

Stochastic Finance (FIN 519)  
Complementary Notes for Textbook ([SCFA](#))

Instructor: Jaehyuk Choi

2023-24 Module 3 (Spring 2024)

# Contents

<b>1</b>	<b>Random Walks and First Step Analysis</b>	<b>6</b>
1.1	First Step Analysis . . . . .	6
1.2	Time and Infinity . . . . .	7
1.3	Tossing an Unfair Coin . . . . .	8
1.4	Numerical Calculation and Intuition . . . . .	9
1.5	First Steps with Generating Functions . . . . .	9
1.6	Exercises . . . . .	10
<b>2</b>	<b>First Martingale Steps</b>	<b>11</b>
2.1	Classic Examples . . . . .	12
2.2	New Martingales from Old . . . . .	13
2.3	Revisiting the Old Ruins . . . . .	14
2.4	Submartingales . . . . .	15
2.5	Doob's Inequalities . . . . .	15
2.6	Martingale Convergence . . . . .	15
2.7	Exercises . . . . .	15
<b>3</b>	<b>Brownian Motion</b>	<b>16</b>
3.1	Covariances and Characteristic Functions . . . . .	17
3.2	Visions of a Series Approximation . . . . .	19
3.3	Two Wavelets . . . . .	19
3.4	Wavelet Representation of Brownian Motion . . . . .	19
3.5	Scaling and Inverting Brownian Motion . . . . .	19
3.6	Exercises . . . . .	20

<b>4</b>	<b>Martingales: The next steps</b>	<b>21</b>
4.1	Foundation Stones . . . . .	21
4.2	Conditional Expectations . . . . .	21
4.3	Uniform Integrability . . . . .	21
4.4	Martingales in Continuous Time . . . . .	21
4.5	Classic Brownian Motion Martingales . . . . .	23
4.6	Exercises . . . . .	25
<b>5</b>	<b>Richness of Paths</b>	<b>26</b>
5.1	Quantitative Smoothness . . . . .	26
5.2	Not Too Smooth . . . . .	26
5.3	Two Reflection Principles . . . . .	26
5.4	The Invariance Principle and Donsker's Theorem . . . . .	28
5.5	Random Walks Inside Brownian Motion . . . . .	29
5.6	Exercises . . . . .	29
<b>6</b>	<b>Itô Integration</b>	<b>30</b>
6.1	Definition of Itô's Integral: First Two Steps . . . . .	31
6.2	Third Step: Itô's Integral as a Process . . . . .	32
6.3	The Integral Sign: Benefits and Costs . . . . .	32
6.4	An Explicit Calculation . . . . .	32
6.5	Pathwise Interpretation of Itô's Integrals . . . . .	33
6.6	Exercises . . . . .	33
<b>7</b>	<b>Localization and Itô's integral</b>	<b>34</b>
7.1	Itô's Integral on $\mathcal{L}_{\text{LOC}}^2$ . . . . .	34
7.2	An Intuitive Representation . . . . .	34
7.3	Why Just $\mathcal{L}_{\text{LOC}}^2$ ? . . . . .	36
7.4	Local Martingales and Honest Ones . . . . .	36
7.5	Alternative Fields and Changes of Time . . . . .	36
7.6	Exercises . . . . .	36

<b>8</b>	<b>Itô's Formula</b>	<b>37</b>
8.1	Analysis and Synthesis . . . . .	37
8.2	First Consequences and Enhancements . . . . .	38
8.3	Vector Extension and Harmonic Functions . . . . .	40
8.4	Functions of Processes . . . . .	41
8.5	The General Itô's Formula . . . . .	41
8.6	Quadratic Variation . . . . .	43
8.7	Exercises . . . . .	43
<b>9</b>	<b>Stochastic Differential Equations</b>	<b>44</b>
9.1	Matching Itô's Coefficients . . . . .	44
9.2	Ornstein–Uhlenbeck Processes . . . . .	45
9.3	Matching Product Process Coefficients . . . . .	46
9.4	Existence and Uniqueness Theorems . . . . .	48
9.5	Systems of SDEs . . . . .	48
9.6	Exercises . . . . .	48
<b>10</b>	<b>Arbitrage and SDEs</b>	<b>49</b>
10.1	Replication and Three Examples of Arbitrage . . . . .	49
10.2	The Black-Scholes Model (WIKIPEDIA) . . . . .	50
10.3	The Black-Scholes Formula (WIKIPEDIA) . . . . .	51
10.4	Two Original Derivations . . . . .	53
10.5	The Perplexing Power of a Formula . . . . .	54
10.6	Exercises . . . . .	54
<b>11</b>	<b>The Diffusion Equation</b>	<b>55</b>
11.1	The diffusion of Mice . . . . .	55
11.2	Solutions of the Diffusion Equation . . . . .	55
11.3	Uniqueness of Solutions . . . . .	55
11.4	How to Solve the Black-Scholes PDE . . . . .	55
11.5	Exercises . . . . .	55

<b>12 Representation Theorems</b>	<b>56</b>
12.1 Stochastic Integral Representation Theorem . . . . .	56
12.2 The Martingale Representation Theorem (WIKIPEDIA) . . . . .	56
12.3 Continuity of Conditional Expectations . . . . .	56
12.4 Representation via Time Change . . . . .	57
12.5 Lévy's Characterization of Brownian Motion . . . . .	58
12.6 Bedrock Approximation Techniques . . . . .	58
12.7 Exercises . . . . .	58
<b>13 Girsanov Theory</b>	<b>59</b>
13.1 Importance Sampling . . . . .	59
13.2 Tilting a Process . . . . .	60
13.3 Simplest Girsanov Theorem . . . . .	61
13.4 Creation of Martingales . . . . .	64
13.5 Shifting the General Drift . . . . .	65
13.6 Exponential Martingales and Novikov's Condition . . . . .	65
13.7 Exercises . . . . .	65
<b>14 Arbitrage and Martingales</b>	<b>66</b>
14.1 Reexamination of the Binomial Arbitrage . . . . .	66
14.2 The Valuation Formula in Continuous Time . . . . .	66
14.3 The Black-Scholes Formula via Martingales . . . . .	71
14.4 American Options . . . . .	72
14.5 Self-Financing and Self-Doubt . . . . .	72
14.6 Admissible Strategies and Completeness . . . . .	72
14.7 Perspective on Theory and Practice . . . . .	72
14.8 Exercises . . . . .	72
<b>15 The Feynman-Kac Connection</b>	<b>73</b>
15.1 First Links . . . . .	73
15.2 The Feynman-Kac Connection for Brownian Motion . . . . .	73
15.3 Lévy's Arcsin Law . . . . .	73
15.4 The Feynman-Kac Connection for Diffusions . . . . .	73

15.5 Feynman-Kac and the Black-Scholes PDEs . . . . .	73
15.6 Exercises . . . . .	73

# 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;

$$S_n = S_0 + X_1 + X_2 + \cdots X_n,$$

where  $X_k = 1$  or  $-1$  with 50%:50% 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*  $\tau$  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 do not know the value of  $\tau$  ( $\tau$  is a probability variable).

### 1.1 First Step Analysis

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

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

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

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

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

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

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 underlying 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,  $P(S_\tau = A | S_0 = 0)$ , 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 compute the expected number of bets,  $\tau$ , until the gambler finishes the game, i.e., when he makes \$A or loses \$B. **SCFA** first proves that the expectation of  $\tau$  (and any power) is finite. (See **SCFA** for detail.)

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

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

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 in a row. On average, how many times do you need to toss a coin? For the solution, see **2016HW 1** in Problems and Solutions.

### 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,

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

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

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  $\varepsilon$ .

**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  $\varepsilon$  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  $6 \times 10^{-6}$  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  $\tau = 0, 1, 2, \dots$  (the meaning of  $\tau$  is also different from the previous sections!). The generating function is in the form of

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

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  $\tau_1$  is the first hitting time from 0 to 1 and  $\tau_2$  is the first hitting time from 1 to 2. Because  $\tau_1$  and  $\tau_2$  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

Exercises **1.1** and **1.3** are available in the **Problems and Solutions**.

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

$$E(M_{n+1} | X_1, X_2, \dots, X_n) = M_n \text{ for all } n \geq 0.$$

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,

$$E(M_{n+1} | M_1, M_2, \dots, M_n) = M_n \text{ for all } n \geq 0.$$

## 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  $\text{Var}(X_n) = \sigma^2$ . 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  $\sigma^2$  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

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

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

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

(Note that the indexing of  $A_n$  here is slightly different from that of **SCFA**.) 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 bounded 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  $\tau$ . 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  $\tau$ -th step is just known. The random variable  $\tau$  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  $\tau$  associated with any stopping strategy you can imagine with common sense is a proper stopping time, so you do not 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  $\tau$  (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**

Exercises **2.1** and **2.4** are available in the **Problems and 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)! We will cover the Bachelier model soon.

BM is defined as below:

**Definition 3.1** A *continuous-time stochastic process*  $\{B_t : 0 \leq t < T\}$  is called a Standard Brownian Motion on  $[0, T)$  if (i)  $B_0 = 0$ , (ii) The increments of  $B_t$ , i.e.,  $B_{t_2} - B_{t_1}$ ,  $B_{t_3} - B_{t_2}$ ,  $\dots$  for  $0 \leq t_1 < t_2 < t_3 < \dots$ , are independent, (iii) the increment,  $B_t - B_s$ , for  $s \leq t$  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 do not 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 BMs 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:

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

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

**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

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

Further, we can show that the multivariate vector  $A\mathbf{V}$  for  $d \times d$  matrix  $A$  has the following mean and covariance (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.,  $\mathbf{V} = A\mathbf{Z}$  for some matrix  $A$ . It will give us two 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 want to select the matrix  $A$  in our favor such that

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

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

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

where  $A$  is an orthonormal matrix ( $AA^T = A^T A = I$ ) and  $\Lambda$  is the diagonal matrix with non-negative eigen-values (the eigen-values are non-negative because  $\Sigma$  is symmetric,  $\Sigma = \Sigma^T$ ).

## Covariance Functions and Gaussian Processes

A process,  $X_t$  is called a Gaussian process if the vectors  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  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  $t$  ( $s \leq t$ ) 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 are standard BMs.

- Scaled process:

$$X_t = \frac{1}{\sqrt{a}} B_{at} \quad \text{for } t \geq 0.$$

For the graphical demonstration of the scaled process, see the **self-similarity** ([WIKIPEDIA](#)).

- **Inverted process:**

$$Y_0 = 0 \quad \text{and} \quad Y_t = t B_{1/t} \quad \text{for} \quad t > 0.$$

- **Time-reversed process (not in SCFA):**

$$Z_t = B_1 - B_{1-t} \quad \text{for} \quad 0 \leq t \leq 1.$$

## 3.6 Exercises

Exercises **3.1–3.4** are available in the **Problems and 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.  $E(X_t | \mathcal{F}_s) = X_s$  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  $\tau$  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  $X_\tau$  (when  $\tau$  is not  $\infty$ ) 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 2.2, which states that the stopped process  $M_{t \wedge \tau}$  derived from the original continuous process  $M_t$  and the stopping time  $\tau$  is also a continuous martingale.

we will skip the rest of the section.

## 4.5 Classic Brownian Motion Martingales

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

1. **Standard BM:**  $B_t$ ,
2.  $B_t^2 - t$ ,
3. **Wald's Martingale (Geometric BM):**  $\exp(\alpha B_t - \alpha^2 t/2)$

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  $\alpha$  is replaced with the volatility  $\sigma$ .

### 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  $\tau$  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 real value  $a$ ,

$$P(\tau_a < \infty) = 1 \quad \text{and} \quad E(e^{-\lambda \tau_a}) = e^{-|a|\sqrt{2\lambda}}. \quad (4.25)$$

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

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



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

Here and in Chapter 2, we are using the following properties of the martingale from the stopped process. If  $M_t$  (or  $M_n$ ) is a martingale and  $\tau$  is a stopping time,

$$M_0 = E(M_{t \wedge \tau}) \quad (\text{for any } t) = \lim_{t \rightarrow \infty} E(M_{t \wedge \tau}) = E(M_\tau).$$

However, if we apply this to the BM martingale,  $M_t = B_t$ , and the same stopping time  $\tau_a$ , we arrive at a contradiction!

$$0 = M_0 = E(M_\tau) = a.$$

For the  $M_0 = E(M_\tau)$  to hold, one condition should be satisfied:

$$|M_{t \wedge \tau}| < C \quad \text{for all } t \text{ and some } C,$$

that is the stopped process  $M_{t \wedge \tau}$  should be bounded.

This condition holds for  $M_t = \exp(\alpha B_t - \alpha^2 t/2)$  because  $M_{t \wedge \tau} \leq \exp(\alpha a)$ . Therefore we could use  $M_0 = E(M_\tau)$  above. But the condition does not hold for  $M_t = B_t$  because  $B_t$  can take any negative value. Therefore,  $M_0 \neq E(M_\tau)$ . If we use the the stopping time from the double barrier (gambler's ruin),  $\tau = \inf\{t : B_t = -B \text{ or } B_t = A\}$ , then  $B_{t \wedge \tau}$  is bounded as  $B_{t \wedge \tau} \leq \max(A, B)$ . That is how we obtained **Theorem 4.5**.

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  $\sigma$ ?

## 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

$$E\left(\frac{1}{\tau_a}\right) = \int_0^\infty E(e^{-\lambda\tau_a}) d\lambda = \int_0^\infty e^{-a\sqrt{2\lambda}} d\lambda = \int_0^\infty u e^{-u} \frac{du}{a^2} = \frac{1}{a^2}, \quad (4.26)$$

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

$$f_{\tau_a}(t) = \frac{|a|}{\sqrt{2\pi t^3}} e^{-a^2/2t} \quad \text{for } t \geq 0 \quad (4.27)$$

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 4.6 is available in the **Problems and 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,  $B_t^* = \max_{0 \leq s \leq t} B_s$ . We'll derive 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 (Reflection Principle of Brownian Paths)** The *reflected process*,

$\tilde{B}_t$  of the standard BM,  $B_t$  defined as

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

for any stopping time  $\tau$  (usually the hitting-time at certain level) is also a standard BM.

The defined process  $\tilde{B}_t$  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  $\tau$ ,

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

The definition makes sense only when we know the value of  $B_\tau$ . 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  $\tilde{B}_t = 2a - B_t$ , which is a BM starting from  $2a$ .

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

$$P(B_t^* > x, B_t < x - y) = P(B_t^* > x, B_t > x + y) = P(B_t > x + y). \quad (5.12)$$

The first equality is directly from the reflection principle. The stopping time  $\tau$  used here is the first hitting-time of the level  $x$ . The condition  $B_t^* > x$  means that the path hit the level  $x$  before time  $t$ , so if  $B_t > x + y$ , the reflected path should satisfy  $\tilde{B}_t < x - y$ . 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 results on the joint density,

$$\begin{aligned} \text{CDF: } P(B_t^* < v, B_t < u) &= \Phi(u/\sqrt{t}) + \Phi((2v - u)/\sqrt{t}) - 1 \\ &= \Phi(u/\sqrt{t}) - \Phi((u - 2v)/\sqrt{t}) \end{aligned} \quad (5.13)$$

$$\text{PDF: } f_{(B_t^*, B_t)}(v, u) = \frac{2(2v - u)}{t^{3/2}} \phi((2v - u)/\sqrt{t}) \quad (5.14)$$

## Density and Distribution of $B_t^*$

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

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

Here, we used the reflection principle,  $P(B_t^* > x, B_t \leq x) = P(B_t^* > x, B_t \geq x)$ . 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  $x$ ,

$$f_{B_t^*}(x) = \frac{2}{\sqrt{t}} \phi\left(\frac{x}{\sqrt{t}}\right) = \sqrt{\frac{2}{\pi t}} e^{-x^2/2t} \quad \text{for } x \geq 0. \quad (5.16)$$

## Density of the Hitting Time $\tau_x$

The differentiation w.r.t.  $t$  gives the density on  $\tau_x$ ,

$$f_{\tau_x}(t) = \frac{x}{t^{3/2}} \phi\left(\frac{x}{\sqrt{t}}\right) \quad \text{for } x \geq 0, t \geq 0. \quad (5.17)$$

Note we saw the result in Eq. (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,

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

From the central limit theorem (CLT) ([WIKIPEDIA](#)), we know that  $S_n/\sqrt{n}$  converges to  $\Phi(0, 1)$  as  $n \rightarrow \infty$ . 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

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

and define a scaled process,

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

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

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

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  $0 \leq t \leq T$ . 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,

$$dS_t = S_t (r dt + \sigma dB_t).$$

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 that

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

This not only holds in the sense of expectation, but also holds with probability 1. Without 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;

$$E(B_t^2 - t) = 0$$

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

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,

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

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

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

The stochastic integral  $I(f)$  has zero mean:

$$E(I(f)) = \lim_{N \rightarrow \infty} \sum_{k=0}^N E(f(B_0, \dots, B_{t_k}, t = t_k)(B_{t_{k+1}} - B_{t_k})) = 0.$$

This is from that  $f(B_0, \dots, B_{t_k}, t = t_k)$  and  $B_{t_{k+1}} - B_{t_k}$  are independent because  $f(B_0, \dots, B_{t_k}, t = t_k)$  depends on  $0 \leq t \leq t_k$ .

## 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

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

and the variance of RHS also yields to the same value

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

Using the following two properties,

$$\begin{aligned} B(t_k) (B(t_{k+1}) - B(t_k)) &= \frac{1}{2} (B^2(t_{k+1}) - B^2(t_k)) - \frac{1}{2} (B(t_{k+1}) - B(t_k))^2 \\ (B(t) - B(s))^2 &= t - s \quad \text{as } t \rightarrow s \quad \text{or} \quad (dB(t))^2 = dt. \end{aligned}$$

we prove the integral identity as

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

## 6.5 Pathwise Interpretation of Itô's Integrals

## 6.6 Exercises

Exercises **6.1–6.3** are available in the **Problems and Solutions**.

# Chapter 7

## Localization and Itô's integral

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

### 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 co-variance function

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

#### 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  $\int_0^t f^2(s)ds \rightarrow \infty$  as  $t \rightarrow \infty$ , 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  $\sigma_d = 1\% / 16 = 0.0625\% = 6.25$  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  $t_1, t_2, \dots, t_n$  and we obtain the corresponding volatilities  $\sigma_1, \sigma_2, \dots, \sigma_n$  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,

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

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,

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

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 $\mathcal{L}_{\text{LOC}}^2$ ?

### 7.4 Local Martingales and Honest Ones

### 7.5 Alternative Fields and Changes of Time

### 7.6 Exercises

Exercise 7.1 is available in the **Problems and Solutions**.

# 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:

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

## 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  $f(x) = x^2/2$  from Theorem 8.1. From Theorem 8.2 we can verify this by the differentiation rule:

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

### 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

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

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

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

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

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

**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

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

is a martingale. It can be reaffirmed using the previous proposition. We can write

$M_t = f(t, B_t)$  where

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

and the function  $f(t, x)$  has properties such that

$$f_t = -\frac{1}{2}\sigma^2 f, \quad f_x = \sigma f, \quad \text{and} \quad f_{xx} = \sigma^2 f.$$

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

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

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 Probability for Brownian Motion with Drift)** If  $X_t = \mu t + \sigma B_t$  and  $\tau$  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 derive 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 the properties,

$$f_t(t, x) = \alpha f(t, x) \quad \text{and} \quad f_x(t, x) = \sigma f(t, x),$$

we can derive 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

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

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)

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

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  $\log X_t$ ,

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

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

$$\begin{aligned} \log(X_t) - \log(X_0) &= \alpha t + \sigma B_t \\ X_t &= X_0 \exp(\alpha t + \sigma B_t) \end{aligned}$$

The GBM is better known in the form of

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

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

$$\begin{aligned} dX_t &= -\frac{X_t}{1+t} dt + \frac{\sigma}{1+t} dB_t \\ (1+t)dX_t &= -X_t dt + \sigma dB_t \end{aligned}$$

So we start with

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

Therefore, we obtain

$$(1+t)X_t - X_0 = \sigma B_t \quad \implies \quad X_t = \frac{\sigma B_t + X_0}{1+t}$$

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

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

So, finally we get

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

## 8.6 Quadratic Variation

## 8.7 Exercises

Exercises **8.2** and **8.4** are available in the **Problems and Solutions**.

# 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,

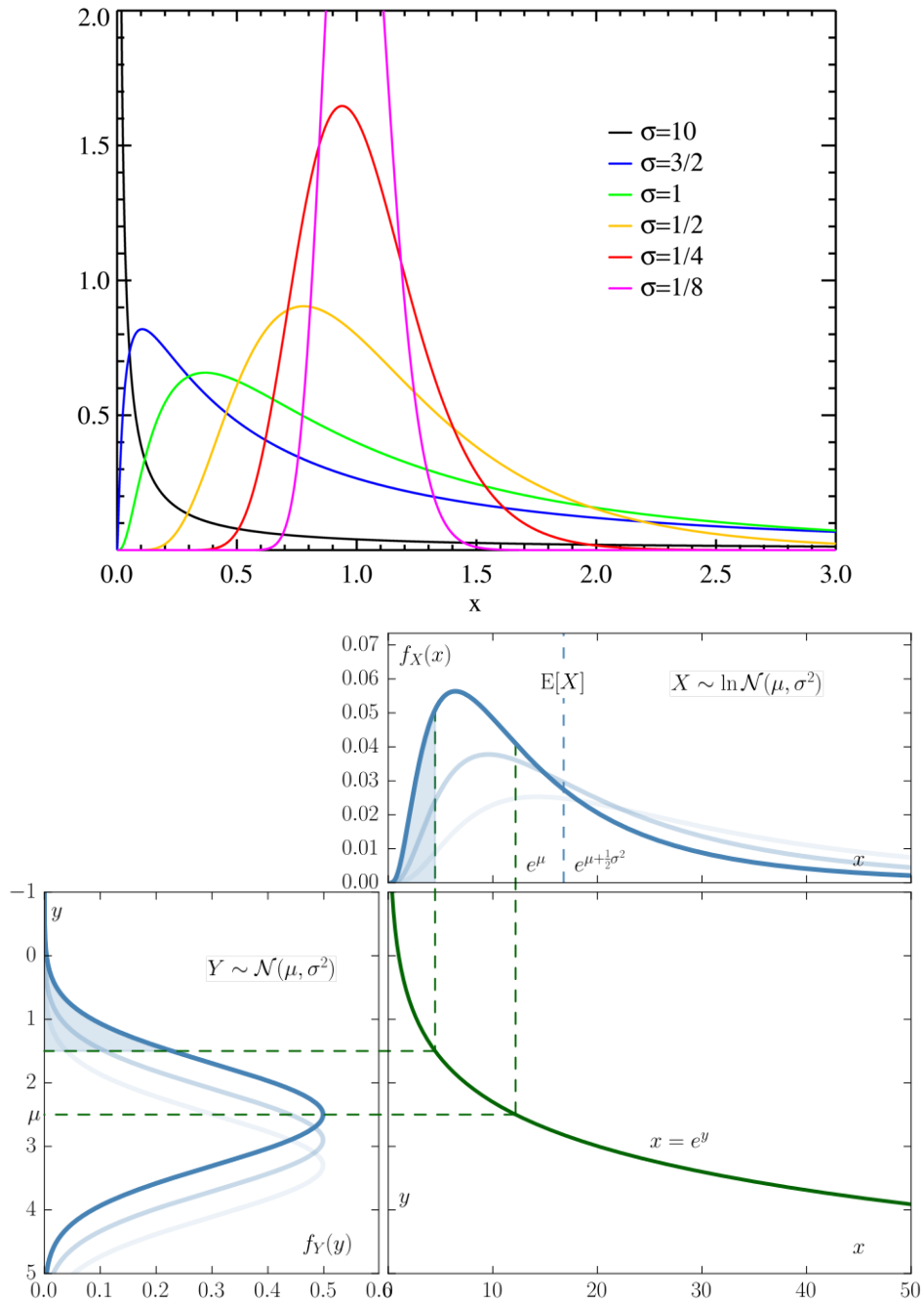
$$X_t = X_0 \exp \left( \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t \right).$$

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  $\sigma^2/2 > \mu > 0$ ,

$$P(X_t < \varepsilon) \rightarrow 1 \quad \text{as} \quad t \rightarrow \infty$$

for very small  $\varepsilon$ . 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  $\sigma dB_t$ . So the process is 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,  $X_\infty$ , 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

$$dx = -\alpha x dt$$

$$\frac{dx}{x} = -\alpha dt$$

$$\log(x) = -\alpha t$$

$$e^{\alpha t} x = x_0$$

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

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

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

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

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

$$\begin{aligned} e^{\alpha t} X_t &= e^{\alpha s} X_s + \frac{\sigma}{\sqrt{2\alpha}} (B'_{e^{2\alpha t}-1} - B'_{e^{2\alpha s}-1}) \\ X_t &= e^{-\alpha(t-s)} X_s + \frac{\sigma e^{-\alpha t}}{\sqrt{2\alpha}} (B'_{e^{2\alpha t}-1} - B'_{e^{2\alpha s}-1}), \end{aligned}$$

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

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

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

$$\begin{aligned} d\left(\frac{X_t}{1-t}\right) &= \frac{dX_t}{1-t} + \frac{X_t dt}{(1-t)^2} = \frac{dB_t}{1-t} \\ \frac{X_t}{1-t} &= \int_0^t \frac{dB_s}{1-s} \quad \Rightarrow \quad X_t = (1-t) \int_0^t \frac{dB_s}{1-s} \end{aligned}$$

Although this is not a complete analytic form, we can derive a few property from it. For example, we can show that  $\text{Cov}(X_s, X_t) = s(1-t)$  (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. Under the stochastic volatility models, the volatility  $\sigma$  follows another stochastic process rather than staying constant. The stochastic volatility models are one method to model the volatility smile observed in the market.

### Example: Stochastic Volatility Models

Heston Model ([WIKIPEDIA](#))

$$\begin{aligned}\frac{dF_t}{F_t} &= \sqrt{V_t} dB_t^1 \\ dV_t &= \kappa(V_\infty - V_t)dt + \alpha\sqrt{V_t}dB_t^2 \\ dB_t^1 dB_t^2 &= \rho dt\end{aligned}$$

Stochastic-Alpha-Beta-Rho (SABR) Model ([WIKIPEDIA](#))

$$\begin{aligned}\frac{dF_t}{F_t^\beta} &= \sigma_t dB_t^1 \\ \frac{d\sigma_t}{\sigma_t} &= \alpha dB_t^2 \\ dB_t^1 dB_t^2 &= \rho dt\end{aligned}$$

## 9.6 Exercises

Exercises **9.1–9.3**, and **9.6** are available in the **Problems and Solutions**.

# Chapter 10

## Arbitrage and SDEs

In this chapter, we derive the Black-Scholes PDE. We also derive 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  $\alpha$  share of stock and  $\beta$  unit of bond can replicate the payoff at  $t = 1$ :

$$4\alpha + \beta = 3 \quad \text{and} \quad \alpha + \beta = 0.$$

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  $r$ :

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

### Arbitrage and Replication

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

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

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,

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

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

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

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

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

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

The the present value of the call options is integrated as

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

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

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

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

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

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

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 the Bachelier model. But they are fairly close, so you can depend on the Bachelier model price,  $C \approx 0.4 \sigma_N \sqrt{T}$ . Be careful not to use the same volatility  $\sigma$ . The approximate relation should be that  $\sigma_N \approx \sigma_{BS} F$ .

**Dependence on Stdev** Nevertheless, the call option depends on the standard deviation,  $\sigma\sqrt{T}$ , not on either  $\sigma$  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 the Bachelier 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  $S_0 N(d_1)$  can be understood as the present price of the asset-or-nothing option, in the same way in the Bachelier model.

**The maximum value of the call option** is  $S_0$  or  $DF$ . It is achieved when  $N(d_1) = N(d_+) = 1$  and  $N(d_2) = N(d_-) = 0$  which is achieved when  $\sigma \rightarrow \infty$  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

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

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

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

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

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

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  $\phi_s$  such that

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

### 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

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

Therefore, we have the martingale representation as

$$X_t = X_0 + \int_0^t \sigma X_s dB_s, \quad \phi_s = \sigma X_0 \exp \left( -\frac{1}{2} \sigma^2 t + \sigma B_t \right)$$

### 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

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

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  $\phi_t$  share of stock at time  $t$  and such replicating strategy  $\phi_t$  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,

$$dX_t = 2B_t dB_t + dt - dt = 2B_t dB_t \quad \text{and} \quad X_0 = E(B_T^2) = T,$$

we obtain the representation

$$X_T - T = \int_0^T 2B_t dB_t$$

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),

$$P(Z > 30) = \frac{1}{N} \sum_1^N 1_{Z_k > 30} \quad \text{for } Z_k \sim N(0, 1)$$

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,

$$\begin{aligned}
E(f(Z)) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(z) e^{-z^2/2} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-(x-\mu)^2/2} e^{-\mu x + \mu^2/2} dx \\
&= E\left(f(X) e^{-\mu X + \mu^2/2}\right) \\
&= E\left(f(Z + \mu) e^{-\mu Z - \mu^2/2}\right) \quad \text{for } X = Z + \mu
\end{aligned}$$

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

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

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

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

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

$$E[f(B_{t_1}, B_{t_2}, \dots, B_{t_n})].$$

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

$B_0 = 0$ . Considering only the exponent of the joint PDF,

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

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

Therefore, we conclude

$$E[f(B_{t_1}, B_{t_2}, \dots, B_{t_n})] = E[f(X_{t_1}, X_{t_2}, \dots, X_{t_n}) M_{t_n}]$$

where  $M_t$  is a martingale,

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

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

$$\begin{aligned}
E[f(X_{t_1}, X_{t_2}, \dots, X_{t_n})] &= E[f(B_{t_1}, B_{t_2}, \dots, B_{t_n}) M_{t_n}] \\
M_t &= \exp\left(+\mu B_t - \frac{1}{2}\mu^2 t\right).
\end{aligned}$$

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 $Q$			Prob Space $P$			
Event	Prob.	Reward	Event	Prob.	Reward	Multiplier ( $M$ )
Win	0.5	+1	Win	$p$	+1	$\alpha \propto 1/p$
Lose	0.5	-1	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 (fair 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  $f(x, t) = n((x - \mu t)/\sigma\sqrt{t})$ . 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  $X_t = B_t + \mu t$ :

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

with the Radon-Nikodym derivative,

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

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

Moreover, the result 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

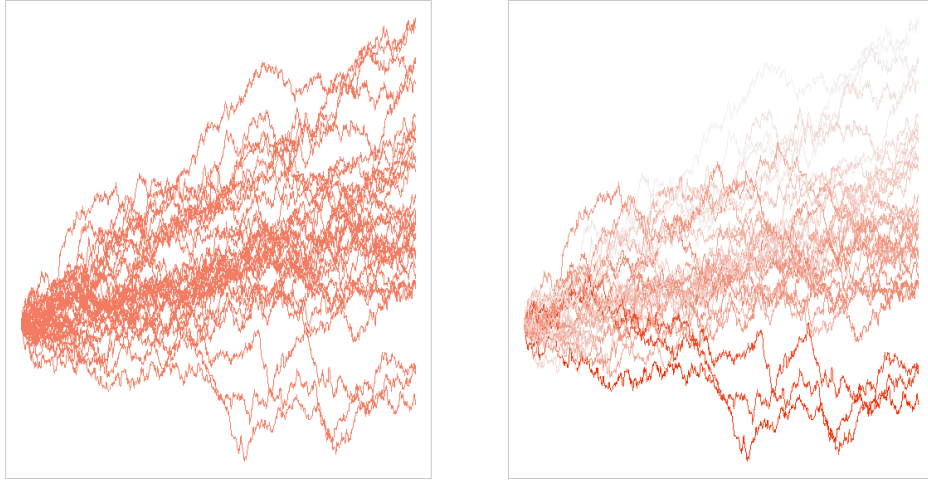
$$E_Q(1_A) = E_P(1_A \cdot M_t) \quad \text{where} \quad M_t = \exp(-\mu B_t - \frac{1}{2}\mu^2 t).$$

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

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

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  $\hat{\mu} = (\mu - r)/\sigma$ . Then, we can define the measure  $Q$  such that

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

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

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

## 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,  $X/\beta_T$ , under the filtration  $\mathcal{F}_t$  is represented as a martingale,

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

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

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

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  $(\mu, \sigma)$  from a security  $S_t$ , the following should be a constant as  $\lambda$ :

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

The parameter  $\lambda$  is known as the market price of risk of  $S_t$ . From the relation,  $\mu = r + \lambda\sigma$ , the market price of risk  $\lambda$  is understood as the extra return expected by investor for the increased risk  $\sigma$  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  $S_t/\beta_t$  is the measure of stock price  $S_t$  in the unit of the saving account  $\beta_t$  ( $\beta_0 = 1$ ). 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

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

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,

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

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  $S_t/\beta_t$  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  $\lambda$  is the market price of risk defined earlier. The  $Q$  measure is the one which  $\lambda$  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

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

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

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

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  $\Delta$  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

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

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 be 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 the Black-Scholes or Bachelier model formulas.

**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  $1 \leq k \leq N$ , the fixed interest rate  $S$  is determined so that the sums of the discount cashflow are equal:

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

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

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

The quantity in the denominator,

$$A(t, T) = \sum_{k=1}^N \Delta \cdot B(t, T + \Delta k),$$

is the discounted cashflows of  $\$ \Delta$  at the payment dates. It is called *annuity*. Annuity works as a numeraire for the forward swap rate. Therefore, under the measure  $Q^A$

associated with the annuity, the forward swap rate seen at  $t$ ,  $S_t = S(t, T)$ , is a martingale:

$$dS_t = \sigma dB_t^A \quad \text{or} \quad \frac{dS_t}{S_t} = \sigma dB_t^A.$$

Now, we can evaluate the price of swaption, i.e., option on swap, using the Black-Scholes or Bachelier model formulas. If you enter into a forward swap contract where you pay a fixed rate of  $K$  and receive the floating rate from  $t = T$ , the value of the swap at  $t = T$  will be expressed in terms of  $A(T, T)$ :

$$\begin{aligned} \text{Swap Value} &= \sum_{k=1}^N \Delta \cdot L(T, T + (k-1)\Delta) \cdot B(T, T + k\Delta) - \sum_{k=1}^N \Delta \cdot K \cdot B(T, T + k\Delta) \\ &= \sum_{k=1}^N \Delta \cdot S_T \cdot B(T, T + k\Delta) - \sum_{k=1}^N \Delta \cdot K \cdot B(T, T + k\Delta) \\ &= (S_T - K) \sum_{k=1}^N \Delta \cdot B(T, T + k\Delta) = (S_T - K)A(T, T) \end{aligned}$$

If you hold a swaption on the same underlying swap (i.e., paying a fixed rate of  $K$  from  $t = T$ ), then the payout of the option at  $t = T$  is given by

$$X(= V_T) = (\text{Swap Value})^+ = (S_T - K)^+ A(T, T),$$

The swaption pricing formula is given as

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

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,

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

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

$$C_0 = S_0 N(d_1) - e^{-rT} K N(d_2)$$



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

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

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

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

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

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