Financial Economics

Topics in Asset Pricing Theory

Jaroslav Borovička[†]

September 11, 2017

[†]E-mail: jaroslav.borovicka@nyu.edu. All errors are mine.

Contents

Chapter 1

Introduction to continuous-time calculus

Literature: Øksendal (2007), Chapters 1–6 contain most of the mathematical foundations. Further topics (Feynman–Kac formula, Kolmogorov equations, Girsanov theorem, etc.) are in Chapter 8. The Appendices in Duffie (2001) provide a somewhat informal treatment of a subset of the topics.

Karlin and Taylor (1981), Chapter 15 contains an excellent treatment of boundary classification of diffusions.

Stokey (2008) provides a new treatment of boundary control problems in macroeconomics, without detailed technicalities of stochastic calculus. It builds on the previous well-known book by Dixit and Pindyck (1994).

1.1 Mathematical preliminaries

Definition 1.1 *Let Ω be a given set. Then a* ***σ-algebra*** *$\mathcal{F}$ on Ω is a family $\mathcal{F}$ of subsets of Ω that satisfies:*

1. $\emptyset \in \mathcal{F}$
2. $F \in \mathcal{F} \implies F^C \in \mathcal{F}$ *(closure to complements)*
3. $A_1, A_2, \ldots \in \mathcal{F} \implies A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ *(closure to countable unions)*

The pair $(\Omega, \mathcal{F})$ is called a ***measurable space****.*

Definition 1.2 *A **probability measure** P on $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0,1]$ such that*

1. $P(\emptyset) = 0,\ P(\Omega) = 1$

2. *If $A_1, A_2, \ldots \in \mathcal{F}$ and $\{A_i\}_{i=1}^{\infty}$ is disjoint ($A_i \cap A_j = \emptyset$ for $i \neq j$) then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.*

*The triple $(\Omega, \mathcal{F}, P)$ is called a **probability space**. It is called a **complete** probability space if $\mathcal{F}$ contains all subsets G of Ω with P-outer measure zero, i.e. with*

$$P^*(G) := \inf\{P(F) : F \in \mathcal{F}, G \subset F\} = 0.$$

Elements of Ω are called **outcomes**, elements of the σ-algebra are called **events**.

Subsets F of Ω which belong to $\mathcal{F}$ will be called **$\mathcal{F}$-measurable**.

Definition 1.3 *Let $\mathcal{U}$ be a family of subsets of Ω. We say that $\mathcal{U}$* ***generates*** *the σ-algebra $\mathcal{F}_{\mathcal{U}}$ if $\mathcal{F}_{\mathcal{U}}$ is the smallest σ-algebra containing $\mathcal{U}$., i.e.*

$$\mathcal{F}_{\mathcal{U}} = \bigcap \{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-algebra on } \Omega \text{ and } \mathcal{U} \subset \mathcal{F}\}.$$

Example 1.4 *Consider $\Omega = \mathbb{R}^n$ (more generally, let Ω be a topological space). Define $\mathcal{U}$ to be the collection of all open subsets of Ω. Then $\mathcal{B} = \mathcal{F}_{\mathcal{U}}$ is the* ***Borel σ-algebra*** *. on Ω, and the elements $B \in \mathcal{B}$ are* ***Borel sets****.*

Remark 1.5 *A typical application in our enviroment will be a σ-algebra generated by observed* ***paths of a Brownian motion****.*

Definition 1.6 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Then the function $Y : \Omega \to \mathbb{R}^n$ is called **$\mathcal{F}$-measurable** if for all open sets (or, equivalently, Borel sets) $U \subset \mathbb{R}^n$:*

$$Y^{-1}(U) := \{\omega \in \Omega : Y(\omega) \in U\} \in \mathcal{F}.$$

If $X : \Omega \to \mathbb{R}^n$ is any function, then the σ-algebra $\mathcal{F}_X$ generated by X is the smallest σ-algebra on Ω containing all the sets

$$X^{-1}(U) : U \subset \mathbb{R}^n \text{ is open.}$$

Random variables, expectations, change of measure

Definition 1.7 *A* ***random variable*** X *is an* $\mathcal{F}$*-measurable function* $X : \Omega \to \mathbb{R}^n$.

Notice that every random variable induces a probability measure μ_X on $\mathbb{R}^n$, defined as

$$\mu_X(B) = P\left(X^{-1}(B)\right).$$

The measure μ_X is called the **distribution** of X (under P).

Definition 1.8 *The number*

$$E[X] := \int_{\Omega} X(\omega)\, dP(\omega) = \int_{\mathbb{R}^n} x d\mu_X(x),$$

if it exists, is called the ***expectation*** *of* X.

Definition 1.9 *Let* $X : \Omega \to \mathbb{R}^n$ *be a random variable and* $p \in [1, \infty)$. *The* L^p***-norm*** *of* X *is defined as*

$$\|X\|_p = \|X\|_{L^p(P)} = \left(\int_\Omega |X(\omega)|^p \, dP(\omega)\right)^{\frac{1}{p}}$$

with the ***sup norm*** *for* $p = \infty$*:*

$$\|X\|_\infty = \|X\|_{L^\infty(P)} = \inf \{N \in \mathbb{R} : |X(\omega)| \leq N \ a.s.\}$$

An L^p***-space*** *is the space of all random variables with a finite* L^p*-norm:*

$$L^p(P) = \left\{X : \Omega \to R^n ; \|X\|_p < \infty\right\}.$$

The expectation above is computed under measure P. Frequently, we will be going from one measure to another. In order to do so, we will be exploiting the Radon–Nikodým theorem.

Definition 1.10 *Two measures P and Q on $(\Omega, \mathcal{F})$ are said to be **equivalent** if $\forall F \in \mathcal{F}$, $Q(F) = 0 \iff P(F) = 0$.*

*Q is said to be **absolutely continuous with respect to** P if $\forall F \in \mathcal{F}$, $P(F) = 0 \implies Q(F) = 0$.*

Theorem 1.11 (Radon–Nikodým) *Let P and Q be two measures on $(\Omega, \mathcal{F})$ such that Q is absolutely continuous with respect to P. Then there exists a unique $\mathcal{F}$-measurable function $Y : \Omega \to \mathbb{R}_+$ such that*

$$Q(F) = \int_F Y \, dP, \quad \forall F \in \mathcal{F}.$$

Proof. Omitted. ■

We can symbolically denote $Y = \frac{dQ}{dP}$, which we will call the **Radon–Nikodým derivative**. The reason is that, symbolically

$$Q(F) = \int_F \frac{dQ}{dP} dP = \int_F dQ.$$

Example 1.12 *Assume that* $(\Omega, \mathcal{F}) = (\mathbb{R}^n, \mathcal{B})$, *i.e., we have a measurable space with Borel* σ*-algebra. Assume that the two measures* P *and* Q *are absolutely continuous with strictly positive densities on* $\mathbb{R}^n$*. Then we have*

$$Q(F) = \int_F q(x)\, dx = \int_F \frac{q(x)}{p(x)} p(x)\, dx$$

so that the Radon–Nikodým derivative is given by $Y(x) = q(x)/p(x)$.

The Radon–Nikodým theorem will also help us with the construction of conditional expectations.

$(\Omega, \mathcal{F}, P)$

Definition 1.13 *Let $\mathcal{H} \subset \mathcal{F}$ be a σ-algebra. The* ***conditional expectation*** *of X conditional on $\mathcal{H}$ is the function $E[X \mid \mathcal{H}] : \Omega \to \mathbb{R}^n$ such that*

$E[Y_t \mid X_0 = x]$

1. *$E[X \mid \mathcal{H}]$ is $\mathcal{H}$-measurable,*

2. *$\int_H E[X \mid \mathcal{H}]\, dP = \int_H X\, dP$ for all $H \in \mathcal{H}$.*

notation: $E[X \mid H] = E[X \mid \mathcal{H}](\omega)$ $\forall \omega \in H$

$E[X] = \int_\Omega X(\omega)\, dP(\omega)$

$H \in \mathcal{H}$

if $H \in \mathcal{H}$ cannot be partitioned: $\int_H E[X \mid \mathcal{H}]\, dP = E[X \mid H] \int_H dP$

$= E[X \mid H]\, P(H)$

Characteristic function

Definition 1.14 *The* ***characteristic function*** *of a random variable* $X : \Omega \to \mathbb{R}^n$ *for* $X = (X_1, \ldots, X_n)'$ *is the function* $\phi_X : \mathbb{R}^n \to \mathbb{C}$ *defined as*

$$\phi_X(u) = E[\exp(i(u \cdot X))] = E[\exp(i(u_1 X_1 + \ldots + u_n X_n))]$$

where $u \in \mathbb{R}^n$ *and* i *is the imaginary unit. In other words,* ϕ_X *is the Fourier transform of* X *(more precisely, of the measure* $P(X \in dx)$*).*

The characteristic function of X uniquely determines its distribution. Also, for $X : \Omega \to \mathbb{R}^n$ with $X \sim N(\mu, \Xi)$, you can verify that

$$\phi_X(u) = \exp\left(i(u \cdot \mu) - \frac{1}{2}u'\Sigma u\right). \tag{1.1}$$

m.g.f $m_X(u) = E[\exp(u \cdot X)]$

Filtrations and stochastic processes

Definition 1.15 *A **filtration** $\{\mathcal{F}_t : t \in \mathcal{T}\}$ on $(\Omega, \mathcal{F})$ is a family of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ such that for all $s, t \in \mathcal{T}$,*

$$s < t \implies \mathcal{F}_s \subset \mathcal{F}_t.$$

Remark 1.16 *Since $\forall t \in \mathcal{T}$, $\mathcal{F}_t$ is a σ-algebra such that $\mathcal{F}_t \subseteq \mathcal{F}$, we can define a conditional expectations operator on $\mathcal{F}_t$ using Definition 1.13. When there is no confusion, we will use the notation*

$$E\left[X \mid \mathcal{F}_t\right] \doteq E_t\left[X\right].$$

Intuitively, we can view a stochastic process $\{X_t : t \in \mathcal{T}\}$ is a parameterized collection of random variables on $(\Omega, \mathcal{F}, P)$ with values in $\mathbb{R}^n$. For each $t \in \mathcal{T}$, X_t is then a random variable that maps $\omega \in \Omega$ into $\mathbb{R}^n$. Also, for every $\omega \in \Omega$, $X_t(\omega)$ is a function that maps t into $\mathbb{R}^n$ (also called a path or a trajectory of the process). However, we would like to also achieve certain consistency conditions across time. We therefore define the process as a measurable function on the product σ-algebra on $\Omega \times \mathcal{T}$.

Definition 1.17 *The* ***product*** $\boldsymbol{\sigma}$***-algebra*** $\mathcal{F} \otimes \mathcal{B}(\mathcal{T})$ *on* $\Omega \times \mathcal{T}$ *is the* σ*-algebra generated by the subsets of the form* $F \times B$ *where* $F \in \mathcal{F}$ *and* $B \in \mathcal{B}$ *where* $\mathcal{B}$ *is the Borel* σ*-algebra on* $\mathcal{T}$*.*

Definition 1.18 *A* ***stochastic process*** *is a function* $X : \Omega \times \mathcal{T} \rightarrow \mathbb{R}^n$ *that is measurable with respect to the product* σ*-algebra* $\mathcal{F} \otimes \mathcal{B}(\mathcal{T})$ *on* $\Omega \times \mathcal{T}$*.*

The stochastic process is a function of two arguments now, $X_t(\omega)$. For a given ω, the function $X_.(\omega)$ is a sample path.

Definition 1.19 *A stochastic process* X *is* ***adapted to filtration*** $\{\mathcal{F}_t\}$ *if,* $\forall t \in \mathcal{T}$*, the function* X_t *is* $\mathcal{F}_t$*-measurable.*

An adapted stochastic process is such that the realization of its path up to time t depends on information revealed up to time t.

Convergence theorems

Billingsley (1979), Section 16.

The following theorems will be useful in the proofs of many of our statements. They describe the properties of limits for expectations of sequences of measurable functions.

When we talk about integrability here, we have in mind Lebesgue integrability, i.e. f is integrable if both f^+ and f^- are integrable, because $\int f d\mu = \int f^+ d\mu - \int f^- d\mu$.

Proposition 1.20 ***(Dominated convergence theorem)*** *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $\{f_n\}$ be a sequence of measurable functions $f_n : \Omega \to \mathbb{R}$. If the sequence converges pointwise almost everywhere to a function f $(f_n \to f)$ and if there is an integrable function g such that $|f_n(\omega)| \leq g(\omega)$ almost everywhere for all n, then f and all f_n are integrable and*

$$\lim_{n \to \infty} \int f_n d\mu = \int f d\mu.$$

Proof. Omitted. ■

Proposition 1.21 ***(Monotone convergence theorem)*** *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and let* $\{f_n\}$ *be a non-decreasing sequence of non-negative measurable functions* $f_n : \Omega \to \mathbb{R}$*. If the sequence converges pointwise almost everywhere to a function* f *(i.e,* $0 \leq f_n \nearrow f$*), then* f *is measurable and*

$$\lim_{n \to \infty} \int f_n d\mu = \int f d\mu.$$

Proof. Omitted. ■

Proposition 1.22 (Fatou's lemma) *For a sequence of nonnegative measurable functions* f_n*,*

$$\int \liminf_n f_n d\mu \leq \liminf_n \int f_n d\mu.$$

Proof. Omitted. ■

1.2 Martingales and stopping times

Definition 1.23 *An n-dimensional process* $X = \{X_t\}_{t\in\mathcal{T}}$ *on* $(\Omega, \mathcal{F}, P)$ *is a* ***martingale with respect to the filtration*** $\{\mathcal{F}_t\}$ *and the probability measure P if:*

1. X_t *is* $\mathcal{F}_t$*-measurable (i.e., X is adapted),*

2. $E[|X_t|] < \infty$ *for all* $t \in \mathcal{T}$*,*

3. $E[X_s \mid \mathcal{F}_t] = X_t$ *for all* $s \geq t$*.*

A ***supermartingale*** *is a process for which condition 3. is replaced by* $E[X_s \mid \mathcal{F}_t] \leq X_t$*, and a* ***submartingale*** *is a process for which condition 3. is replaced by* $E[X_s \mid \mathcal{F}_t] \geq X_t$*.*

A martingale is always defined with respect to a particular filtration.

Example 1.24 *Consider a fair gamble process starting at* $X_0 = 0$*, and with a 0.5 chance of winning \$1, and 0.5 chance of losing \$1 in every period, i.e.,*

$$X_{t+1} = \begin{cases} X_t + 1 & \text{with probability } 0.5 \\ X_t - 1 & \text{with probability } 0.5 \end{cases}.$$

It is straightforward to show that X*, representing accumulated winnings, is a martingale.*

Definition 1.25 *A **stopping time** is an $\mathcal{F}$-measurable function $\tau : \Omega \to \mathcal{T}$ (a random variable) such that*

$$\{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t. \tag{1.2}$$

*A stopping time is said to be **bounded** if there exists a constant $C < \infty$ such that $P(\tau \leq C) = 1$, and is said to be **finite** (almost surely) if $P(\tau < \infty) = 1$.*

A stopping time is a **random variable** – for outcomes $\omega \in \Omega$ in different $F \in \mathcal{F}$, there is potentially a different value of $\tau(\omega)$ — and at each time t, it is known whether the stopping occurred or not.

An example of a stopping time is the time when a process hits a particular boundary.

Example 1.26 *Consider a continuous-time univariate process X with continuous sample paths and the filtration $\{\mathcal{F}_t\}_{t \in \mathcal{T}}$ generated by X. The **first passage time** (or **crossing** time) of threshold $k \in \mathbb{R}$ is the time of first crossing of the threshold:*

$$\tau^k(\omega) = \min\{t : X_t(\omega) = k\}.$$

Notice that the first passage time is a stopping time, since the set in (1.2) is the set of paths that crossed k before t, and this set is in $\mathcal{F}_t$.

Example 1.27 *On the other hand, the **last passage time** is in general not a stopping time, because the set $\left\{\omega \in \Omega : \tau^k(\omega) \leq t\right\}$ depends on the trajectories of the stochastic process in the future (given information at time t, we do not generally know if the process will cross k again in the future), and thus is not in $\mathcal{F}_t$.*

Optional stopping theorem in discrete time

In this section, we work in **discrete time** and want to prove the following statement: Let τ be a stopping time and X a martingale with continuous sample paths. Then the **stopped process** X^τ, defined as

$$X_t^\tau = X_{t\wedge\tau} = X_{\min\{t,\tau\}}$$

is also a martingale.

Observe the construction of the stopped martingale. For a given $\omega \in \Omega$, the path of X is the function $X_t(\omega)$ of time t. Also associated with this ω is the realization of the stopping time $\tau(\omega)$. The path of the stopped process is therefore

$$X_t^\tau(\omega) = \begin{cases} X_t(\omega) & \text{if } t < \tau(\omega) \\ X_{\tau(\omega)}(\omega) & \text{if } t \geq \tau(\omega) \end{cases}.$$

Definition 1.28 *A σ-algebra $\mathcal{F}_{\mathcal{T}}$* ***co-generated by a stopping time*** *τ is the σ-algebra generated by all sets $A \in \mathcal{F}$ such that $\forall t \in \mathcal{T}$, $A \cap \{\tau \leq t\} \in \mathcal{F}_t$.*

In words, the σ-algebra $\mathcal{F}_{\mathcal{T}}$ contains all information that is revealed by the stochastic process until the stopping time occurs. Hence, along paths on which the stopping time $\tau(\omega)$ occurs later, $\mathcal{F}_{\mathcal{T}}$ provide more information. This concept gets rid of the notion of calendar time, and defines time (and hence information revelation) relative to a particular event.

Proposition 1.29 (Optional stopping theorem in discrete time) *Assume that time is discrete,* $\mathcal{T} = \{0, 1, 2, \ldots\}$*. If* $\sigma \leq \tau$ *are two bounded stopping times, then for any discrete-time submartingale* X_n*,* $n = 0, 1, 2, \ldots$*,*

$$X_\sigma \leq E\left[X_\tau \mid \mathcal{F}_\sigma\right] \quad a.s. \tag{1.3}$$

with equality if X *is a martingale. Moreover, an adapted and integrable process* X *is a martingale if and only if*

$$E\left[X_\sigma\right] = E\left[X_\tau\right] \tag{1.4}$$

for any such pair of stopping times.

Proof. See discussion in class. ■

We can now finally prove the desired statement, which is the consequence of the previous proposition.

Corollary 1.30 *If X is a martingale with respect to $\{\mathcal{F}_t\}$ and τ a stopping time, the stopped process X^τ is a martingale with respect to $\{\mathcal{F}_t\}$.*

Proof. See discussion in class. ■

not necessarily bounded

Exercise 1.31 *Consider a sequence of three coin flips, occurring at time $t = 1, 2, 3$ with realizations H (heads) or T (tails). Outcomes ω in this probability space are therefore sequences of type HHH, HHT, etc. Let τ denote the stopping time described as the first time H occurs.*

Show directly from the definition of a σ-algebra co-generated by a stopping time that $\mathcal{F}_\tau$ is constructed from sets

$$\{HHH, HHT, HTH, HTT\}$$
$$\{THT, THH\}, \{TTH\}, \{TTT\}$$

and all their unions, intersections and complements. In particular, explain why the set $\{HHH, HTH\}$ does not belong to that σ-algebra.

***Hint**: It may be useful to plot a tree of ω that branches out as coin flips are realized, and depict $\tau(\omega)$ and the sets measurable with respect to $\mathcal{F}_\tau$ on that tree.*

Example 1.32 *Continuing with Example 1.24, devise now the following strategy. Starting with* $X_0 = 0$*, the agent decides to quit betting when she earns one dollar. Formally, she chooses a stopping time* τ *defined as*

$$\tau(\omega) = \min\{t \geq 0 : X_t(\omega) = 1\}.$$

Hence once the stopped process X^τ *given by* $X_t^\tau(\omega) = X_{\tau(\omega)\wedge t}(\omega)$ *reaches 1, it stays constant.*

Corollary 1.30 implies that this stopped process X^τ *is a martingale, hence* $X_0 = E[X_t^\tau]$*, and this strategy never wins on average.*

Example 1.33 (A doubling strategy) *A similar strategy is a so-called* ***doubling strategy****. Consider an accumulated winings process X with $X_0 = 0$ and*

$$X_{t+1} = \begin{cases} X_t + 2^t & \text{with probability } 0.5 \\ X_t - 2^t & \text{with probability } 0.5 \end{cases}.$$

This stopping time strategy starts with a bet of one dollar. If the player wins, she terminates the game. If she loses, she doubles the bet. Formally, we are interested in the strategy

$$\tau(\omega) = \min\{t \geq 0 : X_t(\omega) = 1\},$$

i.e., the first time the process hits 1.

Notice that in every period, the process is stopped with probability $\frac{1}{2}$. Therefore,

$$P(\tau(\omega) \leq t) = 1 - \left(\frac{1}{2}\right)^t$$

and asymptotically as $t \to \infty$, the agent stops betting with probability one, with one dollar in the pocket. This seems like a profitable strategy.

However, the optional stopping theorem implies the stopped process $X_t^\tau = X_{\min\{t,\tau\}} = X_{t\wedge\tau}$ *is a martingale. Hence* $E\left[X_t^\tau \mid \mathcal{F}_0\right] = X_0^\tau = 0$, *and no finite-horizon strategy wins on average, despite the fact that the limit* $\lim_{t\to\infty} X_t^\tau = 1$ *a.s. (looking like a 'safe win', a.s.).*

The issue is that despite the probability of winning one dollar converges to one, the losses along the remaining paths grow exponentially, so that the martingale property still holds. We will formalize this later using the concept of local martingales.

Optional stopping theorem in continuous time

The continuous-time version of the optional stopping theorem and related results are conceptually the same, they only require additional technical considerations. I state these results here for the sake of completeness but feel free to skip them.

- A crucial simplification lies in restricting ourselves to processes with continuous paths. Because we will be interested in Brownian motions with continuous paths, this does not limit us significantly. See Revuz and Yor (1999) and Karatzas and Shreve (1991) for more detail.
- Continuity of sample paths becomes important because we will be approximating stopping times. When we approximate a stopping time with another one, continuous sample paths imply that we are also correctly approximating the path realizations at the stopping time.

In the previous section, we built the proof of the discrete-time version of the optional stopping theorem on bounded stopping times. Bounded stopping times in discrete time can take only finitely many values. The first result shows that an arbitrary stopping time can be approximated by stopping times that take finitely many values.

Proposition 1.34 *Every stopping time is the decreasing limit of a sequence of stopping times taking only finitely many values.*

Proof. For a stopping time τ construct a sequence $\left\{\tau^k\right\}$ as follows:

$$\begin{aligned} \tau^k &= +\infty \text{ if } \tau \geq k \\ \tau^k &= q2^{-k} \qquad \text{if } \left(q-1\right)2^{-k} \leq \tau < q2^{-k},\ \ q \leq 2^k k \end{aligned}$$

Then every τ^k is a stopping time and the sequence $\left\{\tau^k\right\}$ decreases to τ. ■

Definition 1.35 *A collection of random variables* $\{X_t : t \in \mathcal{T}\}$ *is* ***uniformly integrable*** *if*

$$\sup_{t \in \mathcal{T}} E\left[|X_t| \, 1_{\{|X_t| > x\}}\right] \to 0 \quad \textit{as } x \to \infty.$$

This definition is important because uniform integrability of X will assure that the random variable X_σ where σ is a stopping time is also integrable.

Proposition 1.36 (Optional stopping theorem in continuous tim
If X is a martingale with continuous paths and $\sigma \leq \tau$ are bounded stopping times, then

$$X_\sigma = E\left[X_\tau \mid \mathcal{F}_\sigma\right] \quad a.s.$$

Alternatively, if X is uniformly integrable, then the family $\{X_\sigma\}$ where σ runs through the set of all stopping times is uniformly integrable. If, in addition, $\sigma \leq \tau$, then

$$X_\sigma = E\left[X_\tau \mid \mathcal{F}_\sigma\right] = E\left[X_\infty \mid \mathcal{F}_\sigma\right] \quad a.s.$$

where $X_\infty = \lim_{t \to \infty} X_t$.

Proof. See discussion in class. ■

Proposition 1.37 *A continuous adapted process is a martingale if and only if for every bounded stopping time τ, the random variable X_τ is in L^1 and*

$$E[X_\tau] = E[X_0].$$

Proof. See discussion in class. ■

Corollary 1.38 *If X is a continuous martingale with respect to $\{\mathcal{F}_t\}$ and τ a stopping time, the stopped process X^τ is a martingale with respect to $\{\mathcal{F}_t\}$.*

Proof. Consider the stopping time τ. The stopped process $X_t^\tau = X_{\tau\wedge t}$ is continuous. Consider a bounded stopping time σ, then $\tau \wedge \sigma$ is also bounded. Therefore

$$E[X_\sigma^\tau] = E[X_{\tau\wedge\sigma}] = E[X_0] = E[X_0^\tau].$$

■

Local martingales

Definition 1.39 *An adapted stochastic process* $X : \Omega \times \mathcal{T} \to \mathbb{R}^n$ *is a* ***local martingale*** *with respect to* $\{\mathcal{F}_t\}$ *if there exists a sequence of* $\{\mathcal{F}_t\}$ *stopping times* τ_k, $k = 1, 2, \ldots$ *such that*

1. *the sequence* τ_k *is almost surely increasing, i.e.* $P\left(\tau_k < \tau_{k+1}\right) = 1$,
2. *the sequence* τ_k *almost surely diverging, i.e.,* $P\left(\tau_k \xrightarrow{k\to\infty} \infty\right) = 1$
3. *the stopped process* $X_t^k = X_{t\wedge\tau_k} = X_{\min(t,\tau_k)}$ *is a martingale for every* k.

Observe that every martingale X is trivially a local martingale because every stopped martingale in Condition 3. in the definition above is also a martingale, by Corollary 1.30.

Why do we need the definition of a local martingale? It may be that the process X is such that a small set of paths (a set of measure zero in the limit) diverges to $\pm\infty$, which invalidates the conditional expectation property of a martingale. An example of this was the limiting distribution in the doubling strategy in Example 1.33. By constructing a sequence of stopping times that 'freezes' these paths before they explode, we can make sure that these diverging paths do not distort the computation of the conditional expectation, and the martingale property on the stopped process is preserved.

Example 1.40 *Consider again the process X_t from Example 1.24, and construct a sequence of stopping times*

$$\tau_k = \min\{t \geq 0 : X_t \leq -k\}.$$

The intepretation of the stopping time τ_k is for the agent to stop when cummulative losses reach $-k$.

Observe that the sequence of these stopping times $\{\tau_k\}_{k=1}^{\infty}$ satisfies conditions 1. and 2. in Definition 1.39. It is increasing, because the wealth process has to first cross $-k$ before it crosses $-(k+1)$. It also diverges to infinity because the time to reach the boundary is made arbitrarily long as $k \to \infty$. Condition 3. is satisfied trivially, because X is a martingale and every stopped martingale X^τ is also a martingale, by Corollary 1.30.

While a martingale is a local martingale, the converse is generally not true. To see this, consider again the doubling strategy.

Example 1.41 (A doubling strategy again) *Let us revisit Example 1.33. In that example, the stopped winnings process for the stopping time*

$$\tau(\omega) = \min\{t \geq 0 : X_t(\omega) = 1\}$$

converges to $\lim_{t\to\infty} X_t^{\tau} = 1$ *almost surely, but at the same time there is a set of paths that has a vanishing probability and accumulates larger and larger losses. Hence, for every finite* t*, the process still satisfies the martingale property* $E[X_t^{\tau}] = X_0 = 0$*,* $\forall t \in \mathbb{N}$*.*

Let us now redefine the time axis. Instead of betting at times $t = 1, 2, \ldots$*, we construct these bets at times* $1 - 2^{-t}$*,* $t = 1, 2, \ldots$*. The first bet is at time* $\frac{1}{2}$*, then* $\frac{3}{4}$*,* $\frac{7}{8}$*, etc. Hence bets occur at an increasing rate, all before* $t = 1$*.*

Formally, the new betting process is constructed through a time change of the original winnings process.

t < 1

$$\widetilde{X}_t = \begin{cases} X_n^{\tau} & \text{for } t = 1 - 2^{-n}, \; n \in \mathbb{N} \\ 1 & \text{for } t = 1 \text{ (i.e., } n = \infty), 2, 3, \ldots \end{cases}$$

Observe that the second line is an extension of the original process X^{τ} *using the limiting point. This process is not a martingale on* $[0, 1]$*. While it is still true that* $E\left[\widetilde{X}_t\right] = X_0 = 0$ *for* $t < 1$ *(where* t *is of the form* $t = 1 - 2^{-n}$*), for* $t \geq 1$ *we have* $E\left[\widetilde{X}_1\right] = 1$*.*

The time change can be understood as follows: Because waiting for X_t *to hit* 1 *could take a very long (infinite) time, I will progressively speed time so that the whole history of* X *is played out in* $\widetilde{X}$ *during* $t \in [0, 1)$*. Then, because we know that the process has hit 1 until time* $t = 1$ *with probability one, we can a.s. continuously extend the definition of* $\widetilde{X}_t$ *as* $\widetilde{X}_1 = 1$*.*

*While this looks like a safe construction, it **violates the martingale property**. The process is still a **local martingale**, though. Consider a localizing sequence of stopping times*

$$\tau_k = \min\left\{t \geq 0 : \widetilde{X}_t \leq -k\right\}$$

as the first time the process crosses $-k$. This sequence is increasing, almost surely diverging (because less and less paths ever hit the threshold $-k$ as $-k \to -\infty$), and the stopped process $\widetilde{X}_t^{\tau_k} = \widetilde{X}_{t \wedge \tau_k}$ is a martingale on $t \in [0, 1]$ for every k.

To see the latter, notice that the martingale property holds for $t < 1$ as a stopped martingale is a martingale. So it remains to check $t = 1$, for which we can apply the dominated (bounded) convergence theorem.

Observe that for a given k, the stopped process $\widetilde{X}_t^{\tau_k}$, can be viewed as a sequence of random variables $\left\{\widetilde{X}_t^{\tau_k}\right\}$, $t = 1 - 2^{-n}$. Each of the random variables is bounded (from above by 1, from below by $-k$), and the sequence converges to $\widetilde{X}_1 = 1$ pointwise.

*By the **dominated concergence theorem**, the expectation of the limit must be equal to the limit of the expectations, i.e.,*

$$E\left[\widetilde{X}_1^{\tau_k}\right] = \lim_{n \to \infty} E\left[\widetilde{X}_{1-2^{-n}}^{\tau_k}\right] = 0.$$

Hence, $\widetilde{X}_t^{\tau_k}$ is a martingale. Consequently, conditions of Definition 1.39 are satisfied and the process $\widetilde{X}_t$ is a local martingale.

When is a local martingale a martingale

The discussion in Example 1.41 shows that we must impose some additional restrictions on a local martingale to make it a martingale.

Proposition 1.42 *A nonnegative local martingale X with continuous paths on a given filtration $\{\mathcal{F}_t\}$ (or, more generally, a local martingale bounded from below) is a supermartingale.*

Proof. *See discussion in class.* ■

Proposition 1.43 *A bounded local martingale X on a given filtration $\{\mathcal{F}_t\}$ with continuous paths is a martingale.*

Proof. See discussion in class. ■

Example 1.44 (A doubling strategy yet again) *We observed that the time-changed doubling strategy in Example 1.41 that constructs the bets at accelerating times $t = 1 - 2^{-n}$ and in addition defines the process $\widetilde{X}_t$ at $t = 1$ as $\widetilde{X}_1 = 1$ is a* ***local martingale but not a martingale****.*

Proposition 1.42 also implies that if $\widetilde{X}_t$ was a local martingale bounded from below, it would be a supermartingale, and since $\widetilde{X}_t$ is already bounded from above (stopping when winnings are equal to 1), it would in fact be a martingale.

Economically, imposing a bound from below corresponds to imposing a limit on losses (regardless how generous this limit is). When imposing this limit, not even the time-changed doubling strategies can win on average.

This is a plausible economic assumption that avoids the "$0 \cdot \infty$" problem with diverging paths of decaying probability. Notice that we do not need to specify what the bound exactly is — it is enough to assume that there is one.

1.3 Brownian motion

Definition 1.45 *A* ***k-dimensional Brownian motion*** *is a stochastic process W on $\mathbb{R}^k$ such that*

1. $W_0(\omega) = 0$ *for all* ω,
2. $\forall s, t \in \mathcal{T}$ *for which* $s \leq t$, *the difference* $W_t - W_s \sim N(0, (t-s) I_k)$,
3. *for all* $t_0 < t_1 < t_2 < \ldots t_n \in \mathcal{T}$, *the random variables* $W_{t_j} - W_{t_{j-1}}$, $j \in \{1, \ldots, n\}$ *are independent.*

Said simply, the Brownian motion is a process with independent, normally distributed increments.

Remark 1.46 *Technically, this definition of the Brownian motion does not yield a unique process. However, we can pin the process down uniquely (almost surely) if we choose a modification of the process with continuous sample paths. From now on, we will exclusively work with such a continuous-path modification.*

Observe that the **Brownian motion** (as any other process) naturally **generates a σ-algebra** that includes the realizations of all infinite-horizon paths, and a filtration based on paths up to time t.

In particular, consider the σ-algebra generated by all the sets of the type

$$\left\{\omega \in \Omega : W_s(\omega) \in B, B \in \mathcal{B} \text{ on } \mathbb{R}^k\right\}$$

for all $s \leq t$, and denote it as $\mathcal{F}_t^*$. Further, extend this algebra by adding all subsets of zero probability sets in $\mathcal{F}_t^*$ (in order to complete the probability space), and construct a σ-algebra over this union, denoted $\mathcal{F}_t$. Then $\{\mathcal{F}_t\}$ is the Brownian filtration generated by the Brownian motion W.

Further observe that the Brownian motion has the **Markov property**: $\forall t, s \geq 0$ and for every Borel set $H \in \mathcal{B}$ on $\mathbb{R}^k$

$$P\left(W_{t+s} \in H \mid \mathcal{F}_t\right) = P\left(W_{t+s} \in H \mid W_t\right).$$

Example 1.47 *A Brownian motion is a martingale with respect to its natural filtration. Observe that for* $s < t$

$$E\left[W_t \mid \mathcal{F}_s\right] = E\left[W_t - W_s \mid \mathcal{F}_s\right] + W_s = W_s$$

and

$$\left(E\left[|W_t|\right]\right)^2 \leq E\left[|W_t|^2\right] = nt.$$

Definition 1.48 *The set of points* $\mathcal{P} = \{t_0, \ldots, t_n\}$ *with* $0 = t_0 < t_1 < \ldots < t_n = t$ *is a* ***partition of the interval*** $[0, t]$*. Define*

$$l(\mathcal{P}) = \max \left| t_j - t_{j-1} \right|.$$

to be the ***norm of the partition****.*

Definition 1.49 *Let* $X : \Omega \times \mathcal{T} \to \mathbb{R}$ *be a continuous stochastic process. Then for* $p > 0$ *define the p-th variation process of* X_t *as*

$$\langle X, X \rangle_t^p (\omega) = \lim_{l(\mathcal{P}) \to 0} \sum_{j=0}^{n-1} \left| X_{t_{j+1}}(\omega) - X_{t_j}(\omega) \right|^p$$

where the limit is in probability.

For $p = 1$*, this is the* ***total variation process****, and for* $p = 2$*, this is the* ***quadratic variation process****.*

Exercise 1.50 *It turns out that for the univariate Brownian motion W,*

$$\langle W, W \rangle_t^2 (\omega) = t \quad a.s..$$

Remark 1.51 *Notice that this formula also leads to the conclusion that over a short period of time $(dW_t)^2 = dt$, a result pervasively used in the calculus of diffusions.*

Since the quadratic variation is finite, it can be shown that the total variation of a Brownian motion is infinite. The latter also implies that the paths of a Brownian motion are nowhere differentiable.

Exercise 1.52 *Show that*

$$\forall t > 0 : \langle X, X \rangle_t^1 (\omega) = +\infty$$

and that the paths of a Brownian motion are nowhere differentiable.

Example 1.53 (Doubling strategies with a Brownian motion) *He[...] we provide another example showing that a martingale is a local martingale but the converse is not true. This is a direct counterpart of Example 1.44.*

Consider now a Brownian motion W *and* $\tau = \min\{t \geq 0 : W_t = 1\}$ *the first time the process hits* 1*. Think about this as a stopping rule that says 'Gamble until I earn one dollar, then stop.' Since the Brownian motion will hit* $W_t = 1$ *at some time* $t \geq 0$ *with probability 1, this seems like a profitable strategy.*

Observe that the stopped process $W_{\min\{t,\tau\}} = W_{t\wedge\tau}$ *is a martingale, with zero expectation* $E\left[W_{t\wedge\tau} \mid \mathcal{F}_0\right] = 0$*, but the limit* $\lim_{t\to\infty} W_{t\wedge\tau} = 1$ *a.s. (like a 'safe win', a.s.). Now define a process constructed through a time change of the stopped Brownian motion:*

$$X_t = \begin{cases} W_{\frac{t}{1-t}\wedge\tau} & \text{for } 0 \leq t < 1 \\ 1 & \text{for } 1 \leq t < \infty \end{cases}$$

The process is continuous a.s., but not a martingale.

1.4 Stochastic integration

Riemann (Stieltjes) integral: Construct a partition $\mathcal{P}$ and then show that

$$\lim_{l(\mathcal{P})\to 0} \sum_{j=1}^{n} f\left(\tau_j\right)\left(t_j - t_{j-1}\right) \to \int_0^t f\left(s\right) ds$$

where $\tau_j \in \left[t_{j-1}, t_j\right]$. It does not matter for the construction which τ_j's we choose.

Stochastic (Itô) integral: In what follows, we want to motive the familiar construction

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s.$$

Now the choice of τ_j's will matter.

Example 1.54 *As a motivation, consider the economically interesting example. Let the share price evolve as a Brownian motion and* θ *be the number of shares bought, which can be traded only at a finite number of times* t_j*. Then the evolution of wealth* J_t *is given by*

$$J_T = J_0 + \sum_{j=0}^{n-1} \theta_{t_j} \left(W_{t_{j+1}} - W_{t_j} \right)$$

where θ_{t_j} *is the number of shares purchases at time* t_j*.*

Definition 1.55 *An* ***elementary*** *(also called simple)* ***process*** ϕ *is a process for which there exists a partition* $\mathcal{P}$ *of* $[0,T]$ *such that* $\phi_t = \phi_{t_j}$ *for* $t \in [t_j, t_{j+1})$.

Example 1.56 *Consider a partition* $\mathcal{P}$ *of* $[0,T]$ *and define*

$$\begin{aligned}\phi_t^1(\omega) &= \sum_{j=0}^{n(t)-1} W_{t_j}(\omega)\,\mathbf{1}_{[t_j,t_{j+1}]}(t)\\ \phi_t^2(\omega) &= \sum_{j=0}^{n(t)-1} W_{t_{j+1}}(\omega)\,\mathbf{1}_{[t_j,t_{j+1}]}(t)\end{aligned}$$

where $n(t)$ *is such that* $t \in [t_{n(t)}, t_{n(t)+1})$. *Notice that in order to define the Riemann integral in the deterministic case, we can use both* ϕ^1 *and* ϕ^2.

However, in the stochastic case, this will not work. Observe that

$$E\left[\int_0^T \phi_t^1 dW_t \mid \mathcal{F}_0\right] = E\left[\sum_{j=0}^{n-1} W_{t_j}\left(W_{t_{j+1}} - W_{t_j}\right) \mid \mathcal{F}_0\right] = 0$$
$$E\left[\int_0^T \phi_t^2 dW_t \mid \mathcal{F}_0\right] = E\left[\sum_{j=0}^{n-1} W_{t_{j+1}}\left(W_{t_{j+1}} - W_{t_j}\right) \mid \mathcal{F}_0\right] = T.$$

So despite the fact that both seem to be reasonable approximations of the integral, they give very different answers.

As we will see, the issue why the two expectations above have different limits as $l(\mathcal{P}) \to 0$ is closely related to the fact that W is a process of infinite total variation.

Definition 1.57 *For the class of elementary processes, define the* ***Itô integral*** *as follows:*

$$\int_0^t \phi_s dW_s = \sum_{j=0}^{n(t)-1} \phi_{t_j} \left(W_{t_{j+1}} - W_{t_j}\right) + \phi_{t_{n(t)}} \left(W_t - W_{t_{n(t)}}\right) \quad (1.5)$$

where the last term reflects the interrupted last subinterval of the partition and $n(t)$ *is such that* $t \in [t_{n(t)}, t_{n(t)+1})$.

Remark 1.58 *From the perspective of asset pricing, this is a desirable definition. Observe that* ϕ^1 *is adapted while* ϕ^2 *is not. From the perspective of asset pricing, we can think about (a general)* ϕ^1 *as a portfolio strategy. At time* t_j*, we choose a portfolio* $\phi^1_{t_j}$ *based on information available at time* t_j*, and then hold it to* t_{j+1}*, rebalancing the portfolio at* t_{j+1} *again. This is called a* dynamic strategy *and by refining the partition* $\mathcal{P}$*, we construct such a strategy for infinitesimal rebalancing.*

Sketch of the construction of an Ito integral

Let $\mathcal{L}$ be the set of all processes adapted to the filtration generated by the Brownian motion. Then define:

$$\mathcal{L}^1 = \left\{ f \in \mathcal{L} : \int_0^T |f_t| \, dt < \infty \text{ a.s.} \right\} \tag{1.6}$$

$$\mathcal{L}^2 = \left\{ f \in \mathcal{L} : \int_0^T (f_t)^2 \, dt < \infty \text{ a.s.} \right\} \tag{1.7}$$

$$\mathcal{H}^2 = \left\{ f \in \mathcal{L}^2 : E\left[\int_0^T (f_t)^2 \, dt \right] < \infty \right\} \tag{1.8}$$

Clearly $\mathcal{H}^2 \subset \mathcal{L}^2$. We start with $\mathcal{H}^2$ and then move to $\mathcal{L}^2$. The latter will be somewhat complicated.

The main idea is to argue that for a function $f \in \mathcal{H}^2$, there exists an approximating sequence ϕ^n of elementary functions such that

$$\lim_{n\to\infty} E\left[\int_0^T (\phi_t^n - f_t)^2 \, dt\right] = 0$$

We can then show that the stochastic integral (1.5) for the sequence of elementary functions ϕ^n converges to a limit in L^2, and it is independent of the particular choice of a sequence. Then we associate this limit with the stochastic integral

$$\int_0^T f_t dW_t \tag{1.9}$$

for $f \in \mathcal{H}^2$. Øksendal (2007), Chapter 3.1, has all the details. This allows the following formal definition.

Definition 1.59 *Let $f \in \mathcal{H}^2$. Then the **Itô integral** of f is defined by*

$$\int_0^T f_t(\omega)\, dW_t(\omega) = \lim_{n \to \infty} \int_0^T \phi_t^n(\omega)\, dW_t(\omega) \tag{1.10}$$

where the limit is in $L^2(P)$ sense (set of all random variables with finite second moments), and $\{\phi^n\}$ is a sequence of elementary functions such that

$$E\left[\int_0^T \left(f_t(\omega) - \phi_t^n(\omega)\right)^2 dt\right] \to \infty \quad as\ n \to \infty.$$

Remark 1.60 *The stochastic integral (1.10) is a martingale.*

For $\mathcal{L}^2$ the Itô integral can be defined in an analogous way, using arguments about convergence of step functions $\phi^n \in \mathcal{L}^2$ to $f \in \mathcal{L}^2$ such that $\int_0^T |\phi_t^n - f_t|^2 \, dt \to 0$ in probability. We can then define the integral as

$$\int_0^T f_t(\omega) \, dW_t(\omega) = \lim_{n\to\infty} \int_0^T \phi_t^n(\omega) \, dW_t(\omega) \quad \text{(limit in probability).}$$

However, the resulting integral is only a *local martingale*. This is a much weaker property.

It is also possible to extend Definition 1.59 to multivariate Brownian motions. The integral $\int_0^T v_t dW_t$ where W is an k-dimensional Brownian motion and v is an $n \times k$ dimensional process adapted to the filtration generated by W can be then defined by the componentwise summation

$$\int_0^T v_t(\omega)\, dW_t(\omega) = \sum_{j=1}^{k} [v_t(\omega)]_{\cdot j}\, dW_t^j(\omega)$$

and the individual summands are defined as before.

Example 1.61 *We want to show that when* $W_0 = 0$*, we have*

$$\int_0^T W_t dW_t = \frac{1}{2} W_t^2 - \frac{1}{2} t,$$

using limits with elementary processes.

Properties of the Itô integral

Proposition 1.62 (Itô isometry) *For all $f \in \mathcal{H}^2$, we have*

$$E\left[\left(\int_0^T f_t dW_t\right)^2 \mid \mathcal{F}_0\right] = E\left[\int_0^T (f_t)^2 \, dt \mid \mathcal{F}_0\right]. \tag{1.11}$$

Proof. See discussion in class. ■

1.5 Itô processes and Itô's lemma

Definition 1.63 *An n-**dimensional Itô process** is a process $S : \Omega \times \mathcal{T} \rightarrow \mathbb{R}^n$ such that*

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s \tag{1.12}$$

where $\mu \in (\mathcal{L}^1)^n$, $\sigma \in (\mathcal{L}^2)^{n\times k}$ and W is a k-dimensional Brownian motion. We assume that μ and σ are $\mathcal{F}_t$-adapted where $\{\mathcal{F}_t\}$ is some filtration with respect to which W is a martingale.

*An **Itô diffusion** is an Itô process for which the coefficients satisfy $\mu_s = \mu(X_s)$ and $\sigma_s = \sigma(X_s)$ for all $s \in \mathcal{T}$.*

Remark 1.64 *The process μ is called drift, and σ is called the volatility of the Itô process.*

Remark 1.65 *Sometimes, the Itô process is defined with additional Lipschitz continuity conditions on the parameters. These conditions assure the existence of a unique strong solution for the equation (1.12). We will introduce these conditions in Section 1.8 when we discuss stochatic differential equations explicitly.*

Often, the Itô process is written in the 'differential' form

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

If the processes $\mu, \sigma \in \mathcal{H}^2$, then

$$\begin{aligned} \left.\frac{d}{d\tau} E_t \left[X_\tau\right]\right|_{\tau=t} &= \mu_t \quad \text{a.s.} \\ \left.\frac{d}{d\tau} Var_t \left[X_\tau\right]\right|_{\tau=t} &= \sigma_t \sigma_t' \quad \text{a.s.} \end{aligned}$$

where $E_t[\cdot] = E[\cdot \mid \mathcal{F}_t]$. Informally, we will write

$$\begin{aligned} E_t\left[dX_t\right] &= \mu_t dt \\ Var_t\left[dX_t\right] &= \sigma_t \sigma_t' dt \end{aligned}$$

Definition 1.66 *For an n-dimensional Itô process X, we define*

$$\mathcal{L}(X) = \left\{\theta \in \mathcal{L}^n : \theta'\mu \in \mathcal{L}^1,\, \theta'\sigma \in \left(\mathcal{L}^2\right)^k\right\} \tag{1.13}$$

$$\mathcal{H}^2(X) = \left\{\theta \in \mathcal{L}(X) : E\left[\left(\int_0^T \theta_t'\mu_t' dt\right)^2\right] < \infty,\; \theta'\sigma \in \left(\mathcal{H}^2\right)^k\right\} \tag{1.14}$$

Definition 1.66 can be interpreted as follows. Let θ be a dynamic portfolio strategy and X the associated price process for the assets. Then $\int_0^T \theta_t' dX_t$ is the price process for the value of the portfolio gains. Classes $\mathcal{L}(X)$ and $\mathcal{H}^2(X)$ are therefore classes of portfolio value processes with desirable properties.

Observe that if $\theta \in \mathcal{H}^2(X)$, then the stochastic integral $\int_0^T \theta_t' dX_t$ has finite variance.

Theorem 1.67 (Itô's lemma) *Let X be a univariate Itô process*

$$dX_t = \mu_t dt + \sigma_t dW_t$$

where W is a univariate Brownian motion. Let $f : \mathbb{R}^2 \to \mathbb{R}$ with $f \in C^2(\mathcal{T} \times \mathbb{R})$ (twice continuously differentiable). Then $Y_t = g(t, X_t)$ is an Itô process and

$$dY_t = f_t(t, X_t)\, dt + f_x(t, X_t)\, \mu_t dt + \frac{1}{2} f_{xx}(t, X_t)\, \sigma_t^2 dt + f_x(t, X_t)\, \sigma_t dW_t.$$

Proof. The heuristic proof goes as follows. First consider a 'second-order' Taylor approximation

$$dY_t = f_t dt + f_x dX_t + \frac{1}{2} f_{tt}(dt)^2 + f_{tx} dt dX_t + \frac{1}{2} f_{xx}(dX_t)^2$$

Now observe

$$\begin{aligned} dt dX_t &= dt(\mu_t dt + \sigma_t dW_t) = \mu_t (dt)^2 + \sigma_t dt dW_t \\ (dX_t)^2 &= \mu_t^2 (dt)^2 + 2\mu_t \sigma_t dt dW_t + \sigma_t^2 (dW_t)^2 \end{aligned}$$

We already argued that $(dW_t)^2 = dt$ (a first-order term in t). However, the remaining terms are higher than first order. Since dW_t can be argued to have mean zero and variance dt, the term $dtdW_t$ will be mean zero and variance $(dt)^2$. which is a higher-order stochastic term than dW_t. Therefore, the only term left in the two expressions above is $\sigma_t^2 (dW_t)^2 = \sigma_t^2 dt$. ■

The formula can be extended to multivariate Brownian motions when we note that for two independent Brownian motions W^j and W^k, we have $\left(dW_t^j\right)\left(dW_t^k\right) = 0$.

Theorem 1.68 (Multivariate Itô's lemma) *Let W be a k-dimensiona Brownian motion, X an n-dimensional Itô process*

$$dX_t = \mu_t dt + \sigma_t dW_t$$

and $f : \mathcal{T} \times \mathbb{R}^n \to \mathbb{R}^m$ be from C^2. Then for $Y_t = f(X_t)$, we have for the k-th component Y_t^k

$$dY_t^k = f_t^k dt + f_x^k \mu_t dt + \frac{1}{2} \operatorname{tr}\left[\sigma_t \sigma_t' f_{xx}^k\right] dt + f_x^k \sigma_t dW_t.$$

Theorem 1.69 (Integration by parts) *Suppose the process $f_t(\omega)$ is continuous and of bounded variation with respect to $t \in [0, T]$ for almost all ω. Then*

$$\int_0^T f_t dW_t = f_T W_T - \int_0^T W_t df_t.$$

The assumption of bounded variation of f is crucial, the formula will, for instance, not work for $f_t = W_t$ (see Example 1.61 where we showed that $\int_0^T W_t dW_t = \frac{1}{2}W_t^2 - \frac{1}{2}t$).

Example 1.70 *Observe that we can get the result in Example 1.61 very simply now using Itô's lemma. Compute*

$$d\left(\frac{1}{2}W_t^2\right) = W_t dW_t + \frac{1}{2}dt$$

and integrating up, we get the result right away.

Example 1.71 *We want to compute* $\int_0^T t dW_t$*. It is reasonable to assume that a term like* tW_t *should show up. Therefore take*

$$d\left(tW_t\right) = W_t dt + t dW_t$$

and integrating up, we get

$$\int_0^T t dW_t = TW_T - \int_0^T W_t dt.$$

In many applications, we will want to study processes which are sums of an Itô process and a process A of bounded variation. Assume that we can write

$$dX_t = \mu_t dt + \sigma_t dW_t + dA_t$$

where A is a continuous process of bounded variation. Then

$$dY_t = df\left(t, X_t\right) = \left[f_t + f_x \mu_t + \frac{1}{2} f_{xx} \sigma_t^2\right] dt + f_x \sigma_t dW_t + f_x dA_t.$$

Other generalizations include A with infrequent jumps, such as the Poisson arrival process, including Poisson arrival with a random jump.

Proposition 1.72 *Let W be a Brownian motion with natural filtration $\{\mathcal{F}_t\}$ and $f \in \mathcal{H}^2$ be a process adapted to $\{\mathcal{F}_t\}$. Then*

$$X_t(\omega) = \int_0^t f_s(\omega)\, dW_s$$

is a martingale w.r.t. $\{\mathcal{F}_t\}$ and, for $\lambda, T > 0$,

$$P\left[\sup_{0 \le t \le T} |X_t| \ge \lambda\right] \le \frac{1}{\lambda^2} E\left[\int_0^T |f_s(\omega)|^2\, ds\right].$$

Proof. See discussion in class. ■

Remark 1.73 *If $f \in \mathcal{L}^2$, then X is only a local martingale.*

1.6 Martingale representation theorem

Theorem 1.74 (Itô representation theorem) (Øksendal (2007), Theorem 4.3.3) *Let* $\{\mathcal{F}_t\}$ *be a filtration generated by an k-dimensional Brownian motion and* $F \in L^2(\mathcal{F}_T, P)$ *(a square-integrable (under* P*) random variable measurable w.r.t.* $\mathcal{F}_T$*). Then there exists a unique stochastic process* $f \in \mathcal{H}^2$ *such that*

$$F(\omega) = E[F] + \int_0^T f_t(\omega)\, dW_t.$$

Proof. The proof amounts to showing that the class of processes $\int_0^t f_s(\omega)\, dW_s$ is dense in $L^2(\mathcal{F}_t, P)$. ■

Theorem 1.75 ***(Martingale representation theorem)*** (Øksendal (2007), Theorem 4.3.4) *Let W be a k-dimensional Brownian motion and X a martingale with respect to the natural filtration $\mathcal{F}_t$ of the Brownian motion. Also assume $X_t \in L^2(P)$ for all $t \geq 0$ (square integrability). Then there exists a unique process $g \in \left(\mathcal{H}^2\right)^k$ such that*

$$X_t(\omega) = X_0(\omega) + \int_0^t g_s(\omega)\, dW_s(\omega) \quad a.s., \text{ for all } t \geq 0. \tag{1.15}$$

Proof. See discussion in class. Notice that we cannot just apply Theorem 1.74 to each X_t because the functions f are in general t-specific. But we will show that this is not the case here. ■

Remark 1.76 *If X is only a local martingale, then there exists a process $g \in \left(\mathcal{L}^2\right)^k$ such that the above equation holds.*

Remark 1.77 *Since an Itô can be constructed to have continuous paths, it implies that all martingales adapted to a Brownian filtration have continuous sample paths. This implies that pure randomness (of the martingale type) cannot generate jumps — the Brownian model is a model of continuous information flow. Discontinuous, lumpy information arrivals like earnings announcements at particular dates would have to be modeled using processes other than diffusions.*

Also, we can interpret result (1.15) as stating that the Brownian motion, appropriately scaled over time, spans all martingales. Given the iid nature of Brownian increments, we can interpret g_s *at time* s *as a* time scale *over the next instant.*

1.7 Girsanov's theorem

The martingale representation theorem tells us that (local) martingales can be constructed by integrating up Brownian motions. We would now like to use positive integrals of such Brownian motions to construct changes of measure, in the sense of the Radon–Nikodým derivative. Exponential martingales are just the right thing to use.

Definition 1.78 *A process $\eta \in \left(\mathcal{L}^2\right)^k$ is said to satisfy the* ***Novikov condition*** *if*

$$E\left[\exp\left(\frac{1}{2}\int_0^T \eta_t^2 dt\right)\right] < \infty.$$

Theorem 1.79 *If the process $\eta \in \left(\mathcal{L}^2\right)^k$ satisfies the Novikov condition, then the process ξ^η defined by*

$$\xi_t^\eta = \exp\left(-\int_0^t \eta_s' dW_s - \frac{1}{2}\int_0^t |\eta_s|^2 \, ds\right)$$

is a martingale.

Proof. Using Itô's lemma, we have

$$d\xi_t^\eta = -\xi_t^\eta \eta_t' dW_t$$

which implies that ξ_t^η is a local martingale (this is because $\eta \in \left(\mathcal{L}^2\right)^k$, and a stochastic integral of such a η is only a local martingale in general, see the construction of Itô's integral). However, the Novikov condition assures that the local martingale is also a martingale. ■

This implies that processes from $\mathcal{L}^2$ that satisfy the Novikov condition can be used to construct changes of measure. We have $E\left[\xi_t^\eta\right] = 1$ and $\xi_0^\eta = 1$, so that, for every finite t, we can construct the Radon–Nikodým derivative according to Theorem 1.11:

$$\forall F \in \mathcal{F}_t : Q^\eta\left(F\right) = \int_F \xi_t^\eta dP$$

Given ξ_t^η is strictly positive, the two measures are equivalent.

Theorem 1.80 (Girsanov's theorem) *Consider a process* $\eta \in \left(\mathcal{L}^2\right)^k$ *such that* ξ^η *is a martingale. Then the process*

$$W_t^\eta = W_t + \int_0^t \eta_s ds$$

B.M under P

is a Brownian motion under Q^η.

Moreover, W^η *has a martingale representation property under* Q^η*: for any local* Q^η*-martingale* X_t *adapted to the filtration generated by* W*, there exists a process* $\phi \in \left(\mathcal{L}^2\right)^k$ *such that*

$$X_t = X_0 + \int_0^t \phi_s dW_s^\eta.$$

Remark 1.81 *Thus the drift term* η *from the exponential martingale* ξ^η *serves as a drift adjustment in the change of measure. Notice that* η *does not need to satisfy the Novikov condition — that one is only sufficient for* ξ^η *to be a martingale.*

Corollary 1.82 *Consider an Itô process X on $\mathbb{R}^n$*

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$$

where $\mu \in \left(\mathcal{L}^1\right)^n$ and $\sigma \in \left(\mathcal{L}^2\right)^{n\times k}$ such that

$$\sigma_t \eta_t = \mu_t - \nu_t.$$

If the process ξ^η is a martingale, then X is an Itô process under Q^η, and

$$X_t = X_0 + \int_0^t \nu_s ds + \int_0^t \sigma_s dW_s^\eta.$$

We can therefore go from P to Q^η through changes *only* in drifts not only in the case of diffusions but also in the case of more general Itô processes.

Moreover, the converse is also true. Under any equivalent probability measure w.r.t. which S is a martingale, the diffusion part stays the same, and only the drift changes.

Theorem 1.83 (Diffusion invariance principle) *Consider an Itô process X on $\mathbb{R}^n$*

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s.$$

If X is a martingale under an equivalent probability measure Q, then there exists a Brownian motion W^Q under Q such that

$$X_t = X_0 + \int_0^t \sigma_s dW_s^Q.$$

We can also always find η_t so that the density of $\frac{dQ}{dP}$ is given by ξ_t^η. Of course, we do not need to expect that such a η_t will satisfy the Novikov condition, which is only sufficient but not necessary for ξ_t^η to be a martingale.

A remark on infinite horizon models

The construction of Girsanov's theorem in infinite horizon requires special care, see Duffie (2001), Section 6.N, Huang and Pagès (1992) and Revuz and Yor (1999), Section VIII.1. The problem is that although over finite horizons, P and Q^η are equivalent, this will no longer be true over infinite horizons. However, one can construct a representation as above,

$$W_t^\eta = W_t + \int_0^t \eta_s ds$$

when we restrict the infinite horizon Brownian motions onto any measurable space $(\Omega, \mathcal{F}_t)$ for a given t. At the same time, under Q^η, W^η still has the martingale representation property, i.e. for any local martingale X_t there exists an adapted process ϕ such that $\int_0^t |\phi_s|^2 ds < \infty$ almost surely and such that

$$X_t = X_0 + \int_0^t \phi_s dW_s^\eta.$$

1.8 Stochastic differential equations

In many cases, we will consider models where economic variables of interest are not general Itô processes but solutions to stochastic differential equations. In the one-dimensional case, we can formulate the following definition.

Definition 1.84 *The Itô process X_t satisfies a stochastic differential equation (SDE)*

$$dX_t = \mu\left(t, X_t\right) dt + \sigma\left(t, X_t\right) dW_t$$

with an initial condition X_0 if it satisfies

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

Observe that we have imposed a strong structure. The coefficients μ and σ are not just arbitrary adapted processes (in $\mathcal{L}^1$ and $\mathcal{L}^2$, respectively) but we explicitly model a feedback structure.

Finding explicit solutions to these SDEs requires luck or experience — only very few solutions are actually known. However, we can think about schemes for numerical solutions. However, in order to do so, we would like to know something about existence and uniqueness of the solutions. Many such theorems exist. The following is from Øksendal (2007).

Theorem 1.85 ***(Øksendal (2007), Theorem 5.2.1)*** *Let there be constants $C, D > 0$ such that*

$$|\mu(t,x)| + |\sigma(t,x)| \leq X(1 + |x|)$$

and

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x - y|$$

for any x, y (Lipschitz property). Then, the SDE

$$dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t, \quad X_0 = x_0$$

has a unique continuous solution X_t, which is adapted to the filtration generated by the Brownian motion, and the solution $X \in \mathcal{H}^2$.

Remark 1.86 *The solution constructed on the given filtration generated by the Brownian motion is called a* ***strong solution****. A* ***weak solution****, on the other hand consists of* ***a*** *probability space (potentially different from the one above), and a process X that solves the SDE on the given probability space.*

Example 1.87 *The solution of the SDE*

$$dX_t = \mu dt + \sigma dW_t$$

is

$$X_t = X_0 + \mu t + \sigma W_t$$

which is the arithmetic Brownian motion.

Example 1.88 *Consider*

$$dX_t = \mu X_t dt + \sigma X_t dW_t$$

which has the solution

$$X_t = \exp\left(\left(\mu - \sigma^2/2\right)t + \sigma W_t\right)$$

which is a process called the geometric Brownian motion.

Example 1.89 *Consider the process on* $[0, 1]$

$$\xi_t = \exp\left(-\frac{1}{2}\int_0^t (1-s)^{-3}\,ds - \int_0^t (1-s)^{-3/2}\,dW_s\right), \quad 0 \leq t < 1,\ \xi_0 = 0.$$

Applying the Itô lemma, we get

$$\xi_t = 1 + \int_0^t -(1-s)^{-3/2}\,\xi_s dW_s$$

which can be shown to converge to zero almost surely as ξ_t*. Then, defining* $\xi_1 = 0$ *almost surely, we have a continuous process that is a local martingale (because it is expressed as a stochastic integral) but not a martingale.*

For many economic applications, Theorem 1.85 is too stringent. This is for instance the case for recursive utility models with Duffie and Epstein (1992a,b) where existence has to be proven using other, rather ad-hoc, methods.

1.9 Feynman-Kac formula

The Feynman-Kac formula establishes a link between a class of partial differential equations and stochastic processes driven by Brownian motions. This allows to solve PDEs using simulations of stochastic processes, or, vice versa, solve stochastic differential equations using PDEs.

Consider the PDE

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

with terminal condition $g(x,T) = G(x)$. The Feynman-Kac formula states that the solution $g(x,t)$ can be written as a conditional expectation

$$g(x,t) = E_{x,t}\left[\int_t^T \phi(t,s)\, h(X_s,s)\, ds + \phi(t,T)\, G(X_T)\right]$$

where

$$\phi(t,s) = \exp\left(-\int_t^s r(X_\tau,\tau)\, d\tau\right)$$

for a stochastic process X that satisfies, under the probability measure under which the expectation is taken,

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

and $X_t = x$ is the initial condition.

1.10 Conditional distributions and moments

1.11 Stationary densities

Let us assume we have a Markov Itô process given by

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

We will also assume that the process satisfies certain regularity conditions, see Karlin and Taylor (1981) for more details. We are interested in computing the stationary density for X.

Start by computing the infinitesimal generator for X. The generator $\mathcal{A}$

applied to a twice continuously differentiable function f is defined as

$$\begin{aligned} \mathcal{A}f\left(X_t\right) &= \lim_{\varepsilon\to 0}\frac{E_t\left[f\left(X_{t+\varepsilon}\right)\right]-f\left(X_t\right)}{\varepsilon} = E_t\left[\frac{df\left(X_t\right)}{dt}\right] = \\ &= \frac{1}{dt}\left[E_t\left(\frac{\partial f}{\partial x}\left(X_t\right)dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x\partial x'}\left(dX\right)^2\right)\right] = \\ &= \mu\left(X_t\right)f'\left(X_t\right)+\frac{1}{2}\sigma^2\left(X_t\right)f''\left(X_t\right) \end{aligned}$$

Under stationary distribution with a continuous density $q\left(x\right)dx = dQ\left(x\right)$, we have must have

$$\begin{aligned} \int\mathcal{A}f\left(x\right)dQ\left(x\right) &= \int\left[\mu\left(x\right)f'\left(x\right)+\frac{1}{2}\sigma^2\left(x\right)f''\left(x\right)\right]q\left(x\right)dx = \\ &= \int\mu\left(x\right)f'\left(x\right)q\left(x\right)dx + \int\frac{1}{2}\sigma^2\left(x\right)f''\left(x\right)q\left(x\right)dx = \\ &= \int\mu\left(x\right)f'\left(x\right)q\left(x\right)dx + \left[f'\left(x\right)\frac{1}{2}\sigma^2\left(x\right)q\left(x\right)\right]_{x_{\inf}}^{x_{\sup}} - \\ &\quad -\int\frac{1}{2}\left(\sigma^2\left(x\right)q\left(x\right)\right)'f'\left(x\right)dx \end{aligned}$$

Since this holds for any $f(\cdot)$, take such an f for which

$$f'(x_{\text{sup}}) = f'(x_{\text{inf}}) = 0$$

and thus we get

$$\int \mu(x) f'(x) q(x) \, dx - \int \frac{1}{2} \left(\sigma^2(x) q(x)\right)' f'(x) \, dx = 0$$

Since again, this holds for any $f(\cdot)$, the equality must be pointwise (w.p. 1) and thus

$$\begin{aligned} \mu(x) q(x) - \frac{1}{2} \left(\sigma^2(x) q(x)\right)' &= 0 \\ \frac{2\mu(x) - \sigma^{2\prime}(x)}{\sigma^2(x)} &= \frac{q'(x)}{q(x)} \end{aligned}$$

Integrating this ordinary differential equation, we obtain the stationary density $q(x)$.

Example 1.90 *Consider the process*

$$dX_t = (a + bX_t)\, dt + \sqrt{e} X_t dW_t = \mu(X_t)\, dt + \sigma(X_t)\, dW_t$$

Using

$$\begin{aligned} \mu(x) &= a + bx \\ \sigma^2(x) &= ex^2 \end{aligned}$$

we get

$$\frac{q'(x)}{q(x)} = \frac{2a}{e}\frac{1}{x^2} + \frac{2(b-e)}{e}\frac{1}{x}$$

and thus

$$\begin{aligned} \log q(x) &= \log c_0 - \frac{2a}{e}\frac{1}{x} + \frac{2(b-e)}{e} \log x \\ q(x) &= c_0 x^{\frac{2(b-e)}{e}} \exp\left(-\frac{2a}{e}\frac{1}{x}\right) \end{aligned}$$

where we recognize the density for the inverse gamma distribution

$$q\left(x;\alpha,\beta\right)=\frac{\beta^{\alpha}}{\Gamma\left(\alpha\right)}x^{-\alpha-1}\exp\left(-\frac{\beta}{x}\right)$$

with

$$\begin{aligned}\alpha &= \frac{e-2b}{e}\\ \beta &= \frac{2a}{e}\end{aligned}$$

The conditions are

$$\alpha,\beta>0$$

which translates (together with taking into account $\sqrt{e}$*) to*

$$\begin{aligned}a &> 0\\ e &> 0\\ e &> 2b\end{aligned}$$

1.12 Boundary classification

A more refined analysis, see Karlin and Taylor (1981).

Bibliography

Billingsley, Patrick. 1979. *Probability and Measure.* John Wiley & Sons, Inc.

Dixit, Avinash K. and Robert S. Pindyck. 1994. *Investment under Uncertainty.* Princeton University Press, Princeton, NJ.

Duffie, Darrell. 2001. *Dynamic Asset Pricing Theory.* Princeton University Press, 3rd ed.

Duffie, Darrell and Larry G. Epstein. 1992a. "Asset Pricing with Stochastic Differential Utility." *Review of Financial Studies* 5 (3):411–436.

———. 1992b. "Stochastic Differential Utility." *Econometrica* 60 (2):353–394.

Huang, Chi-Fu and Henri Pagès. 1992. "Optimal Consumption and Portfolio Policies with an Infinite Horizon: Existence and Convergence." *The Annals of Applied Probability* 2 (1):36–64.

Karatzas, Ioannis and Steven E. Shreve. 1991. *Brownian Motion and Stochastic Calculus*. Springer, 2nd ed.

Karlin, Samuel and Howard M. Taylor. 1981. *A Second Course in Stochastic Processes*. Academic Press.

Øksendal, Bernt. 2007. *Stochastic Differential Equations*. Springer Verlag, 6th ed.

Revuz, Daniel and Marc Yor. 1999. *Continuous Martingales and Brownian Motion*. Springer-Verlag, 3rd ed.

Stokey, Nancy L. 2008. *The Economics of Inaction.* Princeton University Press.