Bus Ad 239B–Spring 2003
Lecture Notes

1 The Random Walk Model

The random walk model is a simple model of the evolution of a stock price. The increments in the random walk process are additive, while we normally think that changes in a stock price function multiplicatively. For this reason, we think of the random walk model as representing the natural logarithm of the stock price; equivalently, the stock price is the exponential of the random walk.

In the random walk model, information accrues in small discrete steps. Consider the time interval $[0, T]$. For $n \in \mathbf{N}$, divide the time interval into nT subintervals, each of length $\frac{1}{n}$. At time 0, the random walk process starts out at 0. At the beginning of each interval, toss a coin; if it comes out heads, the random walk process increases by $\frac{1}{\sqrt{n}}$ over the course of the interval; if the coin comes out tails, the random walk process decreases by $\frac{1}{\sqrt{n}}$ over the course of the interval.

Formally, the random walk model is specified as follows. The event space is $\Omega = \{-1, 1\}^{nT}$. Thus, every $\omega \in \Omega$ is a vector of $+1$s and -1s. Observe that Ω is finite, indeed $|\Omega| = 2^{nT}$. The collection of measurable events is $\mathcal{F}$, the collection of all subsets of Ω. The probability measure is $P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{2^{nT}}$; thus, we assign equal probability $\frac{1}{2^{nT}}$ to every $\omega \in \Omega$. We consider two closely-related versions of the random walk process:

$$
\begin{aligned}
X_n(\omega, t) &= \sum_{k=1}^{\lfloor nt \rfloor} \frac{\omega_k}{\sqrt{n}} + \frac{(nt - \lfloor nt \rfloor)\omega_{\lfloor nt \rfloor + 1}}{\sqrt{n}} \\
\hat{X}_n(\omega, t) &= \sum_{k=1}^{\lfloor nt \rfloor} \frac{\omega_k}{\sqrt{n}}
\end{aligned}
$$

When dealing with a fixed n, we will typically omit the subscript n and write the random walk process as $X(\omega, t)$ or $\hat{X}(\omega, t)$.

Each $\omega \in \Omega$ corresponds one of the possible paths the random walk process might follow. $X(\omega, \cdot)$ denotes the function from $[0, T]$ to $\mathbf{R}$ defined by $X(\omega,\cdot)(t) = X(\omega,t)$; this is the *sample path* of the random walk process corresponding to ω. At time 0, we don't know which ω will occur, and thus we don't know which path $X(\omega,\cdot)$ the random walk will follow; we only know that one of the possible paths will occur. When we get to time T, we have been able to observe the full path of the random walk, and thus we know precisely which ω occurred.

The second term in the definition of X_n is a linear interpolation term which makes the paths $X_n(\omega,\cdot)$ into continuous functions; since the paths of Brownian motion are continuous functions, this has the mathematical advantage of putting X_n and Brownian motion into the same space. However, X_n has the disadvantage that for every small $\varepsilon > 0$, the evolution of the path $X_n(\omega,\cdot)$ over the interval $\left[\frac{k}{n}, \frac{k+1}{n}\right)$ is completely known at the time $\frac{k}{n} + \varepsilon$. The paths $\hat{X}_n(\omega,\cdot)$ of $\hat{X}_n$ are step functions, constant across time intervals of the form $\left[\frac{k}{n}, \frac{k+1}{n}\right)$ and discontinuous at the times $\frac{k}{n}$.

Suppose $t \in [0,T]$. The information revealed up to time t is $\omega_1, \omega_2, \ldots, \omega_{\lfloor nt \rfloor}$. Thus, the collection of measurable events at time t is

$$\mathcal{F}_t = \{A \in \mathcal{F} : \omega \in A, \omega'_k = \omega_k \text{ for } k \leq nt \Rightarrow \omega' \in A\}$$

$\hat{X}_n$ is *adapted* to the filtration $\{\mathcal{F}_t\}$, i.e. $\hat{X}_n(\cdot,t)$ is $\mathcal{F}_t$-measurable for all t; X_n is not adapted, because ω_{k+1} is revealed by $X_n\left(\omega, \frac{k}{n} + \varepsilon\right)$ for every positive ε.

The random walk has the following qualitative properties:

1. *Approximate Normality*: Fix $t = \frac{k}{n}$. Let $M(\omega,t)$ be the number of $+1$s in the first k coin tosses. $M(\omega,t)$ has the binomial distribution $b\left(k, \frac{1}{2}\right)$.

$$\begin{aligned} X(\omega,t) &= \frac{M(\omega,t) - (k - (M(\omega,t)))}{\sqrt{n}} \\ &= \frac{2\left(M(\omega,t) - \frac{k}{2}\right)}{\sqrt{n}} \end{aligned}$$

Since the expected value $E(M(\cdot,t)) = \frac{k}{2}$, $E(X(\cdot,t)) = 0$. Since the variance $\mathrm{Var}(M(\cdot,t)) = k\left(\frac{1}{2}\right)\left(1 - \frac{1}{2}\right) = \frac{k}{4}$, $\mathrm{Var}(X(\cdot,t)) = \frac{4 \times \frac{k}{4}}{\left(\sqrt{n}\right)^2} = \frac{k}{n} = t$. By the Central Limit Theorem, the distribution of $X(\cdot,t)$ is very nearly $N(0,t)$, normal with mean zero and variance t, hence standard deviation $\sqrt{t}$.

2. *Independent Increments*: Suppose $t_1 < t_2 < \cdots < t_m$. Then

$$\{\hat{X}(\cdot, t_2) - \hat{X}(\cdot, t_1), \ldots \hat{X}(\cdot, t_m) - \hat{X}(\cdot, t_{m-1})\}$$

   are independent random variables because they're determined by disjoint sets of coin tosses

$$\{\omega_{\lfloor nt_1 \rfloor + 1}, \ldots, \omega_{\lfloor nt_2 \rfloor}\}, \ldots, \{\omega_{\lfloor nt_{m-1} \rfloor + 1}, \ldots, \omega_{\lfloor nt_m \rfloor}\}$$

   The same is true of the increments of X, provided we restrict the times to the form $t_i = \frac{k_i}{n}$.

3. *Tightness*: This is technical and you don't need a full understanding. The random walk paths are obviously continuous, since they are given by functions that are linear on each of the intervals $\left[\frac{k}{n}, \frac{k+1}{n}\right]$. However, as n increases, each of these linear functions, which has slope $\sqrt{n}$, becomes steeper. Roughly speaking, tightness says that the random walk paths nonetheless have continuous limits, with probability one. Technically, the condition is

$$\forall_{\varepsilon > 0} \, \exists_{\delta > 0} \, \forall_n \, P\left(\{\omega : \exists_{s,t} \, |s - t| < \delta, |X_n(\omega, t) - X_n(\omega, s)| > \varepsilon\}\right) < \varepsilon$$

4. *Variation of Paths*: Given a function $f : [0, T] \to \mathbf{R}$, the variation of f is

$$\sup_{m \in \mathbf{N}} \sup_{0 = t_0 < t_1 < \cdots < t_m = T} \sum_{k=1}^{m} |f(t_k) - f(t_{k-1})|$$

   f is said to be of bounded variation if the variation of f is finite. For all ω, the variation of the path $X(\omega, \cdot)$ is $nT\left(\frac{1}{\sqrt{n}}\right) = \sqrt{n}T \to \infty$ as $n \to \infty$. In other words, all of the random walk paths are of variation tending to infinity as $n \to \infty$.

5. *Quadratic Variation*: By analogy with the variation, it would be natural to try to define quadratic variation pathwise: given a function $f : [0, T] \to \mathbf{R}$, we could define the quadratic variation of f to be

$$\sup_{m \in \mathbf{N}} \sup_{0 = t_0 < t_1 < \cdots < t_m = T} \sum_{k=1}^{m} \left(f(t_k) - f(t_{k-1})\right)^2$$

For all ω, if we take $t_k = \frac{k}{n}$, $\sum_{k=1}^{m} (f(t_k) - f(t_{k-1}))^2 = nT\left(\frac{1}{(\sqrt{n})^2}\right) = T$. As you will see in Problem Set 1, problems arise if we attempt to define the quadratic variation one path at a time, in particular if we are allowed to choose the partition $0 = t_0 < t_1 < \cdots < t_m = T$ as a function of ω. Thus, the quadratic variation needs to be defined taking the whole process into account, not one path at a time.

2 The Brownian Motion Model

The Brownian Motion Model is the limit of the random walk model as $n \to \infty$. This can be made precise in a number of ways.[1]

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{T}$ a time set, with either a finite time horizon (i.e. $\mathcal{T} = [0, T]$ for some $T \in \mathbf{R}$) infinite time horizon (i.e. $\mathcal{T} = [0, \infty)$).

A K-dimensional stochastic process is $X : \Omega \times \mathcal{T} \to \mathbf{R}^K$ such that $X(\cdot, t) : \Omega \to \mathbf{R}^K$ is measurable in ω for all $t \in \mathcal{T}$. $X(\omega, \cdot)$ is the function from $\mathcal{T}$ to $\mathbf{R}^K$ defined by $X(\omega, \cdot)(t) = X(\omega, t)$. $X(\omega, \cdot)$ is called a *sample path* of the process; it is one of the (usually infinitely) many possible paths the process could follow.

You will need to distinguish three different measures floating around:

1. P, the probability measure on Ω.

2. λ, Lebesgue measure on $\mathcal{T}$.

3. $P \otimes \lambda$, the product measure on $\Omega \times \mathcal{T}$ generated by P and λ.

For more information, see sections Appendices A.2 and B.2 of Nielsen.

[1] One of the natural ways is Donsker's Theorem. Let $X_n(\omega, t)$ denote the random walk model of Section 1 for a specific $n \in \mathbf{N}$. View $X_n(\omega, \cdot)$ as a random variable taking values in $C([0, T])$, the space of continuous functions from $[0, T]$ into $\mathbf{R}$, with the metric $d(f, g) = \sup_{t \in [0,T]} |f(t) - g(t)|$. Donsker's Theorem asserts that Brownian motion $B(\omega, t)$ is the limit in distribution of X_n as $n \to \infty$. The notion of convergence in distribution of random variables taking values in $C([0, T])$ is the following: for every bounded continous function $F : C([0, T]) \to \mathbf{R}$, $E(F(x_n(\omega, \cdot))) \to E(F(B(\omega, \cdot)))$. It is not hard to see that this is a generalization of the definition of convergence in distribution for random variables taking values in $\mathbf{R}$. For details, see Billingsley [2]. An alternative is to use nonstandard analysis to show that $B(\omega, t)$ can be constructed directly from a so-called "hyperfinite" random walk, as in Anderson [1].

Definition 2.1 A K-dimensional standard Brownian motion is a K-dimensional stochastic process B such that[2]

1. $B(\omega, 0) = 0$ almost surely (i.e. $P(\{\omega : B(\omega, 0) = 0\}) = 1$)

2. *Continuity*: $B(\omega, \cdot)$ is continuous almost surely. If Brownian motion is constructed as a limit of the random walk, this property comes from the tightness property of the random walk.

3. *Independent Increments*: If $0 \leq t_0 < t_1 < \cdots < t_m \in \mathcal{T}$,
$$\{B(\cdot, t_1) - B(\cdot, t_0), \ldots, B(\cdot, t_m) - B(\cdot, t_{m-1})\}$$
is an independent family of random variables. If Brownian motion is constructed as a limit of the random walk, this property comes from the independent increments property of the random walk.

4. *Normality*: If $0 \leq s \leq t$, $B(\cdot, t) - B(\cdot, s)$ is normal with mean $0 \in \mathbf{R}^K$ and covariance matrix $(t - s)I$, where I is the $K \times K$ identity matrix. If Brownian motion is constructed as a limit of the random walk, this property comes from the approximate normality of the random walk.

Theorem 2.2 *There is a probability space on which a K-dimensional standard Brownian motion exists.*

Example 2.3 *Time Change:* Given a K-dimensional standard Brownian motion B, let $Z(\omega, t) = B(\omega, \sigma^2 t)$. Thus, Z is obtained from B by speeding up time by a factor of σ^2. It is easy to see that Z satisfies all the properties of a standard Brownian motion, except that the covariance matrix of $B(\cdot, t) - B(\cdot, s)$ is $\sigma^2(t - s)I$ when $s < t$. If we let $\hat{Z}(\omega, t) = \frac{Z(\omega,t)}{\sigma}$, then $\hat{Z}$ satisfies all the properties of standard Brownian motion. Thus, a constant time change of a standard Brownian motion is a scalar multiple of a (different) standard Brownian motion on the same probability space.

Theorem 2.4 *The sample paths of standard Brownian motion have the following qualitative properties:*

[2] There is some redundancy among the conditions: continuity and independent increments imply normality (though not the specific mean and covariances given here), while independent increments and normality imply continuity.

1. *Almost Sure Unbounded Variation*[3]*:*

$$P(\{\omega : \exists_{s<t}\ B(\omega,\cdot) \textit{ is of bounded variation on } [s,t]\}) = 0$$

2. *Almost Sure Nowhere Differentiability*[4]*:*

$$P(\{\omega : \exists_{t\in\mathcal{T}}\ B(\omega,\cdot) \textit{ is differentiable at } t\}) = 0$$

3. *Iterated Logarithm Laws*

   (a) *Long Run:*

$$P\left(\left\{\omega : \limsup_{t\to\infty} \frac{B(\omega,t)}{\sqrt{2t\ln\ln t}} = 1\right\}\right) = 1$$

   (b) *Short Run:*[5] *For all $t \in \mathcal{T}$*

$$P\left(\left\{\omega : \limsup_{s\searrow t} \frac{B(\omega,s)-B(\omega,t)}{\sqrt{2(s-t)\ln|\ln(s-t)|}} = 1\right\}\right) = 1$$

Remark 2.5 The Iterated Logarithm Laws are key to understanding the qualitative short-run and long-run behavior of Brownian motion. We will model stock prices by processes like $e^{(\mu-\sigma^2/2)t+\sigma B(\omega,t)}$. Consider first the short run. If s is close to t, then $\sqrt{s-t}$ is much bigger than $s-t$. $\ln|\ln s-t|$ goes to infinity as $s \to t$, but the growth rate is *very* slow. The Iterated Logarithm Law tells us that at times s arbitrarily close to t, $B(\omega,s)-B(\omega,t)$ will nearly hit both the upper and lower envelopes $\pm\sqrt{2(s-t)\ln|\ln(s-t)|}$ infinitely often. In particular, in the short run, only the the volatility matters; the drift term $e^{(\mu-\sigma^2/2)t}$ is completely unimportant. On the other hand, in the long run, as $t\to\infty$, $\sqrt{t}\to\infty$ much slower than t; $\ln\ln t\to\infty$, but *very* slowly. We will see that $E(Z(\cdot,t)) = e^{\mu t}$. This explains why we choose to write $\mu-\sigma^2/2$, rather than incorporate the $-\sigma^2/2$ into μ. In the long run, if $\mu>0$, the volatility is overwhelmed in importance by the drift term $e^{\mu t}$.

---

[3]This property can be derived from the variation of the random walk, but the argument is a bit subtle, as the variation is not continuous in $C([0,T])$.

[4]This property shows that, although Brownian motion paths are continuous, they are only barely continuous. A slightly weaker property (Brownian motion paths are almost surely not continuously differentiable on any open interval) follows immediately from almost sure unbounded variation.

[5]In Problem Set 1, you are asked to derive this from the Iterated Logarithm Law in the Long Run.

You will see, in Problem Set 1, that the Quadratic Variation of the Random Walk cannot be defined pathwise; if the partition is allowed to depend in an arbitrary way on the path, the Quadratic Variation need not converge as $n \to \infty$. For the same reason, the Quadratic Variation of Brownian Motion is not defined pathwise. The following theorem says the the Quadratic Variation of Brownian Motion over every interval $[s,t]$ with $s < t$ is $t - s$:

Theorem 2.6 *Let B be a standard 1-dimensional Brownian Motion. Consider a sequence of partitions*

$$s = t_0^n < t_1^n < \cdots < t_{m_n}^n = t$$

indexed by n with

$$\max\left\{\left|t_k^n - t_{k-1}^n\right| : 1 \leq k \leq m_n\right\} \to 0$$

Then

$$\sum_{k=1}^{m_n} \left(B\left(\omega, t_k^n\right) - B\left(\omega, t_{k-1}^n\right)\right)^2 \to t - s \quad a.s.$$

as $n \to \infty$.

The theorem follows from the Strong Law of Large Numbers, using the fact that $B\left(\cdot, t_k^n\right) - B\left(\omega, t_{k-1}^n\right)$ is distributed as $N\left(0, t_k^n - t_{k-1}^n\right)$, so

$$E\left(\left(B\left(\cdot, t_k^n\right) - B\left(\cdot, t_{k-1}^n\right)\right)^2\right) = t_k^n - t_{k-1}^n$$

Proposition 2.7 *(Proposition 1.5 in Neilsen) If B is standard Brownian motion, $\frac{B(\omega,t)}{t} \to 0$ almost surely, i.e.*

$$P\left(\left\{\omega : \lim_{t\to\infty} \frac{B(\omega,t)}{t} = 0\right\}\right) = 1$$

Proof: Notice that from the definition of standard Brownian motion,

$$\mathrm{Var}\left(\frac{B(\cdot,t)}{t}\right) = \frac{\mathrm{Var}B(\cdot,t)}{t^2} = \frac{t}{t^2} \to 0$$

so $\frac{B(\cdot,t)}{t}$ converges to zero in distribution. However, convergence almost surely is stronger than convergence in distribution. Thus, we apply the Iterated Logarithm Law in the Long Run.

$$\limsup_{t\to\infty} \frac{B(\omega,t)}{t} \leq \limsup_{t\to\infty} \frac{B(\omega,t)}{\sqrt{2t\ln\ln t}} \times \limsup_{t\to\infty} \frac{\sqrt{2t\ln\ln t}}{t} = 1 \times 0 = 0$$

almost surely. Since $-B$ is standard Brownian motion,

$$\liminf \frac{B(\omega,t)}{t} = -\limsup_{t\to\infty} -\frac{B(\omega,t)}{t} = 0$$

almost surely. Therefore, $\limsup_{t\to\infty} \frac{B(\omega,t)}{t} = 0$ almost surely. ■

Proposition 2.8 *(Proposition 1.6 in Nielsen) If B is a standard Brownian motion, then the process $\hat{B}$ defined by*

$$\hat{B}(\omega,t) = \begin{cases} tB(\omega,1/t) & \textit{if} \quad t > 0 \\ 0 & \textit{if} \quad t = 0 \end{cases}$$

is also a standard Brownian motion.

Proof: Note that

$$\mathrm{Var}\hat{B}(\cdot,t) = t^2 \mathrm{Var} B(\cdot,1/t) = t^2 \times \frac{1}{t} = t$$

so $\hat{B}(\cdot,t)$ is normal with mean zero and variance t. The other properties of standard Brownian motion follow immediately from the corresponding properties for B.■

3 Generalized Brownian Motion and Correlated Brownian Motion

A generalized Brownian motion allows more freedom than a standard Brownian motion, in the following respects:

1. it may start at an arbitrary level, not just at zero;

2. it may incorporate a deterministic drift term; and

3. the covariance among different components can be nonzero, and the variance of each component may grow at its own deterministic rate.

Definition 3.1 A K-dimensional process Z is a K-dimensional *generalized Brownian motion* if

1. $Z(\omega, 0)$ is deterministic (i.e. independent of ω)

2. $Z(\omega, \cdot)$ is continuous almost surely

3. If $0 \leq t_0 < \cdots < t_n$, then

$$\{Z(\cdot, t_1) - Z(\cdot, t_0), \ldots, Z(\cdot, t_n) - Z(\cdot, t_{n-1})\}$$

   are independent

4. There exist $\mu \in \mathbf{R}^K$ and a matrix $\Sigma \in \mathbf{R}^{K \times K}$ such that if $0 \leq s < t$, $Z(\cdot, t) - Z(\cdot, s)$ is normal with mean $(t-s)\mu$ and covariance matrix $(t-s)\Sigma$.

Notice that since Σ is a covariance matrix, it must be symmetric. μ is called the increment mean vector and Σ the increment covariance matrix.

The following theorem is important because it lets us write a generalized Brownian motion as a linear transformation of a standard Brownian motion, hence allowing us to reduce computations concerning generalized Brownian motions to standard Brownian motions.

Theorem 3.2 *Z is a generalized N-dimensional Brownian motion with positive semidefinite increment covariance matrix Σ if and only if there is a constant $Z_0 \in \mathbf{R}^N$, $\mu \in \mathbf{R}^N$, $\sigma \in \mathbf{R}^{N \times K}$ and a K-dimensional standard Brownian motion B such that*

$$Z(\omega, t) = Z_0 + t\mu + \sigma B(\omega, t)$$

In this case, the increment mean vector of Z is μ and the increment covariance matrix of Z is $\Sigma = \sigma\sigma^T$.

Remark 3.3 Note that we may have $N > K$, $N = K$, or $N < K$. Suppose we think of the components of the generalized Brownian motion as the available securities. If $N > K$, this corresponds to a case in which securities are redundant; there are more securities than there are underlying sources of uncertainty. Σ will be positive semi-definite, but cannot be positive definite. If $N < K$, there are fewer securities than there are underlying sources of uncertainty, and markets will necessarily be incomplete. If $N = K$, then markets are potentially dynamically complete. If $N \leq K$, then Σ may be either positive definite, but it may also be just positive semi-definite. To see that Σ is always at least postive semi-definite, note that

$$x^T \Sigma x = x^T \sigma \sigma^T x = (x^T \sigma)(x^T \sigma)^T = |x^T \sigma|^2 \geq 0$$

Nielsen also defines a *correlated Brownian motion* as a gneralized Brownian motion in which $Z(\cdot, 0) = 0$, $\mu = 0$, and $\Sigma_{ii} = 1$ for each i, i.e. the diagonal elements of Σ are all one. There is a representation theorem for correlated Brownian motions. See Nielsen for details.

Example 3.4 [Black-Scholes Model] (Example 1.7 in Nielsen) There is one stock, whose price is

$$S(\omega, t) = S(0) e^{\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B(\omega, t)}$$

where B is a one-dimensional standard Brownian motion, $S(0) > 0, \mu \in \mathbf{R}$, $\sigma > 0$. Notice that

$$\ln S(\omega, t) = \ln S(0) + \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B(\omega, t)$$

is a generalized Brownian motion. The continuously compounded rate of return over $[t, t + \tau]$ is

$$\begin{aligned}
&\frac{1}{\tau} \ln \left(\frac{S(\omega, t + \tau)}{S(\omega, t)} \right) \\
&= \frac{1}{\tau} \left[\ln S(0) + \left(\mu - \frac{\sigma^2}{2}\right)(t + \tau) + \sigma B(\omega, t + \tau) - \ln S(0) - \left(\mu - \frac{\sigma^2}{2}\right) t - \sigma B(\omega, t) \right] \\
&= \frac{1}{\tau} \left[\left(\mu - \frac{\sigma^2}{2}\right) \tau + \sigma (B(\omega, t + \tau) - B(\omega, t)) \right] \\
&= \left(\mu - \frac{\sigma^2}{2}\right) + \sigma \frac{B(\omega, t + \tau) - B(\omega, t)}{\tau}
\end{aligned}$$

If we define $Z(\omega,\tau) = B(\omega, t+\tau) = B(\omega, t)$, then Z is a standard Brownian motion, so

$$\sigma \frac{B(\omega, t+\tau) - B(\omega, t)}{\tau} \to 0$$

almost surely; thus, the continuously compounded rate of return converges to a constant $\mu - \frac{\sigma^2}{2}$ as $t \to \infty$.

4 Information Structures

Recall that in the random walk, we defined $\mathcal{F}_t$ to be the collection of events definable in terms ofcoin tosses that had occurred up to time $\frac{\lfloor nt \rfloor}{n}$; it simply represents the information available at time t. Equivalently, $\mathcal{F}_t$ is the σ-algebra determined by $\hat{X}$ up to time t, or by X up to time $\frac{\lfloor nt \rfloor}{n}$. We need to extend this definition to continuous-time processes where the probability space is infinite.

Definition 4.1 A *filtration* is a family $(\mathcal{F}_t)_{t \in \mathcal{T}}$ of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ such that $\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s \le t$. A filtration is *augmented* (sometimes called *complete*) if

$$C \subset B, P(B) = 0 \;\; \Rightarrow \;\; \forall_{t \in \mathcal{T}} \, C \in \mathcal{F}_t$$

A stochastic process Z is *adapted* to $(\mathcal{F}_t)_{t \in \mathcal{T}}$ if $Z(\cdot, t)$ is $\mathcal{F}_t$-measurable for all $t \in \mathcal{T}$. Every stochastic process Z generates a filtration: $\mathcal{F}_t$ is, roughly speaking, the σ-algebra of events revealed by Z up to and including time t. More formally, $\mathcal{F}_t$ is the smallest σ-algebra containing

$$\{\{\omega : Z(\cdot, s) \in (a,b)\}, s \le t, a, b \in \mathbf{R}\}$$

Every stochastic process Z is adapted to the filtration it generates. $\hat{X}$ is adapted to the filtration we defined in the random walk model, but X is *not* adapted to that filtration.

Remark 4.2 $\mathcal{F}_t$ is interpreted as the information which has been revealed by time t. Suppose Z is a trading strategy, i.e. $Z(\omega, t)$ specified how many shares of each stock an individual will hold at (ω, t). Then Z must be adapted; the individual can't make decisions based on information that hasn't yet been revealed. There is another reason to insist that trading strategies must be adapted. If we allowed trading strategies that are not adapted, there would be arbitrage. An example of a non-adapted trading strategy would be "buy the stock today if its price will be higher tomorrow, but sell it short today if its price will be lower tomorrow." Its clear that this strategy guarantees a profit, and the profit can be made arbitrarily large by increasing the number of shares that are bought or sold short; thus, if nonadapted trading strategies were allowed, individuals would take actions that would force the price today to change to eliminate the arbitrage, and this price change would reveal the information on which the individuals were basing their trades, enlarging the filtration. Notice, however, that requiring that trading strategies be adapted imposes a fundamental limitation, because it does not allow us to study situations with asymmetric information. In reality, the information possessed varies considerably from one individual to another. Market microstructure focusses on how agents that are better informed than others use that information, and how the information is incorporated into prices as a result. The continuous-time formulation makes it difficult or impossible to address those kinds of questions; in effect, it is assumed that all agents see the same information at any given time.

Definition 4.3 Let $(\Omega, \mathcal{F}, P)$ be a probability space. A stochastic process Z is *measurable* if it is measurable with respect to the product σ-algebra on $\Omega \otimes \mathcal{T}$. Z is *integrable* if $Z(\cdot, t)$ is integrable for all $t \in \mathcal{T}$. Suppose Y is a random variable which is integrable and $\mathcal{G} \subset \mathcal{F}$. The *conditional expectation* of Y with respect to $\mathcal{G}$ is a random variable $W = E(Y|\mathcal{G})$ such that W is $\mathcal{G}$-measurable, and $\int_G W dP(\omega) = \int_G Y dP(\omega)$ for all $G \in \mathcal{G}$. The existence of the conditional expectation is proven using the Radon-Nikodym Theorem; any two conditional expectations agree almost surely.

Remark 4.4 When Ω is finite, as in the random walk model, any σ-algebra $\mathcal{G}$ must be the collection of all unions of elements of a partition of Ω. For example, given $t < T$, we can define the partition of Ω determined by $\omega_1, \omega_2, \ldots, \omega_{\lfloor nt \rfloor}$. The partition sets are sets of the form

$$\{\omega' \in \Omega : \omega'_k = \omega_k \, (1 \leq k \leq nt)\}$$

Define $\omega' \sim_t \omega$ if $\omega'_k = \omega_k$ for $k \leq nt$. This partition generates the σ-algebra $\mathcal{F}_t$ in the sense that $\mathcal{F}_t$ consists precisely of all unions of partition sets. $E(W|\mathcal{F}_t)$ is computed by taking the average value of W over each of the partition sets:

$$E(W|\mathcal{F}_t)(\omega) = \frac{\sum_{\{\omega':\omega'\sim_t\omega\}} W(\omega)}{|\{\omega' : \omega' \sim_t \omega\}|}$$

Definition 4.5 Z is a *martingale* with respect to a filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$ if

1. Z is integrable
2. Z is adapted to $(\mathcal{F}_t)_{t\in\mathcal{T}}$
3. For all $s, t \in \mathcal{T}$ with $s \le t$

$$E\left(Z(\cdot, t)|\mathcal{F}_S\right) = Z(\cdot, s)$$

   almost surely.

Example 4.6 If we let $\mathcal{T} = [0, T]$, the random walk $\hat{X}_n(\omega, t)$ is a martingale with respect to $(\mathcal{F}_t)_{t\in\mathcal{T}}$. To see this, compute

$$\begin{aligned}
&E(\hat{X}_n(\cdot, t)|\mathcal{F}_s)(\omega_0) \\
&= \sum_{\omega\sim_s\omega_0} \frac{\hat{X}_n(\omega, t)}{2^{n(t-s)}} \\
&= \hat{X}_n(\omega_0, s) + \sum_{\omega\sim_s\omega_0} \sum_{k=ns+1}^{nt} \frac{\omega_k}{2^{n(t-s)}\sqrt{n}} \\
&= \hat{X}_n(\omega_0, s) + \sum_{k=ns+1}^{nt} \sum_{\omega\sim_s\omega_0} \frac{\omega_k}{2^{n(t-s)}\sqrt{n}} \\
&= \hat{X}_n(\omega_0, s) + \sum_{k=ns+1}^{nt} \frac{0}{2^{n(t-s)}\sqrt{n}} \\
&= \hat{X}_n(\omega_0, s)
\end{aligned}$$

since each ω_k is 1 exactly half the time and -1 exactly half the time. Note that X_n is not a martingale on $[0, T]$; it is not adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$, and it is not a martingale with respect to the filtration it generates. If we restrict the time set to the set of points $\left\{0, \frac{1}{n}, \ldots, T\right\}$, X_n is a martingale on this restricted time set.

Example 4.7 If B is standard Brownian motion, then B is a martingale with respect to the filtration it generates. This follows from the fact that the increment $B(\cdot,t)-B(\cdot,s)$ has mean zero and is independent of $\mathcal{F}_s$, while $B(\cdot,s)$ is measurable with respect to $\mathcal{F}_s$:

$$\begin{aligned}
&E\left(B(\cdot,t)|\mathcal{F}_s\right)(\omega_0)\\
&\quad= E\left(B(\omega_0,s)+B(\cdot,t)-B(\cdot,s)|\mathcal{F}_s\right)(\omega_0)\\
&\quad= B(\omega_0,s)+E\left(B(\cdot,t)-B(\cdot,s)|\mathcal{F}_s\right)(\omega_0)\\
&\quad= B(\omega_0,s)+0
\end{aligned}$$

5 Wiener Processes

This is section 1.5 in Nielsen.

Definition 5.1 We take as given a filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$. A standard Wiener process is a stochastic process such that

1. $W(\cdot,0)=0$ almost surely.
2. $W(\omega,\cdot)$ is continuous almost surely.
3. W is adapted to $(\mathcal{F}_t)_{t\in\mathcal{T}}$.
4. $W(\cdot,t)-W(\cdot,s)$ is independent of $\mathcal{F}_s$ if $s\leq t$
5. $W(\cdot,t)-W(\cdot,s)$ is normal with mean zero and variance $(t-s)I$

The essential difference between a Wiener process W (as defined by Nielsen) and a standard Brownian motion is that the filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$ is given in advance; it may well be bigger than the filtration generated by W. This is useful for formulating continuous-time finance, since we want to allow for the possibility that there is uncertainty in the model beyond that which is captured in the available securities.

In addition, there are generalized and correlated Wiener processes corresponding to generalized and correlated Brownian motions; see Nielsen for details.

In the lecture, I showed you how to evaluate the integrals in Nielsen's Example 1.15 by completing the square in the exponential, in order to help you do the second question on Problem Set 2. If you succeeded in solving that question, you understand Example 1.15.

You can read or skip Nielsen's Section 1.7, at your option.

6 Stochastic Integrals and Capital Gains

In sections 1.8 and 1.9, Nielsen motivates time integrals and stochastic integrals as ways of generating new kinds of processes. This is certainly true, but there is a much stronger finance motivation for studying them: they are essential to defining the capital gains generated by a trading strategy. In this section, we motivate the stochastic integral by considering the random walk model.

As we have seen, the most common model for stock prices is the exponential of a generalized Brownian motion. Since we're just trying to motivate the stochastic integral, we pretend that the stock price is given by the random walk process X_n, rather than e^{X_n} or e^B. In particular, we assume there is only one stock. Let $\mathcal{T} = \{0, 1/n, \ldots, nT\}$.

Suppose an individual uses the trading strategy $\bar{\Delta}(\omega, t)$. In other words, $\bar{\Delta}$ is a stochastic process which tells the individual how many shares to hold at time t, when the state is ω. We require that $\bar{\Delta}$ be adapted with respect to the filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$. This will be true if and only if $\bar{\Delta}(\omega, t)$ depends only on $\omega_1, \ldots, \omega_{\lfloor nt \rfloor}$. We assume that $\bar{\Delta}(\omega, t)$ is constant on intervals of the form $\left[\frac{k}{n}, \frac{k+1}{n}\right)$; thus, the individual changes his/her portfolio holdings only at times $t \in \mathcal{T}$; this assumption does not alter the set of possible portfolio returns available to the individual.

What is the capital gain generated by the trading strategy $\bar{\Delta}(\omega, t)$? The capital gain between time k/n and $\frac{k+1}{n}$ is

$$\bar{\Delta}\left(\omega, \frac{k}{n}\right)\left(X\left(\omega, \frac{k+1}{n}\right) - X\left(\omega, \frac{k}{n}\right)\right)$$

so the capital gains process up to time $t \in \mathcal{T}$ is

$$\begin{aligned} G(\omega, t) &= \sum_{k=0}^{nt-1} \bar{\Delta}\left(\omega, \frac{k}{n}\right)\left(X\left(\omega, \frac{k+1}{n}\right) - X\left(\omega, \frac{k}{n}\right)\right) \\ &= \sum_{k=0}^{nt-1} \bar{\Delta}\left(\omega, \frac{k}{n}\right)\frac{\omega_{k+1}}{\sqrt{n}} \end{aligned}$$

This is called a Riemann-Stieltjes integral with respect to the integrator $X(\omega, \cdot)$; it is formed by taking values of the integrand $\bar{\Delta}$ and multiplying by changes in the value of the integrator X.

Riemann-Stieltjes integrals are normally defined provided that the integrand is continuous and the integrator is of bounded variation; the integrand $\bar{\Delta}$ is not continuous, but it is a step function, and the Riemann-Stieltjes integral is also defined in this case. The definition of $\int_a^b f(t)dg(t)$ begins with partitions: Suppose $a = t_0 < t_1 < \cdots < t_n = b$, then the Riemann-Stieltjes sum with respect to this partition is

$$\sum_{i=0}^{n-1} f(t_i)(g(t_{i+1}) - g(t_i))$$

The Riemann-Stieltjes integral is defined as the limit of the Riemann-Stieltjes sums as the partition gets finer and finer, provided the limit exists. Note that in defining the Riemann-Stieltjes integral with respect to the random walk, we need to consider partitions finer than the time points $\frac{k}{n}$ at which the random walk coins are tossed. The Riemann-Stieltjes integral makes perfect sense in the random walk model because the integrators (the paths of the random walk) are of bounded variation. It is true that the variation of the paths goes to infinity as n grows, but for each fixed n, every random walk path is piecewise linear, hence of bounded variation.

However, the paths of Brownian motion are almost surely *not* of bounded variation, so one cannot define the capital gain simply by taking a Riemann-Stieltjes integral. Itô finessed this problem by approximating the integrand by simple functions (functions which are piecewise constant over time). The Stieltjes integral makes sense if the integand is a simple function, even if the integrator is not of bounded variation. The properties of Brownian motion allowed Itô to extend this Stieltjes integral from adapted simple stochastic processes to adapted square-integrable stochastic processes. In the next section, we will give Itô's definition of the stochastic integral.[6]

[6] An alternative approach is to use nonstandard analysis and construct Brownian motion as a hyperfinite random walk. The Itô integral with respect to the Brownian motion can be recovered from the Stieltjes integral with respect to the hyperfinite random walk; see Anderson [1] for details.

7 Formal Definition of the Stochastic Integral

We begin with some preliminary material.

Suppose $(A, \mathcal{A}, \mu)$ is a measure space. We have three main examples in mind:

- $(\Omega, \mathcal{F}, P)$, the probability space representing the uncertainty
- $(\mathcal{T}, \mathcal{C}, \lambda)$, the Lebesgue measure space on the time set $\mathcal{T}$
- $(\Omega \otimes \mathcal{T}, \mathcal{F} \otimes \mathcal{C}, P \otimes \lambda)$, the product of the space of uncertainty and time.

Definition 7.1

$$L^1(A) = \{f : A \to \mathbf{R},\ f \text{ measurable},\ \int_A |f| d\mu < \infty\}$$

$$L^2(A) = \{f : A \to \mathbf{R},\ f \text{ measurable},\ \int_A f^2 d\mu < \infty\}$$

We identify two elements f, g of $L^1(A)$ or $L^2(A)$ if $f = g$ except on a set of μ measure zero. Note that if $A = \Omega$, L^2 is the set of random variables with finite variances, and L^1 is the set of random variables with finite means. If $A = \Omega \otimes \mathcal{T}$, $L^1(A)$ and $L^2(A)$ are sets of stochastic processes. By Fubini's Theorem, if Z is a measurable process, then

$$\begin{aligned} & \int_{\Omega \otimes \mathcal{T}} Z^2(\omega, t) d(P \times \lambda) \\ = & \int_{\mathcal{T}} \left(\int_{\Omega} Z^2(\omega, t) dP \right) d\lambda \\ = & \int_{\Omega} \left(\int_{\mathcal{T}} Z^2(\omega, t) d\lambda \right) dP \end{aligned}$$

$L^1(A)$ and $L^2(A)$ are Banach spaces under the norms $\|f\|_1 = \int_A |f| d\mu$ on $L^1(A)$ and $\|f\|_2 = \left(\int_A f^2 d\mu\right)^{1/2}$ on $L^2(A)$. In other words, if we let $d_1(f, g) = \|f - g\|_1$ and $d_2(f, g) = \|f - g\|_2$ be the metrics induced by these norms, then $(L^1(A), d_1)$ and $(L^2(A), d_2)$ are complete metric spaces. A complete metric space is one with the property that every Cauchy sequence converges to an element of the metric space. Thus, if we have a sequence of functions $f_n \in L^2(A)$ and f_n is Cauchy, i.e.

$$\forall_{\varepsilon > 0}\, \exists_N\, \forall_{m,n > N} \|f_m - f_n\|_2 < \varepsilon$$

then

$$\exists_{f \in L^2(A)} \|f_n - f\|_2 \to 0$$

The analogous property is true for $L^1(A)$.

We say that f_n converges to f in probability if, for every $\varepsilon > 0$,

$$P(\{\omega : |f_n(\omega) - f(\omega)| > \varepsilon\}) \to 0$$

We first turn to time integrals. Fix a probability space $(\Omega, \mathcal{F}, P)$ and a filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$. Let $\mathcal{L}^1$ be the set of adapted, measurable processes a such that, for all $t \in \mathcal{T}$,

$$\int_0^t \|a(\omega, s)\| ds < \infty \tag{1}$$

almost surely. Here, $a(\omega, s)$ may either be a scalar or a vector.[7] The condition in Equation (1) says that $a(\omega, \cdot) \in L^1(\mathcal{T})$ for almost all ω; this is weaker than saying $a \in L^1(\Omega \otimes [0, T])$, which requires

$$\int_\Omega \int_0^T \|a(\omega, s)\| ds dP < \infty$$

If $a \in \mathcal{L}^1$, we write $\int_0^t a(\omega, s) ds$ as shorthand for $\int_{[0,t]} a(\omega, s) d\lambda(s)$, the integral of $a(\omega, \cdot)$ over $[0, t]$ with respect to Lebesgue measure.

Proposition 7.2 (Proposition 1.25 in Nielsen) *If $a \in \mathcal{L}^1$, then $\int_0^t a(\omega, s) ds$ is adapted, continuous in t almost surely in ω, and hence measurable.*

Remark 7.3 a itself can be quite discontinuous in time. But since for all t, $a(\omega, \cdot) \in L^1([0, t])$ almost surely,

$$\begin{aligned} &\lim_{\varepsilon \searrow 0} \left(\int_0^{t+\varepsilon} a(\omega, s) ds - \int_0^t a(\omega, s) ds \right) \\ &= \lim_{\varepsilon \searrow 0} \int_t^{t+\varepsilon} a(\omega, s) ds = 0 \end{aligned}$$

because of the countable additivity of Lebesgue measure, the fact that $\lambda(\{t\}) = 0$, and the fact that $a(\omega, \cdot) \in L^2([0, T])$ almost surely. This shows that the

[7] If $a(\omega, s)$ is a scalar, $\|a(\omega, s)\| = |a(\omega, s)|$, the absolute value. If $a(\omega, s)$ is a vector, Nielsen takes $\|a(\omega, s)\| = \|a(\omega, s)\|_2$ to be the Euclidean length of $a(\omega, s)$, one could also take $\|a(\omega, s)\| = \|a(\omega, s)\|_1 = \sum_j |a_j(\omega, s)|$, which meshes better with L^1, but does not change the set of processes in $\mathcal{L}^1$.

integral is almost surely (in ω) continuous (in t). If $b(\omega, \cdot)$ has a discontinuity at t_0, this will typically introduce a kink, but never a discontinuity, in $\int_0^t b(\omega, s)ds$ at $t = t_0$.

We now turn to the definition of the Itô Integral. The physical interpretation of the Itô Integral arises from diffusion processes. The change dW of a Wiener process gives a standard diffusion, occurring at a constant rate. The integrand b specifies how fast the diffusion is occurring at a particular time s and state ω; for a smoke particle being bombarded by air molecules, the rate is a function of the temperature and pressure of the air, and the mass of the smoke particle. In finance, dW represents the volatility of the stock price, while the integrand b represents the portfolio holding. The class $\mathcal{L}^2$ in the following definition is the set of stochastic processes which can be Itô integrated.

Definition 7.4 Fix a filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$ and a K-dimensional Wiener process W with respect to $(\mathcal{F}_t)_{t \in \mathcal{T}}$. Let $\mathcal{L}^2$ denote the set of adapted, measurable processes $b : \Omega \otimes \mathcal{T} \to \mathbf{R}^m$ (where $\mathbf{R}^m$ may denote $\mathbf{R}^1$, $\mathbf{R}^K$, or $\mathbf{R}^{N \times K}$) such that

$$\int_0^t \|b(\omega, s)\|^2 ds < \infty$$

almost surely.[8] Let $\mathcal{H}^2 = \mathcal{L}^2 \cap L^2(\Omega \otimes \mathcal{T})$.

The definition of the Itô proceeds in stages, starting first with simple functions, then extending to $\mathcal{H}^2$, and finally extending to $\mathcal{L}^2$.

Step 1: First suppose $K = 1$, $b : \Omega \times \mathcal{T} \to \mathbf{R}$, and $T \in \mathcal{T}$. Fix $0 = t_0 < t_1 < \cdots < t_n = T$. Assume that $b \in \mathcal{H}^2$, and b is simple[9], i.e.

$$b(\omega, s) = b(\omega, t_k) \text{ for all } s \in [t_k, t_{k+1})$$

Define

$$\int_0^T b dW(\omega) = \sum_{k=0}^{n-1} b(\omega, t_k)(W(\omega, t_{k+1}) - W(\omega, t_k))$$

Observe that this is a Stieltjes integral; it makes sense, even though $W(\omega, \cdot)$ is not of bounded variation, because b is simple.

[8] $\|b(\omega, s)\|$ denotes the Euclidean length of the scalar, vector, or matrix $b(\omega, s)$. For example, if $b(\omega, s)$ is an $N \times K$ matrix, $\|b(\omega, s)\|^2 = \sum_{ij} (b_{ij}(\omega, s))^2$.

[9] Our convention is different from that of Nielsen; his simple functions are left-continuous, while ours are right-continuous.

Lemma 7.5 (Itô Isometry) *If $b \in \mathcal{H}^2$ and b is simple, then*

$$\int_\Omega \left(\int_0^T b dW \right)^2 dP = \int_\Omega \int_0^T |b(\omega, s)|^2 \, ds \, dP$$

In other words,

$$\left\| \int_0^T b dW \right\|_2 = \|b\|_2$$

where the norm on the left side is the norm in $L^2(\Omega)$ and the norm on the right side is the norm in $L^2(\Omega \times [0, T])$.

Proof:

$$\begin{aligned}
&\int_\Omega \left(\int_0^T b \, dW \right)^2 dP \\
&= \int_\Omega \left(\sum_{k=0}^{n-1} b(\omega, t_k) \left(W(\omega, t_{k+1}) - W(\omega, t_k) \right) \right)^2 dP \\
&= \int_\Omega \left[\sum_{k=0}^{n-1} b^2(\omega, t_k) \left(W(\omega, t_{k+1}) - W(\omega, t_k) \right)^2 \right. \\
&\quad \left. + 2 \sum_{j<k} b(\omega, t_j) b(\omega, t_k) \left(W(\omega, t_{j+1}) - W(\omega, t_j) \right) \left(W(\omega, t_{k+1}) - W(\omega, t_k) \right) \right] dP \\
&= \sum_{k=0}^{n-1} \left(\int_\Omega b^2(\omega, t_k) \, dP \right) \left(\int_\Omega \left(W(\omega, t_{k+1}) - W(\omega, t_k) \right)^2 dP \right) \qquad (2) \\
&\quad + 2 \sum_{j<k} \left(\int_\Omega b(\omega, t_j) b(\omega, t_k) \left(W(\omega, t_{j+1}) - W(\omega, t_j) \right) dP \right) \left(\int_\Omega \left(W(\omega, t_{k+1}) - W(\omega, t_k) \right) dP \right) \\
&= \sum_{k=0}^{n-1} \left(\left(\int_\Omega b^2(\omega, t_k) \, dP \right) (t_{k+1} - t_k) \right) + 0 \qquad (3) \\
&= \int_\Omega \sum_{k=0}^{n-1} b^2(\omega, t_k)(t_{k+1} - t_k)) \, dP \\
&= \int_\Omega \int_0^T b^2(\omega, t) \, dt \, dP \\
&= \int_{\Omega \times [0,T]} \int_0^T b^2 \, d(P \otimes \lambda)
\end{aligned}$$

Equation (2) follows because $b(\cdot, t_j)$, $b(\cdot, t_k)$ and $W(\cdot, t_{j+1}) - W(\cdot, t_j)$are independent of $W(\cdot, t_{k+1}) - W(\cdot, t_k)$, while Equation (3) follows from the fact that $W(\cdot, t_{k+1}) - W(\cdot, t_k)$ has mean zero and variance $t_{k+1} - t_k$. ■

Step 2: Extend the Itô Integral to $\mathcal{H}^2$. If $b \in \mathcal{H}^2$, fix n and let $t_k = \frac{k}{n}$, then define

$$b_n(\omega, t) = n \int_{t_{k-1}}^{t_k} b(\omega, s)\, ds \text{ if } t \in [t_k, t_{k+1})$$

For each time interval $[t_k, t_{k+1})$, $b_n(\omega, t)$ is the average of $b(\omega, \cdot)$ over the *previous* interval $[t_{k-1}, t_k)$; this ensures that b_n is simple and adapted. Lusin's Theorem (which states, roughly speaking, that measurable functions are continuous functions on the complement of a set of arbitrarily small measure) then can be used to show that $\|b - b_n\|_2 \to 0$, so the sequence b_n is Cauchy in $L^2(\Omega \times [0, T])$. Thus, given $\varepsilon > 0$, there exists N such that if $m, n > N$, $\|b_m - b_n\|_2 < \varepsilon$. But by the Itô Isometry, if $m, n > N$

$$\begin{aligned} & \left\| \int_0^T b_m\, dW - \int_0^T b_n\, dW \right\|_2 \\ = & \left\| \int_0^T (b_m - b_n)\, dW \right\|_2 \\ = & \|b_m - b_n\|_2 \\ < & \varepsilon \end{aligned}$$

so the sequence $\int_0^T b_m\, dW$ is a Cauchy sequence in $L^2(\Omega)$, hence converges to a unique limit; we define $\int_0^T b\, dW$ to be this limit.

Step 3: Now suppose $b \in \mathcal{L}^2$, so $\int_0^T |b(\omega, s)|^2\, ds < \infty$ almost surely in ω. Let

$$b_n(\omega, s) = \begin{cases} n & \text{if } b(\omega, s) > n \\ b(\omega, s) & \text{if } -n \le b(\omega, s) \le n \\ -n & \text{if } b(\omega, s) < -n \end{cases}$$

Then $b_n \in \mathcal{H}^2$ and that $\int_0^T |b_n - b|^2\, ds \to 0$ almost surely (specifically, for each ω such that $b(\omega, \cdot) \in L^2([0, T])$). One can show that $\int_0^T b_m\, dW$ converges in probability; $\int_0^T b\, dW$ is defined to be the limit.

Step 4: If W is K-dimensional, and $b(\omega, s) \in \mathbf{R}^K$, define

$$\int_0^T b\, dW = \sum_{k=1}^K \int_0^T b_k\, dW_k$$

Notice that if we think of W as the price process of K stocks and b as the portfolio strategy, then $\int_0^T b\,dW$ is the capital gain from the portfolio, the sum of the capital gains on the individual stocks.

If W is K-dimensional, and $b(\omega, s) \in \mathbf{R}^{N\times K}$, define

$$\left(\int_0^T b\,dW\right)_j = \sum_{k=1}^{K} \int_0^T b_{jk}\,dW_k$$

Think of there being N stocks, each of whose price movements is determined by the components of the underlying Wiener process. b_{jk} give the coefficient of stock j on the k^{th} component of the Wiener process and $\int b\,dW$ gives the movement of the N-dimensional vector of stock prices. Note that if b is a K-dimensional vector process, the stochastic integral is a scalar process; if b is an $N \times K$ matrix process, the stochastic integral is an N-dimensional vector process.

The stochastic integral is better behaved mathematically for integrands $b \in \mathcal{H}^2$ than for integrands in $\mathcal{L}^2$. However, $\mathcal{H}^2$ is not closed under the manipulations we need to do in Finance, while $\mathcal{L}^2$ is; hence, we need to consider integrands in $\mathcal{L}^2$.

We have the following facts concerning the Itô Integral for integrands $b \in \mathcal{L}^2$:

- Our definition of $\int_0^T b\,dW$ was given for a single T, and is defined only up to a set of probability zero. Since the set of probability zero can be different for different choices of T, the paths of $\int_0^T b\,dW$ could be badly behaved. Fortunately, it is possible to choose a continuous version of the the integral, i.e. we may assume that except for a set of ω of probability zero, $\int_0^t b(\omega, s)\,dW(\omega, s)$ is continuous in t.

- Linearity:
$$\gamma \int_0^t a\,dW + \delta \int_0^t b\,dW = \int_0^t (\gamma a + \delta b)\,dW$$

- Time consistency: If $0 \le s \le t$, then
$$\int_0^s b\,dW = \int_0^t (\mathbf{1}_{\omega\times[0,s]} b)\,dW$$
where $\mathbf{1}_B$ denotes the indicator function of the set B.

- The Itô Integral is adapted, i.e. $\int_0^t b\,dW$ is an adapted process. This is easily seen to be true for simple processes in $\mathcal{H}^2$, and it is inherited as the integral is defined by limits.

- If Y is a $\mathcal{F}_s$-measurable random variable,

$$\int_s^t (Yb)\,dW = Y \int_s^t b\,dW$$

This would be trivial if the Itô were defined pathwise, but as we have seen, it is not. However, one can verify the property for simple processes in $\mathcal{H}^2$, and verify it is preserved when one takes limits.

The following proposition provides important additional properties of the Itô Integral when the integrand is in $\mathcal{H}^2$.

Proposition 7.6 (Proposition 1.37 in Nielsen) *Let W be a K-dimensional Wiener process.*

1. *If $b : \Omega \times \mathcal{T} \to \mathbf{R}^K$ and $b \in \mathcal{H}^2$, then $\int_0^t b\,dW$ is a martingale.*[10]
2. *If $b, \beta : \Omega \times \mathcal{T} \to \mathbf{R}^K$, and $b, \beta \in \mathcal{H}^2$, then*[11]

$$\begin{aligned}
\mathrm{Cov}\left(\int_s^t b\,dW, \int_s^t \beta\,dW \middle| \mathcal{F}_s \right) &= E\left(\left(\int_s^t b\,dW \right) \left(\int_s^t \beta\,dW \right) \middle| \mathcal{F}_s \right) \\
&= E\left(\int_s^t b \cdot \beta\,du \middle| \mathcal{F}_s \right) \\
&= \int_s^t E\left(b(u) \cdot \beta(u) | \mathcal{F}_s \right)\,du
\end{aligned}$$

[10] If $b \in \mathcal{L}^2$, it is not necessarily the case that $\int b\,dW$ is a martingale; indeed, there is no guarantee that $\int_0^t b\,dW \in L^1(\Omega)$, so the integrals in the definition of a martingale may not even be defined.

[11] Nielsen writes these expressions in integral form, but they are covariances since the means are zero. The k^{th} components of b and β are the coefficients of the stochastic integrals on the k^{th} components of the Wiender process. Because distinct components of the Wiener process are independent, the covariance of W_k and W_ℓ is zero when $k \neq \ell$, so the terms $b_k \beta_\ell$ disappear from the expression, leaving only $b \cdot \beta = b_1\beta_1 + \cdots + b_K\beta_K$.

3. If $b : \Omega \times \mathcal{T} \to \mathbf{R}^{N\times K}$, and $b \in \mathcal{H}^2$, then

$$\begin{aligned}
\mathrm{Cov}\left(\int_s^t b\,dW, \int_s^t b\,dW \middle| \mathcal{F}_s\right) &= E\left(\left(\int_s^t b\,dW\right)\left(\int_s^t b\,dW\right)^T \middle| \mathcal{F}_s\right) \\
&= E\left(\int_s^t bb^T\,du \middle| \mathcal{F}_s\right) \\
&= \int_s^t E\left(b(u)b(u)^T \middle| \mathcal{F}_s\right)\,du
\end{aligned}$$

Thus, bb^T is called the instantaneous covariance matrix of the stochastic integral.

Corollary 7.7 (Corollary 1.38 in Nielsen) *If $b, \beta : \Omega \times \mathcal{T} \to \mathbf{R}^K$, and $b, \beta \in \mathcal{H}^2$, and $0 \le s \le t \le u$, then*

$$\mathrm{Cov}\left(\int_s^t b\,dW, \int_t^u \beta\,dW \middle| \mathcal{F}_s\right) = 0$$

and

$$\mathrm{Cov}\left(\int_s^t b\,dW, \int_t^u \beta\,dW\right) = 0$$

The previous Corollary shows that increments of stochastic integrals over disjoint time intervals are uncorrelated. As the following example shows, they are not generally independent.

Example 7.8 Let W be a 1-dimensional standard Wiener process, and

$$b(\omega, t) = \begin{cases} 1 & \text{if } W(\omega, s) < 1 \text{ for all } s < t \\ 0 & \text{otherwise} \end{cases}$$

Then $Z(\omega, t) = \int_0^t b(\omega, s)dW(\omega, s)$ follows the path $W(\omega, \cdot)$ up until the first time t at which $W(\omega, t) = 1$, at which point it stops. More formally, define $\tau(\omega) = \min\{t : X(\omega, t) = 1\}$; $\tau(\omega)$ is defined almost surely because $X(\omega, \cdot)$ is continuous almost surely. Then

$$Z(\omega, t) = W(\omega, t \wedge \tau(\omega))$$

where $t \wedge s$ denotes $\min\{t, s\}$. Notice that the increments of Z are *not* independent. Indeed, if $0 < s < t$ and $Z(\omega, s) = Z(\omega, s) - Z(\omega, 0) = 1$, then the conditional probability that $Z(\omega, t) - Z(\omega, s) = 0$ is one. On the other hand, if $Z(\omega, s) = Z(\omega, s) - Z(\omega, 0) < 1$, the conditional probability that $Z(\omega, t) - Z(\omega, s) = 0$ is zero.

Example 7.9 You found in Problem Set 2 that

$$\int_0^T \hat{X}_n \, dX_n = \frac{1}{2}\left(X_n^2(\omega, T) - T\right)$$

A slightly more elaborate argument shows that if W is a one-dimensional Wiener process,

$$\int_0^T W \, dW = \frac{1}{2}\left(W^2(\omega, T) - T\right)$$

Remember that the approximations to the integrand W used in defining the integral are always adapted. Because the increments in the Wiener process are normally distributed, whereas the increments in the random walk are always $\pm 1/\sqrt{n}$, the argument needs to rely on the Law of Large Numbers.

Theorem 7.10 (Kunita-Watanabe Theorem, Theorem 1.39 in Nielsen) *Let W be a K-dimensional standard Brownian motion. If Z is a martingale with respect to the filtration generated by W, then there exists $b \in \mathcal{L}^2$ such that*

$$Z(\omega, t) = Z(\omega, 0) + \int_0^t b dW(\omega, s)$$

If $Z(\cdot, T) \in L^2$, then $b \in \mathcal{H}^2$ on $\Omega \times [0, T]$.

Remark 7.11 This is a truly remarkable result.

1. It is hard to give a discrete intuition for it because, in essence, it is a theorem about the filtration generated by a Brownian motion. Note that the theorem implies that Z has a continuous version, so every martingale with respect to the filtration generated by a Brownian motion is continuous. There is something about the filtration that forces the release of new information to be done in a continuous way. You cannot capture sudden events (whether anticipated, such as the press release which follows each meeting of the Federal Reserve Open Market Committee changes the discount rate, or unanticipated, such as a large corporation announcing that it is retracting the last several years of its audited income statements) in a stock price model based on Brownian motion. It should be emphasized that the theorem assumes that Z is a martingale with respect to the filtration *generated* by W; it is not enough that Z be a martingale with respect to the filtration associated with a Wiener process, since that filtration may be larger than the filtration generated by the Wiener process.

2. It is very useful for Finance. Since Z is an Itô process, we can do Itô Calculus on Z. The theorem is essential in proving the Complete Markets Theorem, where it allows us to extract a trading strategy whose value process is a given martingale.

8 Itô Calculus

We want to study stock price processes of the form $e^{Z(\omega,t)}$ where Z is a generalized Brownian motion. In particular, we need to compute

$$\int_0^T \bar{\Delta}(\omega,t)\, de^{Z(\omega,t)}$$

the capital gain generated by a trading strategy $\bar{\Delta}$. Itô's Lemma gives us the key to defining the stochastic integral with respect to processes like $e^{Z(\omega,t)}$.

Fix a K-dimensional standard Wiener process W.

Definition 8.1 An N-dimensional *Itô process* is a stochastic process of the form

$$Z(\omega,t) = Z(\omega,0) + \int_0^t a(\omega,s)\, ds + \int_0^t b(\omega,s) dW(\omega,s) \tag{4}$$

where $a \in \mathcal{L}^1$ is an $N \times 1$ vector-valued process and $b \in \mathcal{L}^2$ is an $N \times K$ matrix-valued process. Note that a and b are allowed to depend on both ω and t. Itô processes are continuous and adapted. Every generalized Wiener process is an Itô process. a is called the *drift*, b the *dispersion*, and bb^T the *instantaneous covariance matrix* of Z. The following symbolic representations are all shorthand for Equation (4):

$$\begin{aligned} Z(t) &= X_0 + \int_0^t a\, ds + \int_0^t b\, dW \\ dZ(t) &= a(t)\, dt + b(t)\, dW(t) \\ dZ &= a\, dt + b\, dW \end{aligned}$$

If $D \subset \mathbf{R}^N$ is and $f : D \to \mathbf{R}$ is C^2, let

$$f_x(x) = \bigtriangledown f|_x = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_N} \right)$$

denote the gradiant of f, viewed as a row vector, and let

$$f_{xx}(x) = Hf|_x = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)$$

denote the Hessian matrix of f.

Theorem 8.2 (Itô's Lemma, Theorem 2.2 in Nielsen) *Let $D \subset \mathbf{R}^N$ be an open set, and Z an N-dimensional Itô process*

$$Z(t) = Z(0) + \int_0^t a\,ds + \int_0^t b\,dW$$

such that

$$P(\{\omega : Z(\omega, t) \in D \textit{ for all } t \in [0, T]\}) = 1$$

and $f : D \to \mathbf{R}$is C^2. Then $f(Z)$ is an Itô process, specifically $f(Z(t)) =$

$$f(Z(0)) + \int_0^t \left[f_x(Z)a + \frac{1}{2}\mathrm{tr}\left(b^T f_{xx}(Z)b\right)\right] ds + \int_0^t f_x(Z)b\,dW \qquad (5)$$

Remark 8.3 By analogy with the Fundamental Theorem of Calculus, the terms involving $f_x(Z)$ are expected, but the term involving $\mathrm{tr}\left(b^T f_{xx}(Z)b\right)$ is at first sight surprising. Note that

$$\mathrm{tr}(b^T f_{xx}(Z)b) = \sum_{i,j=1}^{N} \sum_{k=1}^{K} \frac{\partial^2 f}{\partial x_i \partial x_j}(Z) b_{ik} b_{jk}$$

b_{ik} is the coefficient of X_i on W_k, so $b_{ik}b_{jk}$ is the product of the coefficients of X_i and X_j on the *same* component k of the Wiener process W. Since

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(Z) b_{ik} b_{jk}$$

is integrated with respect to time t, the formula is saying, in effect that $(dW_k)^2 = dt$, but this is just reasserting that the quadratic variation of a Wiener process grows linearly with time. The term arises because the quadratic variation of the Wiener process is not zero, and hence the second order terms in the Taylor expansion of f matter. There are no terms corresponding to $b_{ik}b_{j\ell}$ with $k \neq \ell$, so the formula is saying, in effect, that $(dW_k)(dW_\ell) = 0$ if $k \neq \ell$. Itô's Lemma is often summarized by saying

$$(dW_k)(dW_\ell) = \delta_{k\ell}dt$$

where

$$\delta_{k\ell} = \begin{cases} 1 & \text{if } k = \ell \\ 0 & \text{if } k \neq \ell \end{cases}$$

Example 8.4 [Black-Scholes Stock Price, Example 2.3 in Nielsen] The stock price in the Black-Scholes model is

$$S(t) = S(0)e^{(\mu-\sigma^2/2)t+\sigma W(t)}$$

Let

$$\begin{aligned} Z(t) &= \ln S(t) \\ &= \ln S(0) + \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W(t) \\ &= \ln S(0) + \int_0^t \left(\mu - \frac{\sigma^2}{2}\right) ds + \int_0^t \sigma\, dW(s) \end{aligned}$$

so Z is an Itô process, and

$$\begin{aligned} dZ &= \left(\mu - \frac{\sigma^2}{2}\right) ds + \sigma\, dW \\ S(t) &= e^{Z(t)} \\ dS &= de^Z \\ &= \left[e^Z \left(\mu - \frac{\sigma^2}{2}\right) + \frac{\sigma^2}{2} e^Z\right] dt + e^Z \sigma\, dW \\ &= e^Z \mu\, dt \ + \ e^Z \sigma\, dW \\ &= S\mu\, dt \ + \ S\sigma\, dW \end{aligned}$$

so

$$\frac{dS}{S} = \mu\, dt \ + \sigma\, dW$$

$\frac{dS}{S}$ is the proportional change in S, so the proportional change in S has drift μ and instantaneous variance σ. This provides another explanation of why we write $S = S(0)e^{(\mu-\sigma^2/2)t+\sigma W(t)}$; the $-\sigma^2/2$ is needed to cancel out a $\sigma^2/2$ that comes from Itô's Lemma, resulting in instantaneous drift μ in the proportional change of S.

Plausibility Argument for Itô's Lemma: Here, we give a conceptually simple, but admittedly notationally messy, calculation verifying Itô's Lemma for Stieltjes integrals of simple integrands integrated with respect to a random walk. The intuition behind the standard proof of Itô's Lemma is very close to this argument, but complications arise because the Itô Integral is defined for general integrands by approximation, and because the relation $(\Delta W_j)(\Delta W_k) = \delta_{jk}dt$ is not true over finite time intervals. However, it is easy to see this relation holds for the random walk, The argument given here can be turned into a rigorous proof of Itô's Lemma using nonstandard analysis (Anderson [1]). Let

$$Y(t) = Y(0) + \int_0^t a\,ds + \int_0^T b\,dX$$

where X is a 2-dimensional n-step random walk. In other words, $\Omega = \{-1,1\}^{nT} \times \{-1,1\}^{nT}$, $\omega = (\omega_{\ell k} : \ell = 1,2,\ k = 1,\ldots,nT)$, $N = 2$ and $K = 2$. Suppose also that a and b are simple processes which are measurable in the filtration generated by the random walk; for simplicity, we assume here that a and b are uniformly bounded as $n \to \infty$. $O(h)$ denotes a quantity which is a bounded multiple of h as $h \to 0$, while $o(h)$ denotes a quantity which goes to zero faster than h as $h \to h$. For example, Taylor's Theorem says that if $g : \mathbf{R} \to \mathbf{R}$ is a C^2 function,

$$g(x+h) = g(x) + g'(x)h + \frac{1}{2}g''(x)h^2 + o(h^2)$$

Let

$$\begin{aligned}\Delta Y\left(\omega, \frac{k}{n}\right) &= Y\left(\omega, \frac{k+1}{n}\right) - \left(\omega, \frac{k}{n}\right)\\ &= \begin{pmatrix} \frac{a_1(\omega,k/n)}{n} + \frac{b_{11}(\omega,k/n)\omega_{1(k+1)}+b_{12}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}} \\ \frac{a_2(\omega,k/n)}{n} + \frac{b_{21}(\omega,k/n)\omega_{1(k+1)}+b_{22}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}} \end{pmatrix}\end{aligned}$$

Thus,

$f(Y(\omega,T)) - f(Y(\omega,0))$

$$
\begin{aligned}
&= \sum_{k=0}^{nT-1}\left(f\left(Y\left(\omega,\frac{k+1}{n}\right)\right)-f\left(Y\left(\omega,\frac{k}{n}\right)\right)\right)\\
&= \sum_{k=0}^{nT-1}\bigtriangledown f|_{Y(\omega,k/n)}\cdot\Delta Y\left(\omega,\frac{k}{n}\right)+\frac{1}{2}\sum_{k=0}^{nT-1}\left(\Delta Y\left(\omega,\frac{k}{n}\right)\right)^T Hf|_{Y(\omega,k/n)}\Delta Y\left(\omega,\frac{k}{n}\right)\\
&\quad +no\left(\frac{1}{n}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{k=0}^{nT-1}\bigtriangledown f|_{Y(\omega,k/n)}\Delta Y\left(\omega,\frac{k}{n}\right)\\
&= \sum_{k=0}^{nT-1}\left.\frac{\partial f}{\partial x_1}\right|_{Y(\omega,k/n)}\left(\frac{a_1(\omega,k/n)}{n}+\frac{b_{11}(\omega,k/n)\omega_{1(k+1)}+b_{12}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)\\
&\quad +\sum_{k=0}^{nT-1}\left.\frac{\partial f}{\partial x_2}\right|_{Y(\omega,k/n)}\left(\frac{a_2(\omega,k/n)}{n}+\frac{b_{21}(\omega,k/n)\omega_{1(k+1)}+b_{22}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)\\
&= \int_0^T\left.\frac{\partial f}{\partial x_1}\right|_{Y(\omega,t)}a_1(\omega,t)\,dt+\int_0^T\left.\frac{\partial f}{\partial x_1}\right|_{Y(\omega,t)}b_{11}(\omega,t)\,dX_1(\omega,t)\\
&\quad +\int_0^T\left.\frac{\partial f}{\partial x_1}\right|_{Y(\omega,t)}b_{12}(\omega,t)\,dX_2(\omega,t)+\int_0^T\left.\frac{\partial f}{\partial x_2}\right|_{Y(\omega,t)}a_2(\omega,t)\,dt\\
&\quad +\int_0^T\left.\frac{\partial f}{\partial x_2}\right|_{Y(\omega,t)}b_{21}(\omega,t)\,dX_1(\omega,t)+\int_0^T\left.\frac{\partial f}{\partial x_2}\right|_{Y(\omega,t)}b_{22}(\omega,t)\,dX_2(\omega,t)\\
&= \int_0^T\bigtriangledown f|_{Y(\omega,t)}\,a(\omega,t)\,dt\\
&\quad +\int_0^T\left(\left.\frac{\partial f}{\partial x_1}\right|_{Y(\omega,t)}b_{11}(\omega,t)+\left.\frac{\partial f}{\partial x_2}\right|_{Y(\omega,t)}b_{21}(\omega,t)\right)\,dX_1(\omega,t)\\
&\quad +\int_0^T\left(\left.\frac{\partial f}{\partial x_1}\right|_{Y(\omega,t)}b_{12}(\omega,t)+\left.\frac{\partial f}{\partial x_2}\right|_{Y(\omega,t)}b_{22}(\omega,t)\right)\,dX_2(\omega,t)\\
&= \int_0^T\bigtriangledown f|_{Y(\omega,t)}\,a(\omega,t)\,dt+\int_0^T\bigtriangledown f|_{Y(\omega,t)}\,b(\omega,t)\,dX(\omega,t)
\end{aligned}
$$

Since ω_{1k} and ω_{2k} are independent, the product $\omega_{1k}\omega_{2k}$ equals $+1$ with probability 1/2 and -1 with probability 1/2, so we can form a random walk

$$\bar{X}\left(\omega, \frac{k}{n}\right) = \sum_{j=1}^{k} \frac{\omega_{1j}\omega_2 j}{\sqrt{n}}$$

$\bar{X}$ is a standard random walk, which in the limit is standard Brownian motion.

$$\begin{aligned}
&\frac{1}{2}\sum_{k=0}^{nT-1}\left(\Delta Y\left(\omega,\frac{k}{n}\right)\right)^T Hf|_{Y(\omega,k/n)}\Delta Y\left(\omega,\frac{k}{n}\right) \\
&= \frac{1}{2}\sum_{k=0}^{nT-1}\left(\Delta Y_1\left(\omega,\frac{k}{n}\right),\Delta Y_2\left(\omega,\frac{k}{n}\right)\right)\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}\Big|_{Y(\omega,t)} & \frac{\partial^2 f}{\partial x_1 \partial x_2}\Big|_{Y(\omega,t)} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}\Big|_{Y(\omega,t)} & \frac{\partial^2 f}{\partial x_2^2}\Big|_{Y(\omega,t)} \end{pmatrix}\begin{pmatrix} \Delta Y_1\left(\omega,\frac{k}{n}\right) \\ \Delta Y_2\left(\omega,\frac{k}{n}\right) \end{pmatrix} \\
&= \frac{1}{2}\sum_{k=0}^{nT-1}\frac{\partial^2 f}{\partial x_1^2}\Big|_{Y(\omega,t)}\left(\Delta Y_1\left(\omega,\frac{k}{n}\right)\right)^2 + \frac{\partial^2 f}{\partial x_1^2}\Big|_{Y(\omega,t)}\left(\Delta Y_2\left(\omega,\frac{k}{n}\right)\right)^2 \\
&\quad +2\,\frac{\partial^2 f}{\partial x_1 \partial x_2}\Big|_{Y(\omega,t)}\left(\Delta Y_1\left(\omega,\frac{k}{n}\right)\right)\left(\Delta Y_2\left(\omega,\frac{k}{n}\right)\right)\Bigg) \\
&= \frac{1}{2}\sum_{k=0}^{nT-1}\frac{\partial^2 f}{\partial x_1^2}\Big|_{Y(\omega,t)}\left(\frac{a_1(\omega,k/n)}{n} + \frac{b_{11}(\omega,k/n)\omega_{1(k+1)} + b_{12}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)^2 \\
&\quad + \frac{\partial^2 f}{\partial x_1^2}\Big|_{Y(\omega,t)}\left(\frac{a_2(\omega,k/n)}{n} + \frac{b_{21}(\omega,k/n)\omega_{1(k+1)} + b_{22}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)^2 \\
&\quad +2\,\frac{\partial^2 f}{\partial x_1 \partial x_2}\Big|_{Y(\omega,t)}\left(\frac{a_1(\omega,k/n)}{n} + \frac{b_{11}(\omega,k/n)\omega_{1(k+1)} + b_{12}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right) \\
&\qquad \times\left(\frac{a_2(\omega,k/n)}{n} + \frac{b_{21}(\omega,k/n)\omega_{1(k+1)} + b_{22}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)\Bigg)
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{2}\sum_{k=0}^{nT-1}\frac{\partial^2 f}{\partial x_1^2}\Big|_{Y(\omega,t)}\left(\frac{a_1(\omega,k/n)}{n} + \frac{b_{11}(\omega,k/n)\omega_{1(k+1)} + b_{12}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)^2 \\
&= \frac{1}{2}\sum_{k=0}^{nT-1}\frac{\partial^2 f}{\partial x_1^2}\Big|_{Y(\omega,t)}\Bigg(\frac{O(1)}{n^2} + \frac{O(1)}{n^{3/2}} \\
&\quad +\frac{(b_{11}(\omega,k/n)\omega_{1(k+1)})^2 + (b_{12}(\omega,k/n)\omega_{2(k+1)})^2 + 2b_{11}(\omega,k/n)b_{12}(\omega,k/n)\omega_{1(k+1)}\omega_{2(k+1)}}{n}\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&= O\left(\frac{1}{n^{1/2}}\right) + \frac{1}{\sqrt{n}} \sum_{k=0}^{nT-1} \left.\frac{\partial^2 f}{\partial x_1^2}\right|_{Y(\omega,t)} \frac{b_{11}(\omega,k/n)b_{12}(\omega,k/n)\omega_{1(k+1)}\omega_{2(k+1)}}{\sqrt{n}} \\
&\quad + \frac{1}{2} \sum_{k=0}^{nT-1} \left.\frac{\partial^2 f}{\partial x_1^2}\right|_{Y(\omega,t)} \frac{b_{11}(\omega,k/n)^2 + b_{12}(\omega,k/n)^2}{n} \\
&= O\left(\frac{1}{n^{1/2}}\right) + \frac{1}{\sqrt{n}} \int_0^T \left.\frac{\partial^2 f}{\partial x_1^2}\right|_{Y(\omega,t)} b_{11}(\omega,k/n)b_{12}(\omega,k/n)d\bar{X} \\
&\quad + \frac{1}{2} \int_0^T \left.\frac{\partial^2 f}{\partial x_1^2}\right|_{Y(\omega,t)} \left(b_{11}(\omega,t)^2 + b_{12}(\omega,t)^2\right) dt \\
&= O\left(\frac{1}{\sqrt{n}}\right) + \frac{1}{2} \int_0^T \left.\frac{\partial^2 f}{\partial x_1^2}\right|_{Y(\omega,t)} \left(b_{11}(\omega,t)^2 + b_{12}(\omega,t)^2\right) dt
\end{aligned}
$$

because the stochastic integral with respect to $\bar{X}$ is finite almost surely. Similarly,

$$
\begin{aligned}
&\frac{1}{2} \sum_{k=0}^{nT-1} \left.\frac{\partial^2 f}{\partial x_1^2}\right|_{Y(\omega,t)} \left(\frac{a_2(\omega,k/n)}{n} + \frac{b_{21}(\omega,k/n)\omega_{1(k+1)} + b_{22}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)^2 \\
&\quad = O\left(\frac{1}{\sqrt{n}}\right) + \frac{1}{2} \int_0^T \left.\frac{\partial^2 f}{\partial x_2^2}\right|_{Y(\omega,t)} \left(b_{21}(\omega,t)^2 + b_{22}(\omega,t)^2\right) dt
\end{aligned}
$$

Finally,

$$
\begin{aligned}
&\frac{1}{2} \sum_{k=0}^{nT-1} 2 \left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{Y(\omega,t)} \left(\left(\frac{a_1(\omega,k/n)}{n} + \frac{b_{11}(\omega,k/n)\omega_{1(k+1)} + b_{12}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)\right. \\
&\qquad \left. \times \left(\frac{a_2(\omega,k/n)}{n} + \frac{b_{21}(\omega,k/n)\omega_{1(k+1)} + b_{22}(\omega,k/n)\omega_{2(k+1)}}{\sqrt{n}}\right)\right) \\
&= \sum_{k=0}^{nT-1} \left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{Y(\omega,t)} \left(O\left(\frac{1}{n^{3/2}}\right) + \frac{b_{11}(\omega,k/n)b_{21}(\omega,k/n)(\omega_{1(k+1)})^2}{n}\right. \\
&\quad + \frac{b_{12}(\omega,k/n)b_{22}(\omega,k/n)(\omega_{2(k+1)})^2}{n} \\
&\quad \left. + \frac{(b_{11}(\omega,k/n)b_{22}(\omega,k/n) + b_{12}(\omega,k/n)b_{21}(\omega,k/n))\omega_{1(k+1)}\omega_{2(k+1)}}{n}\right) \\
&= O\left(\frac{1}{\sqrt{n}}\right) + \sum_{k=0}^{nT-1} \left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{Y(\omega,t)} \frac{b_{11}(\omega,k/n)b_{21}(\omega,k/n) + b_{12}(\omega,k/n)b_{22}(\omega,k/n)}{n}
\end{aligned}
$$

$$
\begin{aligned}
&+ \quad \frac{1}{\sqrt{n}} \sum_{k=0}^{nT-1} \left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{Y(\omega,t)} \frac{(b_{11}(\omega,k/n)b_{22}(\omega,k/n)+b_{12}(\omega,k/n)b_{21}(\omega,k/n))\omega_{1(k+1)}\omega_{2(k+1)}}{\sqrt{n}} \\
&= \quad O\left(\frac{1}{\sqrt{n}}\right) + \int_0^T \left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{Y(\omega,t)} (b_{11}(\omega,t)b_{21}(\omega,t)+b_{12}(\omega,t)b_{22}(\omega,t))\,dt \\
&+ \quad \frac{1}{\sqrt{n}} \int_0^T \left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{Y(\omega,t)} (b_{11}(\omega,t)b_{22}(\omega,t)+b_{12}(\omega,t)b_{21}(\omega,t))\,d\bar{X} \\
&= \quad O\left(\frac{1}{\sqrt{n}}\right) + \int_0^T \left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{Y(\omega,t)} (b_{11}(\omega,t)b_{21}(\omega,t)+b_{12}(\omega,t)b_{22}(\omega,t))\,dt
\end{aligned}
$$

Combining the above calculations, and taking the limit as $n \to \infty$, we have

$$
\begin{aligned}
f(Y(\omega,T)) \quad &= \quad f(Y(\omega,0)) \\
&\quad + \int_0^T \bigtriangledown f|_{Y(\omega,t)}\, a(\omega,t)\,dt + \int_0^T \bigtriangledown f|_{Y(\omega,t)}\, b(\omega,t)\,dX(\omega,t) \\
&\quad + \frac{1}{2}\int_0^T \left.\frac{\partial^2 f}{\partial x_1^2}\right|_{Y(\omega,t)} b_{11}(\omega,t)^2 + b_{12}(\omega,t)^2\,dt \\
&\quad + \frac{1}{2}\int_0^T \left.\frac{\partial^2 f}{\partial x_2^2}\right|_{Y(\omega,t)} b_{21}(\omega,t)^2 + b_{22}(\omega,t)^2\,dt \\
&\quad + \frac{1}{2}\int_0^T 2\left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{Y(\omega,t)} (b_{11}(\omega,t)b_{21}(\omega,t)+b_{12}(\omega,t)b_{22}(\omega,t))\,dt \\
&= \quad \int_0^T \Big(\bigtriangledown f|_{Y(\omega,t)}\, a(\omega,t) + \frac{1}{2}\mathrm{tr}(b(\omega,t)^T(Hf|_{Y(\omega,t)})b(\omega,t))\Big)\,dt \\
&\quad + \int_0^T \bigtriangledown f|_{Y(\omega,t)}\, b(\omega,t)\,dW(\omega,t)
\end{aligned}
$$

Proposition 8.5 (Uniqueness of Coefficients of Itô Processes) [12] *Let $a, \alpha \in \mathcal{L}^1$ be N-dimensional, $b, \beta \in \mathcal{L}^2$ be $N \times K$-dimensional, and $X_0, Y_0 \in L^2$ be N-dimensional.*[13] *If*

$$
X_0 + \int_0^t a\,ds + \int_0^t b\,dW = Y_0 + \int_0^t \alpha\,ds + \int_0^t \beta\,dW
$$

[12] This is a slight generalization of Proposition 2.6 in Nielsen.

[13] Nielsen assumes X_0, Y_0 are constants, but there is no reason for them not to be random variables.

for all t, almost surely in ω, then

$$\begin{aligned}
X_0(\omega) &= Y_0(\omega) \;\; P \textit{ almost surely} \\
a(\omega,t) &= \alpha(\omega,t) \;\; \lambda\otimes P \textit{ almost everywhere} \\
b(\omega,t) &= \beta(\omega,t) \;\; \lambda\otimes P \textit{ almost everywhere}
\end{aligned}$$

Proof: If $b,\beta \in \mathcal{H}^2$, this follows immediately from the Itô Isometry. Since we need the result when $b,\beta \in \mathcal{L}^2$, we use Itô's Lemma. It is sufficient to consider the case $N = 1$. Let

$$\begin{aligned}
Z(\omega,t) &= X_0 + \int_0^t a\,ds + \int_0^t b\,dW - \left(Y_0 + \int_0^t \alpha\,ds + \int_0^t \beta\,dW\right) \\
&= X_0 - Y_0 + \int_0^t (a-\alpha)\,ds + \int_0^t (b-\beta)\,dW \\
&= X_0 - Y_0 + \int_0^t \gamma\,ds + \int_0^t \delta\,dW
\end{aligned}$$

where $\gamma = a - \alpha$ and $\delta = b - \beta$. $Z(\omega,\cdot) = 0$ almost surely, so $X_0(\omega) - Y_0(\omega) = Z(\omega,0) = 0$ almost surely. We show $\gamma = 0$ and $\delta = 0$ $\lambda\otimes P$-almost everywhere.

$$\begin{aligned}
0 &= e^{Z(\omega,t)} - 1 \\
&= e^{Z(\omega,0)} + \int_0^t \left[e^0\gamma + \frac{1}{2}e^0\delta^T\delta\right] ds + \int_0^t e^0\delta\,dW - 1 \\
&= \int_0^t \left[\gamma + \frac{1}{2}\delta^T\delta\right] ds + \int_0^t \delta\,dW \\
&= \int_0^t \gamma\,ds + \int_0^t \delta\,dW + \frac{1}{2}\int_0^t \delta^T\delta\,ds \\
&= Z(\omega,t) - Z(\omega,0) + \frac{1}{2}\int_0^t \delta^T\delta\,ds \\
&= \frac{1}{2}\int_0^t \delta^T\delta\,ds
\end{aligned}$$

which implies that $\delta = 0$ ($P\otimes\lambda$-almost everywhere), so $\int_0^t \delta\,dW = 0$ for all t, so $\int_0^t \gamma\,ds = 0$ for all t, so $\gamma = 0$ ($P\otimes\lambda$-almost everywhere). ■

Corollary 8.6 (Proposition 2.7) *If the Itô process*

$$X(t) = X(0) + \int_0^t a\,ds + \int_0^t b\,dW$$

is a martingale with respect to the filtration generated by W, then $a = 0$ $P \otimes \lambda$-almost everywhere.

Remark 8.7 [Caution] The converse is true if $b \in \mathcal{H}^2$, but it is not generally true if $b \in \mathcal{L}^2$.

9 Integrals with respect to Itô Processes

Our basic model for a stock price will be geometric Brownian motion, which is an Itô Process but not a Wiener process. In order to compute the capital gain generated by a portfolio strategy, we need to be able to take Itô integrals with respect to Itô processes. Nielsen *defines* the Itô integral with respect to an Itô process Z as a particular Itô integral with respect to the Wiener process W underlying Z. Here, we show why that is the correct definition. Let

$$Z(t) = Z_0 + \int_0^t a\, ds + \int_0^t b\, dW$$

where

$$\begin{array}{rl} Z_0 \text{ is } \mathcal{F}_0\text{-measurable} & Z_0(\omega) \in \mathbf{R}^N \\ a \in \mathcal{L}^1 & a(\omega, t) \text{ is } N \times 1 \\ b \in \mathcal{L}^2 & b(\omega, t) \text{ is } N \times K \end{array}$$

Suppose that we replace the K-dimensional standard Wiener process W with the random walk X_n, and assume that a, b and γ are simple and adapted with respect to the random walk filtration. To simplify notation, we take $K = 1$. Then if $Z = Z_0 + \int a\, ds + \int b\, dX_n$,

$$Z\left(\omega, \frac{j}{n}\right) = \sum_{k=0}^{j-1} \frac{a\left(\omega, \frac{k}{n}\right)}{n} + \sum_{k=0}^{j-1} \frac{b\left(\omega, \frac{k}{n}\right)\omega_{k+1}}{\sqrt{n}}$$

so

$$\begin{aligned} \Delta Z\left(\omega, \frac{k}{n}\right) &= Z\left(\omega, \frac{k+1}{n}\right) - Z\left(\omega, \frac{k}{n}\right) \\ &= \frac{a\left(\omega, \frac{k}{n}\right)}{n} + \frac{b\left(\omega, \frac{k}{n}\right)\omega_{k+1}}{\sqrt{n}} \end{aligned}$$

$$
\begin{aligned}
\int_0^T \gamma dZ &= \sum_{k=0}^{nT} \gamma\left(\omega, \frac{k}{n}\right) \Delta Z\left(\omega, \frac{k}{n}\right) \\
&= \sum_{k=0}^{nT} \gamma\left(\omega, \frac{k}{n}\right) \left(\frac{a\left(\omega, \frac{k}{n}\right)}{n} + \frac{b\left(\omega, \frac{k}{n}\right)\omega_{k+1}}{\sqrt{n}}\right) \\
&= \sum_{k=0}^{nT} \frac{\gamma\left(\omega, \frac{k}{n}\right) a\left(\omega, \frac{k}{n}\right)}{n} + \sum_{k=0}^{nT} \frac{\gamma\left(\omega, \frac{k}{n}\right) b\left(\omega, \frac{k}{n}\right)\omega_{k+1}}{\sqrt{n}} \\
&= \int_0^T \gamma(\omega, s) a(\omega, s)\, ds + \int_0^T \gamma(\omega, s) b(\omega, s)\, dX_n \\
&= \int_0^T \gamma a\, ds + \int_0^T \gamma b\, dX_n
\end{aligned}
$$

Now, return to the situation in which W is a K-dimensional standard Wiener process. We see that we want

$$\int_0^t \gamma\, dZ = \int_0^t \gamma a\, ds + \int_0^t \gamma b\, dW$$

In order for this to make sense, we need to know that γa is integrable with respect to time and γb is Itô integrable with respect to W. This motivates the following definition:

Definition 9.1 Suppose

$$Z(t) = Z(0) + \int_0^t a\, ds + \int_0^t b\, dW$$

where W is a standard K-dimensional Wiener process, $a \in \mathcal{L}^1$ is N-dimensional and $b \in \mathcal{L}^2$ is $N \times K$-dimensional. Let

$$\mathcal{L}(Z) = \{\gamma : \gamma \text{ is adapted, measurable, } M \times N, \gamma a \in \mathcal{L}^1, \gamma b \in \mathcal{L}^2\}$$

If $\gamma \in \mathcal{L}(Z)$, define

$$\int_0^t \gamma\, dZ = \int_0^t \gamma a\, ds + \int_0^t \gamma b\, dW$$

Remark 9.2 Nielsen states various facts about stochastic integrals with respect to Itô processes.

1. $\int_0^t \gamma\, dZ$ is an Itô Process, hence it is adapted and continuous.

2. γ may "accidentally" be in $\mathcal{L}(Z)$ even if it is not in $\mathcal{L}(Z_i)$ for some i.

3. $\int \gamma\, dZ$ is linear in the integrator Z as well as in the integrand γ.

4. If Y is an $\mathcal{F}_t$-measurable random variable,

$$\int_t^T Y\, dZ = Y \int_t^T dZ = Y(X(T) - X(t))$$

5. If f is C^2on the range of Z and $N = 1$,

$$\mathcal{L}(Z) \subset \mathcal{L}(f(Z))$$

6. Nielsen's section 2.3 contains various versions of Itô's Lemma that you should read on your own.

10 Detrending and Changing the Variance of Itô Processes

This material corresponds to sections 2.4 and 2.5 of Nielsen, but we will approach it in a different way. There are several methods for detrending or altering the variance of an Itô Process.

1. **Doob-Meyer Decomposition:** This is the most natural way to detrend an Itô Process, but it turns out to be less useful in Finance than some of the methods that follow. The Doob-Meyer Decomposition takes a class of processes and represents them as the sum of a process of bounded variation and a martingale. To take a simple example, suppose Z is an Itô Process:

$$Z(\omega, t) = Z(\omega, 0) + \int_0^t a(\omega, s)\, ds + \int_0^t b(\omega, s)\, dW$$

with $a \in \mathcal{L}^1$ and $b \in \mathcal{H}^2$, then $\int_0^t b\, dW$ is a martingale. To see that $\int_0^t a\, ds$ is almost surely of bounded variation, we decompose $a = a_+ -$

a_-, where $a_+(\omega, t) = \max\{a(\omega, t), 0\}$ and $a_-(\omega, t) = -\min\{a(\omega, t), 0\}$. Then

$$\int_0^t a\, ds = \int_0^t a_+\, ds - \int_0^t a_-\, ds$$

and $\int_0^t a_+\, ds$ and $\int_0^t a_-\, ds$ are both nondecreasing, hence $\int_0^t a\, ds$ is of bounded variation.

Proposition 10.1 *Every adapted simple process on the random walk filtration has a Doob-Meyer decomposition.*

Proof: Let $Z(\omega, k/n)$ be an arbitrary adapted simple process on the random walk filtration. Given a node $(\omega, k/n)$, define the nodes $(\omega_+, (k+1)/n)$ and $(\omega_-, (k+1)/n)$ by

$$\begin{array}{ll} (\omega_+)_{k+1} = +1 & (\omega_+)_j = \omega_j \text{ for } j \leq k \\ (\omega_-)_{k+1} = -1 & (\omega_-)_j = \omega_j \text{ for } j \leq k \end{array}$$

In other words, $(\omega_+, (k+1)/n)$ and $(\omega_-, (k+1)/n)$ are the two nodes in the random walk tree that immediately follow the node $(\omega, k/n)$. Define

$$\begin{aligned} a\left(\omega, \frac{k}{n}\right) &= n\left(\frac{Z\left(\omega_+, \frac{k+1}{n}\right) + Z\left(\omega_-, \frac{k+1}{n}\right)}{2} - Z\left(\omega, \frac{k}{n}\right)\right) \\ b\left(\omega, \frac{k}{n}\right) &= \sqrt{n}\left(\frac{Z\left(\omega_+, \frac{k+1}{n}\right) - Z\left(\omega_-, \frac{k+1}{n}\right)}{2}\right) \\ &= \sqrt{n}\left(Z\left(\omega_+, \frac{k+1}{n}\right) - \left(Z\left(\omega, \frac{k}{n}\right) + \frac{a\left(\omega, \frac{k}{n}\right)}{n}\right)\right) \\ &= -\sqrt{n}\left(Z\left(\omega_-, \frac{k+1}{n}\right) - \left(Z\left(\omega, \frac{k}{n}\right) + \frac{a\left(\omega, \frac{k}{n}\right)}{n}\right)\right) \end{aligned}$$

Then

$$Z\left(\omega, \frac{k}{n}\right) = Z(\omega, 0) + \sum_{j=0}^{k-1} \frac{a\left(\omega, \frac{j}{n}\right)}{n} + \sum_{j=0}^{k-1} \frac{b\left(\omega, \frac{j}{n}\right) \omega_{j+1}}{\sqrt{n}}$$

so in the random walk filtration, an arbitrary process can be represented as the sum of a time integral and a martingale. ■

Definition 10.2 Let $\{\mathcal{F}_t\}$ be a filtration. Z is a *submartingale* with respect to $\{\mathcal{F}_t\}$ if, whenever $s \leq t$,

$$X(s) \leq E(X(t)|\mathcal{F}_s)$$

Theorem 10.3 (Doob-Meyer Decomposition) *If $Z(\omega, t)$ is a right-continuous submartingale with respect to a filtration $\{\mathcal{F}_t\}$, then*

$$Z(\omega, t) = A(\omega, t) + M(\omega, t)$$

where A is an adapted nondecreasing process and M is a right-continuous martingale with respect to $\{\mathcal{F}_t\}$.

Remark 10.4 Since A is nondecreasing, it is of bounded variation on finite time intervals. A need not be predictable; predictability in continuous-time is a stronger condition than predictability in the random walk model. In the random walk model, predictability means the value of a process at time k/n is measurable with respect to $\mathcal{F}_{(k-1)/n}$; as $n \to \infty$, the time interval from k/n to $(k+1)/n$ shrinks, so in the limit, the value of a at time t cannot be predicted from information available at times $s < t$.

References

[1] Anderson, Robert M., "A Nonstandard Representation of Brownian Motion and Itô Integration," *Israel Journal of Mathematics* **25**(1976), 15-46.

[2] Billingsley, Patrick, *Convergence of Probability Measures*, John Wiley and Sons, 1968.

[3] Nielsen, Lars Tyge, *Pricing and Hedging of Derivative Securities*,Oxford University Press, 1999.