FBA | QUANTITATIVE FINANCE RESEARCH GROUP

Financial Mathematics

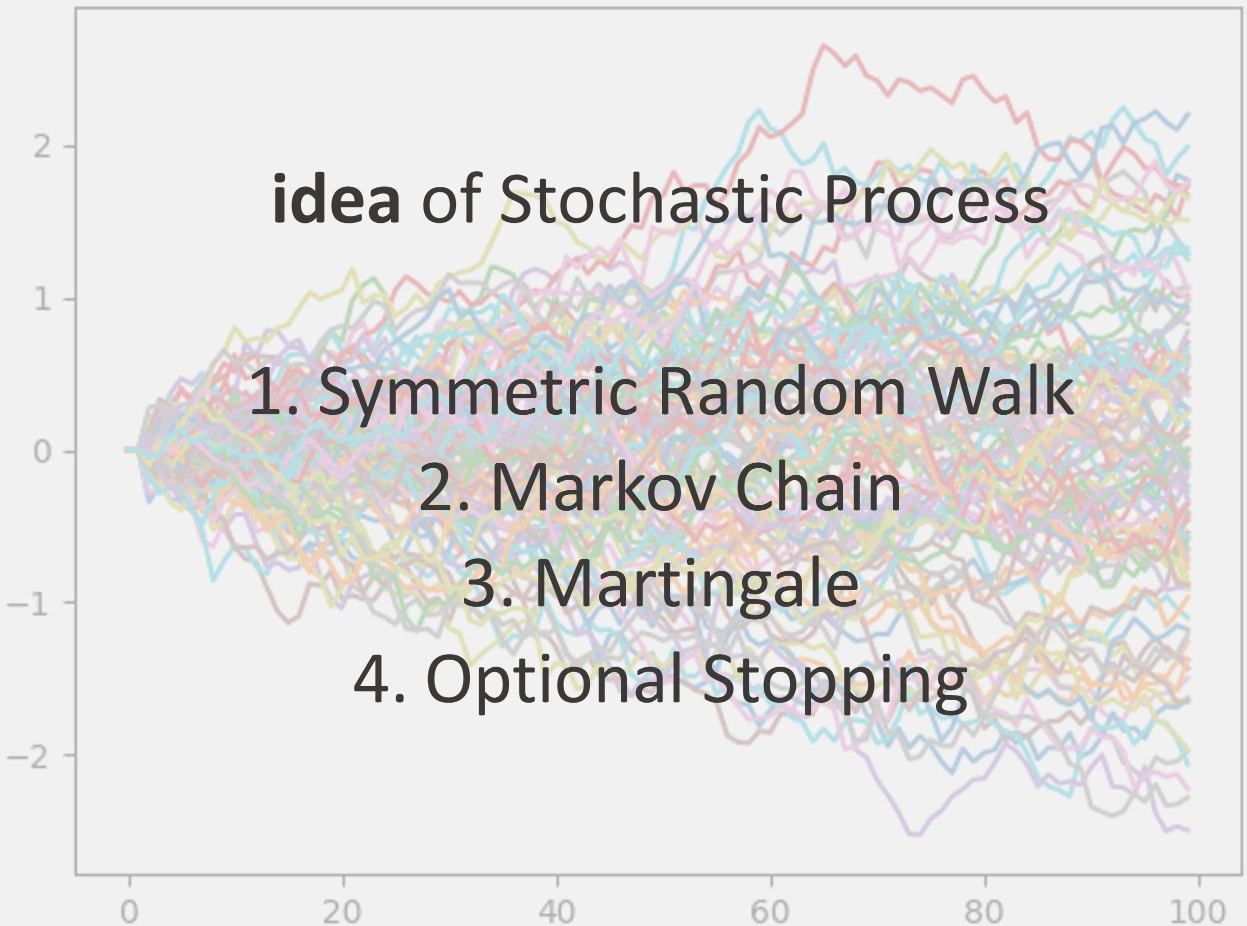
Stochastic Process

2024.03.24

YongJae Song

Ref. MIT OpenCourseWare

Stochastic Process I (Discrete)



idea of Stochastic Process

1. Symmetric Random Walk
2. Markov Chain
3. Martingale
4. Optional Stopping

Symmetric Random Walk

Proposition 2.1. *(i)* $\mathbb{E}[X_k] = 0$ *for all* k.

(ii) (Independent increment) For all $0 = k_0 \le k_1 \le \cdots \le k_r$, *the random variables* $X_{k_{i+1}} - X_{k_i}$ *for* $0 \le i \le r-1$ *are mutually independent.*

(iii) (Stationary) For all $h \ge 1$ *and* $k \ge 0$, *the distribution of* $X_{k+h} - X_k$ *is the same as the distribution of* X_h.

Proof. The proofs are straightforward and are left as an exercise. Note that these properties hold as long as the increments Y_i are identical and independent and have mean 0. □

For two positive integers A and B, what is the probability that the random walk reaches A before it reaches $-B$? Let τ be the first time at which the random walk reaches either A or $-B$. Then $X_\tau = A$ or $-B$. Define

$$f(k) = \mathbf{P}(X_\tau = A \,|\, X_0 = k),$$

and note that our goal is to compute $f(0)$. The recursive formula $f(k) = \frac{1}{2}f(k+1) + \frac{1}{2}f(k-1)$ follows from considering the outcome of the first coin-toss. We also have the boundary conditions $f(A) = 1,\ f(-B) = 0$. If we let $f(-B+1) = \alpha$, then it follows that $f(-B+r) = \alpha r$ for all $r \le A+B$. Therefore, $\alpha = \frac{1}{A+B}$, and it follows that

$$f(0) = \frac{B}{A+B}.$$

Markov Chain

$$\mathbf{P}(X_{n+1} = i \mid X_n, X_{n-1}, \cdots, X_0) = \mathbf{P}(X_{n+1} = i \mid X_n)$$

$$A = \begin{pmatrix} p_{11} & p_{21} & \cdots & p_{m1} \\ p_{12} & p_{22} & \cdots & p_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1m} & p_{2m} & \cdots & p_{mm} \end{pmatrix}.$$

Let $r_{ij}(n) = \mathbf{P}(X_n = j \mid X_0 = i)$ be the *n-th step transition probabilities*. These probabilities satisfy the recurrence relation

$$r_{ij}(n) = \sum_{k=1}^{m} r_{ik}(n-1)p_{kj} \quad for\, n > 1,$$

where $r_{ij}(1) = p_{ij}$. Hence the n-step transition probability matrix can easily be shown to be A^n.

A *stationary distribution* of a Markov chain is a probability distribution over the state space S (where $\mathbf{P}(X_0 = j) = \pi_j$) such that

$$\pi_j = \sum_{k=1}^{m} \pi_k \cdot p_{kj} \quad (\forall j \in S).$$

Martingale

Definition 4.1. A discrete-time stochastic process $\{X_0, X_1, \cdots\}$ is a *martingale* if

$$X_t = \mathbb{E}[X_{t+1}|\mathcal{F}_t],$$

for all $t \geq 0$, where $\mathcal{F}_t = \{X_0, \cdots, X_t\}$ (hence we are conditioning on the initial segment of the process).

This says that our expectated gain in the process is zero at all times. We can also view this definition as a Mathematical formalization of a game of chance being fair.

Proposition 4.2. *For all $t \geq s$, we have $X_s = \mathbb{E}[X_t|\mathcal{F}_s]$.*

Proof. This easily follows from induction. □

Optional Stopping : Small Advantage winning in long-run

Theorem 5.3. *(Doob's optional stopping time theorem, weak form) Suppose that $X_0, X_1, X_2, \cdots$ is a martingale sequence and τ is a stopping time such that $\tau \leq T$ for some constant T. Then $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.*

Proof. Note that

$$X_\tau = X_0 + \sum_{i=0}^{T-1} (X_{i+1} - X_i) \cdot \mathbf{1}_{\{\tau \geq i+1\}}.$$

(we used the fact $\tau \leq T$). Since T is a constant, by linear of expectation we have

$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0] + \sum_{i=0}^{T-1} \mathbb{E}\Big[(X_{i+1} - X_i) \cdot \mathbf{1}_{\{\tau \geq i+1\}}\Big].$$

The main observation is that $\tau \geq i+1$ is determined by $X_0, X_1, \cdots, X_i$. Hence

$$\begin{aligned}
\mathbb{E}\Big[(X_{i+1} - X_i) \cdot \mathbf{1}_{\{\tau \geq i+1\}}\Big] &= \mathbb{E}\Big[\mathbb{E}\left[(X_{i+1} - X_i) \cdot \mathbf{1}_{\{\tau \geq i+1\}} \,|\, \mathcal{F}_i\right]\Big] \\
&= \mathbb{E}\Big[\left(\mathbb{E}[X_{i+1} \,|\, \mathcal{F}_i] - X_i\right) \cdot \mathbf{1}_{\{\tau \geq i+1\}}\Big] \\
&= \mathbb{E}\Big[0 \cdot \mathbf{1}_{\{\tau \geq i+1\}}\Big] = 0.
\end{aligned}$$

Hence $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$. $\square$

Stochastic Process II (Continuous), Standard Brownian Motion

We refer to a particular instance of a path chosen according to the Brownian motion as a *sample Brownian path*.

One way to think of standard Brownian motion is as a limit of simple random walks. To make this more precise, consider a simple random walk $\{Y_0, Y_1, \cdots,\}$ whose increments are of mean 0 and variance 1. Let Z be a piecewise linear function from $[0,1]$ to $\mathbb{R}$ defined as

$$Z\left(\frac{t}{n}\right) = Y_t,$$

for $t = 0, \cdots, n$, and is linear at other points. As we take larger values of n, the distribution of the path Z will get closer to that of the standard Brownian motion. Indeed, we can check that the distribution of $Z(1)$ converges to the distribution of $N(0,1)$, by central limit theorem. More generally, the distribution of $Z(t)$ converges to $N(0,t)$.

Standard Brownian Motion

Here are some facts about the Brownian motion:

1. Crosses the x-axis infinitely often.
2. Has a very close relation with the curve $x = y^2$ (it does not deviate from this curve too much).
3. Is nowhere differentiable.

Note that in real-life we can only observe the value of a stochastic process up to some time resolution (in other words, we can only take finitely many sample points). The fact above implies that standard Brownian motion is a reasonable model, at least in this sense, since the real-life observation will converge to the underlying theoretical stochastic process as we take smaller time intervals, as long as the discrete-time observations behave like a simple random walk.

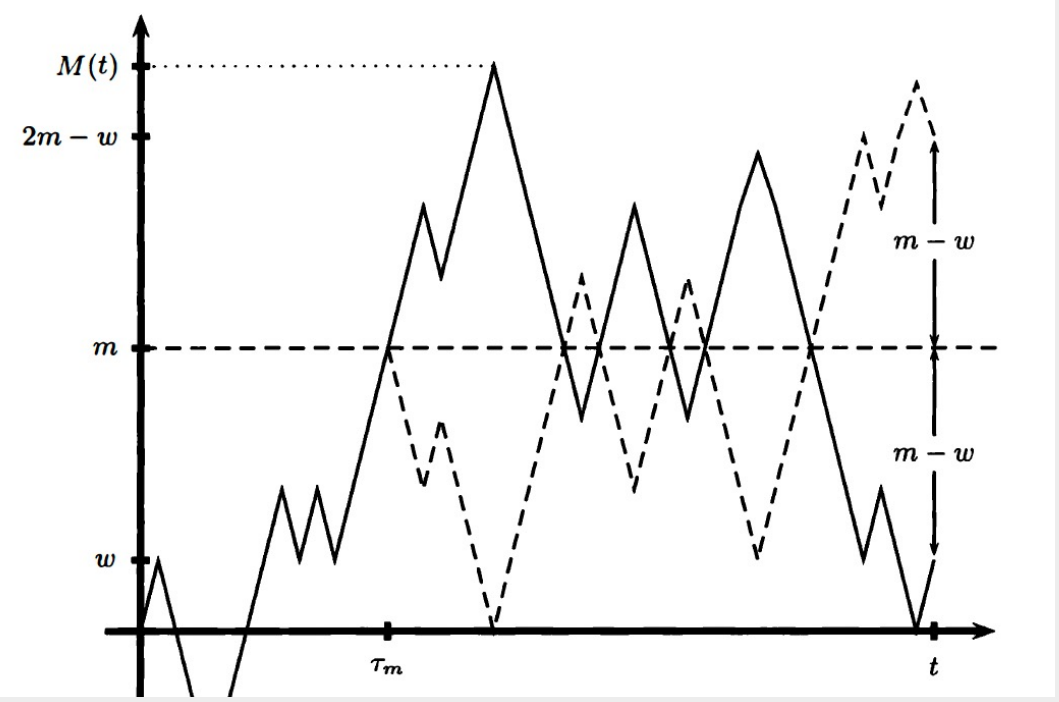
Suppose we use the Brownian motion as a model for daily price of a stock. What is the distribution of the days range? (the max value and min value over a day)

Reflection Principle

Define $M(t) = \max_{0 \le s \le t} B(s)$, and note that $M(t)$ is well-defined since B is continuous and $[0, t]$ is compact. ($\Phi(t)$ is the cumulative distribution function of the normal random variable)

Proposition 2.3. *The following holds:*

$$\mathbf{P}(M(t) \ge a) = 2\mathbf{P}(B(t) > a) = 2 - 2\Phi(\frac{a}{\sqrt{t}}).$$



Reflection Principle (Shreve II)

As Figure 3.7.1 illustrates, for each Brownian motion path that reaches level m prior to time t but is at a level w below m at time t, there is a "reflected path" that is at level $2m - w$ at time t. This reflected path is constructed by switching the up and down moves of the Brownian motion from time τ_m onward. Of course, the probability that a Brownian motion path ends at exactly w or at exactly $2m - w$ is zero. In order to have nonzero probabilities, we consider the paths that reach level m prior to time t and are *at or below* level w at time t, and we consider their reflections, which are *at or above* $2m - w$ at time t. This leads to the key *reflection equality*

$$\mathbb{P}\{\tau_m \leq t, W(t) \leq w\} = \mathbb{P}\{W(t) \geq 2m - w\}, \quad w \leq m, \ m > 0. \qquad (3.7.1)$$

Theorem 3.7.1. *For all $m \neq 0$, the random variable τ_m has cumulative distribution function*

$$\mathbb{P}\{\tau_m \leq t\} = \frac{2}{\sqrt{2\pi}} \int_{\frac{|m|}{\sqrt{t}}}^{\infty} e^{-\frac{y^2}{2}}\, dy, \quad t \geq 0, \qquad (3.7.2)$$

and density

$$f_{\tau_m}(t) = \frac{d}{dt}\mathbb{P}\{\tau_m \leq t\} = \frac{|m|}{t\sqrt{2\pi t}} e^{-\frac{m^2}{2t}}, \quad t \geq 0. \qquad (3.7.3)$$

Reflection Principle (Shreve II)

PROOF: We first consider the case $m > 0$. We substitute $w = m$ into the reflection formula (3.7.1) to obtain

$$\mathbb{P}\{\tau_m \leq t, W(t) \leq m\} = \mathbb{P}\{W(t) \geq m\}.$$

On the other hand, if $W(t) \geq m$, then we are guaranteed that $\tau_m \leq t$. In other words,

$$\mathbb{P}\{\tau_m \leq t, W(t) \geq m\} = \mathbb{P}\{W(t) \geq m\}.$$

Adding these two equations, we obtain the cumulative distribution function for τ_m:

$$\begin{aligned}\mathbb{P}\{\tau_m \leq t\} &= \mathbb{P}\{\tau_m \leq t, W(t) \leq m\} + \mathbb{P}\{\tau_m \leq t, W(t) \geq m\} \\ &= 2\mathbb{P}\{W(t) \geq m\} = \frac{2}{\sqrt{2\pi t}} \int_m^\infty e^{-\frac{x^2}{2t}}\, dx.\end{aligned}$$

We make the change of variable $y = \frac{x}{\sqrt{t}}$ in the integral, and this leads to (3.7.2) when m is positive. If m is negative, then τ_m and $\tau_{|m|}$ have the same distribution, and (3.7.2) provides the cumulative distribution function of the latter. Finally, (3.7.3) is obtained by differentiating (3.7.2) with respect to t.

F

□

Almost Surely,

Proposition 2.4. *For each $t \geq 0$, the Brownian motion is almost surely not differentiable at t.*

Proof. Fix a real t_0 and suppose that the Brownian motion B is differentiable at t_0. Then there exist constants A and ε_0 such that for all $0 < \varepsilon < \varepsilon_0$, $B(t) - B(t_0) \leq A\varepsilon$ holds for all $0 < t - t_0 \leq \varepsilon$. Let $E_{\varepsilon,A}$ denote this event, and $E_A = \bigcap_\varepsilon E_{\varepsilon,A}$. Note that

$$\begin{aligned}\mathbf{P}(E_{\varepsilon,A}) &= \mathbf{P}(E(t) - E(t_0) \leq A\varepsilon \quad \text{for all } 0 < t - t_0 \leq \varepsilon) \\ &= \mathbf{P}(M(\varepsilon) \leq A\varepsilon) = 2(1 - \Phi(A\sqrt{\varepsilon})),\end{aligned}$$

where the right hand side tends to zero as ε goes to zero. Therefore, $\mathbf{P}(E_A) = 0$. By countable additivity, we see that there can be no constant A satisfying above (it suffices to consider integer values of A). □

Dvoretsky, Erdős, and Kakutani in fact proved a stronger statment asserting that the Brownian motion B is nowhere differentiable with probability 1. Hence a sample Brownian path is continuous but nowhere differentiable! The proof is slightly more involved and requires a lemma from probability theory (Borel-Cantelli lemma).

$$\sum Pr < \infty \rightarrow 0, \sum Pr = \infty \rightarrow 1$$

Quadratic Variation (Ito's Chain Rule)

Theorem 2.5. *(Quadratic variation) For a partition* $\Pi = \{t_0, t_1, \cdots, t_j\}$ *of an interval* $[0, T]$*, let* $|\Pi| = \max_i(t_{i+1} - t_i)$*. A Brownian motion* B_t *satisfies the following equation with probability* 1*:*

$$\lim_{|\Pi| \to 0} \sum_i (B_{t_{i+1}} - B_{t_i})^2 = T.$$

Proof. For simplicity, here we only consider partitions where the gaps $t_{i+1} - t_i$ are uniform. In this case, the sum

$$\sum_i (B_{t_{i+1}} - B_{t_i})^2$$

is a sum of i.i.d. random variables with mean $t_{i+1} - t_i$, and finite second moment. Therefore, by the law of large numbers, as $\max\{t_{i+1} - t_i\} \to 0$, we have

$$\sum_i (B_{t_{i+1}} - B_{t_i})^2 = T$$

with probability 1. $\square$

As $\max\{t_{i+1} - t_i\} \to 0$, we see that the above tends to zero. Hence this shows that Brownian motion fluctuates a lot. The above can be summarized by the differential equation $(dB)^2 = dt$. As we will see in the next lecture, this fact will have very interesting implications.

Ito Calculus, Substitution of Second Order

$$df = f'(B_t)dB_t. \tag{1.1}$$

Our new formula at least makes sense, since there is no need to refer to the differentiation $\frac{dB_t}{dt}$ which does not exist. The only problem is that it does not quite work. Consider the Taylor expansion of f to obtain

$$f(x + \Delta x) - f(x) = (\Delta x) \cdot f'(x) + \frac{(\Delta x)^2}{2} f''(x) + \frac{(\Delta x)^3}{6} f'''(x) + \cdots .$$

To deduce Equation (1.1) from this formula, we must be able to say that the significant term is the first term $(\Delta x) \cdot f'(x)$ and all other terms are of smaller order of magnitude. Is this true for $x = B_t$? For $x = B_t$, we have

$$\Delta f = (\Delta B_t) \cdot f'(B_t) + \frac{(\Delta B_t)^2}{2} f''(x) + \frac{(\Delta B_t)^3}{6} f'''(x) + \cdots .$$

Now consider the term $(\Delta B_t)^2$. Since B_t is a Brownian motion, we know that $\mathbb{E}[(\Delta B_t)^2] = \Delta t$. Since a difference in B_t is necessarily accompanied by a difference in t, we see that the second term is no longer negligable. The theory of Ito calculus essentially tells us that we can make the substitution

$(\Delta B_t)^2 = \Delta t$, and the remaining terms are negligable. Hence the equation above becomes

$$\Delta f = (\Delta B_t) \cdot f'(B_t) + \frac{\Delta t}{2} f''(x) + \cdots ,$$

which in terms of infinitesimals becomes

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2} f''(B_t)dt. \tag{1.2}$$

This equation known as the *Ito's lemma* is the main equation of Ito's calculus.

Two Variables

Second-order mixed derivatives:

$$\frac{\partial^2 f}{\partial y\, \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = (f'_x)'_y = f''_{xy} = \partial_{yx} f = \partial_y \partial_x f.$$

More generally, consider a smooth function $f(t, x)$ which depends on two variables, and suppose that we are interested in the differential of $f(t, B_t)$. In classical calculus, we will get

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx,$$

but in Ito calculus, we will have

$$\begin{aligned} df(t, B_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dB_t)^2 \\ &= \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right) dt + \frac{\partial f}{\partial x} dB_t. \end{aligned}$$

Ito Lemma : Definition

Theorem 1.1. *(Ito's lemma) Let $f(t,x)$ be a smooth function of two variables, and let X_t be a stochastic process satisfying $dX_t = \mu_t dt + \sigma_t dB_t$ for a Brownian motion B_t. Then*

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \frac{\partial f}{\partial x} dB_t.$$

Proof. We have

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 \\ &= \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} dB_t + ..dt dB_T + ..(dt)^2. \end{aligned}$$

We can ignore the terms $dt dB_t$ and $(dt)^2$. □

Ito Lemma : Properties [Lognormal, Martingale]

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

(ii) Consider the function $f(x) = \frac{1}{2}x^2$. We see that

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

Equiavlently,

$$\frac{1}{2}B_T^2 = \int_0^T B_t dB_t + \int_0^T \frac{1}{2}dt = \int_0^T B_t dB_t + \frac{T}{2}.$$

This implies that

$$\int_0^T B_t dB_t = \frac{1}{2}B_T^2 - \frac{T}{2}.$$

Note how this 'violates' the fundamental theorem of calculus.

(iii) Let $f(t, x) = exp(\mu t + \sigma x)$. Then

$$df(t, B_t) = (\mu + \frac{1}{2}\sigma^2)f(t, B_t)dt + \sigma f(t, B_t)dB_t.$$

We can now answer the question of finding the stochastic process $X_t(t, B_t)$ such that

$$dX_t = \sigma X_t \cdot dB_t.$$

To do this we can just set $\mu = -\frac{1}{2}\sigma^2$ in the function above, i.e., $X_t(t, B_t) = exp(-\frac{1}{2}\sigma^2 t + \sigma B_t)$.

F B A

Ito Integral for 'nonrandom function of time'

Definition 2.1. Let X_t be a stochastic process. A process Δ_t is called an *adapted process* (with respect to X_t) if for all $t \geq 0$, the random variable Δ_t depends only on X_s for $s \leq t$.

Example 2.2. Let X_t be a stochastic process.

(i) The process $\Delta(t) = X_t$ is an adapted process.

(ii) The process $\Delta(t) = \min\{X_t, c\}$ is an adapted process (where c is a constant).

(iii) The process $\Delta(t) = \max_{0 \leq t \leq T} X_t$ is not an adapted process.

(iv) If τ is a stopping time, then X_τ is an adapted process.

For example, suppose that we model the price of a stock using a stochastic process, and are trying to find a strategy which has positive expected return. Consider a simple strategy where at each time t, we either buy or sell one stock, hence $\Delta_t = 1$ or -1. Our strategy only makes sense if Δ_t is an adapted process, since otherwise it contradicts the fact that we cannot see the future.

Ito Integral for 'Martingale'

Theorem 2.3. *Let B_t be a Brownian motion. Then for all adapted processes $g(t, B_t)$, the integral*

$$\int g(t, B_t)\, dB_s$$

is a martingale, as long as g is a 'reasonable function'. Formally, if $g \in L^2$, i.e.,

$$\int \int_0^t g^2(t, B_t) dt dB_t < \infty.$$

Example 2.4. The process B_t itself is an adapted process. Recall that

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

and $\mathbb{E}[B_t^2] = t$. Hence

$$\mathbb{E}[\int_0^t B_s dB_s] = 0.$$

More generally,

$$\begin{aligned} \mathbb{E}[\int_{t_1}^{t_2} B_s dB_s \,|\, \mathcal{F}_{t_1}] &= \mathbb{E}[\left(\frac{1}{2} B_{t_2}^2 - \frac{t_2}{2}\right) \,|\, \mathcal{F}_1] - \left(\frac{1}{2} B_{t_1}^2 - \frac{t_1}{2}\right) \\ &= \frac{1}{2}(t_2 - t_1) + \frac{1}{2} B_{t_1}^2 - \frac{t_2}{2} - \frac{1}{2} B_{t_1}^2 + \frac{t_1}{2} = 0. \end{aligned}$$

This confirms the theorem above for $\Delta(t) = B_t$.

Here is another useful fact about the Ito integral of an adapted process known as Ito isometry. It can be used to compute the variance of the Ito integral.

Ito Isometry [Compute Mean-Variance for PDF]

Theorem 2.5. *(Ito isometry) Let B_t be a Brownian motion. Then for all adapted processes $\Delta(t)$, we have*

$$\mathbb{E}\left[\left(\int_0^t \Delta(s)dB_s\right)^2\right] = \mathbb{E}\left[\int_0^t \Delta(s)^2 ds\right].$$

Example 2.6. Let $\Delta(t) = 1$. Then the left hand side of the theorem above is

$$\mathbb{E}\left[\left(\int_0^t \Delta(s)dB_s\right)^2\right] = \mathbb{E}[B_t^2] = t,$$

and the right hand side of the above is

$$\mathbb{E}\left[\int_0^t \Delta(s)^2 ds\right] = \mathbb{E}[t] = t.$$

Note that $\int_0^t \Delta(s)dB_s = 0$ by Theorem 2.1 given above. Hence Ito isometry tell us how to compute the variance of this integral .

Measure Change [Drift to Non-Drift]

$\mathbf{P}(A) = \mathbf{P}(X^{-1}(A))$ for all set of paths $A \subset [0,T]^\infty$. We denote a particular realization of a stochastic process as X_ω.

Let Z be a positive random variable satisfying $\int Z(\omega)\, dX(\omega) = 1$. Then we can define a new stochastic process $\tilde{X}$ whose probability distribution is given by

$$\tilde{\mathbf{P}}(A) = \int_A Z(\omega)\, dX(\omega),$$

for all sets A. Since Z is positive, we see that

$$\mathbf{P}(A) > 0 \Leftrightarrow \tilde{\mathbf{P}}(A) > 0 \quad \forall A. \tag{3.1}$$

Hence the set of paths 'observed' under the probability measures $\mathbf{P}$ and $\tilde{\mathbf{P}}$ are the same.

Definition 3.1. We say that two probability distributions $\mathbf{P}$ and $\tilde{\mathbf{P}}$ are *equivalent* if (3.1) holds.

Radon-Nikodym Derivative (Shreve II)

probability distribution of the square of a Brownian motion $B(t)^2$ is not equivalent to the probability distribution of $B(t)$.

The function Z is known as the *Radon-Nikodým derivative* of $\tilde{\mathbf{P}}$ with respect to $\mathbf{P}$, and is denoted as

$$Z = \frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}.$$

Changing measures is of theoretical importance since it provides a tool to understand the relation between two different but equivalent stochastic processes. It is also of practical importance, since converting one probability distribution into another can reveal hidden insights. For example, in finance, we can convert a non-martingale stochastic process into a martingale by changing measure, and this gives a method of pricing financial derivatives.

and we define $\widetilde{\mathbb{P}}$ by (5.2.1). We can then define the *Radon-Nikodým derivative process*

$$Z(t) = \mathbb{E}[Z|\mathcal{F}(t)], \quad 0 \le t \le T. \tag{5.2.6}$$

This process in discrete time is discussed in Section 3.2 of Volume I. The Radon-Nikodým derivative process (5.2.6) is a martingale because of iterated conditioning (Theorem 2.3.2(iii)): for $0 \le s \le t \le T$,

$$\mathbb{E}[Z(t)|\mathcal{F}(s)] = \mathbb{E}\big[\mathbb{E}[Z|\mathcal{F}(t)]\big|\mathcal{F}(s)\big] = \mathbb{E}[Z|\mathcal{F}(s)] = Z(s). \tag{5.2.7}$$

Radon-Nikodym Derivative [Stochastic vs Deterministic] (Shreve II)

$$\widetilde{\mathbb{E}}Y = \mathbb{E}[YZ] = \mathbb{E}\big[\mathbb{E}[YZ|\mathcal{F}(t)]\big] = \mathbb{E}\big[Y\mathbb{E}[Z|\mathcal{F}(t)]\big] = \mathbb{E}[YZ(t)]. \qquad \square$$

Lemma 5.2.2. *Let s and t satisfying $0 \le s \le t \le T$ be given and let Y be an $\mathcal{F}(t)$-measurable random variable. Then*

$$\widetilde{\mathbb{E}}[Y|\mathcal{F}(s)] = \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]. \tag{5.2.9}$$

(i) **(Measurability)** $\mathbb{E}[X|\mathcal{G}]$ *is $\mathcal{G}$-measurable, and*
(ii) **(Partial averaging)**

$$\int_A \mathbb{E}[X|\mathcal{G}](\omega)\, d\mathbb{P}(\omega) = \int_A X(\omega)\, d\mathbb{P}(\omega) \text{ for all } A \in \mathcal{G}. \tag{2.3.17}$$

If $\mathcal{G}$ is the σ-algebra generated by some other random variable W (i.e., $\mathcal{G} = \sigma(W)$), we generally write $\mathbb{E}[X|W]$ rather than $\mathbb{E}[X|\sigma(W)]$.

PROOF: It is clear that $\frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]$ is $\mathcal{F}(s)$-measurable. We must check the partial-averaging property (Definition 2.3.1(ii)), which in this case is

$$\int_A \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]\, d\widetilde{\mathbb{P}} = \int_A Y\, d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F}(s). \tag{5.2.10}$$

Girsanov's Theorem, Levy's Theorem (Shreve II)

Theorem 5.2.3 (Girsanov, one dimension). *Let $W(t)$, $0 \le t \le T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{F}(t)$, $0 \le t \le T$, be a filtration for this Brownian motion. Let $\Theta(t)$, $0 \le t \le T$, be an adapted process. Define*

$$Z(t) = \exp\left\{-\int_0^t \Theta(u)\, dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)\, du\right\}, \tag{5.2.11}$$

$$\widetilde{W}(t) = W(t) + \int_0^t \Theta(u)\, du, \tag{5.2.12}$$

$$d\widetilde{W}(t)\, d\widetilde{W}(t) = \big(dW(t) + \Theta(t)\, dt\big)^2 = dW(t)\, dW(t) = dt.$$

It remains to show that $\widetilde{W}(t)$ is a martingale under $\widetilde{\mathbb{P}}$. We first observe that $Z(t)$ is a martingale under $\mathbb{P}$. With

$$X(t) = -\int_0^t \Theta(u)\, dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)\, du$$

and $f(x) = e^x$ so that $f'(x) = e^x$ and $f''(x) = e^x$, we have

$$\begin{aligned} dZ(t) &= df\big(X(t)\big) \\ &= f'\big(X(t)\big)\, dX(t) + \frac{1}{2} f''\big(X(t)\big)\, dX(t)\, dX(t) \\ &= e^{X(t)}\Big(-\Theta(t)\, dW(t) - \frac{1}{2}\Theta^2(t)\, dt\Big) + \frac{1}{2} e^{X(t)}\Theta^2(t)\, dt \\ &= -\Theta(t) Z(t)\, dW(t). \end{aligned}$$

Risk-Neutral Pricing (Shreve II)

Then there is an adapted process $\Gamma(u)$, $0 \le u \le T$, such that

$$M(t) = M(0) + \int_0^t \Gamma(u)\, dW(u),\ 0 \le t \le T.$$

Corollary 5.3.2. *Let $W(t)$, $0 \le t \le T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{F}(t)$, $0 \le t \le T$, be the filtration generated by this Brownian motion. Let $\Theta(t)$, $0 \le t \le T$, be an adapted process, define*

$$Z(t) = \exp\left\{-\int_0^t \Theta(u)\, dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)\, du\right\},$$

$$\widetilde{W}(t) = W(t) + \int_0^t \Theta(u)\, du,$$

and assume that $\widetilde{\mathbb{E}} \int_0^T \Theta^2(u) Z^2(u)\, du < \infty$. Set $Z = Z(T)$. Then $\mathbb{E}Z = 1$, and under the probability measure $\widetilde{\mathbb{P}}$ given by (5.2.1), the process $\widetilde{W}(t)$, $0 \le t \le T$, is a Brownian motion.

Now let $\widetilde{M}(t)$, $0 \le t \le T$, be a martingale under $\widetilde{\mathbb{P}}$. Then there is an adapted process $\widetilde{\Gamma}(u)$, $0 \le u \le T$, such that

$$\widetilde{M}(t) = \widetilde{M}(0) + \int_0^t \widetilde{\Gamma}(u)\, d\widetilde{W}(u),\ 0 \le t \le T. \tag{5.3.2}$$

T-Forward Measure (Shreve II)

5.6.1 Forward Contracts

Let $S(t)$, $0 \le t \le \overline{T}$, be an asset price process, and let $R(t)$, $0 \le t \le \overline{T}$, be an interest rate process. We choose here some large time $\overline{T}$, and all bonds and derivative securities we consider will mature or expire at or before time $\overline{T}$. As usual, we define the discount process $D(t) = e^{-\int_0^t R(u)du}$. According to the risk-neutral pricing formula (5.2.30), the price at time t of a zero-coupon bond paying 1 at time T is

$$B(t,T) = \frac{1}{D(t)} \widetilde{\mathbb{E}}[D(T)|\mathcal{F}(t)], \quad 0 \le t \le T \le \overline{T}. \tag{5.6.1}$$

Definition 5.6.1. *A* forward contract *is an agreement to pay a specified* delivery price K *at a* delivery date T, *where* $0 \le T \le \overline{T}$, *for the asset whose price at time t is* $S(t)$. *The* T-forward price $\mathrm{For}_S(t,T)$ *of this asset at time* t, *where* $0 \le t \le T \le \overline{T}$, *is the value of* K *that makes the forward contract have no-arbitrage price zero at time* t.

Theorem 5.6.2. *Assume that zero-coupon bonds of all maturities can be traded. Then*

$$\mathrm{For}_S(t,T) = \frac{S(t)}{B(t,T)}, \quad 0 \le t \le T \le \overline{T}. \tag{5.6.2}$$

Ornstein-Uhlenbeck Process

$$dX(t) = -\alpha X(t)dt + \sigma dB(t) \quad \text{with } X(0) = x_0.$$

Ornstein and Uhlenbeck first used (a version of) this equation to study the behavior of gasses. It has been applied (or rediscovered) in a variety of contexts. This SDE exhibits the 'mean reversion' behavior (when $\alpha > 0$).

Coefficient matching method failes for this SDE, so we try a different test function

$$X(t) = a(t)\Big(x_0 + \int_0^t b(s)dB(s)\Big),$$

where $a(0) = 1$. By differentiating each side we get,

$$dX(t) = \frac{a'(t)}{a(t)}X(t)\,dt + a(t)b(t)dB(t),$$

where we assume that $a(t) > 0$ for all t. This should match the given SDE, so we must have

$$-\alpha = \frac{a'(t)}{a(t)} \quad \text{and} \quad \sigma = a(t)b(t).$$

Therefore, $a(t) = e^{-\alpha t}$ and $b(t) = \sigma e^{\alpha t}$. From this, we see that

$$X(t) = x_0 e^{-\alpha t} + \int_0^t \sigma e^{\alpha(s-t)}dB(s).$$

General Methods

1. Finite Difference Method -> Riemann Sum, Taylor's Formula
2. Monte Carlo Pricing -> Random Simulation (FDM Closed Form)
3. Tree Method -> Discrete to Continuous

Heat Equation [PDE]

Our last topic of study is a well-known PDE, heat equation. It is well known that the Black-Scholes equation can be turned into a heat equation after a suitable change of variables.

Let $u(x,t)$ be a function of two varaibles, space and time (denoted x and t) .The following differential equation is known as the one dimensional heat equation (diffusion equation):

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

This is one of the few partial differential equations that is very well understood (and has a closed form solution).

Observation 1. Heat equation is linear, i.e., if $u_1(x,t)$ and $u_2(x,t)$ satisfies the heat equation, then $(u_1 + u_2)(x,t)$ also satisfies the heat equation. More generally if we have a collection of solutions $u_s(x,t)$ indexed by $s \in \mathbb{R}$,then $\int_{-\infty}^{\infty} u_s(x,t) \cdot c(s) ds$ is also a solution (as long as the integral exists and is differentiable up to appropriate order). This means that we can superimpose solutions of 'easy' initial value problems to obtain a solution to a more general initial value problem.

Heat Equation (Willmott)

Observation 2. The 'easy' initial value problem we are going to use is when the initial value is given as a Dirac delta function. Let $\delta(x)$ be the Dirac delta function and suppose that $u_0 = \delta$ so that we are solving

$$u(0, x) = \delta(x).$$

The solution for this initial value problem is known to be

$$u_\delta(x,t) = \frac{1}{2\sqrt{\pi t}} e^{-x^2/(4t)} \quad (\text{for } -\infty < x < \infty, t > 0).$$

Note that the solution 'converges to' the Dirac delta function as t tends to zero. Also note that for fixed value of $t > 0$, this is a probability distribution function of the normal random variable.

Transformation to Constant Coefficient Diffusion Equation

$$V(S,t) = e^{\alpha x + \beta \tau} U(x, \tau),$$

where

$$\alpha = -\tfrac{1}{2}\left(\frac{2r}{\sigma^2} - 1\right), \quad \beta = -\tfrac{1}{4}\left(\frac{2r}{\sigma^2} + 1\right)^2, \quad S = e^x \quad \text{and} \quad t = T - \frac{2\tau}{\sigma^2},$$

then $U(x, \tau)$ satisfies the basic diffusion equation

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2}. \qquad (6.2)$$

Boundary Conditions (Willmott)

The first-order S-derivative term

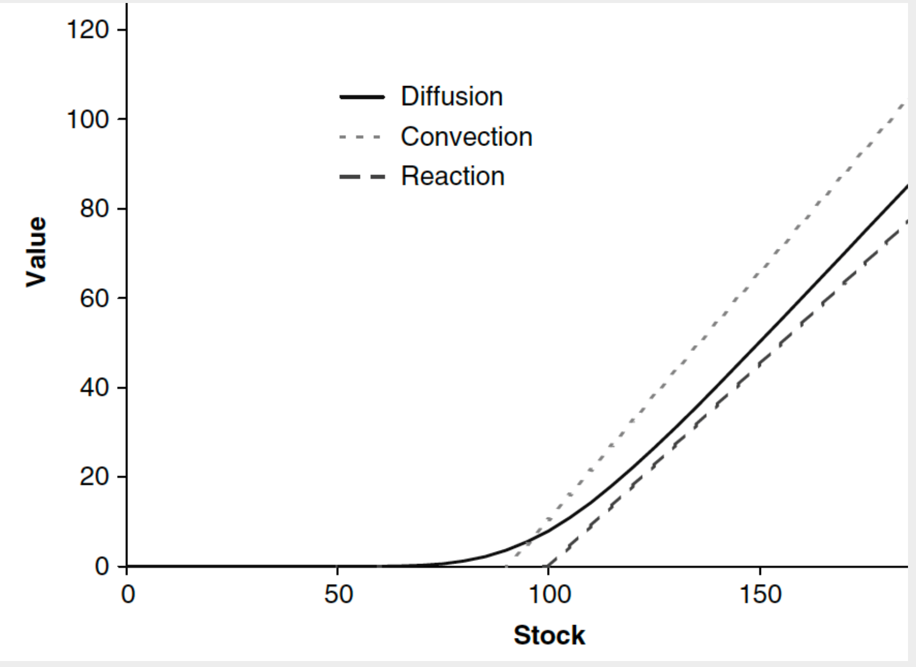
$$rS\frac{\partial V}{\partial S}$$

can be thought of as a convection term. If this equation represented some physical system, such as the diffusion of smoke particles in the atmosphere, then the convective term would be due to a breeze, blowing the smoke in a preferred direction.

The final term

$$-rV$$

is a reaction term. Balancing this term and the time derivative would give a model for decay of a radioactive body, with the half-life being related to r.



FBA | QUANTITATIVE FINAN
RESEARCH GROUP

Black-Scholes Equation (Willmott)

One solution of the Black–Scholes equation is

$$V'(S,t) = \frac{e^{-r(T-t)}}{\sigma S'\sqrt{2\pi(T-t)}} e^{-\left(\log(S/S')+(r-\frac{1}{2}\sigma^2)(T-t)\right)^2/2\sigma^2(T-t)} \tag{6.3}$$

for any S'. (You can verify this by substituting back into the equation, but we'll also be seeing it derived in the next chapter.) This solution is special because as $t \to T$ it becomes zero everywhere, except at $S = S'$. In this limit the function becomes what is known as a **Dirac delta function**. Think of this as a function that is zero everywhere except at one point where it is infinite, in such a way that its integral is one. How is this be of help to us?

Because of the nature of the integrand as $t \to T$ (i.e. that it is zero everywhere except at S' and has integral one), if we choose the arbitrary function $f(S')$ to be the payoff function then this expression becomes the solution of the problem:

$$V(S,t) = \frac{e^{-r(T-t)}}{\sigma\sqrt{2\pi(T-t)}} \int_0^\infty e^{-\left(\log(S/S')+(r-\frac{1}{2}\sigma^2)(T-t)\right)^2/2\sigma^2(T-t)} \,\mathrm{Payoff}(S')\frac{dS'}{S'}.$$

The function $V'(S,t)$ given by (6.3) is called the **Green's function**.

FBA | QUANTITATIVE FINANCE RESEARCH GROUP

Contact fbaquant@gmail.com Website fbaquant.com