exercises 12

3 Brownian Motion **13**

3.1 Covariances and Characteristic Functions 14

3.2 Visions of a Series Approximation 15

3.3 Two Wavelets 16

3.4 Wavelet Representation of Brownian Motion 16

3.5 Scaling and Inverting Brownian Motion 16

3.6 Exercises 16

4 Martingales: The next steps **17**

4.1 Foundation Stones 17

4.2 Conditional Expectations 17
4.3 Uniform Integrability 17
4.4 Martingales in Continuous Time 17
4.5 Classic Brownian Motion Martingles 18
4.6 Exercises 20

5 Richness of Paths **21**
5.1 Quantitative Smoothness 21
5.2 Not Too Smooth 21
5.3 Two Reflection Principles 21
5.4 The Invariance Principle and Donsker's Theorem 23
5.5 Random Walks Inside Brownian Motion 24
5.6 Exercises 24

6 Itô Integration **25**
6.1 Definition of Itô's Integral: First Two Steps 26
6.2 Third Step: Itô's Integral as a Process 27
6.3 The Integral Sign: Benefits and Costs 27
6.4 An Explicit Calculation 27
6.5 Pathwise Interpretation of Itô's Integrals 28

7 Localization and Itô's integral **29**
7.1 Itô's Integral on $\mathcal{L}^2_{\text{LOC}}$ 29
7.2 An Intuitive Representation 29
7.3 Why Just $\mathcal{L}^2_{\text{LOC}}$? 31
7.4 Local Martingales and Honest Ones 31
7.5 Alternative Fields and Changes of Time 31
7.6 Exercises 31

8 Itô's Formula **32**
8.1 Analysis and Synthesis 32
8.2 First Consequences and Enhancements 33
8.3 Vector Extension and Harmonic Functions 35
8.4 Functions of Processes 35
8.5 The General Itô's Formula 35

8.6 Quadratic Variation . . . 37
8.7 Exercises . . . 37

9 Stochastic Differential Equations **38**
9.1 Matching Itô's Coefficients . . . 38
9.2 Ornstein-Uhlenbeck Processes . . . 39
9.3 Matching Product Process Coefficients . . . 40
9.4 Existence and Uniqueness Theorems . . . 42
9.5 Systems of SDEs . . . 42
9.6 Exercises . . . 42

10 Arbitrage and SDEs **43**
10.1 Replication and Three Examples of Arbitrage . . . 43
10.2 The Black-Scholes Model (Wikipedia) . . . 44
10.3 The Black-Scholes Formula (Wikipedia) . . . 45
10.4 Two Original Derivations . . . 47
10.5 The Perplexing Power of a Formula . . . 47
10.6 Exercises . . . 47

11 The Diffusion Equation **48**
11.1 The diffusion of Mice . . . 48
11.2 Solutions of the Diffusion Equation . . . 48
11.3 Uniqueness of Solutions . . . 48
11.4 How to Solve the Black-Scholes PDE . . . 48
11.5 Exercises . . . 48

12 Representation Theorems **49**
12.1 Stochastic Integral Representation Theorem . . . 49
12.2 The Martingale Representation Theorem (Wikipedia) . . . 49
12.3 Continuity of Conditional Expectations . . . 49
12.4 Representation via Time Change . . . 50
12.5 Lévy's Characterization of Brownian Motion . . . 50
12.6 Bedrock Approximation Techniques . . . 50
12.7 Exercises . . . 50

13 Girsanov Theory **51**
13.1 Importance Sampling . . . 51
13.2 Tilting a Process . . . 52
13.3 Simplest Girsanov Theorem . . . 53
13.4 Creation of Martingales . . . 55
13.5 Shifting the General Drift . . . 56
13.6 Exponential Martingales and Novikov's Condition . . . 56
13.7 Exercises . . . 56

14 Arbitrage and Martingales **57**
14.1 Reexamination of the Binomial Arbitrage . . . 57
14.2 The Valuation Formula in Continuous Time . . . 57
14.3 The Black-Scholes Formula via Martingales . . . 61
14.4 American Options . . . 62
14.5 Self-Financing and Self-Doubt . . . 62
14.6 Admissible Strategies and Completeness . . . 62
14.7 Perspective on Theory and Practice . . . 62
14.8 Exercises . . . 62

15 The Feynman-Kac Connection **63**
15.1 First Links . . . 63
15.2 The Feynman-Kac Connection for Brownian Motion . . . 63
15.3 Lévy's Arcsin Law . . . 63
15.4 The Feynman-Kac Connection for Diffusions . . . 63
15.5 Feynman-Kac and the Black-Scholes PDEs . . . 63
15.6 Exercises . . . 63

Chapter 1

Random Walks and First Step Analysis

Random walk is a probability process whose incremental change in unit time is up or down by random;

$$S_n = S_0 + X_1 + X_2 + \cdots X_n,$$
$$\text{where } X_k = 1 \text{ or } -1 \text{ with } 50\%\text{:}50\% \text{ chance.}$$

The process models the wealth of a gambler but it is easier to understand if S_n is the daily closing price of a stock and X_n is the profit and loss (P&L) of the n-th day.

For the rest of this chapter, except §1.5, we are interested in the event of S_n hitting A before hitting $-B$ (the gambler making \$$A$ first before losing \$$B$). Equivalently, the event is the stock price gaining A before losing B (assuming that you set a trading strategy of loss-cutting at $-B$ and profit-realizing at A).

For the purpose, the *stopping time* τ is introduced as the first time n when S_n hits either A or $-B$. So we know that $S_\tau = A$ or $-B$ although we don't know the value of τ (τ is a probability variable).

1.1 First Step Analysis

We first solve the probability of the event, $P(S_\tau = A \mid S_0 = 0)$. Generalizing the problem, let

$$f(k) = P(S_\tau = A \mid S_0 = k)$$

be the probability of the same event with the initial point being $S_0 = k$ rather than 0. The recurrence relation is given as

$$f(k) = \frac{1}{2} f(k-1) + \frac{1}{2} f(k+1) \quad \text{for} \quad -B < k < A \tag{1.1}$$

with the *boundary conditions* $f(A) = 1$ and $f(-B) = 0$. This basically means that $f(k)$ is a linear function.

After some algebra, we get

$$f(k) = \frac{k+B}{A+B}, \quad P(S_\tau = A \mid S_0 = 0) = f(0) = \frac{B}{A+B}.$$

The result is in line with the intuition that the probability goes to 1 when B gets bigger or goes to 0 when A gets bigger.

In relation to finance, almost all probability or expectation values can be thought of as the price of a security or a derivative. In this example, we can think of a derivative that pays \$1 when the underling stock price S_n hits A or expires worthless when S_n hits $-B$. This is a *derivative* security because the payoff is *derived* from the underlying stock S_n. Unlike the usual call or put options, the expiry of this derivative is infinite (sometimes such security is called *perpetual*). The probability we computed above, $P(S_\tau = A \mid S_0 = 0)$, can be understood as the current price of the derivative.

Quiz: (a hedging strategy) Imagine that you (as an investment bank) sold the derivative to investors. How would you *hedge* your position using the underlying stock?

1.2 Time and Infinity

In this section, we computes the expected number of bets, τ, until the gambler finishes the game, i.e., when he makes \$A or loses \$B. **SCFA** first proves that the expectation of τ (and any power) is finite. (See **SCFA** for detail.)

In a similar approach from the previous section, the generalized expectation, $g(k) = E(\tau \mid S_0 = k)$ satisfy the recurrence relation,

$$g(k) = \frac{1}{2}g(k-1) + \frac{1}{2}g(k+1) + 1 \quad \text{for} \quad -B < k < A$$

with the boundary condition, $g(A) = g(-B) = 0$.

Notice that $\frac{1}{2}g(k-1) + \frac{1}{2}g(k+1) - g(k)$ is the convexity (or curvature) operator. From the Taylor expansion, we know for small h,

$$\frac{1}{2}g(x+h) + \frac{1}{2}g(x-h) - g(x) \approx \frac{1}{2}g''(x)h^2.$$

So the recurrence relation above implies that $g(k)$ is a quadratic function on k with the second order coefficient is -1. Therefore we conclude that

$$g(k) = (A-k)(B+k) \quad \text{and} \quad \mathbb{E}(\tau \mid S_0 = 0) = AB$$

This quantity can be also thought of as the price of a financial contract, in which \$1 is accumulated each time unit and the money is paid to the investor when the event is triggered. This type of derivatives are generally called *accumulator*.

SCFA verifies the obtained result for the symmetric case of $A = B$. The standard deviation of S_n is $\sqrt{n}$. (The variance is n.) Since the stdev is the characteristic width (or scale) of the process, we can estimate that the time required for the scale to reach A is A^2, which is consistent with the result.

> **Quiz (a popular interview question)**: Imagine that you keep tossing a fair coin (50% for head and 50% for tail) until you get two heads <u>in a row</u>. On average, how many times do you need to toss a coin?
> See the HW solutions from previous years for the answer.

1.3 Tossing an Unfair Coin

When the probability of X_1 is not fair and instead given as

$$X_n = 1 \text{ or } -1 \text{ with the chance of } p \text{ or } q \text{ respectively} \quad (p+q=1),$$

we can still drive the equivalent results.

After some algebra,

$$f(k) = \frac{(q/p)^{k+B} - 1}{(q/p)^{A+B} - 1} \quad \text{and} \quad P(S_\tau = A | S_0 = 0) = f(0) = \frac{(q/p)^{B} - 1}{(q/p)^{A+B} - 1}.$$

$$\mathbb{E}(\tau \mid S_0 = 0) = \frac{B}{q-p} - \frac{A+B}{q-p}\frac{1-(q/p)^B}{1-(q/p)^{A+B}}$$

One can recover the result of the fair bet case, if p and q are approaching to $\frac{1}{2}$, i.e., $p = \frac{1}{2} + \varepsilon$ and $q = \frac{1}{2} - \varepsilon$ for very small ε.

> **Quiz (numerical implementation)**: If you want to implement the above results, i.e., $f(k)$ and $g(k)$ for a general value of $p = 1/2 + \varepsilon$, you will run into a small issue because you have to write a function for the two cases depending on $\varepsilon = 0$ or $\varepsilon \neq 0$. If ε is very small, then the formula may break. How would you resolve this issue?

1.4 Numerical Calculation and Intuition

I recommend that the students quickly verify the numbers in Table 1.1 using your favorite computer tool (R, Matlab, Python or even a calculator). It is quite noticeable that the probability for a gambler to win \$100 before losing \$100 is only 6×10^{-6} when $p = 0.47$.

1.5 First Steps with Generating Functions

The probability generating function is a powerful trick to obtain a series of values in one go, where the coefficients of the Taylor expansion is the values to seek. This chapter of **SCFA** considers the event of S_n hitting 1 for the first time (no longer the event of hitting A or $-B$) and wants to compute the probability of the event happening at time $\tau = 0, 1, 2, \cdots$ (the meaning of τ is also different from the previous sections!). The generating function is in the form of

$$\phi(z) = E(z^\tau \mid S_0 = 0) = \sum_{k=0}^{\infty} P(\tau = k \mid S_0 = 0)\, z^k,$$

i.e. the coefficient of z^k is the probability of S_n hitting 1 at time $\tau = k$ for the first time.

SCFA obtains the function $\phi(z)$ using the recurrence relation method. One important observation is that $\phi(z)^k$ is the generating function for the event of hitting k, which is from the property that the generating function for the sum of independent random variables is the product of the individual generating functions. For $k = 2$, let τ_1 is the first hitting time from 0 to 1 and τ_2 is the first hitting time from 1 to 2. Because τ_1 and τ_2 are independent (and identical) random variables,

$$E(z^{\tau_1+\tau_2}) = E(z^{\tau_1})E(z^{\tau_2}) = \phi(z)^2.$$

Thus, we end up the recurrent relation

$$\phi(z) = \frac{1}{2}\, z\, \phi(z)^2 + \frac{1}{2} z$$

and the $\phi(z)$ is finally given as

$$\phi(z) = \frac{1 - \sqrt{1 - z^2}}{z}.$$

The root with $+$ sign was excluded because the function has the term of $1/z$ and non-zero constant term (the probability for the negative or zero first hitting time should be zero).

1.6 Exercises

Chapter 2

First Martingale Steps

Martingale is one of the key concepts in stochastic process. Although it is a very formal mathematical concept, it will turn out that many practical results will be derived out of it. For the definition of the martingale, we refer to (WIKIPEDIA):

> In probability theory, a martingale is a model of a fair game where knowledge of past events never helps predict the mean of the future winnings. In particular, a martingale is a sequence of random variables (i.e., a stochastic process) for which, at a particular time in the realized sequence, the expectation of the next value in the sequence is equal to the present observed value even given knowledge of all prior observed values.

In **SCFA**, a stochastic process $\{M_n : 0 \le n\}$ is a *martingale* with respect to another stochastic process $\{X_n : 0 \le n\}$ if (i) the sequence M_n is determined from the past knowledge of $\{X_k : 0 \le k \le n\}$ and (ii) the next expectation value is equal to the present value of M_n (*fundamental martingale identity*),

$$E(M_{n+1} \mid X_1, X_2, \cdots, X_n) = M_n \text{ for all } n \ge 0.$$

In general, however, $\{M_n\}$ is simply a martingale if the next expectation value, conditional on the history of itself, is equal to the present value,

$$E(M_{n+1} \mid M_1, M_2, \cdots, M_n) = M_n \text{ for all } n \ge 0.$$

2.1 Classic Examples

SCFA gives 3 examples of martingales

Example 1 If the X_n are independent random variables with zero mean, the running sum, $S_n = \sum_0^n X_k$ is a martingale. The process S_n was the subject of Chapter 1. So the wealth of a gambler or a stock price are all martingale as long as the game is fair ($E(X_n) = 0$ and the no one can look into the future. In the case of the stock, this assumption is closely related to the efficient market hypothesis (WIKIPEDIA), where the stock prices reflect the market information immediately and fully. Since all the news are *priced in* the stock, the expectation for tomorrow's stock is same as the current value (no one know that tomorrow's news will be good or bad).

This observation gives us a good example of what is **not** a martingale. Imagine that a stock price has a momentum (or a positive auto-correlation) in that the stock price tends to be up (or down) in a day when the price was up (or down) in the previous day, i.e., X_n and X_{n+1} are positively correlated rather than independent. The stock price in that circumstance is not a martingale because one can look into the future (based on the past). Many technical analyses are indeed based on that stock markets have momentum. For a well-known strategy, see the turtle trading rule (download).

Example 2 On top of the assumptions of **Example 1**, let us assume that $\text{Var}(X_n) = \sigma$. Then $M_n = S_n^2 - n\sigma^2$ is also a martingale. See the **SCFA** for the detailed proof. Basically what it tells us is that the squared process S_n^2 increases by the σ^2 on average on each time step, so we need to add the correction term, $-n\sigma^2$ for the process M_n to be a martingale. This is an important precursor to the famous Itô's lemma which we will cover later!

Example 3 If $\{X_n\}$ are non-negative independent random variables with $E(X_n) = 1$, the running product $M_n = X_1 \cdot X_2 \cdots X_n$ is a martingale. See the **SCFA** for the detailed proof. Out of any identical and independent random variables $\{Y_n\}$, we can construct such $\{X_n\}$ by

$$X_n = e^{\lambda Y_n}/\phi(\lambda) \quad \text{where} \quad \phi(\lambda) = E(e^{\lambda Y n})$$

and the resulting martingale is

$$M_n = \exp(\lambda \sum_{k=1}^{n} Y_k)/\phi(z)^n$$

Shortened Notation

This paragraph is about a rather formal mathematical background called *filtration*. While it is an important subject providing a mathematical background for the stochastic process, it is enough to understand what the notation mean in common sense. A filtration, $\{\mathcal{F}_n\}$, can be understood as the set of information available (or events that happened) up to time n. The set $\mathcal{F}_n$ not only contains the event at time n but also all the past events before n. Therefore the contents of $\mathcal{F}_n$ increases as n increases (time passes), i.e., $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. If $\{\mathcal{F}_n\}$ is the filtration that contains information with respect to a stochastic process $\{X_n\}$, i.e., $X_n \in \mathcal{F}_n$, we can shorten many of our previous statements. For example, we can now say a stochastic process $\{M_n\}$ is a martingale with respect to $\{\mathcal{F}_n\}$ and it satisfy

$$E(M_{n+1} \mid \mathcal{F}_n) = M_n \text{ for all } n \geq 0.$$

For a practical purpose, it can not go wrong even if you simply think that $\{\mathcal{F}_n\}$ represents *all* information known to time n, not just the information about $\{X_n\}$.

2.2 New Martingales from Old

The main idea of this section is the Martingale Transform Theorem (Theorem 2.1). Assume that $\{M_n\}$ is a martingale with respect to $\{\mathcal{F}_n\}$ representing the price of a stock (or a gambler's wealth). What if you change the unit of stock every day or the gambler changes the size of bet every time? Let A_n be such multiplier before the outcome of the n-th step. Then, the amount of the wealth will be

$$\widetilde{M}_n = M_0 + A_0(M_1 - M_0) + A_1(M_2 - M_1) + A_2(M_3 - M_2) + \cdots .$$

(<u>Note that the indexing of A_n here is slightly different from that of **SCFA**.</u>) This process $\{\widetilde{M}_n\}$ is called the martingale transform of $\{M_n\}$ by $\{A_n\}$. What the theorem is stating is a common-sense that if the <u>bounded</u> random variable A_n is determined from the information up to the time n (*non-anticipating* to $\{\mathcal{F}_n\}$ or $A_n \in \mathcal{F}_n$), the new process $\{\widetilde{M}_n\}$ is also a martingale. Again, the *no-fortune-telling* condition on $\{A_n\}$ is critical here.

Stopping times provide martingale transforms

In terms of the new martingale $\{\widetilde{M}_n\}$, we can think of a special type of trading (or betting) strategy where $A_k = 1$ if $k \leq \tau$ or 0 otherwise for a random variable τ. It means you have some kind of betting strategy (or trading strategy) such that you stop betting (or investing in stock) after the outcome at the τ-th step is just known. The random variable τ is a *stopping time* only when the stopping decision is made only from the information at each time step, not in the future (no-fortune-telling again!). Using the filtration notation above, we can say

$$\{\tau \leq n\} \in \mathcal{F}_n.$$

It seems quite confusing but what it means in simple words is that, when you are at n-th time step (so you know all information up to time n), you have to know for sure that either you want to stop $\tau = n$ or you already stopped before $\tau < n$ (so $\{\tau \leq n\}$ is already a known event time n). An example of a stopping strategy which is not a stopping time would be something like you stop you stop your bet at time n when you know you'll lose in the next bet, e.g., $M_{n+1} - M_n < 0$, which is obviously when you have a fortune-telling power. So the bottom line is that any τ associated with any stopping strategy you can imagine with common sense is a proper stopping time, so you don't need to worry too much about the stopping time.

In this regard, Theorem 2.2 is trivial from Theorem 2.1. Restating the theorem,

Theorem 2.2 (Stopping Time Theorem) The stopped process $\{M_{n\wedge\tau}\}$ ($n \wedge \tau = \min(n, \tau)$) derived from the original martingale $\{M_n\}$ is also a martingale.

2.3 Revisiting the Old Ruins

Given that we are armed with the knowledge of martingales and stopping times, the author derives the results of Chapter 1 in a much easier and more elegant way. First note that the first hitting time τ (of hitting A or $-B$) is a stopping time indeed. Please read the book for the detailed re-derivation.

2.4 Submartingales

We skip this section.

2.5 Doob's Inequalities

We skip this section.

2.6 Martingale Convergence

We skip this section.

2.7 Exercises

The solution for Exercise 2.1 is in the **SCFA Exercise Solutions**.

Chapter 3

Brownian Motion

Brownian Motion (BM) is the continuous version of the discrete random walk we covered in Chapter 1. Basically it is a stochastic process where normal distributions are repeated so that the stdev is increasing as $\sqrt{t}$. In other books, it is also called *Wiener process* (WIKIPEDIA) after the Mathematician provided the Mathematical background of it.

Steele starts the chapter by stating that Brownian motion is the most important stochastic process, which I can not agree more. Brownian motion will be used a basic building block for about 99% of the stochastic processes that you'll see in financial modeling! So understanding BM is the single most important goal of this course.

BM has been closely linked to finance as well as physics. Although it is often overshadowed by the great success of Black-Scholes-Merton's option pricing theory (1973), a French mathematician, Bachelier (WIKIPEDIA) made a first option pricing theory in his Ph.D. thesis *The Theory of Speculation* (1900) based on BM. And it was 5 years earlier than the Einstein's famous paper on BM (1905)!

BM is defined as below:

Definition 3.1 A *continuous-time stochastic process* $\{B_t : 0 \le t < T\}$ is called a Standard Brownian Motion on $[0,\ T)$ if (i) $B_0 = 0$, (ii) The increments of B_t, i.e., $B_{t_2} - B_{t_1},\ B_{t_3} - B_{t_2},\ \cdots$ for $0 \le t_1 < t_2 < t_3 < \cdots$, are independent, (iii) the increment $B_t - B_s$ for $s \le t$ has the Gaussian distribution with mean 0 and standard deviation $\sqrt{t-s}$ and (iv) $B(t)$ is a continuous function.

The rest of this chapter is focused on how one can represent BM as a (infinite) sum of functions. Although it is an interesting topic (one of my research topic is related to this), we don't see an immediate practical use for our course, so we will skip many of the following sections.

3.1 Covariances and Characteristic Functions

We skip the multivariate Gaussian distribution part for now. We will use some results of this section when we simulate multi-dimensional correlated BM's later. The covariance property of a single variable BM is important.

Multivariate Gaussians

For a d-dimensional random vector $\mathbf{V}$, we define the mean vector and the covariance matrix as:

$$\mathbf{V} = \begin{bmatrix} V_1 \\ \vdots \\ V_d \end{bmatrix}, \quad \mu = E[\mathbf{V}] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix}, \quad \Sigma = \mathrm{Cov}(\mathbf{V}) = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1d} \\ \vdots & \ddots & \vdots \\ \sigma_{d1} & \cdots & \sigma_{dd} \end{bmatrix},$$

$$\text{where} \quad \mu_k = E[V_k] \quad \text{and} \quad \sigma_{kj} = \mathrm{Cov}(V_k, V_j)$$

Definition 3.2 (Multivariate Gaussian) $V \sim N(\mu, \Sigma)$

$$f_V(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu)\right)$$

Gaussian Miracle

One simple case to understand the multivariate Gaussian distribution is a uncorrelated bi-variate Gaussian vector:

$$V = \begin{bmatrix} X \\ Y \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}$$

Since $\det \Sigma = (\sigma_X \sigma_Y)^2$, the density function can be separable:

$$f_V(x,y) = \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}} \cdot \frac{1}{\sigma_Y \sqrt{2\pi}} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}}$$

Characteristic Functions

Proposition 3.1 (Gaussian Characterization) The $d-$dimensional random vector V is a multivariate Gaussian if and only if any linear combinations $\theta^T V = \sum_k \theta_k V_k$ is a univariate Gaussian.

The proof depends on the characteristic function of Gaussian distribution. Although you may not need to fully understand the proof, you should know that

$$E(\theta^T \mathbf{V}) = \theta^T E(\mathbf{V}) = \theta^T \mu \quad \text{and} \quad \mathrm{Var}(\theta^T \mathbf{V}) = \theta^T \Sigma \theta.$$

Further, we can show that the multivariate vector $A\mathbf{V}$ for $d \times d$ matrix A (see the solution of Exercise 3.3)

$$E(A\mathbf{V}) = AE(\mathbf{V}) = A\mu \quad \text{and} \quad \text{Cov}(A\mathbf{V}) = A\,\Sigma\,A^T.$$

Now let us express $\mathbf{V} \sim N(\mu, \Sigma)$ as linear combinations of standard normal vector $\mathbf{Z}$, i.e., $\mathbf{V} = A\mathbf{Z}$, which gives us benefits:

- We can use $\mathbf{V} = A\mathbf{Z}$ for random number generation
- The PDF for $\mathbf{Z}$ is separable as $f_Z(\mathbf{z}) = n(z_1) \cdots n(z_d)$.

We select the matrix A in our favor such that

$$\text{Cov}(\mathbf{V} = A\mathbf{Z}) = A\,I\,A^T = A\,A^T = \Sigma.$$

The matrix A, often called as a square matrix of Σ, is not unique. The most well-known solution is Cholesky decomposition of Σ, $LL^T = \Sigma$ for a lower-triangular matrix L. Another well-known solution is the eigen-decomposition,

$$\Sigma A = A\,\Lambda \quad \Rightarrow \quad \Sigma = A\Lambda A^T = A\sqrt{\Sigma}\left(A\sqrt{\Sigma}\right)^T$$

Here we used $A^T = A^{-1}$ because Σ is symmetric, $\Sigma = \Sigma^T$.

Covariance Functions and Gaussian Processes

A process, X_t is called a Gaussian process if the vectors $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$ form a multivariate Gaussian distribution for any finite set of $\{t_k\}$. On the other hand, the covariance for Brownian motion between time s and t ($s \le t$) is given as

$$\text{Cov}(B_s, B_t) = E(B_t B_s) = E(B_s B_s) + E((B_t - B_s)B_s) = s + 0 = s \wedge t$$

due to the property of the independent increments. The following lemma states that the opposite is also true, i.e, any process with covariance, $s \wedge t$, has independent increments.

Lemma 3.1 If a Gaussian process X_t has $E(X_t) = 0$ and $\text{Cov}(X_s, X_t) = s \wedge t$, X_t has independent increments. Moreover if $X_0 = 0$ and X_t has continuous path, X_t is a standard Brownian motion.

3.2 Visions of a Series Approximation

We skip this section.

3.3 Two Wavelets

We skip this section.

3.4 Wavelet Representation of Brownian Motion

We skip this section.

3.5 Scaling and Inverting Brownian Motion

Proposition 3.2 For any $a > 0$, the following three processes defined by

$$X(t) = \frac{1}{\sqrt{a}} B(at) \text{ for } t \geq 0 \quad \text{(scaled process)},$$

$$Y(0) = 0 \text{ and } Y(t) = t\, B(1/t) \text{ for } t > 0 \quad \text{(inverted process)},$$

$$Z(t) = B(1) - B(1-t) \quad \text{(time-reversed process)}$$

are all standard BM's on $[\,0, \infty)$ for $X(t)$ and $Y(t)$ and on $[\,0, 1]$ for $Z(t)$.

For the graphical demonstration of the scaled process, see the **self-similarity** section in (WikipediA).

3.6 Exercises

Exercise problems from 3.1 to 3.4 are recommended. The solutions are in the **SCFA Exercise Solutions**.

Chapter 4

Martingales: The next steps

This chapter introduces continuous-time martingales, thus parallels with Chapter 2. In a similar way we covered Chapter 2, we will focus on the intuition and skip the sections on the rigorous mathematical definition.

4.1 Foundation Stones

We skip this section.

4.2 Conditional Expectations

We skip this section.

4.3 Uniform Integrability

We skip this section.

4.4 Martingales in Continuous Time

We first introduce filtration under continuous time, $\{\mathcal{F}_t : \ 0 \leq t < \infty\}$. Again, set $\mathcal{F}_t$ represents all the (cumulative) information up to time t, thus $s \leq t$ implies that $\mathcal{F}_s \subset \mathcal{F}_t$. If a continuous filtration $\{\mathcal{F}_t\}$ contains all the information about a continuous stochastic process $\{X_t\}$, we say X_t is $\mathcal{F}_t-$*measurable* or X_t is *adapted* to the filtration $\mathcal{F}_t$. Similarly in Chapter 2, however, it is more convenient to assume that the filtration $\{\mathcal{F}_t\}$ is the filtration which contains *all* information, not just about a stochastic process. Now, the process $\{X_t\}$ is a martingale if

1. $E(|X_t|) < \infty$ for all $0 \leq t < \infty$ and

2. $E(X_t \mid \mathcal{F}_s) = X_s$ for all $0 \leq s \leq t < \infty$.

The Standard Brownian Filtration

This part discusses the minimal filtration Brownian motion is adapted to, i.e., the filtration having just enough information on Brownian motion. In the light of our convenient maximal assumption of the filtration $\mathcal{F}_t$, however, we skip this section.

Stopping Times

The stopping time under continuous time framework is almost same. Again, the stopping time is a stopping strategy under which one can determine stop or not based on the information up to now, not in the future.

A continuous random variable τ is a stopping time with respect to a filtration $\{\mathcal{F}_t\}$ if

$$\{w : \tau(w) \leq t\} \in \mathcal{F}_t \quad \text{for all } t \geq 0.$$

Here w is a particular path or a realization of the process X_t. Also, the stopped variable X_τ (when τ is not ∞) is naturally $X_\tau(w) = X_t(w)$ for $\tau(w) = t$.

Doob's Stopping Time Theorem

Theorem 4.1 This theorem is the continuous-time version of Theorem 2.1, which states that the stopped process $M_{t\wedge\tau}$ derived from the original continuous process M_t and the stopping time τ is also a continuous martingale.

we will skip the rest of the section.

4.5 Classic Brownian Motion Martingles

Theorem 4.4 The following three processes associated with the standard Brownian motion are all continuous martingales

1. B_t,
2. $B_t^2 - t$,
3. $\exp(\alpha B_t - \alpha^2 t/2)$

Note that the three examples above are the continuous-time versions of the examples discussed in Section 2.1. Again, the second case gives the insight, $(dB_t)^2 = dt$, which leads to Itô's lemma. The third one is the geometric Brownian motion (log-normal process) which frequently appears in the derivation of the Black-Scholes-Merton formula if the parameter α is replaced with the volatility σ.

Ruin Probabilities for Brownian Motion

Theorem 4.5 Using the martingale property, we again obtain the same results,

$$P(B_\tau = A) = \frac{B}{A+B} \quad \text{and} \quad E(\tau) = AB,$$

where τ is the first hitting time of double-barrier, $\tau = \inf\{t : B_t = -B \text{ or } B_t = A\}$.

In the proof, we used the property from the second martingale example, $E(B_\tau^2) = E(\tau)$.

Hitting Time of a Level

Now we consider the first hitting time of a one-side barrier, $\tau_a = \min\{t : B_t = a\}$. We obtain the following important result:

Theorem 4.6 For any value real value a,

$$P(\tau_a < \infty) = 1 \quad \text{and} \quad E(e^{-\lambda \tau_a}) = e^{-|a|\sqrt{2\lambda}}.$$

In the proof of the second part, we used that the exponential Brownian motion is a (continuous-time) martingale,

$$1 = E(M_{\tau_a}) = \exp(\alpha a - \alpha^2 \tau_a/2).$$

By taking $\alpha = \mathrm{sign}(a)\sqrt{2\lambda}$, we obtain the second part.

Note that, for a non-negative random variable X, $E(e^{-\lambda X})$ is the Laplace transform of the probability density function of X. The Laplace transform $E(e^{-\lambda X})$ is well-defined for the random variables with infinite moments, thus more useful than the moment generating function $E(e^{\lambda X})$ sometimes.

The second part also gives a pricing formula for a very plausible derivative, which pays \$1 when the underling stock following the standard BM B_t hits the level a. Then the interest rate is r, the second formula gives the present value of the derivative (perpetual digital option)

$$P = E(e^{-r\tau_a}) = e^{-|a|\sqrt{2r}}.$$

The fact that $P = 1$ when $r = 0$ is consistent with the first part, i.e., the probability of hitting the level is 1.

Quiz: how does the price formula modified if the underling stock follows a BM with volatility σ?

First Consequences

The Laplace transform (4.25) of the hitting-time density provides useful insights. For example, we can conclude $E(\tau_a) = \infty$ from that

$$E(\tau_a) = -\frac{d}{d\lambda}E(e^{-\lambda\tau_a})\Big|_{\lambda=0} = \frac{a}{\sqrt{2\lambda}}e^{-a\sqrt{2\lambda}}\Big|_{\lambda=0} = \infty$$

We can also show $E(1/\tau_a) = 1/a^2$ from

$$E\left(\frac{1}{\tau_a}\right) = \int_0^\infty E(e^{-\lambda\tau_a})d\lambda = \int_0^\infty e^{-a\sqrt{2\lambda}}d\lambda = \int_0^\infty ue^{-u}\frac{du}{a^2} = \frac{1}{a^2},$$

where we used the identity $1/t = \int_0^\infty e^{-\lambda t}d\lambda$ and the change of variable $u = a\sqrt{2\lambda}$.

The inverse Laplace transform is actually known, thus we have the analytic expression of the hitting-time density distribution as

$$f_{\tau_a}(t) = \frac{|a|}{\sqrt{2\pi t^3}}e^{-a^2/2t}$$

This is a special case (zero drift) of the Inverse Gaussian distribution family (WIKIPEDIA). We will elegantly derive this result using the reflection principle in Chapter 5.

Looking Back

The author argues that the functional form of the Laplace transform can be guessed from the symmetry argument.

4.6 Exercises

Exercise problem 4.6 is recommended. The solution is provided in the **SCFA Exercise Solutions**.

Chapter 5

Richness of Paths

5.1 Quantitative Smoothness

We skip this section.

5.2 Not Too Smooth

We skip this section.

5.3 Two Reflection Principles

We continue to explore quantitative properties of Brownian motion. In this section we elegantly derive the distributions related to the maximum process, $B_t^* = \max_{0\le s\le t} B_s$. We'll drive the joint distribution of $(B_t^*,\ B_t)$ and the distribution of B_t^* itself. Both results are important for pricing exotic options such as barrier and max options.

At the heart of the derivation is the reflection principle of Brownian motion (Wikipedia). The author first states the principle for the (discrete) random walks, but there's no problem in understanding it directly for the continuous-time processes.

Proposition (5.1) The *reflected process*, $\tilde{B}_t$ of the standard BM, B_t defined as

$$\tilde{B}_t = \begin{cases} B_t & \text{if} \quad t < \tau \\ B_\tau - (B_t - B_\tau) & \text{if} \quad t \ge \tau \end{cases}$$

for any stopping time τ (usually the hitting-time at certain level) is also a Brownian motion.

The defined process B_t^* is flipping the original path of B_t for the portion after the stopping time $t > \tau$. The meaning of principal should be very intuitive even without the mathematical proof. Note that, conditional on $t = \tau$, the two subsequent paths, i.e., the original path B_t and the reflected path $\tilde{B}_t$ are equally probably due to the symmetry and the independence property of BM.

In the same way, we can also reflect the portion before the stopping time τ,

$$\tilde{B}_t = \begin{cases} B_\tau - (B_t - B_\tau) & \text{if} \quad t < \tau \\ B_t & \text{if} \quad t \geq \tau \end{cases}.$$

The definition makes sense only when we know the value of B_τ. In the case of the single barrier hitting-time, we are lucky to have $B_\tau = a$. So the reflected part of the path simply becomes $B_t^* = 2a - B_t$, which is a BM starting from $2a$.

From the reflection principle, we have the following equality of three probabilities, for $x, y \geq 0$,

$$P(B_t^* > x,\ B_t < x - y) = P(B_t^* > x,\ B_t > x + y) = P(B_t > x + y).$$

The first equality is directly from the reflection principle. The stopping time τ used here is the first hitting-time of the level x. The condition $S_t^* > x$ means that the path hit the level x before time t, so if $S_t > x + y$, the reflected path should satisfy $\tilde{S}_t < x - y$. The second equality is trivially due to the continuity of BM. Now we are ready to derive various probability densities.

Joint Distribution of B_t and B_t^*

The joint probability is given as

$$\begin{aligned} P(B_t^* < x,\ B_t < x - y) &= P(B_t < x - y) - P(B_t^* \geq x,\ B_t < x - y) \\ &= P(B_t < x - y) - P(B_t > x + y) \\ &= \Phi((x - y)/\sqrt{t}) + \Phi((x + y)/\sqrt{t}) - 1. \end{aligned}$$

Under the change of variables, $v = x,\ u = x - y$, we have the final result on the joint density,

$$\begin{aligned} \text{CDF:} \quad P(B_t^* < v,\ B_t < u) &= \Phi(u/\sqrt{t}) + \Phi((2v - u)/\sqrt{t}) - 1 \\ &= \Phi(u/\sqrt{t}) - \Phi((u - 2v)/\sqrt{t}) \\ \text{PDF:} \quad f_{(B_t^*, B_t)}(v, u) &= \frac{2(2v - u)}{t^{3/2}} \phi((2v - u)/\sqrt{t}) \end{aligned}$$

Density and Distribution of B_t^*

When $y = 0$, we have the cumulative distribution function,

$$\begin{aligned} P(B_t^* > x) = P(\tau_x < t) &= P(S_t^* > x,\ S_t > x) + P(S_t^* > x,\ S_t \le x) \\ &= P(S_t > x) + P(S_t^* > x,\ S_t \ge x) \\ &= 2P(S_t > x) = P(|S_t| > x) \\ &= 2 - 2\Phi(x/\sqrt{t}). \end{aligned}$$

Equivalently, we have the complementary value,

$$P(B_t^* < x) = P(\tau_x > t) = 2\Phi(x/\sqrt{t}) - 1.$$

The differentiation w.r.t. x gives the density on x,

$$f_{B_t^*}(x) = \frac{2}{\sqrt{t}}\phi\left(\frac{x}{\sqrt{t}}\right) = \sqrt{\frac{2}{\pi t}}\, e^{-x^2/2t} \quad \text{for} \quad x \ge 0$$

Density of the Hitting Time τ_x

The differentiation w.r.t. t gives the density on τ_x,

$$f_{\tau_x}(t) = \frac{x}{t^{3/2}}\phi\left(\frac{x}{\sqrt{t}}\right) \quad \text{for} \quad x \ge 0,\ t \ge 0.$$

Note we saw the result in (4.27) from the inverse Laplace transform!

5.4 The Invariance Principle and Donsker's Theorem

It is worth to mention Donsker's Theorem (Wikipedia), which connects (discrete) random walk and (continuous) Brownian motion.

If $\{X_n\}$ be a sequence of i.i.d. random variables with mean 0 and variance 1 (more general than $X_n = \pm 1$), we have the discrete-time random walk,

$$S_n = \sum_{k=1}^{n} X_n \quad (S_0 = 0).$$

From the central limit theorem (CLT) (Wikipedia), we know that $S_n/\sqrt{n}$ converges to $\Phi(0, 1)$ as $n \to \infty$. Donsker's theorem is basically an extension of the CLT on the whole process of S_n.

Let us extend S_n into a continuous time process by interpolating the points at $t = n$ by

$$S_t^{(n)} = S_n + (t - n)X_{n+1} \quad \text{for} \quad n \le t < n + 1$$

and define a scaled process,

$$B_t^{(n)} = S_{nt}^{(n)}/\sqrt{n}.$$

Donsker's theorem states that the process $B_t^{(n)}$ converges to B_t as $n \to \infty$. The CLT is obviously the special case of Donsker's theorem at $t = 1$,

$$B_1^{(n)} = S_n^{(n)}/\sqrt{n} = S_n/\sqrt{n} \to B_1 \quad \text{as} \quad n \to \infty.$$

It was not a coincidence that we saw same results between random walks and BM for many occasions such as the ruin probability and the expectation for the stopping time.

5.5 Random Walks Inside Brownian Motion

We skip this section.

5.6 Exercises

Chapter 6

Itô Integration

Now that we have standard Brownian motion B_t as an important building block of stochastic processes and knows a few properties, we further move on to other topics; integration related to Brownian motion. To begin with, we want to investigate the integration,

$$I(f) = \int_0^T f(w,t)\, dB_t,$$

where w stands for the whole path of B_t for $0 \le t \le T$. Because this is the continuous-time version of the Martingale transform of B_t by $f(w,t)$, we can already guess that $I(f)$ is also a martingale (under some regularity condition).

The motivation for studying the integration of stochastic processes are natural. It is easy and intuitive to define a new stochastic process via the small increment (differentiation) of another process. A good example is the geometric Brownian Motion from which Black-Scholes formula is derived,

$$dS_t = S_t\,(rdt + \sigma\, dB_t).$$

The integration $\int_0^T dS_t = S_T - S_0$ will lead us to the distribution of the final stock price S_T which we are mostly interested in.

The most important property of Brownian motion regarding a very small time increment is is that

$$(B_t - B_s)^2 \to (t-s) \quad \text{as} \quad t-s \to 0.$$

This not only holds in the sense of expectation, but also holds with probability 1. With loss of generality, we can assume that $B_s = 0$ due to the independence increments of Brownian motion. An intuitive proof is that the distribution $B_t^2 - t$ has mean 0 and variance (or stdev) approaching to 0 as t goes to zero;

$$E(B_t^2 - t) = 0$$
$$\mathrm{Var}(B_t^2 - t) = E(B_t^4) - 2tE(B_t^2) + t^2 = 3t^2 - 2t^2 + t^2 = 2t^2.$$

The result is the well known as the formula, $(dB_t)^2 = dt$, and this is essence of Itô's lemma.

6.1 Definition of Itô's Integral: First Two Steps

In most part of this chapter, the author elaborates the definition of the integral, $I(f)$. We will simply trust our intuition from the non-stochastic calculus,

$$I(f) = \int_0^T f(w,t)\,dB_t = \lim_{N\to\infty} \sum_{k=0}^{N} f(B_0, \cdots, B_{t_k}, t = t_k)(B(t_{k+1}) - B(t_k)),$$

where $\{t_k\}$ are the breaking points of the integration interval $[0, T]$ $(0 = t_0 < \cdots < t_k < \cdots < t_N = T)$.

One trivial example of integration would be the case $f(w,t) = 1$:

$$I(f) = \int_a^b dB_t = B_b - B_a$$

Lemma 6.1 Itô's Isometry (Wikipedia):

$$E\left[\left(\int_0^t f(w,s)dB_s\right)^2\right] = E\left[\int_0^t f^2(w,s)ds\right]$$

The key idea of the proof is from the incremental Independence of Brownian motion. Let's that we divide the interval, $[0,t]$, with $\{t_k\}$ and let the value of $f(w,s)$ up to the time t_k as f_k. Then the computation goes like

$$\begin{aligned}
E\left[\left(\int_0^t f(w,s)dB_s\right)^2\right] &= E\left[\left(\sum_{k=0}^{N-1} f_k \Delta B_{t_k}\right)^2\right] \\
&= E\left[\sum_{k=0}^{N-1} f_k^2 \Delta B_{t_k}^2\right] + E\left[\sum_{k\neq j} f_k f_j \Delta B_{t_k} \Delta B_{t_j}\right] \\
&= E\left[\sum_{k=0}^{N-1} f_k^2 \Delta t_k\right] + 0 \\
&= E\left[\int_0^t f^2(w,s)ds\right].
\end{aligned}$$

6.2 Third Step: Itô's Integral as a Process

6.3 The Integral Sign: Benefits and Costs

6.4 An Explicit Calculation

As a non-trivial integration example, we show that

$$X_t = \int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{t}{2}$$

Note the extra term $-t/2$ appears in the last, which would not be present in the regular calculus. But we are already familiar with the term from the previous martingale theory where we saw that $B_t^2 - t$ is a martingale. The term B_t^2 and t always goes together. Note that the integration formula above is unusually nice because the integration results only depends on the final value of the Brownian motion, B_t, not the whole path history.

As necessary conditions, we first show the first two moments of the two sides match. The first moments are same as zero because both X_t and $B_t^2 - t$ are martingales. We use the Itô's isometry to compute the second moment of LHS

$$\mathrm{Var}(X_t) = E\left[\int_0^t B_s^2 ds\right] = \int_0^t E(B_s^2) ds = \frac{t^2}{2}$$

and the variance of RHS also yields to the same value

$$\mathrm{Var}(\frac{1}{2}B_t^2 - \frac{t}{2}) = \frac{2t^2}{4} = \frac{t^2}{2}.$$

Using the following two properties,

$$B(t_k)\left(B(t_{k+1}) - B(t_k)\right) = \frac{1}{2}\left(B^2(t_{k+1}) - B^2(t_k)\right) - \frac{1}{2}\left(B(t_{k+1}) - B(t_k)\right)^2$$
$$\left(B(t) - B(s)\right)^2 = t - s \quad \text{as} \quad t \to s \quad \text{or} \quad \left(dB(t)\right)^2 = dt.$$

we prove the integral identity as

$$\begin{aligned} X_t = \int_0^t B_s dB_s &\approx \sum_{k=0}^{N-1} B(t_k)\left(B(t_{k+1}) - B(t_k)\right) \\ &= \sum_{k=0}^{N-1} \frac{1}{2}\left(B^2(t_{k+1}) - B^2(t_k)\right) - \sum_{k=0}^{N-1} \frac{1}{2}\left(B(t_{k+1}) - B(t_k)\right)^2 \\ &\approx \frac{1}{2}\left(B^2(t) - B^2(0)\right) - \sum_{k=0}^{N-1}(t_{k+1} - t_k) \\ &= \frac{1}{2}B_t^2 - \frac{t}{2} \end{aligned}$$

6.5 Pathwise Interpretation of Itô's Integrals

Chapter 7

Localization and Itô's integral

7.1 Itô's Integral on $\mathcal{L}^2_{\mathbf{LOC}}$

7.2 An Intuitive Representation

Gaussian Connections

Proposition 7.6 (Gaussian Integrals) If $f(s)$ is a positive continuous function, the process defined as

$$X_t = \int_0^t f(s)dB_s$$

is a mean zero Gaussian process with independent increments and with co-variance function

$$\mathrm{Cov}(X_s, X_t) = \int_0^{s\wedge t} f^2(u)du \quad \left(\mathrm{Var}(X_t) = \int_0^t f^2(u)du\right)$$

Time Change to Brownian Motion: Simplest Case

From the previous proposition, we find an interesting observation that the process X_t becomes a standard Brownian motion when the time variable t is stretched according to the increasing variance.

Corollary 7.1 If $f(s)$ is a positive continuous function such that $\int_0^t f^2(s)ds \to \infty$ as $t \to \infty$, the process

$$Y_t = \int_0^\tau f(s)dB_s \quad \text{where} \quad t = \int_0^\tau f^2(s)ds$$

is a standard Brownian motion under the new time scale t of the *running variance*.

In short form, we can define

$$Y_{t=\int_0^\tau f^2(s)ds} = \int_0^\tau f(s)dB_s$$

Please pay attention to the following two precautions regarding the result. First, although $Y_{t'}$ is a standard BM, the process is not necessarily same as B_t for the corresponding t. Second, if $f(s) = c$ for some constant c, notice the subtle difference between this result and the self-similarly (or scaled process) of BM in Proposition 3.2.

The time change of Brownian motion imply a few interesting things on the volatility.

Annual vs daily volatility: the annualized volatility is used for the pricing formula because $t = 1$ is one year in the formula. However, the daily volatility is more intuitive. Assuming there are 256 trading days in one year (excluding weekends and holidays)

$$\sigma_d = \sigma_y/\sqrt{256} = \sigma_y/16$$

For example, the annualized volatility, $\sigma_y = 1\%$, imply that the standard deviation of the interest rate change in one year is 1% or the interest rate can go up or down by 1% in terms of random walk. Converting to daily volatility, we get $\sigma_d = 1\%/16 = 0.0625\% = 6.25\,\text{bp}$, where 1 bp (basis point, read as *bips* is 0.01%. Daily volatility gives more practical idea about how much the underlying asset should moves daily in order to give the option price in the market.

Boot-strapping of volatility curve: In real financial market, we rarely see a constant volatility on the same underlying asset. The prices of the options with different strike prices (same expiry) imply different volatilities. This is referred as volatility skew of volatility smile and we will study a stochastic process (stochastic volatility model) to explain this. The prices of the ATM options with different expiries also imply different volatilities. Let's assume that we can observe the market prices of the ATM options expiring at $t_1, t_2, \cdots, t_n$ and we obtain the corresponding volatilities $\sigma_1, \sigma_2, \cdots, \sigma_n$ from the ATM implied volatility formula $\sigma = C/0.4\sqrt{t}$. Often, we need to *interpolate* the volatility a maturity in-between $\{t_k\}$. One common method is to linear-interpolate the variance points,

$$V(0) = 0, \quad V(t_1) = \sigma_1^2 t_1, \quad \cdots \quad , V(t_n) = \sigma_n^2 t_n$$

For expiry $t_k < t < t_{k+1}$, the interpolated variance is the linear interpolation between t_n and t_{n+1},

$$V(t) = \frac{t_{k+1} - t}{t_{k+1} - t_k}\sigma_k^2 t_k + \frac{t - t_k}{t_{k+1} - t_k}\sigma_{k+1}^2 t_{k+1}$$

and the corresponding volatility is $\sigma = \sqrt{V(t)/t}$.

The underlying Brownian is $dS_t = f(t)dB_t$ where the volatility function $f(t)$ is a piece-wise constant function,

$$\sigma(0 < t < t_1) = \sigma_1$$
$$\sigma(t_1 < t < t_2) = \sqrt{(\sigma_2^2 t_2 - \sigma_1^2 t_1)/(t_2 - t_1)}$$
$$\cdots$$
$$\sigma(t_k < t < t_{k+1}) = \sqrt{(\sigma_{k+1}^2 t_{k+1} - \sigma_k^2 t_k)/(t_{k+1} - t_k)}$$

For such a process, $dS_t = f(t)dB_t$ in Corollary 7.1, the implied volatility of the option at time t will be $\sigma(t) = \sqrt{(\int_0^t f^2(s)ds)/t}$. Therefore the implied volatility (i.e. the volatility plugged into the formula) is sometimes called *average volatility* and the volatility in the stochastic process is called *instantaneous volatility*.

7.3 Why Just $\mathcal{L}^2_{\mathrm{LOC}}$?

7.4 Local Martingales and Honest Ones

7.5 Alternative Fields and Changes of Time

7.6 Exercises

Chapter 8

Itô's Formula

Theorem 8.1 (Itô's formula - Simplest Case) If $f(x)$ is a function which has a continuous second derivative,

$$f(B_t) = f(0) + \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds$$

or, in stochastic differential equation (SDE) form,

$$df(B_t) = \frac{1}{2}f''(B_t)dt + f'(B_t)dB_t.$$

The value $f(B_t) - f(0)$ is broken down in to two terms:

- $\int_0^t f'(B_s)dB_s$: a martingale with zero mean. This term contains the local variability of $f(B_t)$, *noise* or *risk (uncertainty)*.

- $\frac{1}{2}\int_0^t f''(B_s)ds$: This term contains the drift of $f(B_t)$, *signal* or *return*.

Table 8.2. Box algebra multiplication table

$\cdot$	dt	dB_t
dt	0	0
dB_t	0	dt

8.1 Analysis and Synthesis

The idea behind Theorem 8.1 is from the Taylor expansion of $f(x)$:

$$df(x) = f'(x)dx + \frac{1}{2}f''(x)(dx)^2 + O((dx)^3)$$

Plugging $x = B_t$, we obtain the SDE form:

$$\begin{aligned} df(B_t) &= f'(B_t)dB_t + \frac{1}{2}f''(B_t)(dB_t)^2 + O\big((dB_t)^3\big) \\ &= f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt + O\big((dt)^{3/2}\big) \end{aligned}$$

8.2 First Consequences and Enhancements

In Section 6.4 we proved

$$X_t = \int_0^t B_t \, dB_t = \frac{1}{2}B_t^2 - \frac{1}{2}t.$$

This is easily proved as a case of $f(x) = t^2/2$ from Theorem 8.1. From Theorem 8.2 we can verify this by the differentiation rule:

$$dX_t = d\Big(\frac{1}{2}B_t^2 - \frac{1}{2}t\Big) = B_t dB_t - \frac{1}{2}dt + \frac{1}{2}(dB_t)^2 = B_t dB_t$$

Beyond Space to Space and Time

Theorem 8.2 Itô's formula with Space and Time Variables: For a differentiable function $f(t,x)$, we have the representation

$$f(t,B_t) = f(0,0) + \int_0^t f_x(s,B_s)dB_s + \int_0^t f_t(s,B_s)ds + \frac{1}{2}\int_0^t f_{xx}(s,B_s)ds$$

or

$$df(t,B_t) = \left(f_t(t,B_t) + \frac{1}{2}f_{xx}(t,B_t)\right)dt + f_x(t,B_t)dB_t,$$

where f_x, f_t and f_{xx} are the partial derivatives

$$f_x = \frac{\partial f}{\partial x}, \ f_t = \frac{\partial f}{\partial t} \text{ and } f_{xx} = \frac{\partial^2 f}{\partial x^2}.$$

Again the theorem is based on the Taylor expansion of $f(t,x)$

$$df(t,x) = f_x(t,x)dx + f_t(t,x)dt + \frac{1}{2}f_{xx}(t,x)(dx)^2 + f_{tx}(t,x)\,dt\,dx + \frac{1}{2}f_{tt}(t,x)(dt)^2 + \cdots$$

Plugging $x = B_t$, we obtain the SDE form

$$\begin{aligned} df(t,B_t) &= f_x(t,B_t)dB_t + f_t(t,B_t)dt + \frac{1}{2}f_{xx}(t,B_t)(dB_t)^2 + O\big((dt)^{3/2}\big) \\ &= f_x(t,B_t)dB_t + \left(f_t(t,B_t) + \frac{1}{2}f_{xx}(t,B_t)\right)dt + O\big((dt)^{3/2}\big) \end{aligned}$$

Table 8.2-2. Box algebra multiplication table (modified) from $dB_t \sim (dt)^{1/2}$

$\cdot$	dt	dB_t
dt	$(dt)^2$	$(dt)^{3/2}$
dB_t	$(dt)^{3/2}$	dt

Martingale and Calculus

Proposition 8.1 (Martingale PDE condition). If

$$f_t(t, B_t) + \frac{1}{2} f_{xx}(t, B_t) = 0,$$

then $X_t = f(t, B_t)$ is a local martingale. The SDE of X_t has zero drift term,

$$df(t, B_t) = f_x(t, B_t) dB_t$$

First Examples

We already know that

$$M_t = \exp(-\frac{1}{2}\sigma^2 t + \sigma B_t)$$

is a martingale. It can be reaffirmed using the previous proposition. We can write $M_t = f(t, B_t)$ where

$$f(t, x) = \exp(-\frac{1}{2}\sigma^2 t + \sigma x)$$

and the function $f(t, x)$ has properties such that $f_t = -(1/2)\sigma^2 f$, $f_x = \sigma f$ and $f_{xx} = \sigma^2 f$. Therefore, $f_t = -\frac{1}{2} f_{xx}$ and

$$dM_t = \sigma M_t \, dB_t \quad \text{or} \quad \frac{dM_t}{M_t} = \sigma dB_t.$$

Brownian Motion with Drift: The Ruin Problem

Assume that the gambler's wealth follows a drifted BM

$$X_t = \mu t + \sigma B_t.$$

Then the ruin problem for X_t of calculating $P(X_\tau = A)$ can be simplified by finding a function $h(\cdot)$ satisfying $M_t = h(X_t)$, $h(A) = 1$ and $h(-B) = 0$. By the martingale property, the probability can be computed as

$$h(0) = E(M_0) = E(M_\tau) = P(X_\tau = A)h(A) + P(X_\tau = -B)h(-B) = P(X_\tau = A).$$

Let $f(t, x)$ be a function with separated arguments which satisfy $f(t, B_t) = h(X_t) = h(\mu t + \sigma B_t)$ or $f(t, x) = h(\mu t + \sigma x)$. For $h(X_t)$ to be a martingale, $f(t, x)$ have to satisfy

$$f_t + \frac{1}{2} f_{xx} = \mu h'' + \frac{1}{2}\sigma^2 h' = 0$$

and we find that

$$h(x) = \frac{e^{2\mu B/\sigma^2} - e^{-2\mu x/\sigma^2}}{e^{2\mu B/\sigma^2} - e^{-2\mu A/\sigma^2}}.$$

Finally we state the following Theorem:

Proposition 8.2 (Ruin Probabilitly for Brownian Motion with Drift) If $X_t = \mu t + \sigma B_t$ and τ is the first time X_t hitting A or $-B$, then we have

$$P(X_\tau = A) = \frac{1 - e^{-2\mu B/\sigma^2}}{1 - e^{-2\mu(A+B)/\sigma^2}},$$

which is in a similar form to the result of Section 1.3.

Looking Back – and Down the Street

Exponential Distribution of the Supremum

* Mention the connection to the previous exam problem

Shorthand Notation

8.3 Vector Extension and Harmonic Functions

We skip this section.

8.4 Functions of Processes

Box Calculus And Functions of Geometric Brownian Motion

The process $X_t = \exp(\alpha t + \sigma B_t)$ is known as geometric Brownian motion (GBM) and this is the underlying process of the Black-Scholes-Merton model. Let us drive the SDE of GBM:

$$X_t = X_0 \exp(\alpha t + \sigma B_t), \quad f(t,x) = X_0 \exp(\alpha t + \sigma x)$$

Using that $f_t(t,x) = \alpha f(t,x)$ and $f_x(t,x) = \sigma f(t,x)$, we can drive the differential form of X_t:

$$dX_t = \alpha X_t dt + \sigma X_t dB_t + \frac{1}{2}\sigma^2 X_t (dB_t)^2 = (\alpha + \frac{1}{2}\sigma^2) X_t \, dt + \sigma X_t dB_t$$

or

$$\frac{dX_t}{X_t} = (\alpha + \frac{1}{2}\sigma^2)\, dt + \sigma dB_t$$

8.5 The General Itô's Formula

Theorem 8.4 (Itô's Formula for Standard Processes) For a function $f(t,x)$ and a stochastic process X_t given by

$$dX_t = a(w,t)\, dt + b(w,t)\, dB_t,$$

we have

$$df(t, X_t) = f_t(t, X_t)\,dt + f_x(t, X_t)\,dX_t + \frac{1}{2} f_{xx}(t, X_t)\,dX_t \cdot dX_t$$
$$= \left(f_t(t, X_t) + \frac{1}{2} f_{xx}(t, X_t)\, b^2(w, t) \right) dt + f_x(t, X_t)\,dX_t.$$

The theorem can be proved by carefully applying the chain rule. The key is that the only surviving term of $dX_t \cdot dX_t$ is $b^2(w,t)dB_t \cdot dB_t = b^2(w,t)\,dt$. Notice that $a(w,t)$ is not appearing in the final formula. This General Itô's formula, simply referred to as Itô's formula, is one of the most important in this course!

Now with the help of Theorem 8.4, we can solve GBM from its SDE (the opposite direction)

$$\frac{dX_t}{X_t} = (\alpha + \frac{1}{2}\sigma^2)\,dt + \sigma dB_t.$$

Under the traditional calculus, we know that $\int dx/x = \log x$, so we use $\log x$ as a starting point of our guess. Now we apply Itô's lemma to $\log X_t$,

$$d(\log X_t) = \frac{dX_t}{X_t} - \frac{1}{2}\frac{(dX_t)^2}{X_t^2} = (\alpha + \frac{1}{2}\sigma^2)\,dt + \sigma dB_t - \frac{1}{2}\sigma^2 dt = \alpha dt + \sigma dB_t$$

The RHS of this equation is easily integrable, so we get

$$\log(X_t) - \log(X_0) = \alpha t + \sigma B_t$$
$$X_t = X_0 \exp(\alpha t + \sigma B_t)$$

The GBM is better known in the form of

$$\frac{dX_t}{X_t} = \mu dt + \sigma dB_t$$
$$X_t = X_0 \exp\left((\mu - \frac{1}{2}\sigma^2)t + \sigma B_t\right).$$

As we already know X_t is a martingale only if $\mu = 0$.

Here are two more examples of stochastic integration which can be solved exactly.

Example 1

$$dX_t = -\frac{X_t}{1+t}dt + \frac{\sigma}{1+t}dB_t$$
$$(1+t)dX_t = -X_t\,dt + \sigma dB_t$$

So we start with

$$d\Big((1+t)X_t\Big) = (1+t)\,dX_t + X_t dt + 0\,(dX_t)^2 = \sigma dB_t$$

So, finally we get

$$X_t = \frac{\sigma B_t + c}{1+t}$$

Example 2

$$dX_t = \sigma^2 X_t^3 dt + \sigma X_t^2 dB_t$$

$$\frac{dX_t}{X_t^2} = \sigma^2 X_t \, dt + \sigma dB_t$$

From $-1/x = \int dx/x^2$,

$$-d\Big(\frac{1}{X_t}\Big) = \frac{dX_t}{X_t^2} - \frac{(dX_t)^2}{X_t^3} = \sigma^2 X_t dt + \sigma dB_t - \sigma^2 X_t (dB_t)^2 = \sigma dB_t$$

So, finally we get

$$X_t = \frac{X_0}{1 - X_0 \sigma B_t}$$

8.6 Quadratic Variation

8.7 Exercises

Chapter 9

Stochastic Differential Equations

All stochastic processes that we will meet in finance is in the form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t,$$

where $\mu(t, X_t)$ and $\sigma(t, X_t)$ are short-term (or instantaneous) growth (drift) and variability (volatility) of the underlying asset respectively. It is easy to model a underlying asset using a SDE form. However, not all SDEs are analytically solvable.

9.1 Matching Itô's Coefficients

We already saw that the SDE for GBM,

$$dX_t = \mu X_t dt + \sigma X_t dB_t$$

has the analytic solution,

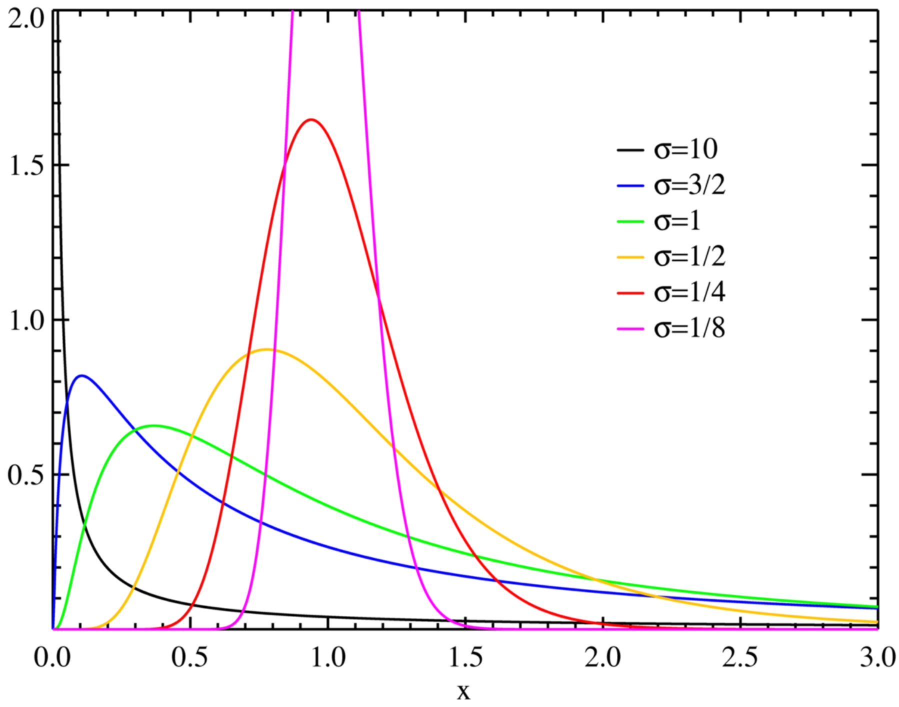
$$X_t = X_0 \exp\Big((\mu - \frac{1}{2}\sigma^2)t + \sigma B_t\Big).$$

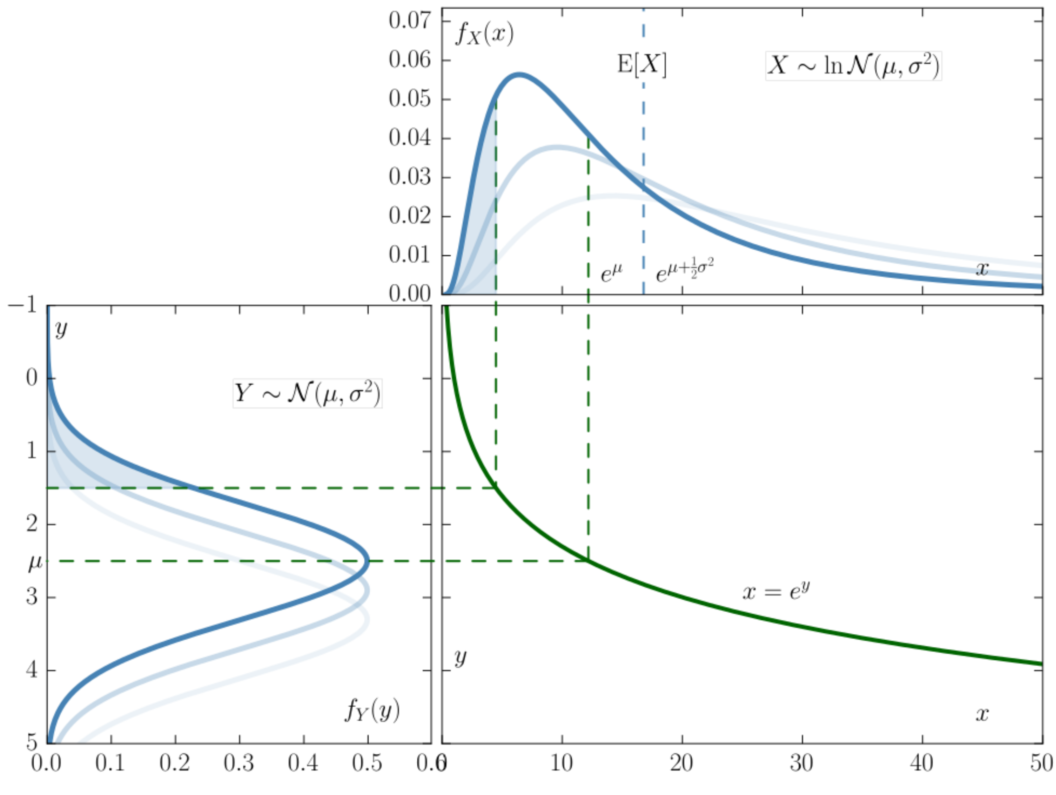
The author is using a method of *matching coefficients*, but our method of *guessing from the traditional calculus* is perhaps better.

The author points out an interesting fact that, if $\sigma^2/2 > \mu > 0$,

$$\text{Prob}\Big(X_t < \varepsilon\Big) \to 1 \quad \text{as} \quad t \to \infty$$

for very small ε. This is a property of lognormal distribution.





9.2 Ornstein-Uhlenbeck Processes

Another example of analytically solvable stochastic process, yet very important in application, is OU process (WIKIPEDIA),

$$dX_t = -\alpha X_t dt + \sigma dB_t \quad \text{for } \alpha, \sigma > 0$$

The drift term $-\alpha X_t\,dt$ is always trying to send X_t back to 0 while the BM term σdB_t. So the process if often called mean-reverting process. The process is sometimes written as

$$dX_t = \alpha(X_\infty - X_t)dt + \sigma dB_t$$

where the constant term, X_∞, is the long-term average or the equilibrium value of X_t. This is same as the original SDE via the change of variable, $X_t \leftarrow X_t - X_\infty$. The OU process is the underlying model for the Vasicek model (WIKIPEDIA) which is a popular interest rate model.

9.3 Matching Product Process Coefficients

Solving the OU SDE

Instead of *matching product* method, we continue use our *guessing from the traditional calculus* method. For this we need to know a bit on ordinary differential equation (ODE). Ignoring the last BM term, we can solve

$$\begin{aligned}
dx &= -\alpha x dt \\
\frac{dx}{x} &= -\alpha dt \\
\log(x) &= -\alpha t \\
e^{\alpha t}x &= x_0
\end{aligned}$$

Finally we have $d\Big(e^{\alpha t}x\Big) = 0$, so $e^{\alpha t}X_t$ is our initial guess. The stochastic differentiation of our guess goes

$$\begin{aligned}
d\Big(e^{\alpha t}X_t\Big) &= \alpha e^{\alpha t}X_t dt + e^{\alpha t}dX_t + \frac{1}{2}0(dX_t)^2 \\
&= \alpha e^{\alpha t}X_t dt - \alpha e^{\alpha t}X_t dt + \sigma e^{\alpha t}dB_t = \sigma e^{\alpha t}dB_t.
\end{aligned}$$

Finally we have a solution in an integration form,

$$e^{\alpha t}X_t = X_0 + \sigma\int_0^t e^{\alpha s}dB_s.$$

Note that this is one of the time change of BM where the volatility term $\sigma e^{\alpha t}$ is a function of time only. The variance of the BM term is

$$\sigma^2\int_0^t e^{2\alpha s}ds = \frac{\sigma^2}{2\alpha}(e^{2\alpha t}-1).$$

The mean and variance of X_t is given as

$$E(X_t) = e^{-\alpha t} X_0, \quad \mathrm{Var}(X_t) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}).$$

So the process X_t converges to $N(0, \sigma^2/2\alpha)$ in a long run.

Using the time change, the solution can be written as

$$X_t = e^{-\alpha t} X_0 + \frac{\sigma e^{-\alpha t}}{\sqrt{2\alpha}} B'_{e^{2\alpha t}-1} \quad \text{for a standard BM, } B'_t$$

A simulation from time s to t ($s < t$) can be done using

$$X_t = e^{-\alpha(t-s)} X_s + \frac{\sigma e^{-\alpha(t-s)}}{\sqrt{2\alpha}} \Big(B'_{e^{2\alpha t}-1} - B'_{e^{2\alpha s}-1} \Big) \quad \text{for} \quad s < t,$$

where

$$B'_{e^{2\alpha t}-1} - B'_{e^{2\alpha s}-1} \sim N(0, e^{2\alpha t} - e^{2\alpha s}).$$

Solving the Brownian Bridge SDE

The Brownian bridge is a Brownian motion which return to 0 at $t = 1$. It can be represented as

$$X_t = B_t - t\, B_1 \quad \text{for} \quad 0 \le t \le 1.$$

The author show the following SDE describes a Browian bridge.

$$dX_t = -\frac{X_t}{1-t} dt + dB_t.$$

The intuition is that, as $t \to 0$ the drift term becomes negatively large, thus bring X_t back to zero. We start out guess from $X_t/(1-t)$.

$$d\Big(\frac{X_t}{1-t}\Big) = \frac{dX_t}{1-t} + \frac{X_t\, dt}{(1-t)^2} = \frac{dB_t}{1-t}$$

$$\frac{X_t}{1-t} = \int_0^t \frac{dB_s}{1-s} \quad \Rightarrow \quad X_t = (1-t) \int_0^t \frac{dB_s}{1-s}$$

Although this is not a complete analytic form, we can derive a few property from it. For example, we can show that $\mathrm{Cov}(X_s, X_t) = s(1-t)$ (see detail in **SCFA**) and we conclude that X_t is indeed a Brownian bridge.

Looking Back – Finding a Paradox

We skip this section.

9.4 Existence and Uniqueness Theorems

We skip this section.

9.5 Systems of SDEs

Instead of the examples in **SCFA**, we show two popular stochastic volatility models as examples. The under the stochastic volatility models, the volatility σ follows another stochastic process rather than staying constant. The stochastic volatility model are one mothod to model the volatility smile observed in the market.

Example: Stochastic Volatility Models

Heston Model (WIKIPEDIA)

$$dF_t = \sqrt{V_t}\, dB_t^1$$
$$dV_t = \kappa(V_\infty - V_t)dt + \alpha\sqrt{V_t}dB_t^2$$
$$dB_t^1\, dB_t^2 = \rho\, dt$$

Stochastic-Alpha-Beta-Rho (SABR) Model (WIKIPEDIA)

$$dF_t = \sigma_t F_t^\beta dB_t^1$$
$$d\sigma_t = \alpha\sigma_t dB_t^2$$
$$dB_t^1\, dB_t^2 = \rho\, dt$$

9.6 Exercises

Please take a look at Exercises 9.1, 9.2 and 9.3 although they are not homeworks.

Chapter 10

Arbitrage and SDEs

In this chapter, we derive the Black-Scholes PDE. We also drive the Black-Scholes formula from the GBM, rather than the PDE. The original derivation from PDE is discussed in Chapter 11, however this is beyond the scope of this course. It is strongly recommended to read this whole chapter as this is the most relevant chapter to finance.

10.1 Replication and Three Examples of Arbitrage

Forward Contracts

Put-Call Parity

Present-value version:

$$C_0 - P_0 = S_0 - e^{-rT}K$$

Forward-value version:

$$C_F - P_F = F - K$$

The Binomial Arbitrage

Under a simple economy with two time step $t = 0$ and 1, there are two financial instruments: a stock and a bond. The stock price $S = \$2$ at $t = 0$ can be either halved to \$1 or doubled to \$4 at $t = 1$. The bond price B is unchanged as \$1 from $t = 0$ to $t = 1$. Now consider a *derivative* which pays \$3 when the stock price is up or \$0 when the stock price is down. We can determine the price of this derivative using the replication argument. Assume that α share of stock and β unit of bond can replicate the payoff at $t = 1$:

$$4\alpha + \beta = 3 \quad \text{and} \quad \alpha + \beta = 1.$$

We solve the solution as $\alpha = 1$ and $\beta = -1$, which means that the derivative can be replicated by buying one share of stock and borrowing \$1 worth of bond. Because the cost of the replication portfolio is \$1, we can also conclude that the price of the derivative at $t = 0$ should be \$1. Here, note that we did not use any information about the probability on the stock price change. Whether the stock has 90% change of going up or 10% of going up, it does not affect the price of the derivative because we obtain the price using replication or no-arbitrage argument. This is the essence of the Black-Scholes option pricing.

10.2 The Black-Scholes Model (Wikipedia)

We assume that the underlying stock follows a GBM process and cash earns interest at rate r:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t, \quad \frac{d\beta_t}{\beta} = r dt.$$

Arbitrage and Replication

Assume that the value of an option is *dynamically* replicated with stock S_t and bond β_t:

$$V_t = a_t S_t + b_t \beta_t,$$

where a_t and b_t is the unit of stock and bond respectively. At the maturity $t = T$, the value V_T should be equal to the final payoff of option, $h(S_T)$. For example,

$$\begin{aligned} f(T, S_T) &= h(S_T) = (S_T - K)^+ \quad \text{for a call option with strike price } K,\\ f(T, S_T) &= h(S_T) = (K - S_T)^+ \quad \text{for a put option with strike price } K. \end{aligned}$$

One important condition that V_T should satisfy is *self-financing condition* (Wikipedia):

$$dV_t = a_t dS_t + b_t d\beta_t.$$

Notice that $V_t = a_t S_t + b_t \beta_t$ does not alway lead to $dV_t = a_t dS_t + b_t d\beta_t$. The condition means that there is no cash in and out of the portfolio when reblancing between stock and cash. For example, you can buy more stock (a_t increasing) by selling some bond (b_t decreasing). You cannot expect that a_t and b_t increase at the same time. Therefore the self-financing condition imposes strong restrictions on a_t and b_t.

Coefficient Matching

The author derives the Black-Scholes PDE from the two equations. From the self-financing condition,

$$dV_t = a_t dS_t + b_t d\beta_t = (a_t \mu S_t + b_t r \beta_t)dt + a_t \sigma S_t dB_t.$$

Now assuming that $V_t = f(t, S_t)$, the replication condition and Itô's calculus lead to

$$dV_t = \left(f_t(t, S_t) + \frac{1}{2} f_{xx}(t, S_t)\sigma^2 S_t^2 + f_x(t, S_t)\mu S_t \right) dt + f_x(t, S_t)\sigma S_t dB_t.$$

Now by matching the coefficients of dt and dB_t, we first obtain the share of stock is given by

$$a_t = f_x(t, S_t), \quad b_t = \frac{1}{r\beta_t} \left(f_t(t, S_t) + \frac{1}{2} f_{xx}(t, S_t)\sigma^2 S_t^2 \right)$$

and finally arrive at the famous Black-Scholes PDE

$$f_t(t, x) = -\frac{1}{2}\sigma^2 x^2 f_{xx}(t, x) - r x f_x(t, x) + r f(t, x).$$

with the terminal condition $f(T, x) = h(x)$.

10.3 The Black-Scholes Formula (Wikipedia)

Although the this derivation based on the probability distribution will be discussed in Chapter 11, we first drive the formula here in order to discuss the consequence.

$$S_T = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma B_T \right) = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}\, z \right) \quad \text{for} \quad z \sim N(0, 1).$$

We first let $z = -d_2$ be the root of the payoff $S_T - K = 0$,

$$S_0 \exp\left((r - \frac{1}{2}\sigma^2)T - \sigma\sqrt{T}\, d_2 \right) = K \quad \Rightarrow \quad d_2 = \frac{\log(S_0\, e^{rT}/K)}{\sigma\sqrt{T}} - \frac{1}{2}\sigma\sqrt{T}.$$

The the present value of the call options is integrated as

$$\begin{aligned}
C_0 &= e^{-rT} \int_{-\infty}^{\infty} (S_T - K)^+ dP(S_T) \\
&= e^{-rT} \int_{-d_2}^{\infty} \frac{dz}{\sqrt{2\pi}}\; S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}\, z - \frac{1}{2}z^2 \right) - K \exp(-\frac{1}{2}z^2) \\
&= e^{-rT} \int_{-d_2}^{\infty} \frac{dz}{\sqrt{2\pi}} S_0 \exp\left(rT - \frac{1}{2}(z - \sigma\sqrt{T})^2 \right) - K \exp(-\frac{1}{2}z^2) \\
&= S_0(1 - N(-d_2 - \sigma\sqrt{T})) - e^{-rT} K\, (1 - N(-d_2)) \\
&= S_0 N(d_2 + \sigma\sqrt{T}) - e^{-rT} K\, N(d_2) \\
&= S_0 N(d_1) - e^{-rT} K\, N(d_2),
\end{aligned}$$

$$\text{where} \quad d_1 = d_2 + \sigma\sqrt{T} \quad \text{or} \quad d_{1,2} = \frac{\log(S_0\, e^{rT}/K)}{\sigma\sqrt{T}} \pm \frac{1}{2}\sigma\sqrt{T}.$$

The formula above is referred to as the Black-Scholes formula (WIKIPEDIA) or Black-Scholes-Merton formula.

Using the forward price of the stock F instead, the option price can be also written as

$$C_0 = D\Big(F\,N(d_+) - K\,N(d_-)\Big) \quad \text{where} \quad d_\pm = \frac{\log(F/K)}{\sigma\sqrt{T}} \pm \frac{1}{2}\sigma\sqrt{T}.$$

Here F is the forward price of the stock and D is the discount factor, i.e., the value of \$1 at the maturity T. Using the interest rate r and the continuous dividend rate q, F and D can be written as

$$F = S_0\,e^{(r-q)T} \quad \text{and} \quad D = e^{-rT}.$$

The above formula is referred to as the Black-76 formula (WIKIPEDIA) or simply the Black formula. The put options are given as

$$P_0 = D\Big(K\,N(-d_-) - F\,N(-d_+)\Big) \quad \text{where} \quad d_\pm = \frac{\log(F/K)}{\sigma\sqrt{T}} \pm \frac{1}{2}\sigma\sqrt{T}.$$

There are a few observations:

ATM price The price of ATM option ($K = F$) under Black-Scholes model is not as simple as that under normal model. But they are fairly close, so you can depend on the normal model price, $C \approx 0.4\,\sigma_N\sqrt{T}$. Be careful not to use the same volatility σ. The approximate relation should be that $\sigma_N \approx \sigma_{BS}F$.

Dependce on Stdev Nevertheless, the call option depends on the standard deviation, $\sigma\sqrt{T}$, not on either σ or T separately (except at the discounting e^{-rT}). So please remember that $\sigma\sqrt{T}$ always go together as one package.

Digital option and Delta From the derivation, it is clear that $e^{-rT}\,N(d_2)$ is the digital call option price. It can be also shown that the delta w.r.t. the spot price S_0 is $N(d_1)$. Unlike in normal model, the digital call option price and the delta are slightly different although $N(d_1)$ and $N(d_2)$ are similar.

Asset-or-nothing option The first term $S_0N(d_1)$ can be understood as the present price of the asset-or-nothing option, in the same way in the normal model.

The maximum value of the call option is S_0 or $D\,F$. It is achieved when $N(d_1) = N(d_+) = 1$ and $N(d_2) = N(d_-) = 0$ which is achieved when $\sigma \to \infty$ for any K.

10.4 Two Original Derivations

The Original Hedged Portfolio Argument

In the original paper, Black and Scholes derived the PDE by constructing a hedged portfolio. As written in the **SCFA**, they were motivated from the industry practice of hedging an option with an appropriate amount of short position in stock. In their paper, the hedged portfolio is one unit of stock, dS_t, hedged with some short option position, $-df(t, S_t)/f_x(t, S_t)$ (read **SCFA** for detail). However, we instead construct one option position $df(t, S_t)$ hedged with some short stock position, $f_x(t, S_t)dS_t$ and drive the same PDE.

The change (i.e., SDE) in the option value $f(t, S_t)$ given as

$$\begin{aligned} df(t, S_t) &= f_x(t, S_t)dS_t + f_t(t, S_t)\,dt + \frac{1}{2}f_{xx}(t, S_t)(dS_t)^2 \\ &= f_x(t, S_t)dS_t + \Big(f_t(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 f_{xx}(t, S_t)\Big)dt. \end{aligned}$$

Now consider the hedged portfolio of a unit of the derivative and some portion of the stock. The unit of stock should be $-f_x(t, S_t)$. In that way the BM (or risky) component of the derivative will be hedged by the risk component $-f_x(t, S_t)dS_t$. Then, the drift of the hedged portfolio

$$df(t, S_t) - f_x(t, S_t)dS_t = \Big(f_t(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 f_{xx}(t, S_t)\Big)dt$$

should match to the compounded growth of the portfolio value from the interest rate r,

$$r\Big(f(t, S_t) - f_x(t, S_t)S_t\Big)dt.$$

So we obtain the same Black-Scholes PDE:

$$f_t(t, S_t) + rS_t f_x(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 f_{xx}(t, S_t) = rf(t, S_t).$$

The Original CAPM Argument

We skip this.

10.5 The Perplexing Power of a Formula

Please read this section.

10.6 Exercises

Chapter 11

The Diffusion Equation

We skip this chapter.

11.1 The diffusion of Mice

11.2 Solutions of the Diffusion Equation

11.3 Uniqueness of Solutions

11.4 How to Solve the Black-Scholes PDE

11.5 Exercises

Chapter 12

Representation Theorems

12.1 Stochastic Integral Representation Theorem

12.2 The Martingale Representation Theorem (WIKIPEDIA)

Theorem 12.3 (Martingale Representation Theorem) For a martingale X_t adapted to the standard BM filtration and a time $t = T$ such that $E(X_T^2) < \infty$, there is a stochastic process ϕ_s such that

$$X_t = X_0 + \int_0^t \phi_s \, dB_s \quad \text{for} \quad 0 \le t \le T.$$

A Special Case Is Already Known

We know that the GBM, $dX_t = \mu X_t dt + \sigma X_t dB_t$, is a martingale when $\mu = 0$ with the solution

$$X_t = X_0 \exp\Big(-\frac{1}{2}\sigma^2 t + \sigma B_t \Big).$$

Therefore, we have the martingale representation as

$$X_t = X_0 + \int_0^t \sigma X_t dB_s, \quad \phi_t = \sigma \exp\Big(-\frac{1}{2}\sigma^2 t + \sigma B_t \Big)$$

12.3 Continuity of Conditional Expectations

The martingale representation theorem can be understood in the context of a random variable X whose value is known at $t = T$, $X \in \mathcal{F}_T$. Without mathematical proof, we can derive a stochastic process of the expectation of X seen at time t:

$$X_t = E(X|\mathcal{F}_t) = E_t(X) \quad \text{with properties} \quad X_T = X, \quad X_0 = E(X).$$

Because X_t is a martingale almost by definition, the martingale representation theorem leads to

$$E(X|\mathcal{F}_t) = E(X) + \int_0^t \phi_s \, dB_s \quad \text{for} \quad 0 \le t \le T.$$

Although it seems quite complicated, the derivative price is a good example to understand X_t. The random variable X is understood as the final payoff which is known at $t = T$ when S_T is known, $X = h(S_T)$. Then the derived process X_t can be understood as the price of the derivative at time t because it is the expectation of the final payoff seen at time t (ignoring interest rate). Extending the martingale representation theorem, we can state that

$$X_T = X_0 + \int_0^T \phi_t \, dS_t.$$

This means that we can replicate the final payout X_T by starting with initial portfolio X_0 and continuously holding ϕ_t share of stock at time t and such replicating strategy ϕ_t exist! Of course, $X_0 = E(h(S_T))$ is the price of the derivative you receive as premium.

A toy example: we can imagine a derivative paying you B_T^2 at expiry $t = T$. We can directly compute X_t:

$$X_t = E(B_T^2|\mathcal{F}_t) = E_t((B_T - B_t)^2 + 2(B_T - B_t)B_t + B_t^2) = (T - t) + 0 + B_t^2.$$

Applying Itô's lemma, $dX_t = 2B_t dB_t + dt - dt$, and $X_0 = E(B_T^2) = T$, we obtain the representation

$$X_t - T = \int_0^t 2B_s dB_s$$

and we know this result very well.

12.4 Representation via Time Change

We skip the rest of chapters.

12.5 Lévy's Characterization of Brownian Motion

12.6 Bedrock Approximation Techniques

12.7 Exercises

Chapter 13

Girsanov Theory

13.1 Importance Sampling

Imagine that, for $Z \sim N(0,1)$, we need to evaluate the following rare event by Monte-Carlo (MC),

$$\text{Prob}(Z > 30) = \frac{1}{N}\sum_{1}^{N} 1_{Z_k > 30} \quad \text{for} \quad Z_k \sim N(0,1)$$

where 1_A is the indicator function for the event A. In fact we already know that the probability is $1 - N(30) = N(-30) \approx 5 \times 10^{-198}$ and it means we have such an event out of 10^{198} Gaussian random numbers, which is impossible to simulate in reality.

It happens that in finance such rare events (a.k.a. tail events) like these are what we are interested in, e.g., company default, abnormal price changes and financial crises.

Shift the Focus to Improve a Monte Carlo

Importance sampling (Wikipedia) is a MC technique to improve efficiency (and increase frequency) by simulating MC under a different probability model and associating it to the original problem.

For the example above, we will shift the mean of the original distribution Z to get $X = Z + \mu \sim N(\mu, 1)$. Although we will put $\mu = 30$ later, we first express the original problem in terms of the modified distribution X by the familiar trick,

$$\begin{aligned} E[f(Z)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(z) e^{-z^2/2}\, dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-(x-\mu)^2/2}\, e^{-\mu x + \mu^2/2}\, dx \\ &= E\big[f(X)\, e^{-\mu X + \mu^2/2}\big] = E\big[f(Z+\mu)\, e^{-\mu Z - \mu^2/2}\big]. \end{aligned}$$

With $\mu = 30$, the original MC problem can be evaluated

$$\begin{aligned}
\text{Prob}(Z > 30) &= \frac{1}{N}\sum_{k=1}^{N} 1_{X_k > 30} \cdot \exp(-30X_k + 450), \quad \text{for} \quad X_k \sim N(30, 1) \\
&= \frac{1}{N}\sum_{k=1}^{N} 1_{Z_k > 0} \cdot \exp(-30Z_k - 450), \quad \text{for} \quad Z_k \sim N(30, 1).
\end{aligned}$$

By shifting the mean from 0 to 30, we obtain more frequent event, $X_k > 30$.

13.2 Tilting a Process

Now let us apply the same trick between a standard BM B_t and a BM with drift $X_t = B_t + \mu t$. We express an expectation on B_t in terms of the probability on X_t. We know

$$\begin{aligned}
E[f(B_t)] &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} f(z)\, n\left(\frac{z}{\sqrt{t}}\right) dz \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} f(x)\, n\left(\frac{x - \mu t}{\sqrt{t}}\right) e^{-\mu(x-\mu t) - \mu^2 t/2}\, dx \\
&= E\big[f(X_t)\, e^{-\mu B_t - \mu^2 t/2}\big].
\end{aligned}$$

This result can be even true for more general expectation depending on the history of X_t,

$$E[f(B_{t_1}, B_{t_1}, \cdots, B_{t_n})].$$

The idea is that the expectation can be evaluated by the multi-variate probability density on the increments of the process, $(B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})$ with $t_0 = 0$ and $B_0 = 0$. Considering only the exponent of the joint PDF,

$$\begin{aligned}
\text{exponent} &= -\frac{1}{2}\sum_{k=1}^{n} \frac{(z_k - z_{k-1})^2}{t_k - t_{k-1}} \\
&= -\frac{1}{2}\sum_{k=1}^{n} \left(\frac{(x_k - x_{k-1})^2}{t_k - t_{k-1}} - 2\mu(x_k - x_{k-1}) + \mu^2(t_k - t_{k-1})\right) \boxed{-\mu x_n + \frac{1}{2}\mu^2 t_n} \\
&= -\frac{1}{2}\sum_{k=1}^{n} \frac{\big((x_k - x_{k-1}) - \mu(t_k - t_{k-1})\big)^2}{t_k - t_{k-1}} \boxed{-\mu(x_n - \mu t_n) - \frac{1}{2}\mu^2 t_n}
\end{aligned}$$

Note that the first term is the joint PDF of $(X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}})$. Therefore, we conclude

$$E[f(B_{t_1}, B_{t_1}, \cdots, B_{t_n})] = E[f(X_{t_1}, X_{t_1}, \cdots, X_{t_n})\, M_{t_n}]$$

where M_t is a martingale,

$$M_t = \exp\Big(-\mu B_t - \frac{1}{2}\mu^2 t\Big).$$

Note that **SCFA** derives the result in the opposite direction,

$$E[f(X_{t_1}, X_{t_1}, \cdots, X_{t_n})] = E[f(B_{t_1}, B_{t_1}, \cdots, B_{t_n})\; M_{t_n}]$$
$$M_t = \exp\Big(+\mu B_t - \frac{1}{2}\mu^2 t\Big).$$

This causes some confusion in terms of the final Girsanov theorem, so I do not follow the convention of **SCFA**.

Functions of a Brownian Path

Hitting Time of a Sloping Line: Direct Approach

13.3 Simplest Girsanov Theorem

Let us consider the gambler's situation again. If a gambler is faced with equal chance of win and lose for \$1 bet, the game is considered fair. If the probability changes to p and $1-p$ for $p < 0.5$, then the odds are against our gambler.

How can we make the game fair to our gambler? One quick solution is the change the reward scheme by introducing the multiplier. It is quite obvious that the multiplier should be inversely proportional to the probability to make the game fair.

Prob Space Q

Event	Prob.	Reward
Win	0.5	$+1$
Lose	0.5	-1

Prob Space P

Event	Prob.	Reward	Multiplier (M)
Win	p	$+1$	$\alpha \propto 1/p$
Lose	$1-p$	-1	$\beta \propto 1/(1-p)$

Although the multiplier should be multiplied to the reward, we can achieve the same by multiplying it to the probability and it turns out to be a very powerful tool. To distinguish the two probability spaces, we denote the original probability one (unfair to gambler) by P and the transformed one (fiar to gambler) by Q. The probabilities for a same event A under P and Q spaces are denoted by $P(A)$ and $Q(A)$. Similarly, the expectations under P and Q spaces are denoted by $E_P(\cdot)$ and $E_Q(\cdot)$ respectively. From these notations, we may express

$$Q(A) = P(A) \cdot M_A = E_P(1_A \cdot M_A).$$

For Q to be qualified as a proper probability space, there are two conditions that the multiplier M should satisfy:

$$M_A > 0 \text{ for all } A \quad \text{and} \quad 1 = \sum_A Q(A) = E_P(M)$$

Given these constraints, the following probability change make the game fair again

$$\alpha = \frac{1}{2p}, \quad \beta = \frac{1}{2(1-p)}.$$

Notice that the multiplier M defines the ratio of the two probabilities $M_A = Q(A)/P(A)$. Under continuous probability space, M is referred to as the Radon-Nikodym derivative (WIKIPEDIA).

$$\frac{dQ(x)}{dP(x)} = M(x)$$

So far we implicitly assumed that probability measure (PDF on the space) and stochastic process (typically BM) can not be separated. For example, a stochastic process imply a PDF, e.g., a standard BM B_t has the PDF $f(x,t) = n(x/\sigma\sqrt{t})$ and a BM with drift, $B_t + \mu t$ has $f(x,t) = n\big((x-\mu t)/\sigma\sqrt{t}\big)$. Before stating Girsanov theorem, however, we need to somehow treat them separated. A question useful for understanding Girsanov theorem is,

> How does a stochastic process change if it is evaluated under a new PDF (probability measure) derived from the original PDF?

From our earlier result is recognized as a probability distortion between B_t and $X_t = B_t + \mu t$:

$$E_Q[f(B_{0\le s\le t})] = E_P[f(X_{0\le s\le t}) \cdot M_t] \quad \text{for}$$

with the Radon-Nikodym derivative,

$$dQ = M_t\, dP \quad \text{where} \quad M_t = \exp\big(-\mu B_t - \frac{1}{2}\mu^2 t\big).$$

Here, M_t satisfies the conditions for the modifiers: (i) $dQ(x) \ge 0$ and (ii) $E_Q[1] = E_P\big[e^{-\mu B_t - \mu^2 t/2}\big] = 1$.

Moreover, the result states effectively states that any probability expectation about the drifted BM, X_t, with the distortion by M_t (P-measure), can be simply replaced with a standard BM, B_t (Q-measure). The most striking case is that, X_t itself becomes B_t under the probability change.

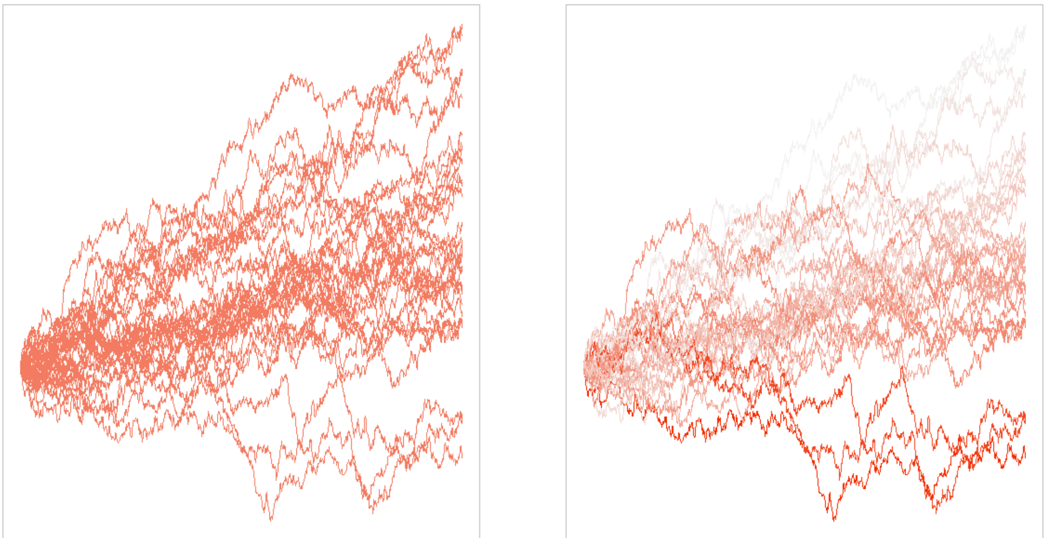
Theorem 13.1 (Simplest Girsanov Theorem: BM with Drift) If B_t is a standard BM under the probability measure P, the the drifted BM $X_t = B_t + \mu t$ is a standard BM under the probability measure Q defined as

$$E_Q(1_A) = E_P(1_A \cdot M_t) \quad \text{where} \quad M_t = \exp(-\mu B_t - \frac{1}{2}\mu^2 t).$$

A quick way of stating Girsanov theorem is that

$$dB_t^Q = dB_t^P + \mu dt$$

where B_t^P and B_t^Q denote the standard BMs under the probability measures P and Q respectively. It is more important to keep in mind that there is a probability measure Q which makes $B_t + \mu t$ a martingale while the formula for the correction term M_t is less focused. The probability measure Q is called an *equivalent martingale measure*.



Graphical representation of Girsanov theorem from (WIKIPEDIA)

13.4 Creation of Martingales

The contents of this section is already stated in **Theorem 13.1**.

Discounted Stock Price/The quest for a New Measure

Under the probability measure P, the stock and bond processes are given as

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t, \quad \frac{d\beta_t}{d\beta_t} = r dt,$$

and the discounted stock price process, $D_t = S_t/\beta_t$, has the SDE

$$\frac{dD_t}{D_t} = (\mu - r)dt + \sigma dB_t.$$

Except for the special case $\mu = r$, D_t is not a martingale under P. However we can *distort* the probability measure so that D_t can be a martingale.

Let $X_t = \hat{\mu}t + B_t$, with $\hat{\mu} = (\mu - \sigma)/\sigma$. Then, we can define the measure Q such that

$$Q(A) = E_P(1_A \cdot M_t), \quad M_t = \exp\left(-\hat{\mu}B_t - \frac{1}{2}\mu^2 t\right).$$

Under the Q measure, X_t is a standard BM under Q and the discounted stock price D_t is a martingale. The measure Q is referred to as the *risk-neutral*. We will discuss this more on Ch. 14.

13.5 Shifting the General Drift

Theorem 13.2 (Removing Drift) Suppose that P is the probability measure associated with standard BM B_t. The BM with a general (time and path-dependent) drift $\mu(w, t)$,

$$X_t = B_t + \int_0^t \mu(w, s)ds,$$

is a standard BM (no drift) under the probability measure Q defined by $Q(A) = E_P(1_A M_t)$ where

$$M_t = \exp\Big(- \int_0^t \mu(w, s)dB_s - \frac{1}{2}\int_0^t \mu^2(w, s)ds\Big)$$

13.6 Exponential Martingales and Novikov's Condition

13.7 Exercises

Chapter 14

Arbitrage and Martingales

14.1 Reexamination of the Binomial Arbitrage

Please read this section along with the paragraph **the Binomial Arbitrage** in Section **10.1**.

14.2 The Valuation Formula in Continuous Time

Present Value and Discounting

From Girsanov theorem, we learn that there is a probability measure Q (risk-neutral measure) under which the discounted stock price becomes a martingale. With the help of the martingale representation theorem, we also know that the expectation of the discounted payoff of a derivative, X/β_T, under the filtration $\mathcal{F}_t$ is represented as a martingale,

$$V_t = \beta_t \, E_Q(X/\beta_T \mid \mathcal{F}_t)$$

This also give the expression for the derivative price at $t = 0$ as

$$V_0 = E_Q(X/\beta_T).$$

The Market Price of Risk

Following the argument of **Option, futures and other derivatives** by John Hull, let us assume there are two securities S_{1t} and S_{2t} following

$$dS_{kt} = \mu_k S_{kt} dt + \sigma_k S_{kt} dW_k \quad \text{for} \quad k = 1, 2.$$

If we construct the portfolio of S_{1t} and S_{2t} in such way that the uncertainty is eliminated:

$$V_t = (\sigma_2 S_{2t})S_{1t} - (\sigma_1 S_{1t})S_{2t}$$

$$dV_t = (\sigma_2 S_{2t})dS_{1t} - (\sigma_1 S_{1t})dS_{2t} = (\mu_1\sigma_2 - \mu_2\sigma_1)S_{1t}S_{2t}dt,$$

then the portfolio V_t should earn the riskfree return r:

$$dV_t = rV_t dt.$$

Equating the two result, we obtain

$$(\sigma_2 - \sigma_1)r = (\mu_1\sigma_2 - \mu_2\sigma_1) \quad \text{or} \quad \frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2}$$

Therefore, we conclude that for any pair of (μ, σ) from a security S_t, the following should be a constant as λ:

$$\frac{\mu - r}{\sigma} = \lambda.$$

The parameter λ is known as the market price of risk of S_t. From the relation, $\mu = r + \lambda\sigma$, the market price of risk λ is understood as the extra return expected by investor for the increased risk σ by a unit amount.

Probability measure associated with numeraire*

Although, risk-neutral measure Q plays an important role in finance, it is not always the most convenient choice for pricing derivatives. We notice that the discounted stock price S_t/β_t is the measure of stock price S_t in the unit of the saving account β_t ($\beta_0 = 1$). We can generalize this concept to use any (risky) security for the measure unit, called *numeraire*.

Assume that two securities S_t and N_t follow GBMs under a real measure P:

$$\frac{dS_t}{S_t} = \mu_S dt + \sigma_S dB_t^P \quad \text{and} \quad \frac{dN_t}{N_t} = \mu_N dt + \sigma_N dB_t^P.$$

Now we are going to measure S_t in the unit of N_t. For the purpose, we derive the SDE for the ratio of two assets S_t and N_t. Through the SDEs of $\log S_t$ and $\log N_t$, we obtain

$$\begin{aligned}
d\log(S_t/N_t) &= d\log S_t - d\log N_t = (\mu_S - \mu_N)dt - \frac{1}{2}(\sigma_S^2 - \sigma_N^2)dt + (\sigma_S - \sigma_N)dB_t^P \\
\frac{d(S_t/N_t)}{S_t/N_t} &= (\mu_S - \mu_N)dt + \frac{1}{2}\Big((\sigma_S - \sigma_N)^2 - (\sigma_S^2 - \sigma_N^2)\Big)dt + (\sigma_S - \sigma_N)dB_t^P \\
&= (\mu_S - \mu_N)dt - \sigma_N(\sigma_S - \sigma_N)dt + (\sigma_S - \sigma_N)dB_t^P \\
&= (\sigma_S - \sigma_N)\left(dB_t^P - \sigma_N dt + \frac{\mu_S - \mu_N}{\sigma_S - \sigma_N}dt\right) \\
&= (\sigma_S - \sigma_N)\left(dB_t^P - \sigma_N dt + \lambda\, dt\right).
\end{aligned}$$

Here, we used the property of market price of risk

$$\mu_S - r = \lambda\sigma_S, \quad \mu_N - r = \lambda\sigma_N \quad \Rightarrow \quad \lambda = \frac{\mu_S - \mu_N}{\sigma_S - \sigma_N}.$$

Under the measure P associated with B_t^P, S_t/N_t is not a martingale. By Girsanov theorem, however, we know that there is a measure Q^N under which S_t/N_t is a martingale, i.e.,

$$dB_t^{Q^N} + \sigma_N dt = dB_t^P + \lambda\, dt.$$

Under the Q^N measure, we can easily obtain the current value of the derivative with the final payout X at time $t = T$ in a similar way. From the martingale property,

$$\frac{V_0}{N_0} = E^{Q^N}\left[\frac{X}{N_T}\right] \quad \Rightarrow \quad S_0 = N_0\, E^{Q^N}\left[\frac{X}{N_T}\right]$$

Now we consider a few important examples of numeraire N_t.

Saving account Let $N_t = \beta_t = e^{rt}$ $(d\beta_t = r\beta_t dt)$. In this case,

$$\mu_N = r \quad \text{and} \quad \sigma_N = 0$$

and the ratio S_t/β_t is the discounted stock price becomes

$$\frac{d(S_t/\beta_t)}{(S_t/\beta_t)} = \sigma_S dB_t^P + (\mu_S - r)dt.$$

The process is a martingale under Q measure

$$dB_t^Q = dB_t^P + \lambda dt \quad \text{for} \quad \lambda = \frac{\mu_S - r}{\sigma_S},$$

where λ is the market price of risk defined earlier. The Q measure is the one which λ is set to zero, i.e. investors are risk-neutral. So it is called risk-neutral measure.

Risky asset Now consider the risky numeraire N_t. From the result on dB_t^Q, we have a simpler expression

$$dB_t^N + \sigma_N dt = dB_t^Q,$$

which tell us another useful Girsanov theorem between Q measure and Q^N measure. Notice that adjusted drift is simply the volatility of the numeraire asset N_t. This is a useful property we will use later to compute the asset-or-nothing portion of the Black-Scholes formula.

***T*-forward measure** In this case, the numeraire is the zero-coupon bond maturing at $t = T$, $B(t,T)$. It turns out that T-forward measure is more practical than the risk neutral measure. Notice that$B(T,T) = 1$. Under this measure, the derivative pricing formula becomes

$$\frac{V_0}{B(0,T)} = E^T\left(\frac{X}{B(T,T)=1}\right) \quad \Rightarrow \quad V_0 = B(0,T)E^T(X).$$

The T-forward measure is also called as forward risk neutral measure. (WIKIPEDIA)

In general, any forward contract **fixed and paid** at time T is martingale under the T-forward measure. If S_t is the current price of a stock (assuming no dividend), the fair forward contract price should be $F_t = S_t/B(t,T)$ from replication argument. Therefore F_t is a martingale under T-forward measure.

Consider a **forward bond contract**. Assume that we enter into a forward contract of buying a bond with maturity Δ at time $t = T$. From replication argument, the fair forward price should be

$$F_t = \frac{B(t,T+\Delta)}{B(t,T)}$$

In a similar way, F_t is a martingale under the T-forward measure.

Delayed T-forward measure and term deposit rate $L(t,T)$ The term deposit L is the annualized interest rate you receive from bank. If you deposit \$ 1 now ($t = 0$), then receive $(1+\Delta L)$ later at $t = \Delta$. From replication, the rate at L at $t = 0$ is determined as

$$1 + \Delta L = \frac{1}{B(0,\Delta)} \quad \Rightarrow \quad L = \frac{1 - B(0,\Delta)}{\Delta \cdot B(0,\Delta)}.$$

To further generalize the deposit contract, imaging you enter into a forward deposit contract: you deposit 1 at $t = T$ and receive $1 + \Delta L$ at time $t = T + \Delta$, but the deposit rate L is pre-determined at $t = 0$. From replication, the fair forward deposit rate seen at time t, $L(t,T)$ should determined as

$$1 + \Delta L(t,T) = \frac{B(t,T)}{B(t,T+\Delta)} \quad \Rightarrow \quad L(t,T) = \frac{B(t,T) - B(t,T+\Delta)}{\Delta \cdot B(t,T+\Delta)}$$

Therefore, the forward deposit rate $L(t,T)$ is a martingale under $(T+\Delta)$-forward measure because it was expressed with numeraire $B(t,T+\Delta)$. The SDE for $L(t,T)$ is given as

$$dL(t,T) = \sigma\, dB_t^{T+\Delta} \quad \text{or} \quad \frac{dL(t,T)}{L(t,T)} = \sigma\, dB_t^{T+\Delta}$$

The most popular term deposit rate has been LIBOR (London Inter-Bank Offer Rate) although it is being replaced with other transparent index after fixing scandals. The cap and floor is a series of options on call and put option (respectively) on LIBOR rate. From the SDE, the option price is quickly computed with Black-Scholes or normal model formula.

Swap rate Interest rate swap is a financial contract exchanging a stream of floating interest payments with a stream of fixed interest payment. If the two parties pay the cashflows at $t = \Delta k$

for $1 \le k \le N$, the fixed interest rate K is determined so that the sums of the discount cashflow are equal:

$$\sum_{k=1}^{N} \Delta L(0,(k-1)\Delta) \cdot B(0,k\Delta) = \sum_{k=1}^{N} \Delta S \cdot B(0,k\Delta)$$

$$S = \frac{\sum_{k=1}^{N} L(0,(k-1)\Delta) \cdot B(0,k\Delta)}{\sum_{k=1}^{N} B(0,k\Delta)}$$

Now we can imagine a forward starting swap from $t = T$. The fair forward swap rate seen at t, $S(t,T)$ is similarly computed as

$$S(t,T) = \frac{\sum_{k=1}^{N} L(t,T+(k-1)\Delta)B(t,T+\Delta k)}{\sum_{k=1}^{N} B(t,T+\Delta k)}$$

The denominator $A(t,T) = \sum_{k=1}^{N} B(t,T+\Delta k)$ is the discounted cashflows of \$ 1 at the payment dates, called annuity. The annuity works as a numeraire for the forward swap rate. Therefore, under the measure Q^A associated with the annuity, the forward swap rate is a martingale:

$$dS(t,T) = \sigma\, dB_t^A \quad \text{or} \quad \frac{dS(t,T)}{S(t,T)} = \sigma\, dB_t^A.$$

The price of the option on swap, i.e., swaption, is evaluated using Black-Scholes or normal model formula. If you enter into a swaption paying a fixed rate of K, then the final payout is given as

$$(S(T,T) - K)^+ A(T,T),$$

The swaption pricing formula is given as

$$V_0 = A(0,T)E^A\left(\frac{(S(T,T)-K)^+A(T,T)}{A(T,T)}\right) = A(0,T)E^A\left((S(T,T)-K)^+\right).$$

Here the annuity $A(0,T)$ is playing the role of discount factor in the black-scholes on a stock.

14.3 The Black-Scholes Formula via Martingales

Revisiting Black-Scholes*

In fact, we derived the Black-Scholes-Merton formula under the risk neutral measure,

$$\frac{C_0}{\beta_0} = E_Q\left[\frac{(S_T-K)^+}{\beta_T}\right] \quad \text{with} \quad \beta_t = e^{rt} \text{ and } \frac{dS_t}{S_t} = rdt + \sigma dB_t.$$

The equivalent martingale measure can resolve the mystery of $N(d_1)$ and $N(d_2)$ in the formula. Recall Black-Scholes-Merton formula,

$$C_0 = S_0N(d_1) - e^{-rT}KN(d_2)$$

$$d_{1,2} = \frac{\log(S_0\, e^{rT}/K)}{\sigma\sqrt{T}} \pm \frac{1}{2}\sigma\sqrt{T},$$

where $z = -d_2$ is the point where the payoff is zero and the integration was done from $z = -d_2$ to ∞. As we know, the call price can be decomposed into the two options,

$$\begin{aligned} C_0 &= e^{-rT} E_Q((S_T - K) \cdot 1_{S_T \geq K}) \\ &= E_Q(\frac{S_T \cdot 1_{S_T \geq K}}{e^{rT}}) - e^{-rT} K E_Q(1_{S_T \geq K}) = D_S - K\, D_1 \end{aligned}$$

where D_1 is the price of a digital (cash-or-nothing) option and D_S is the price of asset-or-nothing option. The two digital options can be priced as

$$\begin{aligned} D_1 &= e^{-rT} E_Q(1_{S_T \geq K}) = e^{-rT} Q(S_T \geq K) = e^{-rT} \int_{-d_2}^{\infty} n(z) dz = e^{-rT} N(d_2) \\ D_S &= E_Q(\frac{S_T \cdot 1_{S_T \geq K}}{e^{rT}}) = S_0 E_{Q^S}(\frac{S_T \cdot 1_{S_T \geq K}}{S_T}) = S_0\, Q^S(S_T \geq K) \\ &= S_0 \int_{-d_2 - \sigma\sqrt{T}}^{\infty} n(z) dz = S_0 N(d_2 + \sigma\sqrt{T}) = S_0 N(d_1). \end{aligned}$$

Here we used the measure Q^S for the valuation of D_S. From $B_t^{Q^S} = B_t^Q - \sigma t$, the integration point $z^Q = -d_2$ under the Q-measure was shifted to $z^S = -d_2 - \sigma\sqrt{T} = -d_1$ ($\sqrt{T} z^S = \sqrt{T} z^Q - \sigma T$) under the Q^S-measure.

14.4 American Options

We skip the remaining sections.

14.5 Self-Financing and Self-Doubt

14.6 Admissible Strategies and Completeness

14.7 Perspective on Theory and Practice

14.8 Exercises

Chapter 15

The Feynman-Kac Connection

We skip this chapter.

15.1 First Links

15.2 The Feynman-Kac Connection for Brownian Motion

15.3 Lévy's Arcsin Law

15.4 The Feynman-Kac Connection for Diffusions

15.5 Feynman-Kac and the Black-Scholes PDEs

15.6 Exercises