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, Nobert 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 $0 = t_0 < t_1 < ... < t_m$ the incrementsis

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

  are independent of each other.
- Normality: each of the increments is normally distributed

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

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

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

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

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

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

Proof. The covariance of W_s and W_t is

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

Or

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

Nondifferentiability of Brownian path

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

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

Total Variation [First Variation]. Let $f(t)$ be a function defined for $0 \le t \le T$. 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 $\Pi = \{t_0, t_1, ..., t_n\} : 0 = t_0 < t_1 < ... < t_n = T$, 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

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

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 concenpt is generally applicapable 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

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

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

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

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

Because $f'(t)$ is continuous, $\int_0^T\left|f'(t)\right|^2 dt$ 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 Π of $[0,T]$, denote by $Q=\sum_{j=0}^{n-1}\left(W_{t_{j+1}}-W_{t_j}\right)^2$,

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

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

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

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

Moreover,

$$Var\left[\left(W_{t_{j+1}}-W_{t_j}\right)^2\right]=\mathbb{E}\left[\left(\left(W_{t_{j+1}}-W_{t_j}\right)^2-\left(t_{j+1}-t_j\right)\right)^2\right]$$

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

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

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

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

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

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

$$\lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \left(W_{t_{j+1}} - W_{t_j}\right)^2 \leq \lim_{\|\Pi\|\to 0} \left[\max_{t_j\in\Pi} \left|W_{t_{j+1}} - W_{t_j}\right|\right] \sum_{j=0}^{n-1} \left|W_{t_{j+1}} - W_{t_j}\right|$$

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

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

In the above derivation, we have used

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

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 $\mathbb{R}^3$:

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

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).

Stochastic Integration

In cintinuous-time finance, a basic task is to compute continuous trading gains/losses. Stochastic integration is such a tool.

Assume that we have a trading strategy H_t which represents the number of shares we hold for an asset at any time t. Assume that the price of the asset moves according to a Brownian motion W_t. Then the proft/loss of this continuous trading is represented by stochastic integral $\int_0^T H(t)\, dW(t)$.

To motivate the above, discretize the time interval $[0,T]$ into n small intervals $0 = t_0 < t_1 < \cdots < t_n = T$. It is easy to see that if we only trade at the grid points t_i, then the trading profit/loss is represented by

$$\sum_i H_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right).$$

If we want to allow continuous trading, then we will have to find the limit of the above sum of random variables. If the limit exist (as the grid becomes finer and finer) in some sense, then we can difine this limit as the stochastic intergral

$$\int_0^T H(t)\,dW(t) := \lim \sum_i H_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right). \text{ [in some sense]}$$

Remark: This definition has a clear finance interpretation: If we interpret H as trading strategy, and $W_{t_{i+1}} - W_{t_i}$ as changes in stock price, the the stochastic integral representations the total trading profit/loss.

Remark: For H_t to be a trading strategy, H_t needs to be $\mathcal{F}_t$ measurable. Otherwise the trading strategies is not executable (why?)

Remark: The convergence of the defining sum is in the sense of L^2-distance. It is also called mean-square, quadratic-mean, second moment, etc. This is a "better" distance to use. I gave a short review on what L^2 -distance is in class, so I will not repeat here.

Definition (Stochastic Integral in Ito sense). If

$$\sum_i H_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right)$$

convergence to a random elelemt (variable) I in the sense of L^2-distance:

$$\lim \sum_{t_i \in \Pi} H_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right) \xrightarrow{L^2} I, \text{ as size of grid } \Pi \to 0. \tag{1}$$

then random variable I is called the Ito Integral, and we define

$$\int_0^T H(t)\,dW(t) \equiv I.$$

Note that, by the definition of L^2-distance, (1) can be written in the more familiar form:

$$\mathbb{E}\left[\sum_{t_i \in \Pi} H_{t_i}\left(W_{t_{i+1}} - W_{t_i}\right) - I\right]^2 \to 0, \text{ as grid size } \Pi \to 0.$$

Hence, it only involves computation of second moment of random variables!

Remark: What is the meaning L^2-distance? Why is L^2-distance a good distance is good to work with. Review earlier lectures.

Remark: The limit is well defined, meaning it does not depend on the partition.

Remark: We can NOT define stochastic integrals this way,

$$\int_0^T H_t dW_t = \int_0^T H_t \left(\frac{dW_t}{dt}\right) dt,$$

because the sample path of BM is not differentiable.

Now we look at a typical example (often asked in the job interviews).

Example (Important!): Compute $\int_0^T W(t)\, dW(t)$ from the definition:

We set up a grid t_i as in the definition. Note that here $H(t) = W(t)$. The left end points are

$$\begin{array}{cc} W(0) & \text{if } 0 \le t < \frac{T}{n} \\ W\left(\frac{T}{n}\right) & \text{if } \frac{T}{n} \le t < \frac{2T}{n} \\ & \cdots \\ W\left(\frac{(n-1)T}{n}\right) & \text{if } \frac{(n-1)T}{n} \le t < T. \end{array}$$

For ease of notation, we set $W_j = W\left(\frac{jT}{n}\right)$. To evaluate the stochastic integral, we need to study

$$\int_0^T W(t)\, dW(t) = \lim_{n\to\infty} \sum_{j=0}^{n-1} W_j \left[W_{j+1} - W_j\right].$$

Now we proceed to evaluate this limit. Note that the sum can be written as

$$\sum_{j=0}^{n-1} W_j \left[W_{j+1} - W_j\right] = \sum_{j=0}^{n-1} W_{j+1} W_j - \sum_{j=0}^{n-1} W_j^2$$

The first term contains the cross product, so we can use formula

$$ab = \frac{1}{2}a^2 + \frac{1}{2}b^2 - \frac{1}{2}(a-b)^2$$

to rewrite it as as

$$\sum_{j=0}^{n-1} W_{j+1} W_j = \frac{1}{2}\sum_{j=0}^{n-1} W_{j+1}^2 + \frac{1}{2}\sum_{j=0}^{n-1} W_j^2 - \frac{1}{2}\sum_{j=0}^{n-1} (W_{j+1} - W_j)^2,$$

Plus in to conclude that

$$\sum_{j=0}^{n-1} W_j \left(W_{j+1} - W_j\right) = \frac{1}{2}\sum_{j=0}^{n-1} \left[W_{j+1}^2 - W_j^2\right] - \frac{1}{2}\sum_{j=0}^{n-1} \left[W_{j+1} - W_j\right]^2.$$

The first summation is a telescoping sum, and the second term is qudratic variation of BM. So we conclude

$$\begin{aligned}\sum_{j=0}^{n-1} W_j \left(W_{j+1} - W_j\right) &= \frac{1}{2}W_n^2 - \frac{1}{2}\sum_{j=0}^{n-1}\left(W_{j+1} - W_j\right)^2 \\ &\rightarrow \frac{1}{2}W^2\left(T\right) - \frac{1}{2}T, \text{ as } n \rightarrow \infty.\end{aligned}$$

Hence we get

$$\int_0^T W\left(t\right) dW\left(t\right) = \frac{1}{2}W^2\left(T\right) - \frac{1}{2}T. \blacksquare$$

In ordinary calculus, if g is a differentiable function with $g\left(0\right) = 0$, then

$$\int_0^T g\left(t\right) dg\left(t\right) = \int_0^T g\left(t\right) g'\left(t\right) dt = \frac{1}{2}g^2\left(t\right)|_0^T = \frac{1}{2}g^2\left(T\right).$$

The extra term $-\frac{1}{2}T$ comes from the nonzero quadratic variation of $W\left(t\right)$.

The above result fot the stochastic integral can be recasted as

$$W^2\left(T\right) = 2\int_0^T W\left(t\right) dW\left(t\right) + \left[W, W\right]\left(T\right).$$

Note: Although the definition of stochastic integrals in (1) works for general integrand H_t, a popular alternative way of achieving the same goal is to define stochastic integral for simple process first, and then use a limiting process to extend the definition to general processes. This approach has the great advantage that there no limiting process is needed for stochastic integrals with integrands of simple processes. *This is what we will do in the follows.*

Ito Integral for simple process. A simple process is such that there exists a partition $\{t_0, t_1, ..., t_n\}$ of $[0, T]$ for which $H\left(t\right)$ is constant in t on each subinterval $[t_j, t_{j+1})$. Assume that $t_k \leq t \leq t_{k+1}$, then the Ito integral for simple process $H\left(t\right)$ is defined as

$$\int_0^t H\left(u\right) dW\left(u\right) = \sum_{j=0}^{k-1} H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right] + H_{t_k}\left[W_t - W_{t_k}\right]. \tag{2}$$

Notice that there is no limit process in the definition.

Theorem. The Ito integral for the simple process is a martingale.

Proof. Suppose $0 \le s \le t \le T$, $s \in [t_l, t_{l+1})$, $t \in [t_k, t_{k+1})$, $t_l < t_k$, then

$$\begin{aligned} I(t) &= \int_0^t H(u)\, dW(u) \\ &= \sum_{j=0}^{l-1} H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right] + H_{t_l}\left[W_{t_{l+1}} - W_{t_l}\right] \\ &\quad + \sum_{j=l+1}^{k-1} H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right] + H_{t_k}\left[W_t - W_{t_k}\right]. \end{aligned}$$

We must show

$$\mathbb{E}\left[I(t)\,|\mathcal{F}(s)\right] = I(s) \equiv \sum_{j=0}^{l-1} H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right] + H_{t_l}\left(W_s - W_{t_l}\right).$$

We examine each of the four terms.

The first term $\sum_{j=0}^{l-1} H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right]$ is $\mathcal{F}(s)$ measurable (why?), so

$$\mathbb{E}\left[\sum_{j=0}^{l-1} H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right]|\mathcal{F}(s)\right] = \sum_{j=0}^{l-1} H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right].$$

The second term satisfies

$$\mathbb{E}\left[H_{t_l}\left[W_{t_{l+1}} - W_{t_l}\right]|\mathcal{F}(s)\right] = H_{t_l}\left(\mathbb{E}\left[W_{t_{l+1}}|\mathcal{F}(s)\right] - W_{t_l}\right) = H_{t_l}\left(W_s - W_{t_l}\right).$$

The third and fourth term has zero conditional expectations: for $t_j \ge t_{l+1} > s$,

$$\begin{aligned} \mathbb{E}\left\{H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right]|\mathcal{F}(s)\right\} &= \mathbb{E}\left\{\mathbb{E}\left[H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right]|\mathcal{F}(t_j)\right]|\mathcal{F}(s)\right\} \\ &= \mathbb{E}\left\{H_{t_j}\left(\mathbb{E}\left[W_{t_{j+1}}|\mathcal{F}(t_j)\right] - W_{t_j}\right)|\mathcal{F}(s)\right\} \\ &= \mathbb{E}\left\{H_{t_j}\left(W_{t_j} - W_{t_j}\right)|\mathcal{F}(s)\right\} \\ &= 0. \end{aligned}$$

Therefore,

$$\mathbb{E}\left\{\sum_{j=l+1}^{k-1} H_{t_j}\left[W_{t_{j+1}} - W_{t_j}\right] + H_{t_k}\left[W_t - W_{t_k}\right]|\mathcal{F}(s)\right\} = 0.$$

Puting together, we obtain the martingale property of $I(t)$.

Remark (Important): Because $I(t)$ is a martingale and $I(0) = 0$, we have $\mathbb{E}I(t) = 0$ for all $t > 0$. It follows that $Var\left[I(t)\right] = \mathbb{E}\left[I^2(t)\right]$.

Theorem (Ito Isometry). The Ito integral for the simple process satisfies

$$\mathbb{E}\left[I(t)\right]^2 = \mathbb{E}\int_0^t H^2(u)\, du.$$

Note that the right-hand side can be interpreted as $\mathbf{E}[H]^2$ for suitably defined $\mathbf{E}$.

Proof. **Denote** $D_j = W_{t_{j+1}} - W_{t_j}$ for $j = 0, ..., k-1$ and $D_k = W_t - W_{t_k}$. Then

$$I(t) = \sum_{j=0}^{k} H_{t_j} D_j \text{ and } I^2(t) = \sum_{j=0}^{k} H_{t_j}^2 D_j^2 + 2 \sum_{0 \le i < j \le k} H_{t_i} H_{t_j} D_i D_j.$$

$$\begin{aligned}
\mathbb{E}\left[H_{t_j}^2 D_j^2\right] &= \mathbb{E}\left\{\mathbb{E}\left[H_{t_j}^2 D_j^2 | \mathcal{F}_{t_j}\right]\right\} \\
&= \mathbb{E}\left\{H_{t_j}^2 \mathbb{E}\left[D_j^2 | \mathcal{F}_{t_j}\right]\right\} \\
&= \mathbb{E}\left[H_{t_j}^2 \left(t_{j+1} - t_j\right)\right],
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left[H_{t_i} H_{t_j} D_i D_j\right] &= \mathbb{E}\left\{\mathbb{E}\left[H_{t_i} H_{t_j} D_i D_j | \mathcal{F}_{t_j}\right]\right\} \\
&= \mathbb{E}\left\{H_{t_i} H_{t_j} D_i \mathbb{E}\left[D_j | \mathcal{F}_{t_j}\right]\right\} \\
&= \mathbb{E}\left\{H_{t_i} H_{t_j} D_i \cdot 0\right\} = 0.
\end{aligned}$$

Therefore,

$$\mathbb{E} I^2(t) = \sum_{j=0}^{k} \mathbb{E}\left[H_{t_j}^2 D_j^2\right] = \sum_{j=0}^{k-1} \mathbb{E}\left[H_{t_j}^2\right]\left(t_{j+1} - t_j\right) + \mathbb{E}\left[H_{t_k}^2\right]\left(t - t_k\right).$$

But H_{t_j} is constant on the interval $[t_j, t_{j+1})$, and hence each term above can be represented by

$$H_{t_j}^2 \left(t_{j+1} - t_j\right) = \int_{t_j}^{t_{j+1}} H^2(u)\, du.$$

Thus

$$\begin{aligned}
\mathbb{E} I^2(t) &= \sum_{j=0}^{k-1} \mathbb{E} \int_{t_j}^{t_{j+1}} H^2(u)\, du + \mathbb{E} \int_{t_k}^{t} H^2(u)\, du \\
&= \mathbb{E}\left[\sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} H^2(u)\, du + \int_{t_k}^{t} H^2(u)\, du\right] \\
&= \mathbb{E}\left[\int_0^t H^2(u)\, du\right].
\end{aligned}$$

Ito Integral for General Integrands

We now move to define Ito integral $\int_0^T H(t)\, dW(t)$ for integrands $H(t)$ that are allowed to vary continuously and also to jump. That is $H(t)$ is no longer a simple process. We do assume that $H(t)$, $t \ge 0$, is adapted to the filtration $\mathcal{F}(t_j)$ and the square-integrability condition

$$\mathbb{E} \int_0^T H^2(t)\, dt < \infty.$$

Suppose there is a sequence $H_n(t)$ of simple processes such that as $n \to \infty$, these processes "converge" to the continuously varying $H(t)$. Here "converge" means that

$$\lim_{n\to\infty} \mathbb{E} \int_0^T |H_n(t) - H(t)|^2 \, dt = 0.$$

For each $H_n(t)$, the Ito integral $\int_0^t H_n(u) \, du$ has already been defined for $0 \le t \le T$. We define the Ito integral for the continuously varying integrand $H(t)$ by the formula

$$\int_0^t H(u) \, dW(u) = \lim_{n\to\infty} \int_0^t H_n(u) \, dW(u), \quad 0 \le t \le T.$$

Here the limit is in L^2 sense.

Theorem. Ito integral $I(t) = \int_0^t H(u) \, dW(u)$ thus defined have the following properties.

(i) (Continuity). As a function of t, the paths of $I(t)$ are continuous.

(ii) (Adaptivity). For each t, $I(t)$ is $F(t)$-measurable.

(iii) (Linearity) If $I_1(t) = \int_0^t H_1(u) \, dW(u)$ and $I_2(t) = \int_0^t H_2(u) \, dW(u)$, then

$$\begin{aligned} I_1(t) + I_2 &= \int_0^t [H_1(u) + H_2(u)] \, dW(u), \\ cI(t) &= \int_0^t cH(u) \, dW(u) \text{ for constant } c \end{aligned}$$

(iv) (Martingale) $I(t)$ is a martingale.

(v) (Ito Isometry) $\mathbb{E} I^2(t) = \mathbb{E} \int_0^t H^2(u) \, du$.

(vi) (Quadratic Variation) $[I, I](t) = \int_0^t H^2(u) \, du$.

Remark: We computed the stochastic integral of $\int_0^T W_t dW_t$ using definition (1). Now we show that our computation also holds for the alternative definition using simple processes. We need to justify $\lim_{n\to\infty} \mathbb{E} \int_0^T |H_n(t) - W(t)|^2 \, dt = 0$. This is because

$$\begin{aligned}\mathbb{E}\int_0^T \left|H_n\left(t\right)-W\left(t\right)\right|^2 dt &= \mathbb{E}\sum_j \int_{t_j}^{t_{j+1}} (W_j - W_s)^2 ds \\ &= \sum_j \int_{t_j}^{t_{j+1}} \mathbb{E}(W_j - W_s)^2 ds \\ &= \sum_j \int_{t_j}^{t_{j+1}} (s-t_j) ds \\ &= \sum_j \frac{1}{2}(t_{j+1}-t_j)^2 \to 0.\end{aligned}$$

1.4 Ito's Formula

Recall the Newton-Leibniz formula in calculus:

$$G(x) - G(a) = \int_a^x g(y)dy. \text{ (here } G(x)' = g(x)).$$

This formula links the operations of differentiation and integration: one is the "inverse" of the other.

- In one direction, if we can find G, then it offers a way to calculate the integral on the right hand side
- In the other direction, we can express function $G(x)$ through its derivatives $g(x)$.

Question: How to prove the Newton-Leibniz formula?

Sketch of Proof (Important). The idea is to use Taylor expansion or its equivalents for partition $a = x_0 < x_1 < ... < x_n = x$:

$$\begin{aligned}G(x) - G(a) &= \sum_{i=0}^{n-1} G\left(x_{i+1}\right) - G\left(x_i\right) \\ &= \sum_{i=0}^{n-1} G'\left(\xi_i^*\right)\left[x_{i+1} - x_i\right] \\ &\to \int_a^x g(y)dy. \text{ (here } G(x)' = g(x)).\end{aligned}$$

Hence

$$G(x) - G(a) = \int_a^x g(y)dy$$

In continuous time finance, we often need to deal with functions of Brownian Motion (or more generally, functions of Ito processes).

Example. The Black-Scholes option pricing formula for European call options, if found, would be a function that depends on the (current) stock price $S(t)$

$$C(S(t), K, T, r, \sigma) \equiv C(S(t)) \text{ [other parameters are omitted]}$$

Hence if we want to derivative the BS formula, we need to be able to handle functions of BM or functions of Ito processes.

The Goal: find an expression for $f(W(t))$ where $f(x)$ is a smooth function, and $W(t)$ is a BM.

Example. In the example in the previous lecture, we obtained

$$W^2(T) = 2\int_0^T W(t)\, dW(t) + [W, W](T).$$

The above formula is special case of the following Ito formula for BM:

Ito's Formula: If function f is twice continuously differentiable, then

$$f(W(T)) - f(W(0)) = \int_0^T f'(W(u))\, dW(u) + \frac{1}{2}\int_0^T f''(W(u))\, du.$$

The integral form has precise mathematical meaning because we have formally defined Ito integral. The differential form only has intuitive but imprecise meaning

Derivation. Let $\Pi = \{t_0, t_1, ..., t_n\}$ be a partition of $[0, T]$. We can express $f(W_T)$

$$f(W_T) - f(W_0) = \sum_{j=0}^{n-1} \left[f\left(W_{t_{j+1}}\right) - f\left(W_{t_j}\right) \right].$$

Since $f(x)$ is twice continuously differentiable, Taylor expansion gives us that

$$f(x) - f(y) = f'(y)(x - y) + \frac{1}{2} f''(y)(x - y)^2 + R(f, x, y).$$

This yields

$$f\left(W_T\right)-f\left(W_0\right)=\sum_{j=0}^{n-1}\left[f\left(W_{t_{j+1}}\right)-f\left(W_{t_j}\right)\right]$$

$$=\sum_{j=0}^{n-1} f'\left(W_{t_j}\right)\left(W_{t_{j+1}}-W_{t_j}\right)+\frac{1}{2}\sum_{j=0}^{n-1} f''\left(W_{t_j}\right)\left[W_{t_{j+1}}-W_{t_j}\right]^2+R$$

If we let $\|\Pi\|\to 0$, then

$$\begin{aligned}
& f\left(W_T\right)-f\left(W_0\right) \\
= & \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1} f(W_{t_j})\left(W_{t_{j+1}}-W_{t_j}\right)+\frac{1}{2}\lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1} f''\left(W_{t_j}\right)\left[W_{t_{j+1}}-W_{t_j}\right]^2 \\
= & \int_0^T f'\left(W_t\right)dW_t+\frac{1}{2}\int_0^T f''\left(W_t\right)dt.
\end{aligned}$$

Note that we have ignored the error term as it can be justified to converge to 0.

A slight extension of the above Ito's formula is the following:

Theorem (Ito formula for Brownian motion). Let $f\left(t,x\right)$ be a function for which the partial derivatives $f_t\left(t,x\right)$, $f_x\left(t,x\right)$, and $f_{xx}\left(t,x\right)$ are defined and continuous, and let $W\left(t\right)$ be a Brownian motion. Then $\forall$ $T\geq 0$,

$$f\left(T,W_T\right)=f\left(0,W_0\right)+\int_0^T f_t\left(t,W_t\right)dt+\int_0^T f_x\left(t,W_t\right)dW_t+\frac{1}{2}\int_0^T f_{xx}\left(t,W_t\right)dt. \tag{3}$$

The only difference in the proof is that we need to expand $f(t,x)$ both with respect to t and x.

Derivation. Let $\Pi=\{t_0,t_1,...,t_n\}$ be a partition of $[0,T]$. For a general function $f\left(t,x\right)$, Taylor's expansion says that

$$\begin{aligned}
& f\left(t_{j+1},x_{j+1}\right)-f\left(x_j,x_j\right) \\
= & f_t\left(t_j,x_j\right)\left(t_{j+1}-t_j\right)+f_x\left(t_j,x_j\right)\left(x_{j+1}-x_j\right) \\
& +\frac{1}{2}f_{xx}\left(t_j,x_j\right)\left(x_{j+1}-x_j\right)^2+f_{tx}\left(t_j,x_j\right)\left(t_{j+1}-t_j\right)\left(x_{j+1}-x_j\right) \\
& +\frac{1}{2}f_{tt}\left(t_j,x_j\right)\left(t_{j+1}-t_j\right)^2+R.
\end{aligned}$$

We replace x_{j+1} and x_j by $W_{t_{j+1}}$ and W_{t_j}, respectively, and sum:

$$f\left(W_T\right)-f\left(W_0\right)=\sum_{j=0}^{n-1}\left[f\left(t_{j+1},W_{t_{j+1}}\right)-f\left(t_j,W_{t_j}\right)\right]$$

$$
\begin{aligned}
&= \sum_{j=0}^{n-1} f_t\left(t_j, W_{t_j}\right)\left(t_{j+1}-t_j\right) + \sum_{j=0}^{n-1} f_x\left(t_j, W_{t_j}\right)\left(W_{t_{j+1}}-W_{t_j}\right) \\
&\qquad + \frac{1}{2}\sum_{j=0}^{n-1} f_{xx}\left(t_j, W_{t_j}\right)\left[W_{t_{j+1}}-W_{t_j}\right]^2 \\
&\qquad + \sum_{j=0}^{n-1} f_{tx}\left(t_j, W_{t_j}\right)\left(t_{j+1}-t_j\right)\left(W_{t_{j+1}}-W_{t_j}\right) \\
&\qquad + \frac{1}{2}\sum_{j=0}^{n-1} f_{tt}\left(t_j, W_{t_j}\right)\left[t_{j+1}-t_j\right]^2 + R
\end{aligned}
$$

As $\|\Pi\| \to 0$, we have

$$
\lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} f_t\left(t_j, W_{t_j}\right)\left(t_{j+1}-t_j\right) = \int_0^T f_t\left(t, W_t\right) dt,
$$

$$
\lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} f_x\left(t_j, W_{t_j}\right)\left(W_{t_{j+1}}-W_{t_j}\right) = \int_0^T f_x\left(t, W_t\right) dW_t,
$$

$$
\lim_{\|\Pi\|\to 0} \frac{1}{2}\sum_{j=0}^{n-1} f_{xx}\left(t_j, W_{t_j}\right)\left[W_{t_{j+1}}-W_{t_j}\right]^2 = \int_0^T f_{xx}\left(t, W_t\right) dt.
$$

Note that, in the third term, it almost like we replace $\left[W_{t_{j+1}}-W_{t_j}\right]^2$ by $\left(t_{j+1}-t_j\right)$. This is where the quadratic variation comes in.

The fourth term and the fifth term all converges to 0: For the fourth term

$$
\begin{aligned}
&\lim_{\|\Pi\|\to 0} \left| \sum_{j=0}^{n-1} f_{tx}\left(t_j, W_{t_j}\right)\left(t_{j+1}-t_j\right)\left(W_{t_{j+1}}-W_{t_j}\right) \right| \\
&\leq \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \left| f_{tx}\left(t_j, W_{t_j}\right)\right| \cdot \left(t_{j+1}-t_j\right) \cdot \left|\left(W_{t_{j+1}}-W_{t_j}\right)\right| \\
&\leq \lim_{\|\Pi\|\to 0} \max_{0\leq k\leq n-1} \left|\left(W_{t_{k+1}}-W_{t_k}\right)\right| \cdot \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \left| f_{tx}\left(t_j, W_{t_j}\right)\right| \cdot \left(t_{j+1}-t_j\right) \\
&= 0 \cdot \int_0^T \left| f_{tx}\left(t, W\left(t\right)\right)\right| dt = 0.
\end{aligned}
$$

The fifth term is treated similarly:

$$\begin{aligned}
& \lim_{\|\Pi\|\to 0}\left|\frac{1}{2}\sum_{j=0}^{n-1} f_{tt}\left(t_j, W_{t_j}\right)\left(t_{j+1}-t_j\right)^2\right| \\
\leq\; & \lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\frac{1}{2}\left|f_{tt}\left(t_j, W_{t_j}\right)\right|\cdot\left(t_{j+1}-t_j\right)^2 \\
\leq\; & \frac{1}{2}\lim_{\|\Pi\|\to 0}\max_{0\leq k\leq n-1}\left|t_{k+1}-t_k\right|\cdot\lim_{\|\Pi\|\to 0}\sum_{j=0}^{n-1}\left|f_{tt}\left(t_j, W_{t_j}\right)\right|\cdot\left(t_{j+1}-t_j\right) \\
=\; & \frac{1}{2}\cdot 0\cdot\int_0^T\left|f_{tt}\left(t, W_t\right)\right|dt = 0.
\end{aligned}$$

The (higher-order) error terms likewise contribute zero to the final answer.■

Differential form of Ito's formula:

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

Note that the differential form's exact meaning is given by the integral form

$$f\left(T, W_T\right) - f\left(0, W_0\right) = \int_0^T f_t\left(t, W_t\right)dt + \int_0^T f_x\left(t, W_t\right)dW_t + \frac{1}{2}\int_0^T f_{xx}\left(t, W_t\right)dt. \tag{4}$$

As a convenience device, people use the following "multiplication table" to help memorize the operation

	dt	$dW(t)$
dt	0	0
$dW(t)$	0	dt

In other words, people write:

$$dW_t dW_t = dt,\; dt dW_t = dW_t dt = 0,\; dt dt = 0.$$

Example: If $f(x) = \frac{1}{2}x^2$, we have

$$\begin{aligned}
\frac{1}{2}W_T^2 &= f\left(W_T\right) - f\left(W_0\right) \\
&= \int_0^T f'\left(W_t\right)dW_t + \frac{1}{2}\int_0^T f''\left(W_t\right)dt \\
&= \int_0^T W_t dW_t + \frac{1}{2}T
\end{aligned}$$

Hence

$$\int_0^T W_t dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}T.$$

Discussions: *It should be noted that, Ito's formula is the key tool for computing the stochastic integrals (as demonstrated by the above example). To do this, you will need to find the appropriate function f such that after applying Ito's formula for $f(W_t)$, the term appears is the stochastic integral you want to compute. If you compare how you compute Riemann integral using Newton-Leibniz formula to stochastic integrals, the requirement for f is the same: it is the anti-derivative of the integrand.*

Note that in some finance books use intuitive treatment of the differential of BM: For example, in Ingersoll (1987), $dW(t)$ is defined as

$$dW(t) = \lim_{\Delta t \to 0} \sqrt{\Delta t}\epsilon, \ \epsilon \sim N(0,1).$$

The above Ito's formula can be extended to the so-called Ito processes.

Ito Processes. Let W_t, $t \geq 0$, be a BM, and let $\mathcal{F}(t)$, $t \geq 0$, be an associated filtration. An Ito process is a stochastic process of the form

$$X_t = X_0 + \int_0^t \sigma_u dW_u + \int_0^t \alpha_u du,$$

where X_0 is nonrandom and α_u and σ_u are adapted stochastic processes.

Lemma. The quadratic variation of the Ito process is

$$[X,X](t) = \int_0^t \sigma_u^2 du.$$

In differential form, we have

$$\begin{aligned} dX_t &= \alpha_t dt + \sigma_t dW_t \\ dX_t dX_t &= \alpha_t^2 dt dt + \sigma_t^2 dW_t dW_t + 2\alpha_t \sigma_t dW_t dt = \sigma_t^2 dt. \end{aligned}$$

Integral with respect to an Ito Process. Let X_t be an Ito process, and let H_t be an adapted process.

We define the integral with respect to an Ito process

$$\int_0^t H_u dX_u = \int_0^t H_u \alpha_u du + \int_0^t H_u \sigma_u dW_u.$$

Remark: What is the quadratic variation of Ito process?

If we treat X_t as the price of an asset and H_t shares invested in the asset, then $\int_0^t H_u dX_u$ can be interpreted as trading gains from the specific trading strategy. Here trading takes place in continuous time.

Theorem (Ito Formula for Ito Process). Let $f(t,x)$ be a function for which the partial derivatives $f_t(t,x)$, $f_x(t,x)$, and $f_{xx}(t,x)$ are defined and continuous, and let X_t be an Ito process. Then for every $T \geq 0$,

$$\begin{aligned} f(T, X_T) &= f(0, X_0) + \int_0^T f_t(t, X_t)\, dt + \int_0^T f_x(t, X_t)\, dX_t \\ &\quad + \frac{1}{2} \int_0^T f_{xx}(t, X_t)\, d[X, X](t) \\ &= f(0, X_0) + \int_0^T f_t(t, X_t)\, dt + \int_0^T f_x(t, X_t)\, \alpha_t dt \\ &\quad + \int_0^T f_x(t, X_t)\, \sigma_t dW_t + \frac{1}{2} \int_0^T f_{xx}(t, X_t)\, \sigma_t^2 dt. \end{aligned}$$

Sketch of Proof.

$$f(T, X_T) - f(0, X_0) = \sum_{j=0}^{n-1} \left[f\left(t_{j+1}, X_{t_{j+1}}\right) - f\left(t_j, X_{t_j}\right) \right]$$

$$\begin{aligned} = \quad & \sum_{j=0}^{n-1} f_t(t_j, X(t_j))(t_{j+1} - t_j) + \sum_{j=0}^{n-1} f_x(t_j, X(t_j))(X(t_{j+1}) - X(t_j)) \\ & + \frac{1}{2} \sum_{j=0}^{n-1} f_{xx}\left(t_j, X_{t_j}\right) \left[X_{t_{j+1}} - X_{t_j}\right]^2 \\ & + \sum_{j=0}^{n-1} f_{tx}\left(t_j, X_{t_j}\right)(t_{j+1} - t_j)\left(X_{t_{j+1}} - X_{t_j}\right) \\ & + \frac{1}{2} \sum_{j=0}^{n-1} f_{tt}\left(t_j, X_{t_j}\right)[t_{j+1} - t_j]^2 + \text{ higher-order terms.} \end{aligned}$$

$$\lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} f_x\left(t_j, X_{t_j}\right)\left(X_{t_{j+1}} - X_{t_j}\right) = \int_0^T f_x(t, X_t)\, dX_t,$$

$$\lim_{\|\Pi\|\to 0}\frac{1}{2}\sum_{j=0}^{n-1} f_{xx}\left(t_j, X_{t_j}\right)\left[X_{t_{j+1}}-X_{t_j}\right]^2=\int_0^T f_{xx}\left(t, X_t\right) d\left[X, X\right](t).$$

In differential form, we have

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

The Ito's formula for Ito processes can be used to evaluate/manipulate stochastic integrals (with respect BM or Ito process). These are standard operations and appear frequently in job interviews.

Example: Compute the stochastic integral

$$\int_0^T \left[\sin(W_t)\right]^2 dW_t.$$

Solutions: We need to consider:

$$f'(x)=\sin^2(x)=\frac{1-\cos(2x)}{2}=\frac{1}{2}-\frac{1}{2}\cos(2x).$$

which means

$$\begin{aligned} f(x) &= \frac{1}{2}x-\frac{1}{4}\sin(2x). \\ f''(x) &= \sin(2x) \\ \frac{1}{2}W_t-\frac{1}{4}\sin(2W_t)-\left[\frac{1}{2}W_0-\frac{1}{4}\sin(2W_0)\right]-\frac{1}{2}\int_0^t \sin(2W_u)du &= \int_0^t \left[\sin(W_u)\right]^2 dW_u \\ \frac{1}{2}W_t-\frac{1}{4}\sin(2W_t)-\frac{1}{2}\int_0^t \sin(2W_u)du &= \int_0^t \left[\sin(W_u)\right]^2 dW_u. \end{aligned}$$

Example: Compute

$$\int_0^T e^{-t}\sin(W_t)dW_t.$$

Solutions: To use Ito's formula, we need to chose $f_x(t,x)$ to be

$$f_x(t,x)=e^{-t}\sin(x)$$

to match the integrans. It follows that

$$f\left(t,x\right)=-e^{-t}\cos(x),$$

and hence

$$\begin{aligned} f_{xx}\left(t,x\right) &= e^{-t}\cos\left(x\right), \\ f_t &= e^{-t}\cos(x). \end{aligned}$$

Now Ito's formula gives

$$\begin{aligned} f\left(T,W_T\right)-f\left(0,W_0\right) &= \int_0^T f_t\left(t,W_t\right)dt+\int_0^T f_x\left(t,W_t\right)dW_t+\frac{1}{2} \\ -e^{-T}\cos(W_T)-\left[-e^{-0}\cos(W_0)\right] &= \int_0^T e^{-t}\cos(W_t)dt+\left[\int_0^T e^{-t}\sin(W_t)dW_t\right. \\ -e^{-T}\cos(W_T)-\left[-e^{-0}\cos(W_0)\right]-\int_0^T e^{-t}\cos(W_t)dt-\frac{1}{2}\int_0^T e^{-t}\cos\left(W_t\right)dt &= \left[\int_0^T e^{-t}\sin(W_t)dW_t\right] \\ -e^{-T}\cos(W_T)+1-\int_0^T e^{-t}\cos(W_t)dt-\frac{1}{2}\int_0^T e^{-t}\cos\left(W_t\right)dt &= \left[\int_0^T e^{-t}\sin(W_t)dW_t\right] \\ -e^{-T}\cos(W_T)+1-\frac{3}{2}\int_0^T e^{-t}\cos(W_t)dt &= \left[\int_0^T e^{-t}\sin(W_t)dW_t\right]. \end{aligned}$$

Example: Compute

$$\mathbb{E}\left[\int_0^T\cos(W_t)dW_t\right]^3.$$

Solutions: This represents one of the more difficult job interview questions. The key is to represent the quantity inside the expectation using Ito's integral. Set

$$Y_t=\int_0^t\cos\left(W_s\right)dW_s.$$

The above is equivalent to

$$dY_t=\cos\left(W_t\right)dW_t.$$

To apply Ito, we define $f(x) = x^3$, and

$$Z_t = f(Y_t).$$

Ito's formula (for Ito diffusion) gives

$$\begin{aligned} dZ_t &= f_x\left(Y_t\right)dY_t \ + \frac{1}{2}f_{xx}\left(Y_t\right)\left(dY_t\right)^2 \\ &= 3Y_t^2\cos\left(W_t\right)dW_t + \frac{1}{2}6Y_t\left[\cos\left(W_t\right)\right]^2 dt. \end{aligned}$$

The integral form is:

$$Y_t^3 - Y_0^3 = \int_0^t 3Y_s^2\cos\left(W_s\right)dW_s + \int_0^t \frac{1}{2}6Y_s\left[\cos\left(W_s\right)\right]^2 ds.$$

It follows that (note that $Y_0 = 0$)

$$\begin{aligned} \mathbb{E}Y_t^3 &= \mathbb{E}\left[\int_0^t 3Y_s^2\cos\left(W_s\right)dW_s + \int_0^t 3Y_s\left[\cos\left(W_s\right)\right]^2 ds\right] \\ &= \int_0^t 3\mathbb{E}\left\{Y_s\left[\cos\left(W_s\right)\right]^2\right\}ds. \end{aligned}$$

Note: sometimes you can actually compute the expectation inside, for some problems.

Example: Let X_t satisfy

$$dX_s = \alpha X_s ds + \beta X_s dW_s.$$

Find an expression for X_t. [Note: this example is important for finance applications]

1.5 Applications of Ito's Formula

Geometric Brownian Motion:

$$dS_t = \alpha_t S_t dt + \sigma_t S_t dW_t.$$

Consider a function $f\left(S_t\right) = \log\left(S_t\right)$, then

$$\begin{aligned} df\left(S_t\right) &= f'\left(S_t\right)dS_t + \frac{1}{2}f''\left(S_t\right)dS_t dS_t = \left[\alpha_t - \frac{1}{2}\sigma_t^2\right]dt + \sigma_t dW_t \\ f\left(S_T\right) - f\left(S_0\right) &= \log\frac{S_T}{S_0} = \int_0^T\left[\alpha_t - \frac{1}{2}\sigma_t^2\right]dt + \int_0^T \sigma_t dW_t \\ S_T &= S_0\exp\left\{\int_0^T\left[\alpha_t - \frac{1}{2}\sigma_t^2\right]dt + \int_0^T \sigma_t dW_t\right\}. \end{aligned}$$

For constant α and σ, we have$S_T = S_0\exp\left\{\left(\alpha - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right\}.$

Definition of Stochastic Differentiations. For a given pair of functions $\mu(t,x)$ and $\sigma(t,x)$, stochastic differentiation referes equations of the following form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{5}$$

with the initial condition

$$X_0 = x_0.$$

A solution to the stochastic differential equation above means an adaptive process X_t satisfying the above equation.

Important: Recall that the meaning of differentials are defined by the corresponding integrals. So solving SDE (5) is to solve the corresponding Stochastic Integral Equation:

$$X_t - X_0 = \int_0^t \mu\left(s, X_s\right) ds + \int_0^t \sigma(s, X_s)dW_s.$$

The key to solve this equation is to eliminate the appearance of X_t on the right-hand side.

Ornstein-Uhlenbeck Process:

$$dX_t = \left(\alpha - \beta X_t\right) dt + \sigma dW_t.$$

This process has been used to model interest rate behavior by Vasicek.

Define $f\left(t, X_t\right) = e^{\beta t} X_t$, then

$$\begin{aligned} df\left(t, X_t\right) &= \beta e^{\beta t} X_t dt + e^{\beta t} dX_t = \beta e^{\beta t} X_t dt + e^{\beta t}\left(\alpha - \beta X_t\right) dt + e^{\beta t}\sigma dW_t \\ &= \alpha e^{\beta t} dt + e^{\beta t}\sigma dW_t \end{aligned}$$

$$\begin{aligned} f\left(t, X_t\right) &= f\left(0, X_0\right) + \int_0^t \alpha e^{\beta t} du + \int_0^t e^{\beta t}\sigma dW_u \\ e^{\beta t} X_t &= X_0 + \alpha \frac{1}{\beta} e^{\beta u}|_0^t + \int_0^t e^{\beta u}\sigma dW_u \\ X_t &= e^{-\beta t} X_0 + \frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right) + \sigma e^{-\beta t}\int_0^t e^{\beta u} dW_u \\ \mathbb{E}\left[X_t\right] &= e^{-\beta t} X_0 + \frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right). \end{aligned}$$

Very often, Ito's formula can be used to calculate the moments of Ito processes, as the next example shows.

Important: When solving SDE, the stochastic integrals NOT INVOLVING the unknown process X_t is considered DONE, just as when computing Stochastic Integrals the Riemann Integrals is considered done.

Cox-Ingersoll-Ross Model: A popular model for interest rate

$$dX_t = (\alpha - \beta X_t)\, dt + \sigma \sqrt{X_t} dW_t.$$

Define $Y_t = f(t, X_t) = e^{\beta t} X_t$. Then

$$\begin{aligned} df(t, X_t) &= f_t(t, X_t)\, dt + f_x(t, X_t)\, dX_t + \frac{1}{2} f_{xx}(t, X_t)\, dX_t dX_t \\ &= \alpha e^{\beta t} dt + \sigma e^{\beta t} \sqrt{X_t} dW_t. \end{aligned}$$

Integration of both sides yields

$$\begin{aligned} f(t, X_t) &= e^{\beta t} X_t = X_0 + \int_0^t \alpha e^{\beta u} du + \int_0^t \sigma e^{\beta u} \sqrt{X_u} dW_u \\ &\implies \mathbb{E}[X_t] = e^{-\beta t} X_0 + \frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right). \end{aligned}$$

$$dY_t = df(t, X_t) = \alpha e^{\beta t} dt + \sigma e^{\frac{\beta t}{2}} \sqrt{e^{\beta t} X_t} dW_t = \alpha e^{\beta t} dt + \sigma e^{\frac{\beta t}{2}} \sqrt{Y_t} dW_t,$$

$$\begin{aligned} d\left[Y_t^2\right] &= 2Y_t dY_t + \frac{1}{2} 2 dY_t dY_t = 2Y_t \left[\alpha e^{\beta t} dt + \sigma e^{\frac{\beta t}{2}} \sqrt{Y_t} dW_t\right] \\ &\quad + \left[\sigma e^{\frac{\beta t}{2}} \sqrt{Y_t} dW_t\right]\left[\sigma e^{\frac{\beta t}{2}} \sqrt{Y_t} dW_t\right] \\ &= 2\alpha e^{\beta t} Y_t dt + 2\sigma e^{\frac{\beta t}{2}} Y_t^{\frac{3}{2}} dW_t + \sigma^2 e^{\beta t} Y_t dt. \end{aligned}$$

Therefore,

$$Y_t^2 = Y_0^2 + \left(2\alpha + \sigma^2\right) \int_0^t e^{\beta u} Y_u du + 2\sigma \int_0^t e^{\frac{\beta u}{2}} Y_u^{\frac{3}{2}} dW_u,$$

$$\begin{aligned} \mathbb{E}\left[Y_t^2\right] &= Y_0^2 + \left(2\alpha + \sigma^2\right) \int_0^t e^{\beta u} \mathbb{E}[Y_u]\, du \\ &= X_0^2 + \left(2\alpha + \sigma^2\right) \int_0^t e^{\beta u} \left[X_0 + \frac{\alpha}{\beta}\left(e^{\beta u} - 1\right)\right] du \\ &= X_0^2 + \frac{2\alpha + \sigma^2}{\beta}\left(X_0 - \frac{\alpha}{\beta}\right)\left(e^{\beta t} - 1\right) + \frac{2\alpha + \sigma^2}{2\beta} \cdot \frac{\alpha}{\beta}\left(e^{2\beta t} - 1\right). \end{aligned}$$

$$\begin{aligned} \mathbb{E}\left[X_t^2\right] &= e^{-2\beta t} \mathbb{E}\left[Y_t^2\right] = e^{-2\beta t} X_0^2 + \frac{2\alpha + \sigma^2}{\beta}\left(X_0 - \frac{\alpha}{\beta}\right)\left(e^{-\beta t} - e^{-2\beta t}\right) \\ &\quad + \frac{2\alpha + \sigma^2}{2\beta} \cdot \frac{\alpha}{\beta}\left(1 - e^{-2\beta t}\right). \end{aligned}$$

$$\begin{aligned} Var\left[X_t\right] &= \mathbb{E}\left[X_t^2\right] - \left(\mathbb{E}\left[X_t\right]\right)^2 \\ &= \frac{\sigma^2}{\beta} X_0 \left(e^{-\beta t} - e^{-2\beta t}\right) + \frac{\alpha\sigma^2}{2\beta^2}\left(1 - 2e^{-\beta t} + e^{-2\beta t}\right). \end{aligned}$$

In particular,

$$\lim_{t\to\infty} Var\left[X_t\right] = \frac{\alpha\sigma^2}{2\beta^2}.$$

The Black-Scholes Option Pricing Model

- Securities traded in the model
  - A riskless bond, with a price process
    $$dB_t = rB_t dt,$$
    where r is the *continuously compounded* interest rate. The solution to the above ODE with initial condition B_0 is $B_t = B_0 e^{rt}$.
  - A stock whose price follows Geometric BM:
    $$dS_t = \mu S_t dt + \sigma S_t dW_t.$$
  - A European call option with payoff $Y_T = (S_T - K)^+$.
- Assumptions: Perfect market, no taxes, transactions costs, no short sale restrictions, and all assets are perfectly divisible
- Goal: To price the European option based on the no-arbitrage pricing principle
- Understanding stock price dynamics
  - The BS model assumes the geometric BM. Why not Brownian motion with drift? The reason is that
    $$dS_t = \mu dt + \sigma dW_t$$
    $$\begin{aligned} \int_0^t dS_u &= \int_0^t \mu du + \int_0^t \sigma dW_u \\ S_t - S_0 &= \mu t + \sigma W_t \Rightarrow S_t \sim N\left(S_0 + \mu t, \sigma^2 t\right), \end{aligned}$$
    Undesirable, because S_t can take negative values
  - For a geometric Brownian motion
    $$dS_t = \mu S_t dt + \sigma S_t dW_t,$$
    Consider the dynamics of $\log S_t$. By Ito's lemma, we have
    $$\begin{aligned} d \log S_t &= \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t \\ \log S_t &\sim N\left(\log S_0 + \left(\mu - \frac{1}{2}\sigma^2\right) t, \sigma^2 t\right) \end{aligned}$$

Deriving the Black-Scholes PDE (Method I):

The key idea is that we can create an option by *continuously trading stock and bond*.

Assume we have found the option pricing formula $C(S,t)$ for which $C(S_t,t)$ will give the current price of the option at time t. We look for a continuous trading strategy (a,b) of stock and bond that produces a portolio whose value Y_t always matches the options value $C(S_t,t)$ at all time t.

- The strategy (a,b) means that

$$a_t S_t + b_t B_t = Y_t.$$

When this is true, we have

$$dY_t = a_t dS_t + b_t dB_t = (a_t \mu S_t + b_t B_t r)\, dt + a_t \sigma S_t dW_t. \tag{6}$$

- On the other hand, Ito's formula gives us the dynamics of the above value from the option's side:

$$dY_t = \left[C_t(S_t,t) + C_S(S_t,t)\,\mu S_t + \frac{1}{2} C_{SS}(S_t,t)\,\sigma^2 S_t^2 \right] dt + C_S(S_t,t)\,\sigma S_t dW_t. \tag{7}$$

To show that indeed we can have strategy (a,b) exist, we can match the coefficients in both dt and dW_t terms for the two expressions of the dynamics of $Y_t = C(S_t,t)$:

For dW_t term:

$$a_t \sigma S_t = C_S(S_t,t)\,\sigma S_t,$$

which yields

$$a_t = C_S(S_t,t).$$

For dt term, we have

$$a_t \mu S_t + b_t B_t r = C_t(S_t,t) + C_S(S_t,t)\,\mu S_t + \frac{1}{2} C_{SS}(S_t,t)\,\sigma^2 S_t^2.$$

Using the expression for a_t, we get

$$C_S(S_t,t)\,\mu S_t + b_t B_t r = C_t(S_t,t) + C_S(S_t,t)\,\mu S_t + \frac{1}{2} C_{SS}(S_t,t)\,\sigma^2 S_t^2,$$

so we get b_t as

$$b_t = [B_t r]^{-1}\left[C_t(S_t,t) + \frac{1}{2}C_{SS}(S_t,t)\,\sigma^2 S_t^2\right].$$

Finally, note that the option price satisfies

$$C(S_t,t) = a_t S_t + b_t B_t.$$

Hence this expression and the one above yields

$$C(S_t,t) = C_S(S_t,t)\,S_t + r^{-1}\left[C_t(S_t,t) + \frac{1}{2}C_{SS}(S_t,t)\,\sigma^2 S_t^2\right]$$

The above equation leads to the so-called Black-Scholes PDE

$$C_t(S,t) + rSC_S(S,t) + \frac{1}{2}\sigma^2 S^2 C_{SS}(S,t) \ = rC(S,t),$$

or more compactly

$$C_t + rSC_S + \frac{1}{2}\sigma^2 S^2 C_{SS} \ = rC. \tag{8}$$

This PDE has the boundary (or so-called terminal condition) condition, due to the European call option's payoff function

$$C(S,T) = (S-K)^+. \tag{9}$$

To summarize, we have used the replication (or dynamic trading) to derive the PDE that the European call option has to satisfy. That is, we have converted a finance problem into a PDE problem. To derive the celebrated Black-Scholes option pricing formula, all we need to solve mathematically the above PDE (8) with the appropriate boundary condition (9).

Now we state the Black-Scholes formula, for which we will give several derivations in the rest of this class. The solution to the PDE is (S denotes S_t for simplicity)

$$\begin{aligned} C(S,t) &= SN(d_1) - Ke^{-r(T-t)}N(d_2), \\ d_{1,2} &= \frac{\ln(S/K) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \end{aligned} \tag{10}$$

w here $N(\cdot)$ is normal CDF.

- Note that the expected stock return, μ, does not appear in the Black-Scholes formula.

Solving the Black-Scholes PDE (Method 1)

The Black-Scholes PDE (8) with the boundary condition (9) can be solved by transforming it into the well-known heat equations (which we have discussed in the Review class). Here is the sketch on how to do it.

(1). We can use the log transformation to eliminate the the appearance of S and S^2 in the coefficients of C_S and C_{SS}. This will help to convert the Black-Scholes PDE into a PDE with *constant coefficients*. More specifically, we can define a new variable $x = \log\left(\frac{S}{K}\right)$, and $\tau = T - t$, and define

$$Z(x,\tau) = C(Ke^x, T-\tau).$$

This will convert the (8) and (9) into

$$\frac{\partial Z}{\partial x} - \frac{1}{2}\sigma^2 \frac{\partial^2 Z}{\partial x^2} + \left(\frac{\sigma^2}{2} - r\right)\frac{\partial Z}{\partial x} + rZ = 0 \tag{11}$$

and

$$Z(x,0) = C\left(Ke^x, T\right). \tag{12}$$

(2). The above PDE (11) and (12) is much closer to the heat equation in physics. It can be converted into the standard heat equation by another transformation. We can define a new function

$$u(x,\tau) = e^{\alpha x + \beta\tau} Z(x,\tau),$$

for appropprioately chosen α and β. With this new function, we can express the PDE (11) and (12) as

$$\frac{\partial u}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} + A\frac{\partial u}{\partial x} + Bu = 0 \tag{13}$$

and

$$u(x,0) = e^{\alpha x} Z(x,0) = e^{\alpha x} C(Ke^x, T), \tag{14}$$

where

$$\begin{aligned} A &= \alpha\sigma^2 + \frac{\sigma^2}{2} - r \\ B &= (1+\alpha)r - B - \frac{\alpha^2\sigma^2 + \alpha\sigma^2}{2}. \end{aligned}$$

We can choose α and β to make $A = B = 0$:

$$\alpha = \frac{r}{\sigma^2} - \frac{1}{2}, \beta = \frac{r}{2} + \frac{\sigma^2}{8} + \frac{r^2}{2\sigma^2}.$$

Such choices leads to the following standard PDE

$$\frac{\partial u}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} = 0, \tag{15}$$

with the boundary condition

$$u(x,0) = e^{\alpha x} Z\left(x,0\right) = e^{(\frac{r}{\sigma^2}-\frac{1}{2})x} C(Ke^x, T). \tag{16}$$

(3). The above heat equation (15) has the standard solution (from the Review class) given by

$$u\left(x,\tau\right) = \frac{1}{\sqrt{2\sigma^2\pi\tau}}\int_{-\infty}^{\infty} e^{-\frac{(x-s)^2}{2\sigma^2\tau}} u\left(s,0\right) ds.$$

Carrying out this integral using (16) gives the celebrated Black-Scholes option pricing formula (10)! To this end, just note that

$$\begin{aligned} u(x,0) &= e^{(\frac{r}{\sigma^2}-\frac{1}{2})x} C(Ke^x, T) \\ &= e^{(\frac{r}{\sigma^2}-\frac{1}{2})x} \left(Ke^x - K\right)^+ \\ &= \begin{cases} e^{(\frac{r}{\sigma^2}-\frac{1}{2})x} K\left(e^x - 1\right) & \text{for } x > 0 \\ 0 & \text{else} \end{cases} \end{aligned}$$

Feyman-Kac Theorem

- The well-known Feyman-Kac Theorem gives stochastic representation for the solution to the Black-Scholes type PDE. It provides a link between PDE and SDE. This result is very important for finance applications.

- We now assume a slightly more general stock dynamics thank in BS model. Suppose the stock price is given by the following SDE

$$dS_t = \mu\left(S_t,t\right)dt + \sigma\left(S_t,t\right)dW_t$$

and

$$B_t = B_0 \exp\left[\int_0^t r\left(S_u,u\right)du\right].$$

- Consider a derivative security with payoff $Y_T = g\left(S_T\right)$. Then the Black-Scholes' argument shows that the price function $C(x,t)$ for which $Y_t = C\left(S_t,t\right)$ solves the following PDE

$$C_t\left(x,t\right) + C_x\left(x,t\right) r\left(x,t\right) + \frac{1}{2} C_{xx}\left(x,t\right) \sigma^2\left(x,t\right) = r\left(x,t\right) C\left(x,t\right)$$

with the (terminal) boundary condition

$$C(x, T) = g(x).$$

- Feyman-Kac formula relates SDE to PDE
- A SDE is an equation of the form

$$dX_u = \mu(u, X_u)\, du + \sigma(u, X_u)\, dW_u,\ \ X_t = x$$

where $\mu(u, x)$ and $\sigma(u, x)$ are given functions, called the drift and diffusion, respectively

The solution to the SDE is a stochastic process X_T, defined for $T \geq t$, such that

$$X_t = x,$$
$$X_T = X_t + \int_t^T \mu(u, X_u)\, du + \int_t^T \sigma(u, X_u)\, dW_u.$$

Under mild conditions on the functions $\mu(u, X_u)$ and $\sigma(u, X_u)$, there exists a unique process X_T, $T \geq t$, satisfying the above two equation. However, this process can be difficult to determine explicitly because it appears on both the left- and right-hand sides of the equation.

Theorem (Feynman-Kac, Simple). Consider the stochastic differential equation

$$dX_u = \mu(u, X_u)\, du + \sigma(u, X_u)\, dW_u.$$

Let $h(y)$ be a Borel-measurable function. Fix $T > 0$, and let $t \in [0, T]$ be given. Define the function

$$g(t, x) = \mathbb{E}^{t,x}[h(X_T)],$$

the expectation of $h(X_T)$, where X_T solves the above SDE with initial condition $X_t = x$ (We assume that $\mathbb{E}^{t,x}|h(X(T))| < \infty$ for all t and x.) Then $g(t, x)$ satisfies the partial differential equation

$$g_t(t, x) + \mu(t, x)\, g_x(t, x) + \frac{1}{2}\sigma^2(t, x)\, g_{xx}(t, x) = 0$$

and the terminal condition $g(T, x) = h(x)$ for all x.

- **Theorem (Discounted Feynman-Kac).** Consider the stochastic differential equation

$$dX_u = \mu(u, X_u)\, du + \sigma(u, X_u)\, dW_u.$$

Let $h(y)$ be a Borel-measurable function and let r be constant. Fix $T > 0$, and let $t \in [0, T]$ be given. Define the function

$$f(t,x) = \mathbb{E}^{t,x}\left[e^{-r(T-t)}h(X_T)\right].$$

(We assume that $\mathbb{E}^{t,x}|h(X_T)| < \infty$ for all t and x.) Then $f(t,x)$ satisfies the partial differential equation

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

and the terminal condition

$$f(T,x) = h(x) \text{ for all } x.$$

- Before turning to a proof of Feyman-Kac formula, let's first use it to solve the Black-Scholes PDE to obtain the Black-Scholes option formula.

By the Fayman-Kac formula, the solution to the Black-Scholes PDE for european call option price is represented by

$$C(S,t) = \mathbb{E}^{S,t}\left[e^{-r(T-t)}(S(T)-K)^+\right] \tag{17}$$

where S_t is a stochastic process satisfying

$$dS_t = rS_t dt + \sigma S_t dW_t.$$

Remark: By Feyman-Kac Theorem, this S_t does not have to be related to the original S_t which represents the stock price, and we could have used a different symbol, say X_t.

Note that this process S_t is a geometric BM, and the above expectation (17) can be evaluated explicitely. To this end, set $Y_t = \log S_t$, and the Y_t follows

$$dY_t = (r - \frac{1}{2}\sigma^2)dt + \sigma W_t$$

Or, equivalently

$$Y_T - Y_t = (r - \frac{1}{2}\sigma^2)(T-t) + \sigma[W_T - W_t]$$

This means that Y_T has normal distribution with mean $\log S + (r - \frac{1}{2}\sigma^2)(T-t)$, and variance of $\sigma^2(T-t)$ [here Z is the standard normal random variable]:

$$Y_T \sim \log S + (r - \frac{1}{2}\sigma^2)(T-t) - \sigma\sqrt{T-t}Z$$

It follows that

$$
\begin{aligned}
& \mathbb{E}^{S,t}\left[e^{-r(t-T)}(S_T-K)^+\right] \\
= \;& e^{-r(T-t)}\mathbb{E}^{\log(S),t}(\exp(Y_T)-K)^+ \\
= \;& e^{-r(T-t)}\int\left[e^{\left[\log S+(r-\frac{1}{2}\sigma^2)(T-t)\right]-\sigma\sqrt{T-t}z}-K\right]^+\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\
= \;& e^{-r(T-t)}\int_{z<d^-}\left(e^{\left[\log S+(r-\frac{1}{2}\sigma^2)(T-t)\right]-\sigma\sqrt{T-t}z}-K\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\
= \;& e^{-r(T-t)}\int_{z<d^-}e^{\left[\log S+(r-\frac{1}{2}\sigma^2)(T-t)\right]-\sigma\sqrt{T-t}z}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\
& -Ke^{-r(T-t)}\int_{z<d^-}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz.
\end{aligned}
$$

It all amounts to evaluate the two integrals.

First note that the second integral gives the desired second term in BS-formula. The first integral is

$$
\begin{aligned}
& e^{-r(T-t)}\int_{z<d^-}e^{\left[\log S+(r-\frac{1}{2}\sigma^2)(T-t)\right]-\sigma\sqrt{T-t}z}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\
= \;& e^{-r(T-t)}\int_{z<d^-}\frac{1}{\sqrt{2\pi}}e^{\left[\log S+(r-\frac{1}{2}\sigma^2)(T-t)\right]-\sigma\sqrt{T-t}z-\frac{z^2}{2}}dz
\end{aligned}
$$

The exponential in the integrand is (just after completing the square):

$$
e^{\left[\log S+(r-\frac{1}{2}\sigma^2)(T-t)\right]-\sigma\sqrt{T-t}z-\frac{z^2}{2}}=e^{\left[\log S+r(T-t)\right]-\frac{(z+\sigma\sqrt{T-t})^2}{2}}.
$$

Hence the first term is

$$
\begin{aligned}
& e^{-r(T-t)}\int_{z<d^-}\frac{1}{\sqrt{2\pi}}e^{\left[\log S+r(T-t)\right]-\frac{(z+\sigma\sqrt{T-t})^2}{2}}\;dz \\
= \;& e^{-r(T-t)}\int_{y<d^-+\sigma\sqrt{T-t}}\frac{1}{\sqrt{2\pi}}e^{\left[\log S+r(T-t)\right]-\frac{y^2}{2}}\;dy \\
= \;& e^{-r(T-t)}e^{\left[\log S+r(T-t)\right]}\int_{z<d^+}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\;dy \\
= \;& SN(d^+).
\end{aligned}
$$

This completes the derivation of the Black-Scholes Option Pricing Formula

$$
C(S,t)=SN(d^+)-Ke^{-r(T-t)}N(d^-),
$$

where

$$
d^{\pm}=\frac{\log(S/K)+(r\pm\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.
$$

Now we turn to the proof of the two Feyman-Kac Theorems.

- **Sketch of Proof of Feyman-Kac (Simple).** First we note that $g(t, X_t)$ is a martingale. Given $0 \leq s \leq t \leq T$, then

$$\mathbb{E}[g(t, X_t) | \mathcal{F}_s] = \mathbb{E}[\mathbb{E}[h(X_T) | \mathcal{F}_t] | \mathcal{F}_s] = \mathbb{E}[h(X_T) | \mathcal{F}_s] = g(s, X_s).$$

Because $g(t, X_t)$ is a martingale, then the net dt term in the differential $dg(t, X_t)$ must be zero. We know

$$\begin{aligned} dg(t, X_t) &= g_t dt + g_x dX_t + \frac{1}{2} g_{xx} dX_t dX_t = g_t dt + g_x [\mu dt + \sigma dW_t] + \frac{1}{2} g_{xx} \sigma^2 dt \\ &= \left[g_t + \mu g_x + \frac{1}{2} \sigma^2 g_{xx} \right] dt + \sigma g_x dW. \end{aligned}$$

Setting the dt term to zero and putting back the argument (t, X_t), we obtain

$$g_t(t, X_t) + \mu(t, X_t) g_x(t, X_t) + \frac{1}{2} \sigma^2(t, X_t) g_{xx}(t, X_t) = 0$$

along every path of X. Therefore,

$$g_t(t, x) + \mu(t, x) g_x(t, x) + \frac{1}{2} \sigma^2(t, x) g_{xx}(t, x) = 0$$

at every point (t, x) that can be reached by (t, X_t).

- The general principle behind the proof of the Feynman-Kac theorem is: (i) find the martingale; (ii) take the differential; and (iii) set the dt term equal to zero.

Sketch of Proof of Discounted Feyman-Kac Theorem. Let X_t be the solution to the above SDE starting at time zero. Then

$$f(t, X_t) = \mathbb{E}\left[e^{-r(T-t)} h(X_T) | \mathcal{F}_t \right].$$

However, it is not the case that $f(t, X_t)$ is a martingale. Indeed, if $0 \leq s \leq t \leq T$, then

$$\begin{aligned} & \mathbb{E}[f(t, X_t) | \mathcal{F}_s] = \mathbb{E}\left[\mathbb{E}\left[e^{-r(T-t)} h(X_T) | \mathcal{F}_t \right] | \mathcal{F}_s \right] \\ = \quad & \mathbb{E}\left[e^{-r(T-t)} h(X_T) | \mathcal{F}_s \right], \end{aligned}$$

which is not the same as

$$f(s, X_s) = \mathbb{E}\left[e^{-r(T-s)} h(X_T) | \mathcal{F}_s \right].$$

because of the differing discount terms.

- Instead, we have

$$e^{-rt} f(t, X_t) = \mathbb{E}\left[e^{-rT} h(X_T) | \mathcal{F}_t \right].$$

Then

$$
\begin{aligned}
& \mathbb{E}\left[e^{-rt} f\left(t, X_t\right) | \mathcal{F}_s\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-rT} h\left(X_T\right) | \mathcal{F}_t\right] | \mathcal{F}_s\right] \\
= \; & \mathbb{E}\left[e^{-rs}\mathbb{E}\left[e^{-r(T-s)} h\left(X_T\right) | \mathcal{F}_t\right] | \mathcal{F}_s\right] \\
= \; & e^{-rs}\mathbb{E}\left[e^{-r(T-s)} h\left(X_T\right) | \mathcal{F}_s\right] = e^{-rs} f\left(s, X_s\right).
\end{aligned}
$$

As a result,

$$
\begin{aligned}
& d\left[e^{-rt} f\left(t, X_t\right)\right] \\
= \; & e^{-rt}\left[-r\right] f dt + e^{-rt} f_t dt + e^{-rt} f_x dX_t + e^{-rt}\frac{1}{2} f_{xx} dX_t dX_t \\
= \; & e^{-rt}\left[\left(-r\right) f dt + f_t dt + f_x\left(\mu dt + \sigma dW_t\right) + \frac{1}{2} f_{xx}\sigma^2 dt\right] \\
= \; & e^{-rt}\left[-rf + f_t + \mu f_x + \frac{1}{2} f_{xx}\sigma^2\right] dt + e^{-rt} f_x \sigma dW.
\end{aligned}
$$

Setting the dt term to zero, we obtain the desired PDE.

Relationship to risk-neutral pricing and change of measure.

2.3 Risk-Neutral Pricing and Girsanov Theorem

Risk-neutral pricing is also called martingale pricing.

Upto this point, we have derived the Black-Scholes pricing formula for Euroeapn call options. The entire derivation can be summarized as follows:

Step 1. The stock follows

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t$$

- Assume the option price formula is given by $C(S_t, t)$ for a function $C(x, t)$. Also assume the option can be replicated by a (dynamically traded) portfolio of stock and bond using self-financing strategy a_t and b_t.

Step 2. Use Ito's formula on the option price $C(S_t, t)$ and the replication portfolio to obtain their dynamics. Match the two coefficients for the dt and dW_t terms. This will give you two equations to determine a_t and b_t. The fact that a_t and b_t exist means replication is possible!

Step 3. The trading strategy a_t and b_t yields the Black-Scholes PDE

$$C_t(x,t) + r(x,t)C_x(x,t) + \frac{1}{2}\sigma^2(x,t)C_{xx}(x,t) = r(x,t)C(x,t)$$

with the (terminal) boundary condition

$$C(x,T) = g(x).$$

Step 4. Solve the above Black-Scholes PDE. This can be done use pure PDE methods, or use Feyman-Kac Theorem. The Feyman-Kac method says that, we can read off from the PDE the two coefficients to construct auxilary process X_t :

$$dX_t = r(X_t,t)dt + \sigma(X_t,t)dW_t.$$

Then the solutions to the pricing PDE can be given by a conditional expectation of X_T

$$\begin{aligned} C(x,t) &\equiv \mathbb{E}^{x,t}\left[e^{-r(X_t,t)(T-t)}(X_T-K)^+\right] \\ &\equiv \mathbb{E}\left[e^{-r(X_t,t)(T-t)}(X_T-K)^+|X_t=x\right]. \end{aligned}$$

Step 5. The above expectation can be evaluated easily to yield the pricing formula for some choices of μ and σ, as in Black-Scholes formula.

The ideas behind the above procedures are generalized to martingale (risk-neutral) pricing.

The Basic Framework of Martingale (Risk-Neutral) Pricing

Let V_T be an $\mathcal{F}_T$-measurable random variable and represents the payoff at T of a derivative security. We allow stochastic interest rate r_t and let D_t denote the discount process by $dD_t = -r_tD_tdt$.

The martingale pricing methodology consists of:

(i). Finding a measure Q such that the *discounted stock price* D_tS_t is a martingale under Q.

(ii). Then the pricing is done by martingale property under Q:

$$\begin{aligned} D_tV_t &= \mathbb{E}^Q\left[D_TV_T|\mathcal{F}_t\right] \\ V_t &= \mathbb{E}^Q\left[\frac{D_T}{D_t}V_T\middle|\mathcal{F}_t\right] \end{aligned}$$

$$V_t = \mathbb{E}^Q \left[e^{-\int_t^T r_s ds} V_T \middle| \mathcal{F}_t \right]$$

This method is called martingale (or risk-neutral) pricing.

Remark: *The motivation for the above framework can be seen by examining carefully our derivation of Black-Scholes formula using replication portfolios. Assume that we can find a replication portfolion X_t of stock and bonds using self-financing strategies for the derivative with payoff V_T at time T. Then the fact that $D_t X_t$ is a martingale under some measure Q implies*

$$D_t X_t = \mathbb{E}^Q \left[D_T X_T | \mathcal{F}_t \right] = \mathbb{E}^Q \left[D_T V_T | \mathcal{F}_t \right].$$

The value X_t of the replicating portfolio is the capital needed at time t to successfully complete the hedge of the short position in the derivative security with payoff V_T. Hence we call this the price V_t of the derivative security at time t, and obtain from above

$$V_t = \mathbb{E}^Q \left[\frac{D_T}{D_t} V_T | \mathcal{F}_t \right].$$

All this is nice, BUT, unless we can do the following the framework is not operational:

- How to find the Q measure?
- How to compute the expectations under Q?
- How the P and Q measures are related?

The key is Girsanov's Theorem on changes of measure (sometimes called Cameron-Martin-Girsanov's Theorem).

To see why changing the underlying probability measure can change the dynamics/distribution of a random variable/process, consider the following example:

Let (Ω, F, P) be the probability on which we have a Brownian Motion W_t :

$\Omega = \mathbb{R}$

F = filtration generated by W_t

$P =$ a probability measure that is standard normal distribution with density $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$.

To be specific, consider the random variable (the case of stochastic process is similar) X, defined as

$$X(\omega) = \omega.$$

The distribution for X is then

$$\begin{aligned} F(a) &= P(\{\omega : X(\omega) < a\}) \\ &= P(\{\omega : \omega < a\}) \\ &= \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= N(0,1). \end{aligned}$$

It is also easy to verify that

$$X(\omega) + 2 \text{ has distribution of } N(2,1).$$

If we change P to Q with $\frac{1}{\sqrt{2\pi}}e^{-(x-2)^2/2}$, then, under Q, $X(\omega) + 2$

$$\begin{aligned} Q(a) &= Q(\{\omega : X(\omega) + 2 < a\}) \\ &= Q(\{\omega : \omega < a - 2\}) \\ &= \int_{-\infty}^{a-2} \frac{1}{\sqrt{2\pi}} e^{-(x-2)^2/2} dx \\ &= \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-(y)^2/2} dy \\ &= N(0,1). \end{aligned}$$

Hence switching to measure from P to Q can kill the mean of a random variable $X + 2$. Girsanov's Theorem is just extending this observation to the case of stochastic processes.

Girsanov Theorem [one dimensional]. *Let W_t, $0 \le t \le T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}_t$ be a filtration for the Brownian motion. Let η_t be an adapted process. Define*

$$\xi_t = \exp\left\{-\int_0^t \eta_s dW_s - \frac{1}{2}\int_0^t \eta_s^2 ds\right\}$$

and assume that

$$\mathbb{E}\int_0^T \eta_s^2 \xi_s^2 ds < \infty.$$

Set $\xi = \xi_T$. *Then* $\mathbb{E}[\xi] = 1$. Define Q *by*

$$Q(A) = \int_A \xi(\omega)\, dP(\omega).$$

Then the process $\tilde{W}_t$ *defined by*

$$\tilde{W}_t = \int_0^t \eta_s ds + W_t \text{ [Note: differential form: } d\widetilde{W_t} = \eta_t dt + dW_t\text{]}$$

is a Brownian motion under Q*.*

An Application of Girsanov's Theorem.

- Consider a stock price process (extended geometric BM), under measure P,

$$dS_t = \alpha_t S_t dt + \sigma_t S_t dW_t.$$

  Note: this SDE has explicite solution

$$S_t = S_0 \exp\left\{\int_0^t \sigma_s dW_s + \int_0^t \left[\alpha_s - \frac{1}{2}\sigma_s^2\right] ds\right\}.$$

- Let r_t be the riskless interest rate and define the discount process

$$D_t = e^{-\int_0^t r_s ds}$$

  and note that $dD_t = -r_t D_t dt$.

- The discounted stock price satisfies

$$\begin{aligned}
d\left[D_t S_t\right] &= dD_t S_t + D_t dS_t + dD_t dS_t \\
&= -r_t D_t S_t dt + D_t\left[\alpha_t S_t dt + \sigma_t S_t dW_t\right] \\
&= \left[\alpha_t - r_t\right] D_t S_t dt + \sigma_t D_t S_t dW_t \\
&= \sigma_t D_t S_t \left[\frac{\alpha_t - r_t}{\sigma_t} dt + dW_t\right] \\
&= \sigma_t D_t S_t \left[\eta_t dt + dW_t\right]
\end{aligned}$$

  where $\eta_t = \frac{\alpha_t - r_t}{\sigma_t}$ is the so-called market price of risk.

- Now, we can define the "equivalent martingale" (risk-neutral) measure Q

  - Define measure Q (and $\tilde{W}_t$) required by the Girsanov's theorem by setting

$$\eta_t = \frac{\alpha_t - r_t}{\sigma_t}$$

    and

$$d\tilde{W}_t = \eta_t dt + dW_t.$$

    Then

$$d\left[D_t S_t\right] = \sigma_t D_t S_t d\tilde{W}_t.$$

- So under Q, the discounted stock price $D_t S_t$ is a martingale

$$D_t S_t = S_0 + \int_0^t \sigma_u D_u S_u d\tilde{W}_u.$$

- Now pricing can be done using the martingale property of discounted stock price:

$$D_t V_t = \mathbb{E}^Q \left[D_T V_T | \mathcal{F}_t \right] = \mathbb{E}^Q \left[D_T V_T | \mathcal{F}_t \right].$$

$$V_t = \mathbb{E}^Q \left[\frac{D_T}{D_t} V_T \middle| \mathcal{F}_t \right] = \mathbb{E}^Q \left[e^{-\int_t^T r_u du} V_T \middle| \mathcal{F}_t \right].$$

This can be done by evaluating the conditional expectation of the last expression.

Remark: The Q is called the risk-neutral measure, because under Q, S_t has mean rate of return of r_t:

$$\begin{aligned} dS_t &= \alpha_t S_t dt + \sigma_t S_t dW_t \\ &= \alpha_t S_t dt + \sigma_t S_t \left[d\tilde{W}_t - \eta_t dt \right] \\ &= r_t S_t dt + \sigma_t S_t d\tilde{W}_t. \end{aligned}$$

Remark: About Girsanov's Theorem. To get some intuit feel about the Theorem, let's consider a special case.

With constant σ and r,

$$\begin{aligned} S_T &= S_t \exp\left\{ \sigma \tilde{W}_T + \left(r - \frac{1}{2}\sigma^2 \right) T \right\} \\ &= S_t \exp\left\{ \sigma \left[\tilde{W}_T - \tilde{W}_t \right] + \left(r - \frac{1}{2}\sigma^2 \right) (T - t) \right\}. \end{aligned}$$

Let $\tau = T - t$, $Y = -\frac{\tilde{W}_T - \tilde{W}_t}{\sqrt{T-t}} \sim N(0,1)$. Then

$$S_T = S_t \exp\left\{ -\sigma\sqrt{\tau} Y + \left(r - \frac{1}{2}\sigma^2 \right) \tau \right\}.$$

Then the price of the option at t is given as

$$\begin{aligned} & C(t, S_t) = \mathbb{E}^Q \left[e^{-r(T-t)} (S_T - K)^+ | \mathcal{F}_t \right] \\ = \; & \mathbb{E}_t^Q \left[e^{-r\tau} \left(S_t \exp\left\{ -\sigma\sqrt{\tau} Y + \left(r - \frac{1}{2}\sigma^2 \right) \tau \right\} - K \right)^+ \right] \\ = \; & \int_{-\infty}^{+\infty} e^{-r\tau} \left(S_t \exp\left\{ -\sigma\sqrt{\tau} y + \left(r - \frac{1}{2}\sigma^2 \right) \tau \right\} - K \right)^+ \cdot \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{y^2}{2} \right\} dy. \end{aligned}$$

The integrand is positive if $S_t e^{-\sigma\sqrt{\tau} y + \left(r - \frac{1}{2}\sigma^2 \right) \tau} > K$

$$-\sigma\sqrt{\tau} y > \log\left(\frac{K}{S_t} \right) - \left(r - \frac{1}{2}\sigma^2 \right) \tau$$

$$y < \frac{-\log\left(\frac{K}{S_t}\right) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} = \frac{\log\left(\frac{S(t)}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} = d_2.$$

So the integral becomes

$$\begin{aligned}
& C\left(t, S_t\right) \\
= \quad & \int_{-\infty}^{d_2} e^{-r\tau}\left(S_t \exp\left\{-\sigma\sqrt{\tau}y + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\} - K\right) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
= \quad & \frac{S_t}{\sqrt{2\pi}} \int_{-\infty}^{d_2} e^{\left\{-\frac{y^2}{2} - \sigma\sqrt{\tau}y - \frac{1}{2}\sigma^2\tau\right\}} dy - \frac{e^{-r\tau}K}{\sqrt{2\pi}} \int_{-\infty}^{d_2} e^{-\frac{y^2}{2}} dy \\
= \quad & \frac{S_t}{\sqrt{2\pi}} \int_{-\infty}^{d_2} e^{\left\{-\frac{\left(y+\sigma\sqrt{\tau}\right)^2}{2}\right\}} dy - e^{-r\tau} K N\left(d_2\right) \\
= \quad & \frac{S_t}{\sqrt{2\pi}} \int_{-\infty}^{d_2+\sigma\sqrt{\tau}} e^{\left\{-\frac{z^2}{2}\right\}} dz - e^{-r\tau} K N\left(d_2\right) \\
= \quad & S_t N\left(d_2 + \sigma\sqrt{\tau}\right) - e^{-r\tau} K N\left(d_2\right).
\end{aligned}$$