

## **Introduction to Stochastic Calculus**

Yuewu Xu, Ph.D.

School of Business

Fordham University

Fall, 2022

## Overview

- Continuous-time finance is one of the most exciting and most important developments of modern finance. It has fundamentally changed how risks are priced and managed in the real world.
- Main asset pricing ideas from single-period models:
  - Mean-variance portfolio theory: Markowitz (1952).
  - Capital asset pricing model: Sharpe (1964) and Lintner (1965).
  - Arbitrage pricing theory: Ross (1976, 1977).
- Continuous-time finance models extend the insights of single-period models and provide more realistic modeling of financial markets.
- Discrete vs. continuous-time models
  - Multi-period models can be studied in discrete or continuous time.
  - Each approach has its own dis/advantages.
  - Discrete-time models are easier to understand, but sometimes are more difficult to solve.
  - Continuous-time models rely on more complicated mathematics, but can lead to more elegant and powerful solutions.
- Potential topics covered
  - Introduction to stochastic calculus.
  - Option pricing: Black-Scholes model, risk-neutral pricing.
  - Term structure of interest rates (time permits)
  - Credit risk models (time permits).
- Teaching philosophy
  - Focusing on main ideas and sketch of proofs.
  - Focusing on understanding the results: why it is true, and how to use it in practice.
  - Measure-theoretic type of proofs and conditions kept at minimum (if you are interested, read the textbook).

## Part 1. Introduction to Stochastic Calculus

Main topics

- Brownian motion
- Stochastic integration
- Ito's formula
- Applications of Ito's formula

### 1.2 Brownian Motion

Brownian motion is one of the most widely studied continuous-time stochastic processes and is a major building block for continuous-time asset pricing models

A little history: Scatter a few grains of pollen on the surface of an apparently still beaker of liquid. Under a microscope you will observe that each grain is not still but jitters about on the liquid surface. This was first noticed by a botanist called Robert Brown in 1827. He was looking for microscopic life in a drop of water when he noticed that small grains in the water were jiggling around in a strange way - almost as if they were alive! This type of random motion is called Brownian motion after him.

Einstein showed that continual collision with water molecules causes Brownian motion. When he investigated Brownian motion at the beginning of the 20th century not all scientists believed in molecules, and Einstein was trying to demonstrate that they really did exist.

Einstein (1905) listed the following three properties of BM:

- (i) The sample paths must be continuous (based on physics);
- (ii) The increments follow a normal distribution with a variance that is proportion to the time elapsed (based on CLT);

- (iii) The increments of BM are independent, i.e., pollen grain has no memory.

Einstein could not prove that such process exists. In 1920s, Norbert Wiener proved BM exists and hence it is also called Wiener process.

**Definition of Brownian Motion.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space. A family of random variables  $W_t$  indexed by time  $t$  (assume that  $W_0 \equiv 0$ ) is called a Brownian Motion if it satisfies

- Continuous sample path: For each  $\omega \in \Omega$ , the function  $W_t(\omega)$  is a continuous function of  $t \geq 0$
- Independent increments: for all  $0 = t_0 < t_1 < \dots < t_m$  the increments

$$W_{t_1} - W_0, W_{t_2} - W_{t_1}, \dots, W_{t_m} - W_{t_{m-1}}$$

are independent of each other.

- Normality: each of the increments is normally distributed

$$\mathbb{E}[W_{t_{i+1}} - W_{t_i}] = 0 \text{ and } \mathbb{E}[W_{t_{i+1}} - W_{t_i}]^2 = t_{i+1} - t_i.$$

A filtration for Brownian motion is defined as  $\mathcal{F}_t = \sigma(W_s, s \leq t)$ . That is, all observable event is based on observing the BM before  $t$ . Adaptivity means that  $W(t)$  is  $\mathcal{F}_t$ -measurable. Independence of future increments means that for  $0 \leq t < u$ ,  $W(u) - W(t)$  is independent of  $\mathcal{F}_t$ .

A few important properties of Brownian motion

(1). Brownian motion is a martingale.

**Proof.** Let  $0 \leq s < t$  be given. Then

$$\begin{aligned} & \mathbb{E}[W_t | \mathcal{F}_s] \\ = & \mathbb{E}[W_t - W_s + W_s | \mathcal{F}_s] \\ = & \mathbb{E}[W_t - W_s | \mathcal{F}_s] + \mathbb{E}[W_s | \mathcal{F}_s] \quad (\text{linearity property}) \\ = & \mathbb{E}[W_t - W_s | \mathcal{F}_s] + W_s \quad (W_s \text{ is known given } \mathcal{F}_s) \\ = & \mathbb{E}[W_t - W_s] + W_s \quad (\text{independence of increments}) = W_s. \end{aligned}$$

This property means that the best forecast of tomorrow's value is today's value. This is the idea behind the random walk model

(2).  $W_t^2 - t$  is a martingale.

**Proof.** Let  $0 \leq s < t$  be given. Then

$$\begin{aligned}
 \mathbb{E}[W_t^2 - t | \mathcal{F}_s] &= \mathbb{E}[(W_t - W_s + W_s)^2 - t | \mathcal{F}_s] \\
 &= \mathbb{E}[(W_t - W_s)^2 + 2(W_t - W_s)W_s + W_s^2 - t | \mathcal{F}_s] \\
 &= \mathbb{E}[(W_t - W_s)^2 | \mathcal{F}_s] + 2\mathbb{E}[(W_t - W_s)W_s | \mathcal{F}_s] \\
 &\quad + \mathbb{E}[W_s^2 - t | \mathcal{F}_s] \\
 &= \mathbb{E}[(W_t - W_s)^2] + 2\mathbb{E}[W_t - W_s | \mathcal{F}_s]W_s + W_s^2 - t \\
 &= (t - s) + W_s^2 - t = W_s^2 - s.
 \end{aligned}$$

(3). For  $0 \leq s < t$ ,  $\text{cov}(W_s, W_t) = s$ .

**Proof.** The covariance of  $W_s$  and  $W_t$  is

$$\begin{aligned}
 \mathbb{E}[W_s W_t] &= \mathbb{E}[\mathbb{E}[W_s W_t | \mathcal{F}_s]] = \mathbb{E}[W_s \mathbb{E}[W_t | \mathcal{F}_s]] \\
 &= \mathbb{E}[W_s^2] = s.
 \end{aligned}$$

Or

$$\begin{aligned}
 \mathbb{E}[W_s W_t] &= \mathbb{E}[W_s (W_t - W_s + W_s)] \\
 &= \mathbb{E}[W_s (W_t - W_s) + W_s^2] \\
 &= \mathbb{E}[W_s] \mathbb{E}[W_t - W_s] + \mathbb{E}[W_s^2] \\
 &= 0 + s = s.
 \end{aligned}$$

### Nondifferentiability of Brownian path

One important property of BM is that its sample path is not differentiable as a function of  $t$ .

- This can be understood intuitively from the fact  $W(t) \sim N(0, t)$ .
- The nondifferentiability result is seen from the fact that BM has unbounded total variation.

**Total Variation [First Variation].** Let  $f(t)$  be a function defined for  $0 \leq t \leq T$ . If there exists a finite  $M > 0$  such that

$$\sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)| < M,$$

for all grid  $\Pi = \{t_0, t_1, \dots, t_n\} : 0 = t_0 < t_1 < \dots < t_n = T$ , then  $f$  is said to have finite Total Variation; The smallest value  $M$  is called the Total Variation of  $f$ .

We often say the size of the grid is

$$\|\Pi\| = \max_{j=0, \dots, n-1} (t_{j+1} - t_j).$$

**Example 1:** What is the total variation of  $f(x) = x$  for  $x \in [0, 2]$ .

**Example 2:** What is the total variation of  $f(x) = x^2$  for  $x \in [-2, 2]$ .

**Example 3:** What is the total variation of  $f(x) = \sin(x)$  for  $x \in [0, 2\pi]$ .

The Geometric interpretation of Total variation.

It should be noted that the total variation concept is generally applicable to "smooth" functions. For "non-smooth" functions such as the sample path of Brownian Motion, the total variation can be shown to be infinite. In such cases, we need the concept of "quadratic variation".

**Quadratic Variation.** The quadratic variation of  $f$  up to time  $T$  is

$$[f, f](T) = \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2.$$

**FACT:** If  $f$  has a continuous derivative, then its quadratic variation is zero.

$$\begin{aligned} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2 &= \sum_{j=0}^{n-1} |f'(t_j^*)|^2 (t_{j+1} - t_j)^2 \\ &\leq \|\Pi\| \cdot \sum_{j=0}^{n-1} |f'(t_j^*)|^2 (t_{j+1} - t_j), \quad t_j^* \in [t_j, t_{j+1}], \end{aligned}$$

$$\begin{aligned} [f, f](T) &\leq \lim_{\|\Pi\| \rightarrow 0} \left[ \|\Pi\| \cdot \sum_{j=0}^{n-1} |f'(t_j^*)|^2 (t_{j+1} - t_j) \right] \\ &= \lim_{\|\Pi\| \rightarrow 0} \|\Pi\| \cdot \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} |f'(t_j^*)|^2 (t_{j+1} - t_j) \\ &= \lim_{\|\Pi\| \rightarrow 0} \|\Pi\| \cdot \int_0^T |f'(t)|^2 dt = 0. \end{aligned}$$

Because  $f'(t)$  is continuous,  $\int_0^T |f'(t)|^2 dt$  is finite.

**Theorem [Quadratic Variation of BM].** Let  $W$  be a BM, then  $[W, W](T) = T$  for all  $T \geq 0$ .

**Proof.** For any finite partition  $\Pi$  of  $[0, T]$ , denote by  $Q = \sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})^2$ ,

and we want to show that  $Q \rightarrow T$  as  $\|\Pi\| \rightarrow 0$ . We can show that

$$\mathbb{E}[Q] = T \text{ and } \lim_{\|\Pi\| \rightarrow 0} \mathbb{E}[Q - T]^2 = 0.$$

[Note: This means that  $Q$  converges to  $T$  in  $\mathcal{L}^2$ -norm]. We have

$$\begin{aligned} \mathbb{E}[(W_{t_{j+1}} - W_{t_j})^2] &= \text{Var}[W_{t_{j+1}} - W_{t_j}] = t_{j+1} - t_j \\ \Rightarrow \mathbb{E}[Q] &= \mathbb{E}\left[\sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})^2\right] = \sum_{j=0}^{n-1} (t_{j+1} - t_j) = T. \end{aligned}$$

Moreover,

$$\begin{aligned} \text{Var}[(W_{t_{j+1}} - W_{t_j})^2] &= \mathbb{E}\left[\left((W_{t_{j+1}} - W_{t_j})^2 - (t_{j+1} - t_j)\right)^2\right] \\ &= \mathbb{E}\left[(W_{t_{j+1}} - W_{t_j})^4\right] - 2(t_{j+1} - t_j) \mathbb{E}\left[(W_{t_{j+1}} - W_{t_j})^2\right] \\ &\quad + (t_{j+1} - t_j)^2 = 2(t_{j+1} - t_j)^2. \end{aligned}$$

Note  $\mathbb{E} \left[ (W_{t_{j+1}} - W_{t_j})^4 \right] = 3 (t_{j+1} - t_j)^2$ . Therefore,

$$\begin{aligned} Var(Q) &= \sum_{j=0}^{n-1} Var \left[ (W_{t_{j+1}} - W_{t_j})^2 \right] = \sum_{j=0}^{n-1} 2 (t_{j+1} - t_j)^2 \\ &\leq \sum_{j=0}^{n-1} \|\Pi\| 2 (t_{j+1} - t_j) = 2 \|\Pi\| T. \end{aligned}$$

Therefore,  $\lim_{\|\Pi\| \rightarrow 0} Var(Q) = 0$  and  $\lim_{\|\Pi\| \rightarrow 0} Q = \mathbb{E}(Q) = T$ . These two properties imply the result in the Theorem ■

**Corollary.**  $[W, W](T) = T \Rightarrow TV(W) = \infty$  almost surely.

**Proof.** We prove by contradiction. Suppose  $TV(W) < \infty$ , then

$$\begin{aligned} \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})^2 &\leq \lim_{\|\Pi\| \rightarrow 0} \left[ \max_{t_j \in \Pi} |W_{t_{j+1}} - W_{t_j}| \right] \sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}| \\ &\leq \lim_{\|\Pi\| \rightarrow 0} \max_{t_j \in \Pi} |W_{t_{j+1}} - W_{t_j}| \sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}| \\ &\leq \lim_{\|\Pi\| \rightarrow 0} \max_{t_j \in \Pi} |W_{t_{j+1}} - W_{t_j}| TV(W) = 0. \end{aligned}$$

This contradicts with  $[W, W](T) = T$ . ■

In the above derivation, we have used

$$\begin{aligned} \mathbb{E} \left[ (W_{t_{j+1}} - W_{t_j})^2 \right] &= t_{j+1} - t_j, \\ Var \left[ (W_{t_{j+1}} - W_{t_j})^2 \right] &= 2 (t_{j+1} - t_j)^2. \end{aligned}$$

**Review:**  $L^2$  distance, Second Moment, etc. Convergence in  $L^2$ .

Now we provide a brief introduction to the so-called  $L^2$ -distance, which will be needed for studying stochastic integrals.

The  $L^2$ -distance is an extension to the Euclidean distance in  $\mathbb{R}^3$ :

$$\|x - y\| = \sqrt{\sum_{i=1}^3 (x_i - y_i)^2} \text{ for } x = (x_1, x_2, x_3), \text{ and } y = (y_1, y_2, y_3).$$



And we have the familiar concept of convergence in terms of the Euclidean distance.

The  $L^2$ -distance is a natural extension of the Euclidean distance to the infinite dimensional space (space of functions):

$$\|f - g\|_{L^2} = \sqrt{\int |f(z) - g(z)|^2 dz}.$$

To reconcile this with the Euclidean distance, just note that integration sign is just a continuous sum. The  $L^2$  distance enjoys most of the properties of the Euclidean distance.

For random variables  $X$  and  $Y$ , the  $L^2$ -distance can be defined as

$$\|X - Y\|_{L^2} = \sqrt{\mathbb{E}|X - Y|^2}.$$

Hence for the study of random variables, the  $L^2$ -distance is important, (which is equivalent to analyzing Variances).