Stochastic Financial Models

J.R. Norris

December 4, 2019

Contents

1 Utility and mean-variance analysis

1.1 Contingent claims and utility functions

We can model the opportunities available to an investor, and subject to chance, by random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In this interpretation, we often refer to random variables as *contingent claims*. A contingent claim X delivers for consumption an amount $X(\omega)$ which is determined by chance $\omega \in \Omega$. The probability measure $\mathbb{P}$ is thought of as encoding the beliefs of the investor about the avaliable contingent claims. By a *utility function* we mean a non-decreasing function

$$U : \mathbb{R} \to [-\infty, \infty).$$

We think of $U(x)$ as quantifying the satisfaction obtained by the investor on consuming an amount x. For a contingent claim, we will sometimes consider its expected utility. Whenever we mention the expected utility $\mathbb{E}(U(X))$ of a contingent claim X, it is to be understood that X has the property that $\mathbb{E}(U(X)^+) < \infty$, so that $\mathbb{E}(U(X))$ is well defined and $\mathbb{E}(U(X)) < \infty$. We will often assume that the investor acts to maximize expected utility. Thus

$$Y \text{ is preferred to } X \qquad \text{if and only if} \qquad \mathbb{E}(U(X)) \leqslant \mathbb{E}(U(Y)).$$

Note that the case of equality is included. If $\mathbb{E}(U(X)) = \mathbb{E}(U(Y))$, then we say that the investor is *indifferent* between X and Y. We say that the investor is *risk neutral* if, for any integrable random variable X, the investor is indifferent between $\mathbb{E}(X)$ and X. We say that the investor is *risk averse* if, for any integrable random variable X, the investor prefers $\mathbb{E}(X)$ to X.

Recall that U is *concave* if, for all $x, y \in \mathbb{R}$ and all $p \in (0, 1)$,

$$(1-p)U(x) + pU(y) \leqslant U((1-p)x + py).$$

If this relation holds with $<$ in place of $\leqslant$ whenever $x \neq y$, then U is said to be *strictly concave*.

Proposition 1.1. *An investor with utility function U is risk averse if and only if U is concave.*

Proof. Suppose that the investor is risk averse. Consider a random variable X taking the value x with probability $1-p$ and y with probability p. Then the investor prefers $\mathbb{E}(X)$ to X, that is to say,

$$(1-p)U(x) + pU(y) = \mathbb{E}(U(X)) \leqslant U(\mathbb{E}(X)) = U((1-p)x + py).$$

Hence U is concave.

On the other hand, suppose that U is concave. Let X be an integrable random variable with mean $\mathbb{E}(X)$. Then, by Jensen's inequality, $\mathbb{E}(U(X)^+) < \infty$ and

$$\mathbb{E}(U(X)) \leqslant U(\mathbb{E}(X))$$

so $\mathbb{E}(X)$ is preferred to X. Hence the investor is risk averse. $\square$

For $\gamma \in (0, \infty)$, the *CARA utility function of parameter* γ is given by

$$U(x) = \mathrm{CARA}_\gamma(x) = -\exp(-\gamma x).$$

For $R \in (0, 1) \cup (1, \infty)$, the *CRRA utility function of parameter* R is given by

$$U(x) = \mathrm{CRRA}_R(x) = \begin{cases} \dfrac{x^{1-R}}{1-R}, & \text{if } x > 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

The *CRRA utility function of parameter* 1 is given by

$$U(x) = \mathrm{CRRA}_1(x) = \begin{cases} \log x, & \text{if } x > 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

Here CARA stands for *constant absolute risk aversion* and CRRA stands for *constant relative risk aversion.* These names are explained by the following calculations, which we do not attempt to make rigorous.

Consider an investor with utility function U and a small contingent claim X. For a given constant w, we seek to determine whether the investor prefers w or $w + X$. Since X is small, by Taylor's theorem,

$$U(w + X) \approx U(w) + U'(w)X + \frac{1}{2}U''(w)X^2$$

so

$$\mathbb{E}(U(w + X)) \approx U(w) + U'(w)\mathbb{E}(X) + \frac{1}{2}U''(w)\mathbb{E}(X^2)$$

and so $w + X$ is preferred to w if and only if

$$\frac{2\mathbb{E}(X)}{\mathbb{E}(X^2)} \geqslant -\frac{U''(w)}{U'(w)}.$$

The right-hand side $-U''(w)/U'(w)$ is called the *Arrow–Pratt coefficient of absolute risk aversion.* Note that this coefficient takes the constant value γ when $U = \mathrm{CARA}_\gamma$.

By the same argument, the investor prefers $w(1 + X)$ to w if and only if

$$\frac{2\mathbb{E}(X)}{\mathbb{E}(X^2)} \geqslant -\frac{wU''(w)}{U'(w)}.$$

The new right-hand side $-wU''(w)/U'(w)$ is called the *Arrow–Pratt coefficient of relative risk aversion.* This takes the constant value R when $U = \mathrm{CRRA}_R$.

1.2 Reservation prices and marginal prices

Consider an investor with concave utility function U. Fix a set $\mathcal{A}$ of *available claims* such that $\mathbb{E}(U(X)^+) > \infty$ for all $X \in \mathcal{A}$. Suppose that the investor is able to choose any contingent claim in $\mathcal{A}$, and that $\mathbb{E}(U(X))$ is maximized over $\mathcal{A}$ at X^*. Let Y be another contingent claim and let $\pi \in \mathbb{R}$. The investor will wish to buy the claim Y for price π if, for some $X \in \mathcal{A}$,

$$\mathbb{E}(U(X + Y - \pi)) > \mathbb{E}(U(X^*)).$$

The *reservation bid price* $\pi_b(Y)$ for Y is defined to be the supremum of such prices π. On the other hand, the investor will wish to sell the claim Y for price π if, for some $X \in \mathcal{A}$,

$$\mathbb{E}(U(X - Y + \pi)) > \mathbb{E}(U(X^*)).$$

The *reservation ask price* $\pi_a(Y)$ for Y is defined to be the infimum of such prices π.

Proposition 1.2 (Ask above, bid below)**.** *Assume that the set of available claims $\mathcal{A}$ is convex. Then, for any contingent claim Y,*

$$\pi_b(Y) \leqslant \pi_a(Y).$$

Proof. It will suffice to show that there is no price π at which the investor will both buy and sell. To show this, suppose for a contradiction that there exist $X_a, X_b \in \mathcal{A}$ such that

$$\mathbb{E}(U(X_a - Y + \pi)) > \mathbb{E}(U(X^*)), \quad \mathbb{E}(U(X_b + Y - \pi)) > \mathbb{E}(U(X^*)).$$

Since $\mathcal{A}$ is convex, $X = (X_a + X_b)/2 \in \mathcal{A}$. Since U is concave,

$$\frac{U(X_a - Y + \pi) + U(X_b + Y - \pi)}{2} \leqslant U(X).$$

On taking expectations, we obtain the following contradiction

$$\mathbb{E}(U(X^*)) < \frac{\mathbb{E}(U(X_a - Y + \pi)) + \mathbb{E}(U(X_b + Y - \pi))}{2} \leqslant \mathbb{E}(U(X)) \leqslant \mathbb{E}(U(X^*)).$$

□

Assume now $\mathcal{A}$ is an affine space and that U is differentiable and strictly concave. Then the minimizing contingent claim $X^* \in \mathcal{A}$ is almost surely unique. We define the *marginal price* $\pi_m(Y)$ of a contingent claim Y by

$$\pi_m(Y) = \frac{\mathbb{E}(U'(X^*)Y)}{\mathbb{E}(U'(X^*))}.$$

The following calculation, which we do not make rigorous, explains this definition.

First note that, for any contingent claim $\xi \in \mathcal{A} - \mathcal{A}$, the map $t \mapsto \mathbb{E}(U(X^* + t\xi))$ achieves it minimum at X^*. Hence

$$0 = \left.\frac{d}{dt}\right|_{t=0} \mathbb{E}(U(X^* + t\xi)) = \mathbb{E}(U'(X^*)\xi).$$

It is plausible that there is a differentiable path $t \mapsto X^*(t)$ in $\mathcal{A}$ such that

$$\mathbb{E}(U(X^*(t) + tY - \pi_b(tY))) = \mathbb{E}(U(X^*)).$$

Then $X^*(0) = X^*$. Set

$$\xi = \left.\frac{d}{dt}\right|_{t=0} X^*(t), \quad \pi = \left.\frac{d}{dt}\right|_{t=0} \pi_b(tY).$$

It is plausible that $\xi \in \mathcal{A} - \mathcal{A}$ and that we can differentiate in t to obtain

$$0 = \left.\frac{d}{dt}\right|_{t=0} \mathbb{E}(U(X^*(t) + tY - \pi_b(tY))) = \mathbb{E}(U'(X^*)(\xi + Y - \pi)) = \mathbb{E}(U'(X^*)(Y - \pi)).$$

The same calculation can be done for ask prices. Hence we obtain

$$\pi_m(Y) = \left.\frac{d}{dt}\right|_{t=0} \pi_a(tY) = \left.\frac{d}{dt}\right|_{t=0} \pi_b(tY).$$

The marginal price is thus the unique price above which the investor will sell, and below which the investor will buy, a small multiple of the claim Y.

1.3 Single-period asset price model

By a *single-period asset price model* we mean a pair of random variables (S_0, S_1) in $\mathbb{R}^d$. For $n = 0, 1$ and $i = 1, \dots, d$, we interpret S_n^i as the price of (a unit of) asset i at time n. We often add to the model a *numéraire*, that is to say a pair of random variables (S_0^0, S_1^0) in $(0, \infty)$, interpreted as a further asset whose price is always positive. Write $\bar{S}_n = (S_n^0, S_n)$. We call $(\bar{S}_0, \bar{S}_1)$ a *single-period asset price model with numéraire*. We will assume that S_0 and S_0^0 are non-random unless otherwise stated. We usually consider the special case where the numéraire is a *riskless bond*, that is, when $S_0^0 = 1$ and $S_1^0 = 1 + r$ for some constant $r > -1$. Then r is called the *interest rate*. It may be that there is a random variable $\rho \geqslant 0$ with the property that $S_0^i = \mathbb{E}(S_1^i \rho)$ for all i. In this case, we call ρ a *state-price density*.

We consider the problem of an investor with wealth w_0 at time 0. In the case without numéraire, for each $\theta \in \mathbb{R}^d$, the investor is able to buy θ^i units of asset i at time 0, subject to the constraint $\theta.S_0 = w_0$. The components θ^i are allowed to be negative. Having chosen such a *portfolio* θ, the investor will then have wealth $\theta.S_1$ at time 1.

In the case with numéraire, the portfolio is described by a vector $\bar{\theta} = (\theta^0, \theta) \in \mathbb{R}^{d+1}$, which is chosen subject to the constraint $\bar{\theta}.\bar{S}_0 = w_0$. The investor's wealth at time 1 is then given by $\bar{\theta}.\bar{S}_1$.

1.4 Portfolio selection using the mean-variance criterion

Consider first a single-period asset price model (S_0, S_1). Assume that S_0 is non-random and that S_1 has mean μ and variance V, with S_0, μ linearly independent and V invertible. Let w_0 and w_1 be given. We seek a portfolio $\theta \in \mathbb{R}^d$ to satisfy the following mean-variance criterion

$$\begin{array}{ll} \text{minimize} & \operatorname{var}(\theta.S_1) \\ \text{subject to} & \theta.S_0 = w_0, \quad \mathbb{E}(\theta.S_1) = w_1. \end{array}$$

We have

$$\mathbb{E}(\theta.S_1) = \theta.\mu, \quad \operatorname{var}(\theta.S_1) = \theta.(V\theta)$$

so our problem is to

$$\begin{array}{ll} \text{minimize} & \theta.(V\theta) \\ \text{subject to} & \theta.S_0 = w_0, \quad \theta.\mu = w_1. \end{array}$$

We use the method of Lagrange multipliers. Fix $\lambda_0, \lambda_1 \in \mathbb{R}$ and minimize

$$L(\theta, \lambda) = \frac{1}{2}\theta.(V\theta) - \lambda_0\theta.S_0 - \lambda_1\theta.\mu.$$

The minimizing θ satisfies

$$0 = \frac{\partial}{\partial\theta_i}L(\theta, \lambda) = (V\theta)^i - \lambda_0 S_0^i - \lambda_1\mu^i$$

so

$$\theta = A(\lambda_0 S_0 + \lambda_1\mu), \quad A = V^{-1}.$$

We then choose λ_0, λ_1 to satisfy the constraints

$$w_0 = \theta.S_0 = a\lambda_0 + b\lambda_1, \quad w_1 = \theta.\mu = b\lambda_0 + c\lambda_1$$

where

$$a = S_0.(AS_0), \quad b = S_0.(A\mu) = \mu.(AS_0), \quad c = \mu.(A\mu).$$

Set

$$\Delta = \det\begin{pmatrix} a & b \\ b & c \end{pmatrix} = ac - b^2.$$

Then the minimizing portfolio is given by

$$\theta^* = \theta^*(w_1) = \frac{cw_0 - bw_1}{\Delta}AS_0 + \frac{aw_1 - bw_0}{\Delta}A\mu$$

with variance

$$\operatorname{var}(\theta^*.S_1) = \frac{aw_1^2 - 2bw_0w_1 + cw_0^2}{ac - b^2}.$$

The set

$$\{\theta^*(w_1) : w_1 \in \mathbb{R}\}$$

is called the *mean-variance-efficient frontier.* Note that the minimal variance is quadratic in the target wealth w_1. This is further minimized over w_1 at $w_1 = (b/a)w_0$, with minimal variance w_0^2/a. We will denote this *minimum variance portfolio* by

$$\theta^*_{\min} = \frac{w_0}{a} AS_0.$$

Now let us add to the model a riskless bond of interest rate r. This allows the investor to keep some wealth in the bond, investing less in the market, or to borrow using the bond and invest more in the market. We again seek the portfolio which minimizes the variance of wealth at time 1, subject to the constraints of given wealth w_0 at time 0 and expected wealth w_1 at time 1. The problem is now to find $\bar{\theta} = (\theta^0, \theta) \in \mathbb{R}^{d+1}$ to

$$\begin{array}{ll} \text{minimize} & \theta.(V\theta) \\ \text{subject to} & \theta^0 + \theta.S_0 = w_0, \quad \theta^0(1+r) + \theta.\mu = w_1. \end{array}$$

On eliminating θ^0 using the first constraint, we are left with the single constraint

$$\theta.(\mu - (1+r)S_0) = w_1 - (1+r)w_0.$$

Hence, by Lagrange multipliers, or otherwise, the minimizing portfolio $\bar{\theta}^*$ is given by

$$\theta^* = \lambda\theta^*_{\mathrm{m}}, \quad \theta^*_{\mathrm{m}} = A(\mu - (1+r)S_0), \quad A = V^{-1}$$

with λ and θ^{*0} determined by the constraints. In particular, we find that

$$\lambda = \frac{w_1 - (1+r)w_0}{(1+r)^2 a - 2(1+r)b + c}.$$

The portfolio θ^*_{m} is called the *market portfolio.*

1.5 Portfolio selection using a CARA utility function

We use the same asset price models as in the preceding section, but now make the additional assumption that S_1 is Gaussian. Fix $\gamma \in (0, \infty)$ and consider the CARA utility function

$$U(x) = \mathrm{CARA}_\gamma(x) = -\exp\{-\gamma x\}.$$

We discuss first the case without riskless asset. An investor with wealth w_0 at time 0 may choose from the set of contingent claims

$$\mathcal{A} = \{\theta.S_1 : \theta \in \mathbb{R}^d, \theta.S_0 = w_0\}.$$

Thus, to maximize expected utility, we seek a portfolio $\theta \in \mathbb{R}^d$ to

$$\begin{array}{ll} \text{maximize} & \mathbb{E}(U(\theta.S_1)) \\ \text{subject to} & \theta.S_0 = w_0. \end{array}$$

Since S_1 is Gaussian, we have

$$\mathbb{E}(U(\theta.S_1)) = -\exp\{-\gamma\theta.\mu + \gamma^2\theta.(V\theta)/2\}$$

so our problem is to

$$\begin{array}{ll} \text{maximize} & \theta.\mu - \frac{1}{2}\gamma\theta.(V\theta) \\ \text{subject to} & \theta.S_0 = w_0. \end{array}$$

We use the method of Lagrange multipliers. Fix $\lambda \in \mathbb{R}$ and maximize

$$L(\theta, \lambda) = \theta.\mu - \frac{1}{2}\gamma\theta.(V\theta) - \lambda\theta.S_0.$$

The maximizing portfolio θ^* satisfies

$$0 = \frac{\partial}{\partial\theta_i}L(\theta, \lambda) = \mu_i - \gamma(V\theta)^i - \lambda S_0^i$$

so

$$\theta^* = \gamma^{-1}A(\mu - \lambda S_0), \quad A = V^{-1}.$$

On choosing λ to fit the constraints, we obtain

$$\theta^* = \theta^*_{\min} + (A\mu - (b/a)AS_0)/\gamma$$

where a and b are as in the preceding subsection.

We turn to the case with riskless asset $S_0^0 = 1$ and $S_1^0 = 1 + r$. The investor now seeks a portfolio $\bar{\theta} \in \mathbb{R}^d$ to

$$\begin{array}{ll} \text{maximize} & \mathbb{E}(U(\bar{\theta}.\bar{S}_1)) \\ \text{subject to} & \bar{\theta}.\bar{S}_0 = w_0. \end{array}$$

Now

$$\bar{\theta}.\bar{S}_0 = \theta^0 + \theta.S_0$$

and

$$\mathbb{E}(U(\bar{\theta}.\bar{S}_1)) = -\exp\{-\gamma(\theta^0(1+r) + \theta.\mu) + \gamma^2\theta.(V\theta)/2\}$$

so our problem is to

$$\begin{array}{ll} \text{maximize} & \theta^0(1+r) + \theta.\mu - \frac{1}{2}\theta.(V\theta) \\ \text{subject to} & \theta^0 + \theta.S_0 = w_0. \end{array}$$

We use the constraint to eliminate θ^0 in the objective function and then differentiate to see that the maximizing portfolio $\bar{\theta}^*$ is given by

$$\theta^* = \gamma^{-1}\theta^*_{\mathrm{m}}, \quad \theta^*_{\mathrm{m}} = A(\mu - (1+r)S_0), \quad A = V^{-1}, \quad \theta^{*0} = w_0 - \theta^*.S_0.$$

1.6 Capital-asset pricing model

We continue our study of the single-period asset price model with riskless bond. Define

$$\beta^i = \frac{\operatorname{cov}(S_1^i, \theta_{\mathrm{m}}^*.S_1)}{\operatorname{var}(\theta_{\mathrm{m}}^*.S_1)}, \quad \mu^{\mathrm{m}} = \theta_{\mathrm{m}}^*.\mu, \quad S_0^{\mathrm{m}} = \theta_{\mathrm{m}}.S_0.$$

We call β^i the *beta* or *sensitivity* of asset i.

Proposition 1.3. *For $i = 1, \ldots, d$, we have*

$$\mu^i - (1+r)S_0^i = \beta^i(\mu^{\mathrm{m}} - (1+r)S_0^{\mathrm{m}}).$$

Proof. For $\theta = \theta_{\mathrm{m}} = A(\mu - (1+r)S_0)$, we have

$$\mu^{\mathrm{m}} - (1+r)S_0^{\mathrm{m}} = \theta.(\mu - (1+r)S_0) = \theta.(V\theta) = \operatorname{var}(\theta.S_1)$$

so

$$\mu^i - (1+r)S_0^i = e_i.(V\theta) = \operatorname{cov}(S_1^i, \theta.S_1) = \beta^i \operatorname{var}(\theta.S_1) = \beta^i(\mu^{\mathrm{m}} - (1+r)S_0^{\mathrm{m}}).$$

□

We might suppose, given the appearance of the market portfolio θ_{m}^* as optimal, that the aggregate of all portfolios held by all investors, namely the capitalization-weights of the relevant market index, offers an observable version of θ_{m}^*. Write

$$S_1^i = (1+R^i)S_0^i, \quad S_1^{\mathrm{m}} = (1+R^{\mathrm{m}})S_0^{\mathrm{m}}.$$

We can write the formula just shown in terms of the returns R^i and R^{m} as

$$\mathbb{E}(R^i) = r + \tilde{\beta}^i(\mathbb{E}(R^{\mathrm{m}}) - r), \quad \tilde{\beta}^i = \frac{\operatorname{cov}(R^i, R^{\mathrm{m}})}{\operatorname{var}(R^{\mathrm{m}})}.$$

The sensitivities $\tilde{\beta}^i$ could be estimated from historical data, which would give a pricing formula for the expected returns $\mathbb{E}(R^i)$. If we could also estimate $\mathbb{E}(R^i)$, this might indicate where the market was undervaluing or overvaluing an asset. However, the precise estimation of expected returns is generally agreed to be infeasible.

2 Martingales

2.1 Conditional probabilities and expectations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Recall that, given an event B of positive probability, we define the conditional probability $\mathbb{P}(.|B)$ by

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad A \in \mathcal{F}.$$

The associated expectation is written $\mathbb{E}(.|B)$. If a random variable X is integrable (with respect to $\mathbb{P}$), then X is also integrable with respect to $\mathbb{P}(.|B)$ and the following formula holds

$$\mathbb{E}(X|B) = \frac{\mathbb{E}(X 1_A)}{\mathbb{P}(B)}.$$

In this course we will use also some further notions of conditional probability and conditional expectation. These are associated not to an event B but to a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$. In general, these notions have to be approached indirectly, by the following theorem.

Theorem 2.1. *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let X be an integrable random variable. Then there exists an integrable random variable Y with the following properties:*

(a) *Y is $\mathcal{G}$-measurable,*

(b) $\mathbb{E}(Y 1_A) = \mathbb{E}(X 1_A)$ *for all* $A \in \mathcal{G}$.

Moreover, if Y' is another integrable random variable satisfying (a) *and* (b), *then $Y' = Y$ almost surely. Moreover, the same statements are valid if we replace 'integrable' by 'non-negative', meaning $[0, \infty]$-valued, throughout.*

We call any such random variable Y (*a version of*) *the conditional expectation of X given $\mathcal{G}$* and we write

$$Y = \mathbb{E}(X|\mathcal{G}) \quad \text{almost surely.}$$

In the case where $\mathcal{G} = \sigma(Z)$ for some random variable Z, we also write $Y = \mathbb{E}(X|Z)$ almost surely. In the case where $X = 1_A$ for some event A, we also write $Y = \mathbb{P}(A|\mathcal{G})$ almost surely.

The following argument proves the uniqueness statement in Theorem 2.1, and also establishes a useful monotonicity property. Let X' be another integrable random variable such that $X \leqslant X'$ almost surely. Suppose that Y' is an integrable random variable which satisfies (a) and (b) with respect to X'. Set $A = \{Y \geqslant Y'\}$ and consider the non-negative random variable $Z = (Y - Y')1_A$. Then, since $A \in \mathcal{G}$,

$$\mathbb{E}(Y 1_A) = \mathbb{E}(X 1_A) \leqslant \mathbb{E}(X' 1_A) = \mathbb{E}(Y' 1_A)$$

so $\mathbb{E}(Z) = \mathbb{E}(Y 1_A) - \mathbb{E}(Y' 1_A) \leqslant 0$ and so $Z = 0$ almost surely, which implies that $Y \leqslant Y'$ almost surely. In the case $X = X'$, we deduce by symmetry that $Y = Y'$ almost surely, as claimed.

We will omit proof of the existence statement in Theorem 2.1. A proof may be found in lecture notes for Advanced Probability. However, in the case where $\mathcal{G}$ is discrete, we can construct a suitable random variable Y directly, as we now show. Consider the case where

$$\mathcal{G} = \sigma(B_n : n \in \mathbb{N}) = \left\{ \bigcup_{n \in I} B_n : I \subseteq \mathbb{N} \right\}$$

for some sequence $(B_n : n \in \mathbb{N})$ of disjoint events whose union is Ω. Given an integrable random variable X, set

$$Y = \sum_{n \in \mathbb{N}} \mathbb{E}(X|B_n) 1_{B_n}$$

where we make the convention that $\mathbb{E}(X|B_n) = 0$ whenever $\mathbb{P}(B_n) = 0$. Then Y is $\mathcal{G}$-measurable and, by monotone convergence,

$$\mathbb{E}(|Y|) = \mathbb{E}\left(\sum_{n \in \mathbb{N}} |\mathbb{E}(X|B_n)| 1_{B_n} \right) = \sum_{n \in \mathbb{N}} \mathbb{E}(|\mathbb{E}(X|B_n)| 1_{B_n}) \leqslant \sum_{n \in \mathbb{N}} \mathbb{E}(|X| 1_{B_n}) = \mathbb{E}(|X|) < \infty.$$

so Y is integrable. Moreover, for $A = \cup_{n \in I} B_n \in \mathcal{G}$, by dominated convergence,

$$\mathbb{E}(Y 1_A) = \mathbb{E}\left(Y \sum_{n \in I} 1_{B_n} \right) = \sum_{n \in I} \mathbb{E}(Y 1_{B_n}) = \sum_{n \in I} \mathbb{E}(X|B_n) \mathbb{P}(B_n) = \sum_{n \in I} \mathbb{E}\left(X 1_{B_n} \right) = \mathbb{E}(X 1_A)$$

where we used $|Y|$ and $|X|$ as dominating random variables for the second and last equalities respectively. Hence Y has both properties (a) and (b) of Theorem 2.1, and we have shown that

$$\mathbb{E}(X|\mathcal{G}) = \sum_{n \in \mathbb{N}} \mathbb{E}(X|B_n) 1_{B_n} \quad \text{almost surely.}$$

We now examine some general properties of conditional expectation.

Proposition 2.2. *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let X and W be integrable random variables. Then*

(i) $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$,

(ii) *if X is $\mathcal{G}$-measurable, then $\mathbb{E}(X|\mathcal{G}) = X$ almost surely,*

(iii) *if X is independent of $\mathcal{G}$, then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$ almost surely,*

(iv) *if $X \geqslant 0$ almost surely, then $\mathbb{E}(X|\mathcal{G}) \geqslant 0$ almost surely,*

(v) $\mathbb{E}(\alpha X + \beta W|\mathcal{G}) = \alpha \mathbb{E}(X|\mathcal{G}) + \beta \mathbb{E}(W|\mathcal{G})$ *almost surely, for all $\alpha, \beta \in \mathbb{R}$.*

Proof. Properties (i) to (iii) follow easily from the defining properties (a) and (b), while (iv) follows from the monotonicity property shown above (take $X = 0$ and $X' = X$). To show (v), set $Y = \alpha \mathbb{E}(X|\mathcal{G}) + \beta \mathbb{E}(W|\mathcal{G})$. Then Y is $\mathcal{G}$-measurable and, for all $A \in \mathcal{G}$,

$$\mathbb{E}(Y 1_A) = \alpha \mathbb{E}(\mathbb{E}(X|\mathcal{G}) 1_A) + \beta \mathbb{E}(\mathbb{E}(W|\mathcal{G}) 1_A) = \alpha \mathbb{E}(X 1_A) + \beta \mathbb{E}(W 1_A) = \mathbb{E}((\alpha X + \beta W) 1_A).$$

Hence $Y = \mathbb{E}(\alpha X + \beta W|\mathcal{G})$ almost surely, as claimed. $\square$

Proposition 2.3 (Tower property)**.** *Let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$ with $\mathcal{H} \subseteq \mathcal{G}$. Let X be an integrable random variable. Then*

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}) = \mathbb{E}(X|\mathcal{H}) \quad \textit{almost surely.}$$

Proof. Choose a version Y of $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H})$. Then Y is integrable and $\mathcal{H}$-measurable. Moreover, for all $A \in \mathcal{H}$,

$$\mathbb{E}(Y1_A) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})1_A) = \mathbb{E}(X1_A)$$

where the second equality holds because $A \in \mathcal{G}$. Hence $Y = \mathbb{E}(X|\mathcal{H})$ almost surely. $\square$

Proposition 2.4 (Taking out what is known)**.** *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Let X be an integrable random variable and let Z be a $\mathcal{G}$-measurable random variable such that ZX is integrable. Then*

$$\mathbb{E}(ZX|\mathcal{G}) = Z\mathbb{E}(X|\mathcal{G}) \quad \textit{almost surely.}$$

Proof (not examinable in this course). Assume for now that $X \geqslant 0$. Choose a version $Y \geqslant 0$ of $\mathbb{E}(X|\mathcal{G})$. Consider first the case when $Z = 1_B$ for some $B \in \mathcal{G}$. Then ZY is $\mathcal{G}$-measurable and, for all $A \in \mathcal{G}$,

$$\mathbb{E}(ZY1_A) = \mathbb{E}(Y1_{A\cap B}) = \mathbb{E}(X1_{A\cap B}) = \mathbb{E}(ZX1_A).$$

This extends to all simple $\mathcal{G}$-measurable random variables by linearity. Suppose now that Z is any non-negative $\mathcal{G}$-measurable random variable. Define, for $n \geqslant 1$,

$$Z_n = \left(2^{-n}\lfloor 2^n Z\rfloor\right) \wedge n.$$

Then Z_n is simple and $\mathcal{G}$-measurable for all n so, for all $A \in \mathcal{G}$,

$$\mathbb{E}(Z_nY1_A) = \mathbb{E}(Z_nX1_A).$$

Also $0 \leqslant Z_n \uparrow Z$ as $n \to \infty$. Since $X \geqslant 0$ and $Y \geqslant 0$, we can pass to the limit $n \to \infty$ using monotone convergence to deduce that, for all $A \in \mathcal{G}$

$$\mathbb{E}(ZY1_A) = \mathbb{E}(ZX1_A).$$

In particular, by taking $A = \Omega$, we see that if ZX is integrable then so is ZY. For a general $\mathcal{G}$-measurable Z with ZX integrable, we can apply the preceding to $Z^{\pm} = (\pm Z) \vee 0$ to see that ZY is integrable and, by subtraction, for all $A \in \mathcal{G}$,

$$\mathbb{E}(ZY1_A) = \mathbb{E}(ZX1_A).$$

Hence $ZY = \mathbb{E}(ZX|\mathcal{G})$ almost surely, as claimed. $\square$

Proposition 2.5 (Averaging over independent variables)**.** *Let X_1, X_2 be random variables taking values in the measurable spaces $(E_1, \mathcal{E}_1), (E_2, \mathcal{E}_2)$ respectively. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and suppose that X_1 is $\mathcal{G}$-measurable, while X_2 is independent of $\mathcal{G}$. Let F be a non-negative measurable function on $E_1 \times E_2$. Then we can define a non-negative measurable function f on E_1 by $f(x) = \mathbb{E}(F(x, X_2))$ and we have*

$$\mathbb{E}(F(X_1, X_2)|\mathcal{G}) = f(X_1), \quad \textit{almost surely.}$$

Proof (not examinable in this course). Define

$$\mathcal{A} = \{B_1 \times B_2 : B_1 \in \mathcal{E}_1 \text{ and } B_2 \in \mathcal{E}_2\}.$$

Then $\mathcal{A}$ is a π-system (see Probability and Measure) and $\mathcal{A}$ generates the product σ-algebra $\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2$. It is straightforward to check the desired conclusion when $F = 1_B$ with $B \in \mathcal{A}$. Denote by $\mathcal{D}$ the set of elements of $\mathcal{E}$ such that the conclusion holds for $F = 1_B$. Then $\mathcal{A} \subseteq \mathcal{D}$ and it can be checked that $\mathcal{D}$ is a d-system. Hence $\mathcal{D}$ is the whole of $\mathcal{E}$. The conclusion extends to the case where F is a simple function by linearity, and then to the case where F is any non-negative measurable function by monotone convergence. □

2.2 Definitions

Suppose given a measurable space $(\Omega, \mathcal{F})$. A *filtration* is a family of σ-algebras $(\mathcal{F}_n)_{n\geqslant 0}$ on Ω such that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \mathcal{F}$ for all n. A *random process* is a family of random variables $(X_n)_{n\geqslant 0}$ on $(\Omega, \mathcal{F})$. We say that $(X_n)_{n\geqslant 0}$ is *adapted* (to $(\mathcal{F}_n)_{n\geqslant 0}$) if X_n is $\mathcal{F}_n$-measurable for all n. Define

$$\mathcal{F}_n^X = \sigma(X_k : 0 \leqslant k \leqslant n).$$

We call $(\mathcal{F}_n^X)_{n\geqslant 0}$ the *natural filtration* of $(X_n)_{n\geqslant 0}$. We think of a filtration as representing the emergence of information over time. In particular $\mathcal{F}_n^X$ contains all the events determined by the process $(X_n)_{n\geqslant 0}$ up to time n.

Suppose now that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, which is equipped with a filtration $(\mathcal{F}_n)_{n\geqslant 0}$. We say that a random process $(X_n)_{n\geqslant 0}$ is a *martingale* if, for all n,

(a) X_n is $\mathcal{F}_n$-measurable,

(b) $\mathbb{E}(|X_n|) < \infty$,

(c) $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$ almost surely.

When (b) holds, we say that $(X_n)_{n\geqslant 0}$ is *integrable*. Condition (c) is known as the *martingale property*. Thus, in short, a martingale is an adapted integrable process which satisfies the martingale property.

If conditions (a) and (b) hold and, for all n,

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leqslant X_n \quad \text{almost surely}$$

then we say that $(X_n)_{n \geqslant 0}$ is a *supermartingale*. Similarly, if conditions (a) and (b) hold and, for all n,

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geqslant X_n \quad \text{almost surely}$$

then we say that $(X_n)_{n \geqslant 0}$ is a *submartingale*.

Any martingale is also a martingale in its natural filtration. When we refer to a martingale without specifying a filtration, the natural filtration is understood by default.

All the above notions have continuous-time analogues, which are mostly obtained simply by writing t for n, with the convention that $t \in [0, \infty)$ where we had $n \in \mathbb{Z}^+$. Given a filtration $(\mathcal{F}_t)_{t \geqslant 0}$, we say that $(X_t)_{t \geqslant 0}$ is a *martingale* if it is adapted, integrable and satisfies, for all $s, t \geqslant 0$ with $s \leqslant t$,

$$\mathbb{E}(X_t|\mathcal{F}_s) = X_s \quad \text{almost surely.}$$

The supermartingale and submartingale properties are modified similarly. We say that $(X_t)_{t \geqslant 0}$ is *continuous* if the map

$$t \mapsto X_t(\omega) : [0, \infty) \to \mathbb{R}$$

is continuous for all ω. These notions will be relevant when we come to look at Brownian motion, which is itself a continuous martingale.

2.3 Examples

Let $(X_n)_{n \geqslant 1}$ be a sequence of independent identically distributed random variables. Set $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and for $n \geqslant 1$ set $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Set $S_0 = 0$ and $Z_0 = 1$ and define for $n \geqslant 1$

$$S_n = \sum_{k=1}^n X_k, \quad Z_n = \prod_{k=1}^n X_k.$$

In the case where X_1 is integrable with $\mathbb{E}(X_1) = 0$, the process $(S_n)_{n \geqslant 0}$ is a martingale. On the other hand, in the case where $X_1 \geqslant 0$ with $\mathbb{E}(X_1) = 1$, the process $(Z_n)_{n \geqslant 0}$ is a martingale. We call $(S_n)_{n \geqslant 0}$ an *additive martingale* and $(Z_n)_{n \geqslant 0}$ a *multiplicative martingale*.

Further examples arise in connection with Markov chains. Let $(X_n)_{n \geqslant 0}$ be a Markov chain with countable state-space S and transition matrix P. Set $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$. Given a bounded or non-negative function f on S, we set

$$Pf(x) = \sum_{y \in S} p_{xy} f(y).$$

Then

$$\mathbb{E}(f(X_{n+1})|\mathcal{F}_n) = Pf(X_n) \quad \text{almost surely.}$$

We say that f is *subharmonic* if $f(x) \leqslant Pf(x)$ for all x.

Suppose that f is subharmonic and define $M_n = f(X_n)$. Then

$$\mathbb{E}(M_{n+1}|\mathcal{F}_n) = Pf(X_n) \geqslant f(X_n) = M_n \quad \text{almost surely}$$

so $(M_n)_{n \geqslant 0}$ has the submartingale property.

2.4 Optional stopping

We say that a random time $T : \Omega \to \{0, 1, \dots\} \cup \{\infty\}$ is a *stopping time* if, for all n,

$$\{T \leqslant n\} \in \mathcal{F}_n.$$

Theorem 2.6 (Optional stopping)**.** *Let $(M_n)_{n\geqslant 0}$ be a martingale and let T be a bounded stopping time. Then $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.*

Proof. We can choose n so that $T \leqslant n$. Then

$$M_T = (M_T - M_{T-1}) + \cdots + (M_1 - M_0) + M_0 = M_0 + \sum_{k=1}^{n} (M_k - M_{k-1}) 1_{\{k \leqslant T\}}.$$

Since T is a stopping time, $\{k \leqslant T\} = \{T \leqslant k-1\}^c \in \mathcal{F}_{k-1}$ for all k. Then, since $(M_n)_{n\geqslant 0}$ is a martingale,

$$\mathbb{E}(M_k 1_{\{k\leqslant T\}}) = \mathbb{E}(M_{k-1} 1_{\{k\leqslant T\}}).$$

Hence

$$\mathbb{E}(M_T) = \mathbb{E}(M_0) + \sum_{k=1}^{n} \mathbb{E}((M_k - M_{k-1}) 1_{\{k\leqslant T\}}) = \mathbb{E}(M_0).$$

□

This is also called *Doob's optional sampling theorem*. We have stated the basic form of the result. We now show an extension to unbounded stopping times under additional conditions.

Theorem 2.7. *Let $(M_n)_{n\geqslant 0}$ be a martingale and let T be an almost surely finite stopping time. Suppose that, for some constant $C < \infty$, one of the following two conditions holds:*

(a) $|M_n| \leqslant C$ *for all* $n \leqslant T$,

(b) $\mathbb{E}(T) < \infty$ *and* $|M_n - M_{n-1}| \leqslant C$ *for all* $n \in \{1, \dots, T\}$.

Then $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.

Proof. Since T is a stopping time, $T \wedge n$ is a bounded stopping time for all n so, by the optional stopping theorem,

$$\mathbb{E}(M_{T\wedge n}) = \mathbb{E}(M_0).$$

Hence, it will suffice to show that $\mathbb{E}(M_{T\wedge n}) \to \mathbb{E}(M_T)$ as $n \to \infty$. Since T is almost surely finite, we have $M_{T\wedge n} \to M_T$ almost surely as $n \to \infty$. Condition (a) implies that $|M_{T\wedge n}| \leqslant C$ for all n, so the desired limit holds by bounded convergence. On the other hand, condition (b) implies that $|M_{T\wedge n}| \leqslant |M_0| + CT$ for all n, so the desired limit holds by dominated convergence. □

It is instructive to consider two simple examples where the optional stopping theorem does not apply. Consider first the additive martingale $S_n = \sum_{k=1}^n X_k$ associated to a sequence of independent identically distributed random variables $(X_n)_{n\geqslant 1}$, with

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2.$$

Set

$$T = \min\{n \geqslant 0 : S_n = 1\}.$$

Then $(S_n)_{n\geqslant 0}$ is a simple symmetric random walk, which is recurrent, so $T < \infty$ almost surely. But

$$\mathbb{E}(S_T) = 1 \neq 0 = S_0$$

so the conclusion of the optional stopping theorem is false.

Consider now the multiplicative martingale $Z_n = \prod_{k=1}^n X_k$ associated to a sequence of independent identically distributed random variables $(X_n)_{n\geqslant 1}$, with

$$\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 2) = 1/2.$$

Set

$$T = \min\{n \geqslant 0 : Z_n = 0\}.$$

Then $\mathbb{E}(T) = 2$. But

$$\mathbb{E}(Z_T) = 0 \neq 1 = Z_0$$

so the conclusion of the optional stopping theorem is again false. In this and the preceding example, one of the conditions in hypothesis (b) above holds, but not the other.

Here is a further variant of the optional stopping theorem.

Theorem 2.8. *Let $(M_n)_{n\geqslant 0}$ be a martingale and let T be a stopping time. Then the stopped process $(M_{T\wedge n})_{n\geqslant 0}$ is also a martingale.*

Proof. Note that

$$M_{T\wedge n} = M_0 + \sum_{k=1}^{n} (M_k - M_{k-1}) 1_{\{T\geqslant k\}}$$

and, since T is a stopping time, $\{T \geqslant k\} = \{T \leqslant k-1\}^c \in \mathcal{F}_{k-1}$. Since $(M_n)_{n\geqslant 0}$ is adapted and integrable, this makes it clear that $M_{T\wedge n}$ is $\mathcal{F}_n$-measurable and integrable for all n. Now

$$M_{T\wedge(n+1)} - M_{T\wedge n} = (M_{n+1} - M_n) 1_{\{T\geqslant n+1\}}$$

so, by taking out what is known,

$$\mathbb{E}(M_{T\wedge(n+1)} - M_{T\wedge n} | \mathcal{F}_n) = 1_{\{T\geqslant n+1\}} \mathbb{E}(M_{n+1} - M_n | \mathcal{F}_n) = 0.$$

Hence $(M_{T\wedge n})_{n\geqslant 0}$ has the martingale property. □

We say that a random process $(H_n)_{n\geqslant 1}$ is *previsible* if H_n is $\mathcal{F}_{n-1}$-measurable for all n.

Theorem 2.9 (Martingale transform). *Let $(M_n)_{n\geqslant 0}$ be a martingale and let $(H_n)_{n\geqslant 1}$ be a bounded previsible process. Define a new random process $(Y_n)_{n\geqslant 0}$ by*

$$Y_0 = 0, \quad Y_n = \sum_{k=1}^{n} H_k(M_k - M_{k-1}), \quad n \geqslant 1.$$

Then $(Y_n)_{n\geqslant 0}$ is also a martingale.

Proof. Since $(M_n)_{n\geqslant 0}$ is adapted and integrable, it is clear from the definition that $(Y_n)_{n\geqslant 0}$ is also adapted and integrable. Now

$$Y_{n+1} - Y_n = H_{n+1}(M_{n+1} - M_n)$$

and H_{n+1} is $\mathcal{F}_n$-measurable so, by taking out what is known,

$$\mathbb{E}(Y_{n+1} - Y_n|\mathcal{F}_n) = H_{n+1}\mathbb{E}(M_{n+1} - M_n|cF_n) = 0.$$

Hence $(Y_n)_{n\geqslant 0}$ also has the martingale property. □

The process $(Y_n)_{n\geqslant 0}$ is called the *martingale transform* of $(M_n)_{n\geqslant 0}$ by $(H_n)_{n\geqslant 1}$. The proof follows the same argument as the preceding result, noting that it remains valid on substituting $1_{\{T\geqslant k\}}$ by H_k.

The optional stopping theorem and its variants may be interpreted in terms of an investor who is able to buy and sell a risky asset. We model the price of the asset by the martingale $(M_n)_{n\geqslant 0}$. Suppose that the investor holds one unit of the asset at time 0 and wishes to sell it by time n. The investor is free to choose when to sell, based on the information available at the present time, that is, at any stopping time T of the filtration $(\mathcal{F}_n)_{n\geqslant 0}$. Then the optional stopping theorem shows that the expected return $\mathbb{E}(M_T)$ is the same no matter what stopping time is chosen. More generally, we may suppose that the investor is free to buy or sell any number of units in each time period, again based on information available at the start of the period. The amount H_k held between times $k-1$ and k will then be $\mathcal{F}_{k-1}$-measurable, so defining a previsible process, and the total gain by time n is then given by Y_n. Theorem 2.9 shows that no bounded previsible strategy results in an expected gain or loss.

3 Pricing contingent claims

3.1 Multi-period asset price model

Consider a discrete-time random process $(\bar{S}_n)_{0\leqslant n\leqslant T}$ in $\mathbb{R}^{d+1}$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and adapted to a filtration $(\mathcal{F}_n)_{0\leqslant n\leqslant T}$. Write

$$\bar{S}_n = (S_n^0, S_n^1, \dots, S_n^d) = (S_n^0, S_n).$$

We assume that $(S_n^0)_{0\leqslant n\leqslant T}$ is a *numéraire*, that is to say, $S_n^0 > 0$ for all n. Then we call $(\bar{S}_n)_{0\leqslant n\leqslant T}$ an *asset price model.*

We use this set-up to model the evolution over time of $d+1$ assets, interpreting S_n^i as the price at time n of the ith asset. There are T time periods: 0 to 1, 1 to 2 and finally $T-1$ to T. Thus $S_n^i - S_{n-1}^i$ is the change in price of the ith asset over the nth time period.

We will be interested not so much in the absolute prices S_n^i, but in the *discounted prices* X_n^i, given by

$$X_n^i = S_n^i / S_n^0$$

and we write

$$\bar{X}_n = (X_n^0, X_n^1, \dots, X_n^d) = (1, X_n).$$

Often, $(S_n^0)_{0\leqslant n\leqslant T}$ will be interpreted as a bond or bank account. In this case, we will write

$$S_n^0 = (1 + r_n) S_{n-1}^0$$

and call r_n the *interest rate.*

We often take $(\mathcal{F}_n)_{0\leqslant n\leqslant T}$ to be the filtration generated by $(\bar{S}_n)_{0\leqslant n\leqslant T}$ and take $\mathcal{F} = \mathcal{F}_T$. In that case, if say $(S_n^0)_{0\leqslant n\leqslant T}$ is a given deterministic process, and the prices of the other *risky assets* take only a countable set of values, then we will be able to specify $\mathbb{P}$ completely by specifying the probabilities $\mathbb{P}(S_0 = s_0, S_1 = s_1, \dots, S_T = s_T)$ for all possible sequences $(s_n)_{0\leqslant n\leqslant T}$ in $\mathbb{R}^d$.

Let $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ be a random process in $\mathbb{R}^{d+1}$. Write

$$\bar{\theta}_n = (\theta_n^0, \theta_n^1, \dots, \theta_n^d) = (\theta_n^0, \theta_n).$$

We use such a process to model a portfolio held by an investor, where θ_n^i is the number of units of asset i held in the nth time period, allowing the possibility that this number may take any real value. We say that a portfolio is *self-financing* if

$$\bar{\theta}_n . \bar{S}_n = \bar{\theta}_{n+1} . \bar{S}_n \quad \text{for } n = 1, \dots, T-1.$$

This expresses that no funds are added or withdrawn from the portfolio at time n, it is simply rebalanced between the $d+1$ assets. We associate to a self-financing portfolio its *value process* $(V_n)_{0\leqslant n\leqslant T}$ given by

$$V_0 = \bar{\theta}_1 . \bar{X}_0, \quad V_n = \bar{\theta}_n . \bar{X}_n \quad \text{for } n = 1, \dots, T.$$

Thus V_n is the total (discounted) value of the portfolio at time n.

We say that a portfolio is *previsible* if $\bar{\theta}_n$ is $\mathcal{F}_{n-1}$-measurable for all n. This is a natural condition, since it expresses that $\bar{\theta}_n$ depends only on what is known at the start of the nth period, that is, at time $n-1$.

Proposition 3.1. *Let $(\theta_n)_{1\leqslant n\leqslant T}$ be a previsible process in $\mathbb{R}^d$ and let $V_0 \in \mathbb{R}$. Then there is a unique previsible process $(\theta_n^0)_{1\leqslant n\leqslant T}$ in $\mathbb{R}$ such that $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ is a self-financing portfolio with initial value V_0. Moreover, the value process of this portfolio is given by*

$$V_T = V_0 + \sum_{n=1}^{T} \theta_n.(X_n - X_{n-1}).$$

Proof. The equations $\bar{\theta}_1.\bar{X}_0 = V_0$ and $\bar{\theta}_n.\bar{S}_n = \bar{\theta}_{n+1}.\bar{S}_n$ for $n = 1, \dots, T-1$, which express that the portfolio $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ has initial value V_0 and is self-financing, may be written in the form

$$\theta_1^0 + \theta_1.X_0 = V_0, \quad \theta_n^0 S_n^0 + \theta_n.S_n = \theta_{n+1}^0 S_n^0 + \theta_{n+1}.S_n, \quad \text{for } n = 1, \dots, T-1$$

and then solved uniquely to obtain a previsible process $(\theta_n^0)_{1\leqslant n\leqslant T}$. Then, since $X_n^0 = 1$ for all n, the value of the resulting portfolio changes by $\theta_n.(X_n - X_{n-1})$ over the nth time period, giving the claimed formula for V_T by induction. □

3.2 Examples of contingent claims

By a *contingent claim* of *maturity* T we mean any non-negative $\mathcal{F}_T$-measurable random variable. We think of the contingent claim C as a contract which pays the investor the amount C at time T. The fact that the contract can be exercised only at the given time T can be emphasised by referring to it as a *European option*.

Here are some simple examples in the case $d = 1$ of one risky asset:

(a) $(S_T - K)^+$ is the *call* with *strike price* K, which can be understood as the right but not the obligation to buy one unit of the asset at time T for price K,

(b) $(S_T - K)^-$ is the *put* with *strike price* K, which can be understood as the right but not the obligation to sell one unit of the asset at time T for price K.

Calls and puts are also called *options*, emphasising the option to buy or sell. The following are examples of *exotic options*, depending on the entire path $(S_n)_{0\leqslant n\leqslant T}$. They are both examples of *barrier options* which are *knocked out* or *knocked in* when the price hits or passes a given barrier B.

(c) The *up-and-out call* is a call which is knocked out at level B, given by

$$C = \begin{cases} (S_T - K)^+, & \text{if } \max_{0\leqslant n\leqslant T} S_n < B, \\ 0, & \text{otherwise.} \end{cases}$$

(d) The *down-and-in put* is a put which is knocked in at level B, given by

$$C = \begin{cases} (S_T - K)^-, & \text{if } \min_{0 \leqslant n \leqslant T} S_n \leqslant B, \\ 0, & \text{otherwise.} \end{cases}$$

3.3 Equivalent probability measures

We say that a probability measure $\tilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ is *equivalent* to $\mathbb{P}$, and we write $\tilde{\mathbb{P}} \sim \mathbb{P}$, if there exists a non-negative random variable ρ such that

(a) $\mathbb{P}(\rho > 0) = 1$,

(b) $\tilde{\mathbb{P}}(A) = \mathbb{E}(\rho 1_A)$ for all $A \in \mathcal{F}$.

Note that $\tilde{\mathbb{P}}(A) = 0$ whenever $\mathbb{P}(A) = 0$. We leave as an exercise to check that any two non-negative random variables satisfying (b) are equal $\mathbb{P}$-almost surely. We call ρ a *density* for $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$ and write

$$d\tilde{\mathbb{P}}/d\mathbb{P} = \rho \quad \text{almost surely.}$$

By a standard argument of measure theory, the expectation $\tilde{\mathbb{E}}$ associated with $\tilde{\mathbb{P}}$ is given by

$$\tilde{\mathbb{E}}(X) = \mathbb{E}(\rho X)$$

for all non-negative random variables X. We leave as an exercise to check that the relation $\tilde{\mathbb{P}} \sim \mathbb{P}$ is symmetric and transitive, with

$$d\mathbb{P}/d\tilde{\mathbb{P}} = 1/\rho \quad \text{almost surely.}$$

3.4 Arbitrage

By an *arbitrage* for $(\bar{S}_n)_{0 \leqslant n \leqslant T}$ we mean a previsible self-financing portfolio $(\bar{\theta}_n)_{1 \leqslant n \leqslant T}$ with initial value $V_0 = 0$ such that

$$V_T \geqslant 0 \quad \text{almost surely}$$

and

$$V_T > 0 \quad \text{with positive probability.}$$

An arbitrage thus delivers a positive return with no downside risk. The value process $(V_n)_{0 \leqslant n \leqslant T}$ depends on $(\bar{S}_n)_{0 \leqslant n \leqslant T}$ only through $(X_n)_{0 \leqslant n \leqslant T}$, and it depends on $(\bar{\theta}_n)_{1 \leqslant n \leqslant T}$ only through V_0 and $(\theta_n)_{1 \leqslant n \leqslant T}$. We will therefore allow ourselves to refer to $(\theta_n)_{1 \leqslant n \leqslant T}$ as an arbitrage for $(X_n)_{0 \leqslant n \leqslant T}$ when V_T satisfies the given conditions.

Note that the notion of arbitrage depends only on the equivalence class of the probability measure $\mathbb{P}$.

If there is no arbitrage for $(\bar{S}_n)_{0 \leqslant n \leqslant T}$, then we say that $(\bar{S}_n)_{0 \leqslant n \leqslant T}$ is *arbitrage free*. This is considered a reasonable assumption, on the grounds that the other investors in the market would not collectively give away such a return, even if they held different views about the evolution of the asset prices.

The following proposition shows a generic way in which an arbitrage-free asset price model can arise.

Proposition 3.2. *Let $(X_n)_{0 \leqslant n \leqslant T}$ be a martingale in $\mathbb{R}^d$. Then $(X_n)_{0 \leqslant n \leqslant T}$ is arbitrage free.*

Proof. Let $(\theta_n)_{1 \leqslant n \leqslant T}$ be a previsible process in $\mathbb{R}^d$ and set

$$V_n = \sum_{k=1}^{n} \theta_k.(X_k - X_{k-1}).$$

Suppose that $V_T \geqslant 0$ almost surely. We have $\mathbb{E}(V_T|\mathcal{F}_T) = V_T$ almost surely. Suppose in a backwards induction for $n \leqslant T$ that $V_n = \mathbb{E}(V_T|\mathcal{F}_n)$ almost surely. We have

$$V_n = V_{n-1} + \theta_n.(X_n - X_{n-1}).$$

Fix $R < \infty$ and set $A = \{|\theta_n| \leqslant R \text{ and } |V_{n-1}| \leqslant R\}$. Then $A \in \mathcal{F}_{n-1}$ and

$$1_A V_n = 1_A V_{n-1} + 1_A \theta_n.(X_n - X_{n-1}).$$

Since $V_n = \mathbb{E}(V_T|\mathcal{F}_n) \geqslant 0$ almost surely and $1_A|V_{n-1}| \leqslant R$ and $1_A|\theta_n| \leqslant R$, we can take the conditional expectation on $\mathcal{F}_{n-1}$, taking out what is known, to obtain

$$1_A \mathbb{E}(V_n|\mathcal{F}_{n-1}) = 1_A V_{n-1} + 1_A \theta_n \mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}) = 1_A V_{n-1} \quad \text{almost surely.}$$

But R was arbitrary and $|\theta_n|$ and $|V_{n-1}|$ are almost surely finite, so we must have

$$V_{n-1} = \mathbb{E}(V_n|\mathcal{F}_{n-1}) = \mathbb{E}(V_T|\mathcal{F}_{n-1}) \quad \text{almost surely}$$

and the induction proceeds. In particular, we see that $\mathbb{E}(V_T|\mathcal{F}_0) = V_0 = 0$ almost surely, so $\mathbb{E}(V_T) = 0$ and so $V_T = 0$ almost surely. Hence there is no arbitrage. □

A simpler version of the preceding argument shows that there is no bounded arbitrage. For, if $(\theta_n)_{1 \leqslant n \leqslant T}$ is a bounded previsible process, then the martingale transform

$$V_T = \sum_{n=1}^{T} \theta_n.(X_n - X_{n-1})$$

is integrable and $\mathbb{E}(V_T) = 0$. Hence, if $V_T \geqslant 0$ almost surely, then $V_T = 0$ almost surely, so $(\theta_n)_{1 \leqslant n \leqslant T}$ is not an arbitrage. A more elaborate argument was used in the proof to show that even unbounded processes cannot give an arbitrage.

3.5 Characterization of a single-period model with no arbitrage

We explore the implications of the no arbitrage assumption in the special case $T = 1$. Thus we consider a pair of random variables $(\bar{S}_0, \bar{S}_1)$ in $\mathbb{R}^{d+1}$. We make the usual assumption that $S_0^0 > 0$ and $S_0^1 > 0$. A previsible self-financing portfolio with $V_0 = 0$ is then specified by the choice of a single $\mathcal{F}_0$-measurable random variable $\theta = \theta_1$ in $\mathbb{R}^d$, and then $V_1 = \theta.Y$, where $Y = X_1 - X_0$ and $X_n = S_n/S_n^0$ for $n = 0, 1$. Then θ is an arbitrage if $\theta.Y \geqslant 0$ almost surely and $\theta.Y > 0$ with positive probability. When $\mathcal{F}_0 = \{\emptyset, \Omega\}$, the following result characterizes the arbitrage free case.

Proposition 3.3. *Let Y be a random variable in $\mathbb{R}^d$. Then the following are equivalent:*

(a) *there exists no $\theta \in \mathbb{R}^d$ such that $\theta.Y \geqslant 0$ almost surely and $\theta.Y > 0$ with positive probability,*

(b) *there is an equivalent probability measure $\tilde{\mathbb{P}}$ for which Y is integrable with $\tilde{\mathbb{E}}(Y) = 0$.*

Proof. Suppose that (a) holds. Then, for all $\theta \in \mathbb{R}^d$, if $\theta.Y \neq 0$ with positive probability, then $\theta.Y > 0$ with positive probability.

It will suffice to consider the case where all exponential moments of Y are finite. In general, we can replace $\mathbb{P}$ by the equivalent probability measure $\tilde{\mathbb{P}}$ given by $d\tilde{\mathbb{P}}/d\mathbb{P} \propto e^{-|Y|^2}$, for which the exponential moments of Y are all finite. We will assume that this has been done and drop the tildes.

Define $\phi : \mathbb{R}^d \to (0, \infty)$ by $\phi(\theta) = \mathbb{E}(e^{\theta.Y})$. Then ϕ is differentiable and

$$\phi'(\theta) = \mathbb{E}(Y e^{\theta.Y}).$$

We will show that ϕ *achieves a minimum value at some* $\theta^* \in \mathbb{R}^d$. We can define an equivalent probability measure $\tilde{\mathbb{P}}$ by $d\tilde{\mathbb{P}}/d\mathbb{P} = e^{\theta^*.Y}/Z$, where $Z = \phi(\theta^*)$. Then

$$\tilde{\mathbb{E}}(|Y|) = \mathbb{E}(|Y|e^{\theta^*.Y})/Z < \infty, \quad \tilde{\mathbb{E}}(Y) = \mathbb{E}(Y e^{\theta^*.Y})/Z = \phi'(\theta^*)/Z = 0.$$

Hence (b) holds.

Write $\mathbb{R}^d$ as an orthogonal direct sum $E_0 \oplus E_1$, where

$$E_0 = \{\theta \in \mathbb{R}^d : \theta.Y = 0 \text{ almost surely}\}.$$

Then $\phi(\theta_0 + \theta_1) = \phi(\theta_1)$ for all $\theta_0 \in E_0$ and $\theta_1 \in E_1$. Hence it will suffice to show that ϕ achieves a minimum on E_1. Since $\phi(0) = 1$, it will then suffice to show that $\phi(\theta) \geqslant 1$ for all $\theta \in E_1$ with $|\theta|$ sufficiently large.

Set $\psi(t) = 0 \vee t \wedge 1$ and define $f(\theta) = \mathbb{E}(\psi(\theta.Y))$. Then ψ is Lipschitz on $\mathbb{R}$ of constant 1 and f is continuous on $\mathbb{R}^d$ by bounded convergence. Define $S = \{\theta \in E_1 : |\theta| = 1\}$. For $\theta \in S$, we have $\mathbb{P}(\theta.Y \neq 0) > 0$ so $\mathbb{P}(\theta.Y > 0) > 0$ and so $f(\theta) > 0$. Set

$$\varepsilon = \frac{1}{2} \inf\{f(\theta) : \theta \in S\}.$$

Then $\varepsilon > 0$ because S is compact. Then, for all $\theta \in S$,

$$\mathbb{P}(\theta.Y \geqslant \varepsilon) = \mathbb{P}(\theta.Y - \varepsilon \geqslant 0) \geqslant \mathbb{E}(\psi(\theta.Y - \varepsilon)) \geqslant \mathbb{E}(\psi(\theta.Y)) - \varepsilon = f(\theta) - \varepsilon \geqslant \varepsilon$$

Hence we obtain $\phi(t\theta) \geqslant \varepsilon e^{t\varepsilon} \geqslant 1$ for all $t \geqslant (1/\varepsilon)\log(1/\varepsilon)$, as required.

Suppose on the other hand that (b) holds and let $\theta \in \mathbb{R}^d$. Then

$$\tilde{\mathbb{E}}(\theta.Y) = \theta.\tilde{\mathbb{E}}(Y) = 0$$

so, if $\theta.Y \geqslant 0$ almost surely, then $\theta.Y = 0$ almost surely. Hence (a) holds. $\square$

3.6 Fundamental theorem of asset pricing

Let $(\bar{S}_n)_{0\leqslant n\leqslant T}$ be a standard asset price model in $\mathbb{R}^{d+1}$. Recall that $X_n = (X_n^1, \ldots, X_n^d)$ where $X_n^i = S_n^i/S_n^0$ for all i. We say that a probability measure $\tilde{\mathbb{P}}$ on $\mathcal{F}$ is an *equivalent martingale measure* if $\tilde{\mathbb{P}} \sim \mathbb{P}$ and $(X_n)_{0\leqslant n\leqslant T}$ is a martingale under $\tilde{\mathbb{P}}$. The term *risk-neutral measure* is also used in this case.

Theorem 3.4. *The following are equivalent:*

(a) $(\bar{S}_n)_{0\leqslant n\leqslant T}$ *has no arbitrage,*

(b) $(\bar{S}_n)_{0\leqslant n\leqslant T}$ *has an equivalent martingale measure.*

The fact that (b) impies (a) follows from Proposition 3.2. A proof of the reverse implication given in the Appendix.

3.7 Completeness

Let C be a time-T contingent claim. Write $D = C/S_T^0$ for its discounted value. We say that C is *attainable* or *replicable* if there is a previsible self-financing portfolio $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ such that $C = \bar{\theta}_T.\bar{S}_T$. Equivalently, C (or we sometimes say D) is attainable if there exists an $\mathcal{F}_0$-measurable random variable V_0 and a previsible process $(\theta_n)_{1\leqslant n\leqslant T}$ in $\mathbb{R}^d$ such that

$$D = V_0 + \sum_{n=1}^{T} \theta_n.(X_n - X_{n-1}).$$

Then V_0 is called the *fair price* for C and $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ is called a *replicating portfolio*. We also call $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ a *hedging portfolio*: it allows the seller of the contingent claim C to deliver it at time T by investing V_0 suitably in the available assets at time 0, and shifting the portfolio among the assets according to $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ until time T. Note that anyone who sells the claim C at time 0 for a different price to V_0 allows other investors to make risk-free profits by replicating C or $-C$ in the market. If all contingent claims are attainable, then we say that the asset price model $(\bar{S}_n)_{0\leqslant n\leqslant T}$ is *complete*. We will identify a class of complete models in the next section.

Proposition 3.5. *Assume that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \sigma(\bar{S}_1, \dots, \bar{S}_T)$.*

(a) *Let C be a non-negative attainable time-T contingent claim and suppose that $\tilde{\mathbb{P}}$ is an equivalent martingale measure. Then the fair price V_0 for C is given by*

$$V_0 = \tilde{\mathbb{E}}(D), \quad D = C/S_T^0.$$

(b) *If $(\bar{S}_n)_{0\leqslant n\leqslant T}$ is complete, and the numéraire is non-random, then there is at most one equivalent martingale measure.*

Proof. Since C is attainable, there is a previsible process $(\theta_n)_{1\leqslant n\leqslant T}$ such that

$$D = V_0 + \sum_{n=1}^{T} \theta_n (X_n - X_{n-1}).$$

Then, by the argument used in the proof of Proposition 3.2,

$$\tilde{\mathbb{E}}(D|\mathcal{F}_0) = V_0 \quad \text{almost surely.}$$

Since $\mathcal{F}_0$ is trivial, this proves (a). Then, if the numéraire is non-random, we have $\tilde{\mathbb{E}}(C) = V_0 S_T^0$, so $\tilde{\mathbb{E}}(C)$ does not depend on the choice of $\tilde{\mathbb{P}}$. If $(\bar{S}_n)_{0\leqslant n\leqslant T}$ is complete, then this is true of all non-negative time-T contingent claims, in particular for $C = 1_A$ for all $A \in \mathcal{F}_T$. Hence $\tilde{\mathbb{P}}$ is unique. $\square$

3.8 Binomial model

Fix parameters $r, a, b \in (-1, \infty)$, $p \in (0,1)$ and $S_0 \in (0, \infty)$, with $a < b$. We say that $(S_n^0, S_n)_{0\leqslant n\leqslant T}$ is a *binomial model* with interest rate r and parameters $a < b$ and p if

$$S_n^0 = (1+r)^n, \quad S_n = S_0 \prod_{k=1}^{n} (1 + R_k)$$

where $R_1, \dots, R_T$ are independent identically distributed random variables with

$$\mathbb{P}(R_1 = a) = 1 - p, \quad \mathbb{P}(R_1 = b) = p.$$

This is also called the *Cox–Ross–Rubinstein model.* Assume that $\mathcal{F} = \mathcal{F}_T = \sigma(R_1, \dots, R_T)$.

Proposition 3.6. *A binomial model with interest rate r and parameters $a < b$ and p has an arbitrage unless $r \in (a, b)$.*

Proof. Consider the self-financing portfolio $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ with $V_0 = 0$, $\theta_1 = 1$ and $\theta_n = 0$ for $n \geqslant 2$. Then

$$V_T = X_1 - X_0 = \frac{S_0(1+R_1)}{1+r} - S_0 = S_0 \frac{R_1 - r}{1+r}.$$

If $r \leqslant a$, then $V_T \geqslant 0$ and $V_T = S_0(b-r)/(1+r) > 0$ with probability p, so $(\bar{\theta}_n)_{1\leqslant n\leqslant T}$ is an arbitrage. On the other hand, if $r \geqslant b$, then $V_T \leqslant 0$ and $V_T = -S_0(r-a)(1+r) < 0$ with probability $1-p$, so $(-\bar{\theta}_n)_{1\leqslant n\leqslant T}$ is an arbitrage. $\square$

Proposition 3.7. *Let $(S_n^0, S_n)_{0\leqslant n\leqslant T}$ be a binomial model with interest rate r and parameters $a < b$ and p. Suppose that $r \in (a,b)$ and define $p^* \in (0,1)$ by*

$$p^* = \frac{r-a}{b-a}.$$

Consider the equivalent probability measure $\mathbb{P}^*$ on $\mathcal{F}$ given by

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \left(\frac{p^*}{p}\right)^{U_T}\left(\frac{1-p^*}{1-p}\right)^{D_T}$$

where $U_T = (T+S_T)/2$ and $D_T = (T-S_T)/2$. Then, under $\mathbb{P}^*$, the random variables $R_1,\dots,R_T$ are independent and identically distributed, with

$$\mathbb{P}(R_1 = a) = 1-p^*, \quad \mathbb{P}(R_1 = b) = p^*.$$

Moreover, the discounted price process $(X_n)_{0\leqslant n\leqslant T}$ is a martingale under $\mathbb{P}^*$.

Proof. Note that $r = (1-p^*)a + p^*b$. We have

$$X_n = \frac{S_n}{S_n^0} = S_0\prod_{k=1}^n\left(\frac{1+R_k}{1+r}\right)$$

and

$$\mathbb{E}^*\left(\frac{1+R_1}{1+r}\right) = \frac{(1-p^*)(1+a)+p^*(1+b)}{1+r} = \frac{1+(1-p^*)a+p^*b}{1+r} = 1.$$

Since the random variables $R_1,\dots,R_T$ are independent and identically distributed under $\mathbb{P}^*$, this implies that $(X_n)_{0\leqslant n\leqslant T}$ is a $\mathbb{P}^*$-martingale. □

Note that Propositions 3.2 and 3.7 show that the binomial model has no arbitrage for $r \in (a,b)$. We will show in Proposition 3.9 that the binomial model is also complete. Hence, the fair price V_0 of any contingent claim of the form $C = f(S_T)$ is given by

$$V_0 = \frac{\mathbb{E}^*(C)}{(1+r)^T} = (1+r)^{-T}\sum_{k=0}^T\binom{T}{k}(1-p^*)^{T-k}p^{*k}f(S_0(1+a)^{T-k}(1+b)^k).$$

More generally, given a contingent claim $C = f(S_0, S_1, \dots, S_T)$, its fair price at time 0 is given by

$$V_0 = \frac{\mathbb{E}^*(C)}{(1+r)^T} = (1+r)^{-T}\sum f(s_0, s_1,\dots,s_T)\mathbb{P}^*(S_1 = s_1,\dots,S_T = s_T)$$

where we sum over the 2^T possible paths $(s_n)_{0\leqslant n\leqslant T}$ starting from S_0 and

$$\mathbb{P}^*(S_1 = s_1,\dots,S_T = s_T) = (1-p^*)^{T-k}p^{*k}$$

where $k = |\{n \in \{1, \dots, T\} : s_n = (1+b)s_{n-1}\}|$. It is sometimes efficient to organize this calculation over a tree. Set

$$f_T(s_0, \dots, s_T) = f(s_0, \dots, s_T)$$

and carry out the backwards recursion for $n \leqslant T-1$

$$f_n(s_0, \dots, s_n) = (1-p^*) f_{n+1}(s_0, \dots, s_n, (1+a)s_n) + p^* f_{n+1}(s_0, \dots, s_n, (1+b)s_n).$$

The nodes of the tree are the partial paths $(s_0, \dots, s_n)$, each with two upward edges to $(s_0, \dots, s_n, (1+a)s_n)$ and $(s_0, \dots, s_n, (1+b)s_n)$. We start with the values at the leaves $(s_0, \dots, s_T)$ and compute down the tree to the root s_0.

Proposition 3.8. *For $n = 0, 1, \dots, T$, almost surely,*

$$\mathbb{E}^*(f(S_0, \dots, S_T)|\mathcal{F}_n) = f_n(S_0, \dots, S_n)$$

and in particular

$$\mathbb{E}^*(C) = f_0(S_0).$$

Proof. The claim holds for $n = T$. Let $n \leqslant T-1$ and suppose for a reverse induction that the claim holds for $n+1$. Fix $s_0, \dots, s_n$ and define $A = \{S_0 = s_0, \dots, S_n = s_n\}$. Then

$$\begin{aligned}
\mathbb{E}^*(f(S_0, \dots, S_T)|A) &= \mathbb{E}^*(f_{n+1}(S_0, \dots, S_n, S_{n+1})|A) \\
&= (1-p^*) f_{n+1}(s_0, \dots, s_n, (1+a)s_n) + p^* f_{n+1}(s_0, \dots, s_n, (1+b)s_n) \\
&= f_n(s_0, \dots, s_n) = \mathbb{E}^*(f_n(S_0, \dots, S_n)|A).
\end{aligned}$$

Now $f_n(S_0, \dots, S_n)$ is $\mathcal{F}_n$-measurable, and every element of $\mathcal{F}_n$ is a finite union of such sets A. Hence, almost surely,

$$\mathbb{E}^*(f(S_0, \dots, S_T)|\mathcal{F}_n) = f_n(S_0, \dots, S_n)$$

and the induction proceeds. □

We now show completeness for the binomial model by identifying a replicating portfolio for a general contingent claim $C = f(S_0, \dots, S_T)$. Define

$$\Delta_n(s_0, \dots, s_{n-1}) = \frac{f_n(s_0, \dots, s_{n-1}, (1+b)s_{n-1}) - f_n(s_0, \dots, s_{n-1}, (1+a)s_{n-1})}{(1+r)^{T-n}(b-a)s_{n-1}}.$$

Proposition 3.9. *The contingent claim $C = f(S_0, \dots, S_T)$ has a replicating portfolio $(\theta_n)_{1 \leqslant n \leqslant T}$ given by*

$$\theta_n = \Delta_n(S_0, \dots, S_{n-1}).$$

Proof. Define, for $n = 0, 1, \ldots, T$,

$$V_n = \frac{\mathbb{E}^*(C|\mathcal{F}_n)}{(1+r)^T}.$$

Fix a path $(s_0, s_1, \ldots, s_{n-1})$ starting from S_0, set $\phi(x) = f_n(s_0, \ldots, s_{n-1}, (1+x)s_{n-1})$ and consider the events

$$\Omega_0 = \{S_1 = s_1, \ldots, S_{n-1} = s_{n-1}\}, \quad \Omega_a = \Omega_0 \cap \{R_n = a\}, \quad \Omega_b = \Omega_0 \cap \{R_n = b\}.$$

On Ω_a, we have

$$V_n = \frac{\phi(a)}{(1+r)^T}, \quad V_{n-1} = \frac{f_{n-1}(s_0, \ldots, s_{n-1})}{(1+r)^T} = \frac{(1-p^*)\phi(a) + p^*\phi(b)}{(1+r)^T}$$

so

$$V_n - V_{n-1} = \frac{p^*(\phi(a) - \phi(b))}{(1+r)^T}$$

and

$$X_n - X_{n-1} = \frac{1+a}{1+r}X_{n-1} - X_{n-1} = \frac{(a-r)s_{n-1}}{(1+r)^n}$$

so

$$\theta_n(X_n - X_{n-1}) = \left(\frac{a-r}{b-a}\right)\frac{\phi(b) - \phi(a)}{(1+r)^T} = V_n - V_{n-1}.$$

By a similar calculation, the same identity holds on Ω_b, and hence everywhere, since $s_1, \ldots, s_{n-1}$ are arbitrary. On summing over n, we obtain

$$D = V_0 + \sum_{n=1}^{T} \theta_n(X_n - X_{n-1}).$$

showing that $(\theta_n)_{1\leqslant n\leqslant T}$ is a replicating portfolio for C, as claimed. □

3.9 Joint distribution of a simple random walk and its maximum

A reflection trick allows to obtain the following result.

Proposition 3.10. *Let $(W_n)_{0\leqslant n\leqslant T}$ be a simple random walk on the integers starting from 0, with $\mathbb{P}(W_1 = 1) = p$. Set $M_T = \max_{0\leqslant n\leqslant T} W_n$. Then, for all integers $m, k \geqslant 0$, with $k \leqslant T$ and $2k - T \leqslant m \leqslant k$,*

$$\mathbb{P}(M_T = m \text{ and } W_T = 2k - T) = \left(\binom{T}{k-m} - \binom{T}{k-m-1}\right)p^k(1-p)^{T-k}.$$

Proof. Write E for the set of paths taken by the random walk. Thus

$$E = \{w = (w_n)_{0 \leqslant n \leqslant T} : w_0 = 0 \text{ and } w_n = w_{n-1} \pm 1 \text{ for } n = 1, \dots, T\}.$$

Consider the map $\phi : E \to E$ which reflects about level m the portion of the path (if any) after it first hits level m. Then $\phi \circ \phi$ is the identity, so ϕ is a bijection. Define

$$A = \{w \in E : \max_{0 \leqslant n \leqslant T} w_n \geqslant m \text{ and } w_T = 2k - T\}.$$

Then

$$\phi(A) = \{w \in E : w_T = 2m - 2k + T\}.$$

Hence

$$|A| = |\phi(A)| = \binom{T}{k-m}.$$

But the walk takes every path in A with equal probability $p^k(1-p)^{T-k}$. Hence

$$\mathbb{P}(M_T \geqslant m \text{ and } W_T = 2k - T) = \binom{T}{k-m} p^k (1-p)^{T-k}$$

and the result follows by subtracting the corresponding formula for $m + 1$. □

In the special case where $(1 + a)(1 + b) = 1$, we can realise the binomial model as a function of a simple random walk $(W_n)_{0 \leqslant n \leqslant T}$ on the integers by setting

$$S_n = S_0(1 + b)^{W_n}.$$

Assume that the interest rate $r \in (a, b)$ and set $p^* = (r - a)/(b - a)$ as usual. Proposition 3.10 then gives an explicit formula for the fair price at time 0 of any contingent claim of the form

$$C = F\left(S_T, \max_{0 \leqslant n \leqslant T} S_n\right)$$

which is given by

$$V_0 = \frac{\mathbb{E}^*(C)}{(1+r)^T} = \sum_{k=0}^{T} \sum_{\substack{m=0 \\ m \geqslant 2k-T}}^{k} \frac{f(k,m)}{(1+r)^T} \left(\binom{T}{k-m} - \binom{T}{k-m-1} \right) p^{*k} (1-p^*)^{T-k}$$

where $f(k, m) = F(S_0(1 + b)^{2k-T}, S_0(1 + b)^m)$.

4 Dynamic programming

4.1 Bellman equation

The trees in an orchard are arranged in a rectangular grid. The numbers of apples on each tree are shown in the following array.

2	5	6	1	9	4	3	3	2	9
5	3	8	2	1	4	7	7	1	1
4	9	2	1	4	5	5	7	4	3
1	5	3	3	3	2	4	5	3	7
8	3	4	5	1	2	1	4	1	1
0	2	5	7	8	1	3	1	9	2
3	1	5	6	2	9	4	1	1	1
7	2	3	2	4	5	1	6	5	9
4	3	5	6	1	1	1	2	2	3
8	8	4	5	2	5	7	7	4	2
3	4	2	4	1	9	9	7	1	1

You start in the leftmost column of the array, at the tree with no apples. You now move one-by-one across the columns, from left to right. You may choose at each step whether to go the tree in the same row or the tree in the row above, or the tree in the row below. Thus at your first step, you can choose to go to a tree with 1 apple, 2 apples or 3 apples; if you choose the tree with 3 apples, then next step you get to choose a tree with 3, 4 or 5 apples. Supposing that you keep all the apples from every tree that you visit in this way, how many apples would you collect if you used the best route? If you were allowed to select the tree that you started at, which one would it be?

55	53	46	39	38	29	21	14	11	9
63	51	48	40	31	29	25	18	10	1
62	58	42	35	34	30	23	18	11	3
59	57	49	37	33	25	22	16	10	7
65	56	52	46	26	24	17	15	8	1
56	55	53	48	41	25	18	12	11	2
58	54	53	47	35	33	24	15	10	1
62	55	50	39	37	29	21	20	14	9
61	53	49	44	33	26	21	16	11	3
65	57	48	43	38	32	25	18	7	2
60	52	45	42	37	36	27	14	3	1

A simple idea gives an efficient approach this problem. Label the tree in the nth column and xth row by (n, x), and write $a(n, x)$ for the number of apples on that tree. Write $V(n, x)$ for the maximal number of apples which can be collected starting from that tree along any allowed path. Then $V(10, x) = a(10, x)$ for all x and

$$V(n, x) = a(n, x) + \max\{V(n+1, x-1), V(n+1, x), V(n+1, x+1)\}, \quad n = 1, \dots, 9$$

where we set $V(n,0) = V(n,12) = 0$. Then we can compute V by a backwards recursion, as shown in the second array. This not only tells us that the best route from $(1,6)$ collects 56 apples, but also reveals that route as $+-00-0-00$, where $\pm$ means move up or down and 0 means stay in the same row. Moreover, we see that the trees $(1,5)$ and $(1,10)$ are the best starting points.

We now formulate the idea just used in a more general context. Fix a state-space E, an action-space A and an integer $T \geqslant 1$. Suppose given a measurable map

$$F : \{0,1,\dots,T-1\} \times E \times A \times [0,1] \to E.$$

Let $(\varepsilon_n)_{1\leqslant n\leqslant T}$ be a sequence of independent uniform random variables in $[0,1]$ and set $\mathcal{F}_n = \sigma(\varepsilon_1,\dots,\varepsilon_n)$. Fix an initial time $k \in \{0,1,\dots,T\}$. By an *adapted control*, we mean a sequence of random variables $u = (u_n)_{k\leqslant n\leqslant T-1}$ in A such that u_n is $\mathcal{F}_n$-measurable for all n. Given, an initial state $x \in E$, and an adapted control u we can define a random process $(X_n)_{k\leqslant n\leqslant T}$ by setting

$$X_k = x, \quad X_{n+1} = F(n, X_n, u_n, \varepsilon_{n+1}), \quad n = k,\dots,T-1.$$

We will sometimes write $X_n = X_n^u(k,x)$ to make explicit the dependence on (k,x) and on the control u.

Consider the problem of choosing u to optimize the expected reward

$$V^u(k,x) = \mathbb{E}\left(\sum_{n=k}^{T-1} r(n, X_n^u(k,x), u_n) + R(X_T^u(k,x))\right).$$

where r and R are given non-negative measurable functions on $\{0,\dots,T-1\} \times E \times A$ and E respectively. Define the *value function* V on $\{0,1,\dots,T\} \times E$ by

$$V(k,x) = \sup_u V^u(k,x)$$

where the supremum is taken over all adapted controls. If $V(k,x) = V^u(k,x)$, then we say that u is an *optimal control* from (k,x).

Proposition 4.1 (Bellman equation)**.** *Define a function v on $\{0,1,\dots,T\} \times E$ by the following backwards recursion*

$$\begin{aligned} v(T,x) &= R(x), \\ v(n,x) &= \sup_{a\in A}\{r(n,x,a) + Pv(n,x,a)\}, \quad n = 0,1,\dots,T-1 \end{aligned} \tag{1}$$

where

$$Pv(n,x,a) = \mathbb{E}(v(n+1, F(n,x,a,\varepsilon_{n+1}))).$$

Suppose that there is a measurable function $a : \{0,1,\dots,T-1\} \times E \to A$ such that

$$v(n,x) = r(n,x,a(n,x)) + Pv(n,x,a(n,x)), \quad n = 0,1,\dots,T-1.$$

Then $V = v$. Moreover, for all $k \in \{0,1,\dots,T-1\}$ and $x \in E$, we can define recursively an optimal control from (k,x) by

$$u_n^* = a(n, X_n^{u^*}(k,x)).$$

Proof. It will suffice to consider the case $k = 0$. Fix an adapted control $u = (u_n)_{0 \leqslant n \leqslant T-1}$ and a starting state $x \in E$. Write $X_n = X_n^u(0, x)$. By Proposition 2.1,

$$\mathbb{E}(v(n+1, X_{n+1})|\mathcal{F}_n) = Pv(n, X_n, u_n) \quad \text{almost surely.}$$

Consider the process

$$M_n = \sum_{k=0}^{n-1} r(k, X_k, u_k) + v(n, X_n).$$

Then, almost surely,

$$\begin{aligned}\mathbb{E}(M_{n+1}|\mathcal{F}_n) &= \sum_{k=0}^{n-1} r(k, X_k, u_k) + r(n, X_n, u_n) + Pv(n, X_n, u_n) \\ &\leqslant \sum_{k=0}^{n-1} r(k, X_k, u_k) + v(n, X_n) = M_n\end{aligned}$$

with equality if $u = u^*$. Hence

$$V^u(0, x) = \mathbb{E}(M_T) \leqslant M_0 = v(0, x) = V^{u^*}(0, x).$$

Since u was arbitrary, this shows also that $V(0, x) = v(0, x)$. □

A analogous result holds by a similar argument when we regard r and R as costs and seek to minimize the expected total cost, and also when there are a mixture of costs and rewards, with some additional care about integrability. It is sometimes convenient to allow a time-dependent state-space E_n and a time-and-state-dependent action-space $A_{n,x}$. The same argument applies. We used independent uniform random variables $(\varepsilon_n)_{1 \leqslant n \leqslant T}$ as a source of randomness, but the argument applies equally with any sequence of independent random variables.

4.2 American calls and puts

An *American call* of expiry T confers the right but not the obligation to buy one unit of the underlying asset $(S_t)_{0 \leqslant n \leqslant T}$ for price K at any stopping time $\tau \leqslant T$ chosen by the holder. An *American put* of expiry T confers the right but not the obligation to sell one unit of the underlying asset $(S_t)_{0 \leqslant n \leqslant T}$ for price K at any stopping time $\tau \leqslant T$ chosen by the holder. Let us assume an interest rate of r. Then the American call is the family of time-T contingent claims

$$\{C_\tau : \tau \text{ a stopping time }, \tau \leqslant T\}$$

while the American put is the family

$$\{P_\tau : \tau \text{ a stopping time }, \tau \leqslant T\}$$

where

$$C_\tau = (1+r)^{T-\tau}(S_\tau - K)^+, \quad P_\tau = (1+r)^{T-\tau}(K - S_\tau)^+.$$

We will consider the pricing of these options when $(S_n)_{0\leqslant n\leqslant T}$ is a binomial model of parameters $a < b$ with interest rate $r \in (a,b)$. Write, as usual, $\mathbb{P}^*$ for the equivalent martingale measure, corresponding to the parameter choice $p^* = (r-a)/(b-a)$. Since the binomial model is complete, the investor can hedge all contingent claims C with $\mathbb{E}^*(C) = 0$ in the market. Hence he will choose from each of the families above the claim with maximal expectation under $\mathbb{P}^*$.

Consider first the American call. Fix a stopping time $\tau \leqslant T$ and fix $n \in \{0,1,\dots,T\}$. Note that

$$\mathbb{E}^*(S_T|\mathcal{F}_n) = (1+r)^{T-n}S_n, \quad \text{almost surely.}$$

Consider the event

$$A = \{S_\tau \geqslant K \text{ and } \tau = n\}.$$

Then

$$\begin{aligned}\mathbb{E}^*(C_T 1_{\{\tau=n\}}) \geqslant \mathbb{E}^*((S_T - K)1_A) &= \mathbb{E}^*(((1+r)^{T-n}S_n - K)1_A)\\ &\geqslant \mathbb{E}^*((1+r)^{T-n}(S_n - K)1_A) = \mathbb{E}^*(C_\tau 1_{\{\tau=n\}}).\end{aligned}$$

On summing over n, we see that $\mathbb{E}^*(C_T) \geqslant \mathbb{E}^*(C_\tau)$. Hence the choice $\tau = T$ is always optimal, making the American and European calls equivalent.

In order to find the fair price V_0 of the American put, we must solve the following optimal stopping problem: maximize

$$\mathbb{E}^*\left((1+r)^{T-\tau}(K - S_\tau)^+\right)$$

over all stopping times $\tau \leqslant T$. At each time $n \in \{0,\dots,T-1\}$, the investor has two possible actions: to stop, receiving reward $(1+r)^{T-n}(K - S_n)$, or to continue. If he has not already stopped, at time T he receives a reward of $(K - S_T)^+$. Define

$$E_n = \{S_0(1+a)^{n-k}(1+b)^k : k = 0,\dots,n\}.$$

The Bellman equation for this problem is then

$$\begin{aligned}&v(T,x) = (K-x)^+, \quad \text{for } x \in E_T,\\ &v(n,x) = \max\{(1+r)^{T-n}(K-x), (1-p^*)v(n+1,x(1+a)) + p^*v(n+1,x(1+b))\},\\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{for } x \in E_n \text{ and } n = 0,\dots,T-1.\end{aligned}$$

This can be solved by backwards recursion to find $v(0,S_0)$. Then

$$V_0 = \sup_\tau \mathbb{E}^*((1+r)^{T-\tau}(K - S_\tau)^+) = v(0,S_0)$$

and the optimal stopping time τ is the smallest $n \in \{0,1,\dots,T\}$ such that

$$(1+r)^{T-n}(K - S_n)^+ = v(n,S_n).$$

5 Brownian motion

5.1 Definition and basic properties

A real-valued random process $(B_t)_{t\geqslant 0}$ is said to be a *Brownian motion* if $B_0 = 0$ and

(a) for all $s, t \geqslant 0$, the random variable $B_{s+t} - B_s$ is Gaussian, of mean 0 and variance t, and is independent of $\sigma(B_r : r \leqslant s)$,

(b) for all $\omega \in \Omega$, the map $t \mapsto B_t(\omega) : [0, \infty) \to \mathbb{R}$ is continuous.

We sometimes replace the condition $B_0 = 0$ by $B_0 = x$, when we will refer to $(B_t)_{t\geqslant 0}$ as a Brownian motion starting from x. We often write $\mathbb{P}_x$ and $\mathbb{E}_x$ in this case as a reminder. If the starting point is not mentioned, then the default is always to start at 0.

Proposition 5.1. *Let $(B_t)_{t\geqslant 0}$ be a continuous random process starting from 0. Then the following are equivalent:*

(a) *$(B_t)_{t\geqslant 0}$ is a Brownian motion,*

(b) *$(B_t)_{t\geqslant 0}$ is a zero-mean Gaussian process with $\mathbb{E}(B_sB_t) = s \wedge t$ for all $s, t \geqslant 0$.*

Proof. Suppose (a) holds. To see that $(B_t)_{t\geqslant 0}$ is zero-mean Gaussian, it suffices to note that, for $0 = t_0 \leqslant t_1 \leqslant \dots t_n$, the random variable $(B_{t_1}, \dots, B_{t_n})$ is a linear function of the independent zero-mean Gaussian random variables $(B_{t_1} - B_{t_0}, \dots, B_{t_n} - B_{t_{n-1}})$ and hence is Gaussian. For the covariance, we have, for $s \leqslant t$,

$$\mathbb{E}(B_sB_t) = \mathbb{E}(B_s^2) + \mathbb{E}(B_s(B_t - B_s)) = s$$

since $B_s \sim N(0, s)$ and $B_t - B_s$ is independent of B_s. Hence (b) holds.

Suppose on the other hand that (b) holds. Then, for $s, t \geqslant 0$, $B_{s+t} - B_s$ is a zero-mean Gaussian and

$$\mathbb{E}((B_{s+t} - B_s)^2) = (s + t) - 2s + s = t$$

so $B_{s+t} - B_s \sim N(0, t)$. To show that $B_{s+t} - B_s$ is independent of $\sigma(B_r : r \leqslant s)$, it suffices by a result from Probability and Measure to show that it is independent of $(B_{r_1}, \dots, B_{r_n})$ for all n and all $0 \leqslant r_1 \leqslant \dots \leqslant r_n \leqslant s$. Then, given that $(B_t)_{t\geqslant 0}$ is zero-mean Gaussian, it suffices to note that

$$\mathbb{E}(B_{r_k}(B_{t+s} - B_s)) = 0$$

for all k. Hence $(B_t)_{t\geqslant 0}$ is a Brownian motion. □

The following proposition is left as an exercise.

Proposition 5.2 (Scaling property)**.** *Let $(B_t)_{t\geqslant 0}$ be a Brownian motion and let $c \in (0, \infty)$. Set $\tilde{B}_t = c^{-1}B_{c^2t}$. Then $(\tilde{B}_t)_{t\geqslant 0}$ is also a Brownian motion.*

Proposition 5.3. *Let $(B_t)_{t\geqslant 0}$ be a Brownian motion. Then $(B_t)_{t\geqslant 0}$ exits every finite interval almost surely.*

Proof. Fix an interval I, of length L say. Consider the sequence of independent events $(A_n)_{n\geqslant 1}$ given by

$$A_n = \{|B_n - B_{n-1}| > L\}.$$

Then $\mathbb{P}(A_n) = \mathbb{P}(|B_1| > L) > 0$ for all n. By independence, this implies that $\mathbb{P}(\cup_n A_n) = 1$. But, on A_n, if $(B_t)_{t\geqslant 0}$ has not left I by time $n-1$, it must do so by time n. Hence $(B_t)_{t\geqslant 0}$ leaves I almost surely. □

Write $\mathcal{F}_t = \sigma(B_s : s \in [0,t])$. A random time $T : \Omega \to [0,\infty]$ is a *stopping time* if

$$\{T \leqslant t\} \in \mathcal{F}_t \quad \text{for all } t.$$

Write $\mathcal{F}_T$ for the set of events $A \in \mathcal{F}_\infty$ such that

$$A \cap \{T \leqslant t\} \in \mathcal{F}_t \quad \text{for all } t.$$

We will need the following proposition which is proved in Advanced Probability.

Proposition 5.4 (Strong Markov property)**.** *Let $(B_t)_{t\geqslant 0}$ be a Brownian motion and let T be an almost surely finite stopping time. Define $(\tilde{B}_t)_{t\geqslant 0}$ by*

$$\tilde{B}_t = B_{T+t} - B_T.$$

Then $(\tilde{B}_t)_{t\geqslant 0}$ is also a Brownian motion and is independent of $\mathcal{F}_T$.

Proposition 5.5. *Let $(B_t)_{t\geqslant 0}$ be a Brownian motion and let $a \in \mathbb{R}$. Define*

$$T_a = \inf\{t \geqslant 0 : B_t = a\}.$$

Then T_a is an almost surely finite stopping time.

Proof. It suffices to consider the case $a > 0$. Since $(B_t)_{t\geqslant 0}$ is continuous, we have

$$\{T_a \leqslant t\} = \{B_t = a\} \cup \bigcap_{r<a, r\in\mathbb{Q}} \bigcup_{s<t, s\in\mathbb{Q}} \{B_s > r\} \in \mathcal{F}_t.$$

Hence T_a is a stopping time. Set $T = T_1 \wedge T_{-1}$. Then T is a stopping time and, since Brownian motion leaves every interval, $T < \infty$ almost surely. By symmetry

$$\mathbb{P}(B_T = 1) = \mathbb{P}(B_T = -1) = 1/2.$$

Then, by the strong Markov property, the sequence of integers hit by $(B_t)_{t\geqslant 0}$, omitting immediate repeats, is a simple symmetric random walk. Since the random walk is recurrent, it follows that $T_a < \infty$ almost surely. □

5.2 Brownian motion as a limit of random walks

Brownian motion can be understood as a limit object, associated to random walks having steps of mean 0 and variance 1. We take a limit which scales down the time-step and the spatial steps of the walk in a coordinated way, so that the variance of the walk at time 1 is always 1. It turns out that, in this limit, all such random walks 'look almost the same' and all 'approximate' the behaviour of Brownian motion. The following theorem is a combination of Wiener's Theorem and Donsker's Invariance Principle, which are both proved in Advanced Probability.

Theorem 5.6. *Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is not discrete. Let m be a probability measure on $\mathbb{R}$ of mean 0 and variance 1. Then there exist a random process $(B_t)_{t\geqslant 0}$ and, for each $k \in \mathbb{N}$, a random process $(W_t^{(k)})_{t\geqslant 0}$ such that:*

(a) *$(B_t)_{t\geqslant 0}$ is a Brownian motion,*

(b) *$(W_{n/k}^{(k)})_{n\geqslant 0}$ is a random walk with step distribution m, and $(W_t^{(k)})_{t\geqslant 0}$ is the linear interpolation of its values on $\{n/k : n \in \mathbb{Z}^+\}$,*

(c) *$W_t^{(k)}/\sqrt{k} \to B_t$ as $k \to \infty$ uniformly on compacts in t almost surely.*

5.3 Change of probability measure

We now show that the distribution of a Brownian motion with constant drift is absolutely continuous with respect to Wiener measure and we identify its density. This is a special case of the Cameron–Martin theorem.

Proposition 5.7. *Let $T \geqslant 0$ and $c \in \mathbb{R}$. Let $B = (B_t)_{t\in[0,T]}$ be a Brownian motion and define $\tilde{B} = (\tilde{B}_t)_{t\in[0,T]}$ by $\tilde{B}_t = B_t + ct$. Then, for all measurable sets $A \subseteq C[0,T]$, we have*

$$\mathbb{P}(\tilde{B} \in A) = \mathbb{E}(1_{\{B\in A\}} e^{cB_T - c^2T/2}).$$

Proof. Consider the set $\mathcal{A}$ of all subsets of $C[0,T]$ of the form

$$A = \{x \in C[0,T] : x_{t_k} - x_{t_{k-1}} \in I_k \text{ for } k = 1, \dots, n\}$$

where $n \in \mathbb{N}$ and $0 = t_0 \leqslant t_1 \leqslant \dots \leqslant t_n = T$ and $I_1, \dots, I_n$ are intervals in $\mathbb{R}$. By a standard argument from Probability and Measure, it will suffice to prove the formula in the case $A \in \mathcal{A}$.

For $X \sim N(0,s)$ and $\tilde{X} = X + cs$, we have $\tilde{X} \sim N(cs, s)$ so, for any interval I,

$$\begin{aligned}\mathbb{P}(\tilde{X} \in I) &= \int_I \frac{1}{\sqrt{2\pi s}} e^{-(x-cs)^2/(2s)} dx \\ &= \int_I \frac{1}{\sqrt{2\pi s}} e^{-x^2/(2s)} e^{cx - c^2 s/2} dx = \mathbb{E}(1_{\{X\in I\}} e^{cX - c^2 s/2}).\end{aligned}$$

Set $s_k = t_k - t_{k-1}$ and $X_k = B_{t_k} - B_{t_{k-1}}$ and $\tilde{X}_k = \tilde{B}_{t_k} - \tilde{B}_{t_{k-1}}$. Then $X_1, \dots, X_n$ are independent, with $X_k \sim N(0, s_k)$ and $\tilde{X}_k = X_k + cs_k$. Hence

$$\mathbb{P}(\tilde{B} \in A) = \mathbb{P}\left(\prod_{k=1}^{n} 1_{\{\tilde{X}_k \in I_k\}}\right) = \prod_{k=1}^{n} \mathbb{P}(\tilde{X}_k \in I_k)$$
$$= \prod_{k=1}^{n} \mathbb{E}(1_{\{X_k \in I_k\}} e^{cX_k - c^2 s_k/2}) = \mathbb{E}\left(\prod_{k=1}^{n} 1_{\{X_k \in I_k\}} e^{cX_k - c^2 s_k/2}\right) = \mathbb{E}\left(1_{\{B \in A\}} e^{cB_T - c^2 T/2}\right)$$

where we used that $X_1 + \cdots + X_n = B_T$ and $s_1 + \cdots + s_n = T$. □

5.4 Reflection principle

The reflection trick used in the proof of Proposition 3.10 has the following continuum analogue.

Proposition 5.8 (Reflection principle)**.** *Let $(B_t)_{t \geqslant 0}$ be a Brownian motion and let $a \geqslant 0$. Set*

$$T_a = \inf\{t \geqslant 0 : B_t = a\}$$

and define $(\tilde{B}_t)_{t \geqslant 0}$ by

$$\tilde{B}_t = \begin{cases} B_t, & \text{if } t \leqslant T_a, \\ 2a - B_t, & \text{if } t > T_a. \end{cases}$$

Then $(\tilde{B}_t)_{t \geqslant 0}$ is also a Brownian motion.

Proof. Write T for T_a in the proof. We know that $T < \infty$ almost surely and T is a stopping time. Define processes $Y = (Y_t)_{t \geqslant 0}$ and $Z = (Z_t)_{t \geqslant 0}$ by

$$Y_t = B_{T \wedge t}, \quad Z_t = B_{T+t} - B_T.$$

Then Y_t is $\mathcal{F}_T$-measurable for all t and, by the strong Markov property, Z is a Brownian motion independent of $\mathcal{F}_T$. So $-Z$ is also a Brownian motion independent of $\mathcal{F}_T$. Hence (Y, Z) and $(Y, -Z)$ have the same distribution on $C[0,\infty) \times C_0[0,\infty)$.

Consider the measurable map $F : C[0,\infty) \times C_0[0,\infty) \to C[0,\infty)$ given by

$$F(y, z) = \begin{cases} y(t), & \text{if } t \leqslant \tau(y), \\ y(\tau(y)) + z(t - \tau(y)), & \text{if } t > \tau(y) \end{cases}$$

where $\tau(y) = \inf\{t \geqslant 0 : y(t) = a\}$. Then $F(Y, Z) = B$ and $F(Y, -Z) = \tilde{B}$, so B and $\tilde{B}$ have the same distribution. □

The reflection principle facilitates some useful calculations for Brownian motion. For example, consider the maximum process

$$M_t = \sup_{0 \leqslant s \leqslant t} B_s.$$

Then, for all $a \geqslant 0$ and $x \leqslant a$,

$$\{M_t \geqslant a \text{ and } B_t \leqslant x\} = \{\tilde{B}_t \geqslant 2a - x\}$$

so

$$\mathbb{P}(M_t \geqslant a \text{ and } B_t \leqslant x) = \mathbb{P}(B_t \geqslant 2a - x).$$

In particular, taking $x = a$, we see that

$$\mathbb{P}(M_t \geqslant a) = \mathbb{P}(B_t \geqslant a) + \mathbb{P}(M_t \geqslant a \text{ and } B_t \leqslant a) = 2\mathbb{P}(B_t \geqslant a).$$

Hence M_t has the same distribution as $|B_t|$. We can moreover compute the moment generating function. For

$$\mathbb{E}(e^{uM_1}) = 2\int_0^\infty e^{ux}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx = 2e^{u^2/2}\int_0^\infty \frac{1}{\sqrt{2\pi}}e^{-(x-u)^2/2}dx = 2e^{u^2/2}\Phi(u)$$

and so, by scaling,

$$\mathbb{E}(e^{uM_t}) = \mathbb{E}(e^{u\sqrt{t}M_1}) = 2e^{u^2t/2}\Phi(u\sqrt{t}).$$

5.5 Hitting probabilities

Proposition 5.9. *Let $(B_t)_{t\geqslant 0}$ be a Brownian motion and let $a \geqslant 0$. Then T_a has a density function h_a on $[0,\infty)$ given by*

$$h_a(t) = \frac{a}{\sqrt{2\pi t^3}}e^{-a^2/(2t)}.$$

Proof. By the reflection principle,

$$\mathbb{P}(T_a \leqslant t) = \mathbb{P}(M_t \geqslant a) = 2\mathbb{P}(B_t \geqslant a).$$

But

$$\mathbb{P}(B_t \geqslant a) = \mathbb{P}(\sqrt{t}B_1 \geqslant a) = \int_{a/\sqrt{t}}^\infty \frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy$$

and we can now differentiate in t to see that T_a has the claimed density. $\square$

5.6 Transition density for killed Brownian motion

Proposition 5.10. *Let $x, a \in \mathbb{R}$ with $x \leqslant a$. Let $(B_t)_{t\geqslant 0}$ be a Brownian motion starting from x. Then, for all non-negative measurable functions f,*

$$\mathbb{E}_x(f(B_t)1_{\{T_a>t\}}) = \int_{-\infty}^a f(y)p_t^a(x,y)dy$$

where

$$p_t^a(x,y) = p_t(x,y) - p_t(x, 2a - y).$$

Proof. By a standard argument, it suffices to consider the case where $f = 1_{(-\infty,b]}$ for some $b \leqslant a$. Set $M_t = \sup_{s \leqslant t} B_s$. Then, using the reflection principle,

$$\begin{aligned}
\mathbb{E}_x(f(B_t)1_{\{T_a>t\}}) &= \mathbb{P}_x(M_t < a \text{ and } B_t \leqslant b) \\
&= \mathbb{P}_x(B_t \leqslant b) - \mathbb{P}_x(M_t \geqslant a \text{ and } B_t \leqslant b) \\
&= \mathbb{P}_x(B_t \leqslant b) - \mathbb{P}_x(B_t \geqslant 2a - b) \\
&= \int_{-\infty}^{b} p_t(x,y)dy - \int_{-\infty}^{b} p_t(x, 2a-y)dy = \int_{-\infty}^{a} f(y)p_t^a(x,y)dy.
\end{aligned}$$

□

By change of measure we can deduce an analogous result for Brownian motion with drift. Fix $c \in \mathbb{R}$ and set

$$\tilde{B}_t = B_t + ct, \quad \tilde{T}_a = \inf\{t \geqslant 0 : \tilde{B}_t = a\}.$$

Then

$$\mathbb{E}_x(f(\tilde{B}_t)1_{\{\tilde{T}_a>t\}}) = \mathbb{E}_x(f(B_t)e^{c(B_t-x)-c^2t/2}1_{\{T_a>t\}}) = \int_{-\infty}^{a} f(y)\tilde{p}_t^a(x,y)dy.$$

where

$$\tilde{p}_t^a(x,y) = e^{c(y-x)-c^2t/2}p_t^a(x,y).$$

6 Black–Scholes model

6.1 Black–Scholes pricing formula

By a *Black–Scholes model* we mean any pair of processes $(S_t^0)_{t\geqslant 0}$ and $(S_t)_{t\geqslant 0}$ of the form

$$S_t^0 = e^{rt}, \quad S_t = S_0 e^{\sigma B_t + \mu t}$$

where $(B_t)_{t\geqslant 0}$ is a Brownian motion and where $r, \mu \in \mathbb{R}$ and $\sigma, S_0 \in (0, \infty)$. We interpret $(S_t^0)_{t\geqslant 0}$ as the price of a *riskless bond of interest rate* r. We interpret $(S_t)_{t\geqslant 0}$ as the price of a *risky asset*. We call μ the *drift* and σ the *volatility*.

The process $(e^{\sigma B_t - \sigma^2 t/2})_{t\geqslant 0}$ is a martingale. Hence for the special choice of drift

$$\mu^* = r - \sigma^2/2.$$

the discounted asset price $(e^{-rt}S_t)_{t\geqslant 0}$ is a martingale.

Proposition 6.1. *Let $(S_t^0, S_t)_{t\geqslant 0}$ be a Black–Scholes model of interest rate r, drift μ and volatility σ. Fix $T \in (0, \infty)$ and consider the equivalent probability measure $\mathbb{P}^*$ given by*

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = e^{\lambda B_T - \lambda^2 T/2}, \quad \sigma\lambda = \mu^* - \mu = r - \sigma^2/2 - \mu.$$

Then, under $\mathbb{P}^*$, the discounted asset price $(e^{-rt}S_t)_{0\leqslant t\leqslant T}$ is a martingale.

Proof. Set $B_t^* = B_t - \lambda t$. By Proposition 5.7, under $\mathbb{P}^*$, $(B_t)_{0\leqslant t\leqslant T}$ is a Brownian motion with drift λ, so $(B_t^*)_{0\leqslant t\leqslant T}$ is a Brownian motion. Now

$$\sigma B_t^* + \mu^* t = \sigma B_t - \sigma\lambda t + \mu^* t = \sigma B_t + \mu t$$

so

$$e^{-rt}S_t = S_0 e^{\sigma B_t + \mu t - rt} = S_0 e^{\sigma B_t^* + \mu^* t - rt} = S_0 e^{\sigma B_t^* - \sigma^2 t/2}.$$

Hence, under $\mathbb{P}^*$, $(e^{-rt}S_t)_{0\leqslant t\leqslant T}$ is a martingale. □

We will abuse notation in writing $\mathbb{P}^*$ instead of $\mathbb{P}$ when considering a Black–Scholes model with drift μ^*.

By a *time-T contingent claim*, we mean an $\mathcal{F}_T$-measurable random variable C, considered as an amount payable to the investor at time T. Let us assume that an investor is free to trade in the asset and the bond. The *Black–Scholes price* V_0 for the claim C is defined by

$$V_0 = e^{-rT}\mathbb{E}^*(C). \tag{2}$$

An argument which we will not develop shows that V_0 is the unique fair price for C, in the sense that any other price presents an opportunity to make risk-free profits by buying or selling the claim and trading continuously in the asset and the bond. We have not made precise what it means for the investor to trade freely in the asset and the bond. However, the following examples give some plausibility to the assertion that V_0 is the unique fair price for C. In the case $C = S_T$, the investor can replicate C by buying one unit of the asset at time 0, so S_T must have time-0 price S_0. On the other hand, in the case of a constant claim $C = K$, the investor can replicate C by investing $e^{-rT}K$ in the bond at time 0, so K (at time T) must have time-0 price $e^{-rT}K$. In both cases, the price is the Black–Scholes price.

More generally, by a *simple replicable claim* we mean any claim C such that

$$e^{-rT}C = C_0 + \sum_{k=1}^{n} \theta_k (X_{t_k} - X_{t_{k-1}})$$

for some constant C_0, some $n \in \mathbb{N}$ and $0 = t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n = T$, with θ_k a bounded $\mathcal{F}_{t_{k-1}}$-measurable random variable for all k. Here X_t is the discounted asset price $e^{-rt}S_t$. We leave as an exercise to show that any such claim C can be replicated at time T for cost C_0 at time 0. Since $(X_t)_{0 \leqslant t \leqslant T}$ is a martingale under $\mathbb{P}^*$, we can take expectations to obtain

$$V_0 = e^{-rT}\mathbb{E}^*(C) = C_0.$$

Hence V_0 is the unique fair price for C.

In fact, by the Brownian martingale representation theorem, every integrable $\mathcal{F}_T$-measurable contingent claim is the limit in probability of simple replicable claims, and this can be used to justify the validity of the pricing formula in general.

6.2 Black–Scholes PDE

For some simple types of option, the Black–Scholes pricing theory can be cast in terms of PDE. In this section, we show how the stochastic theory reduces to a PDE for options depending only on the final asset price. Consider a Black–Scholes model with interest rate r and risky asset $(S_t)_{t \geqslant 0}$ having drift μ and volatility σ. We derive first the form of the backward equation for the transition density of $(S_t)_{t \geqslant 0}$. Consider the case $S_0 = s$. Then

$$\log S_t = \log s + \sigma B_t + \mu t$$

so $\log S_t$ has density $p(\sigma^2 t, x(t), .)$ on $\mathbb{R}$, where

$$p(t, x, z) = \frac{1}{\sqrt{2\pi t}} e^{-|x-z|^2/(2t)}, \quad x(t) = \log s + \mu t.$$

Write $\rho(t, s, .)$ for the density function of S_t on $(0, \infty)$ when $S_0 = s$. We will write $\dot{\rho}$ and ρ' for derivatives in the first and second argument respectively. We have

$$y\rho(t, s, y) = p(\sigma^2 t, x(t), z), \quad z = \log y$$

so

$$sy\rho'(t, s, y) = p'(\sigma^2 t, x(t), z), \quad sy(\rho' + s\rho'')(t, s, y) = p''(\sigma^2 t, x(t), z)$$

and so

$$\dot{\rho}(t, s, y) = y^{-1}\left(\sigma^2 \dot{p} + \mu p'\right) = y^{-1}\left(\frac{1}{2}\sigma^2 p'' + \mu p'\right) = \frac{1}{2}\sigma^2 s^2 \rho''(t, s, y) + \left(\mu + \frac{1}{2}\sigma^2\right) s\rho'(t, s, y).$$

Proposition 6.2. *Let F be a continuous function on $(0, \infty)$ of polynomial growth. For $t \in [0, T]$ and $s \in (0, \infty)$, write $V(t, s)$ for the time-t value of the time-T contingent claim $F(S_T)$, conditional on $S_t = s$, as given by the Black–Scholes pricing formula*

$$V(t, s) = e^{-r(T-t)}\mathbb{E}^*(F(S_T)|S_t = s) = e^{-r(T-t)}\mathbb{E}(F(se^{\sigma B_{T-t}+\mu^*(T-t)})).$$

Then V is continuous on $[0, T] \times (0, \infty)$ and $C^{1,2}$ on $[0, T) \times (0, \infty)$ and satisfies the Black–Scholes PDE

$$\mathcal{L}V = \dot{V} + \frac{1}{2}\sigma^2 s^2 V'' + rsV' - rV = 0$$

with terminal value $V(., T) = F$.

Proof. Set

$$v(t, s, y) = e^{-r(T-t)}\rho(T - t, s, y)$$

where ρ is the transition density for the case of drift $\mu^* = r - \sigma^2/2$. Then

$$\dot{v}(t, s, y) = e^{-r(T-t)}(r\rho - \dot{\rho})(T - t, s, y)$$

and

$$\dot{\rho} = \frac{1}{2}\sigma^2 s^2 \rho'' + rs\rho'$$

so

$$\dot{v} = rv - \frac{1}{2}\sigma^2 s^2 v'' - rsv'$$

and so

$$\mathcal{L}v = \dot{v} + \frac{1}{2}\sigma^2 s^2 v'' + rsv' - rv = 0.$$

Now

$$V(t, s) = \int_0^\infty F(y)v(t, s, y)dy$$

and the growth condition on F, combined with super-polynomial decay in $v(t, s, .)$, allows to differentiate under the integral sign to obtain

$$\mathcal{L}V(t, s) = \int_0^\infty F(y)\mathcal{L}v(t, s, y)dy = 0.$$

Finally, continuity of V on $[0, T] \times (0, \infty)$ follows from continuity of Brownian motion by dominated convergence. □

6.3 Binomial approximation to Black–Scholes

Recall Theorem 5.6 on the convergence of random walks to Brownian motion. In the special case where the step distribution m of the random walk is uniform on $\{-1, 1\}$, this guarantees the existence of random processes $(B_t)_{t\geqslant 0}$ and, for each $k \in \mathbb{N}$, $(W_t^{(k)})_{t\geqslant 0}$ such that:

(a) $(B_t)_{t\geqslant 0}$ is a Brownian motion,

(b) $(W^{(k)}_{n/k})_{n\geqslant 0}$ is a simple symmetric random walk on the integers, and $(W^{(k)}_t)_{t\geqslant 0}$ is the linear interpolation of its values on $\{n/k : n \in \mathbb{Z}^+\}$,

(c) $W^{(k)}_t/\sqrt{k} \to B_t$ as $k \to \infty$ uniformly on compacts in t almost surely.

Given $\mu, r \in \mathbb{R}$ and $\sigma \in (0, \infty)$, for k sufficiently large, we can define $a_k < r_k < b_k$ by

$$1 + a_k = \exp\left(-\frac{\sigma}{\sqrt{k}} + \frac{\mu}{k}\right), \quad 1 + b_k = \exp\left(\frac{\sigma}{\sqrt{k}} + \frac{\mu}{k}\right), \quad 1 + r_k = \exp\left(\frac{r}{k}\right).$$

Then, given $S_0 \in (0, \infty)$, set

$$S^{(k)}_t = S_0 \exp\left(\frac{\sigma W^{(k)}_t}{\sqrt{k}} + \mu t\right), \quad S_t = S_0 \exp(\sigma B_t + \mu t), \quad S^{(k)0}_t = S^0_t = \exp(rt).$$

Then

(a) $(S^0_t, S_t)_{t\geqslant 0}$ is a Black–Scholes model of drift μ, volatility σ and interest rate r,

(b) $(S^{(k)0}_{n/k}, S^{(k)}_{n/k})_{n\geqslant 0}$ is a binomial model of parameters $a_k < r_k < b_k$ and $p = 1/2$,

(c) $S^{(k)}_t \to S_t$ as $k \to \infty$ uniformly on compacts in t almost surely.

Hence we can obtain the Black–Scholes model as a limit of binomial models. Moreover, as we now show, the equivalent martingale measure for the binomial model converges to that for the Black–Scholes model.

Fix a time interval $[0, T]$ and a convergent sequence $\lambda_k \to \lambda$ in $\mathbb{R}$. For convenience, we will assume now that T is an integer. Set

$$p_k = \frac{1}{2}\left(1 + \frac{\lambda_k}{\sqrt{k}}\right).$$

We can and do choose k sufficiently large that $p_k \in (0, 1)$. Set

$$U^{(k)}_T = \frac{kT + W^{(k)}_T}{2}, \quad D^{(k)}_T = \frac{kT - W^{(k)}_T}{2}$$

and define

$$Z^{(k)}_T = \left(1 - \frac{\lambda_k}{\sqrt{k}}\right)^{D^{(k)}_T} \left(1 + \frac{\lambda_k}{\sqrt{k}}\right)^{U^{(k)}_T}, \quad Z_T = \exp\left(\lambda B_T - \frac{\lambda^2 T}{2}\right).$$

We can define equivalent probability measures $\tilde{\mathbb{P}}^{(k)}$ and $\tilde{\mathbb{P}}$ on $\mathcal{F}$ by

$$\frac{d\tilde{\mathbb{P}}^{(k)}}{d\mathbb{P}} = Z^{(k)}_T, \quad \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = Z_T.$$

Then, by Proposition 5.7,

(a) under $\tilde{\mathbb{P}}$, $(S_t^0, S_t)_{0\leqslant t\leqslant T}$ is a Black–Scholes model of drift $\mu + \lambda\sigma$, volatility σ and interest rate r,

(b) under $\tilde{\mathbb{P}}^{(k)}$, $(S_{n/k}^{(k)0}, S_{n/k}^{(k)})_{0\leqslant n\leqslant kT}$ is a binomial model of parameters $a_k < r_k < b_k$ and p_k.

The convergence $W_T^{(k)}/\sqrt{k} \to B_T$ allows us to show that $Z_T^{(k)} \to Z_T$ almost surely. Then, since $Z_T^{(k)} \geqslant 0$ and $Z_T \geqslant 0$ and $\mathbb{E}(Z_T^{(k)}) = \mathbb{E}(Z_T) = 1$, we have also $Z_T^{(k)} \to Z_T$ in $L^1(\mathbb{P})$. It follows that, for any bounded continuous map G on $C[0,T]$, as $k \to \infty$,

$$\tilde{\mathbb{E}}^{(k)}\left(G(S^{(k)})\right) \to \tilde{\mathbb{E}}\left(G(S)\right).$$

Define λ_k^* and λ^* by

$$\frac{1}{2}\left(1 + \frac{\lambda_k^*}{\sqrt{k}}\right) = p_k^* = \frac{r_k - a_k}{b_k - a_k}, \quad \mu + \sigma\lambda^* = \mu^* = r - \frac{\sigma^2}{2}$$

and write $\mathbb{P}^{(k)*}$ and $\mathbb{P}^*$ for the corresponding equivalent measures, which are the martingale measures for the binomial model and the Black–Scholes model, respectively. It is left as an exercise to check that $\lambda_k \to \lambda$, so

$$\mathbb{E}^{(k)*}\left(G(S^{(k)})\right) \to \tilde{\mathbb{E}}^*\left(G(S)\right)$$

giving an exact tie-up between the discrete-time and continuous-time theories.

6.4 Computational methods for option prices

The Black–Scholes pricing formula

$$V_0 = e^{-rT}\mathbb{E}^*(C)$$

expresses a quantity we need to know for trading in terms of an integral with respect to Wiener measure. For some important options, there is a closed form expression for V_0. We will derive some such expressions in later sections. For now, we consider some general ways to compute numerically the value of the integral. It will be convenient to assume throughout that $\mu = \mu^* = r - \sigma^2/2$ and drop the stars.

Note first that, while C is typically expressed in terms of the asset price process $(S_t)_{0\leqslant t\leqslant T}$, this can be written in the form $S_t = S_0 e^{\sigma B_t + \mu t}$, where $(B_t)_{t\geqslant 0}$ is a Brownian motion. So, we can assume that $C = F(B)$ for some function F on $C[0,T]$. Indeed, for terminal-value options, we can assume that $C = f(B_T)$ for some function f on $\mathbb{R}$. Hence, in the simplest case, we have

$$V_0 = e^{-rT}\int_{\mathbb{R}} f(\sqrt{T}y)\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy$$

and we could compute V_0 by a trapezoidal approximation.

There are at least two contexts where this will not be an efficient way to proceed. First, going beyond the scope of our prior discussion, we may have multiple assets, leading to a multi-dimensional integral

$$V_0 = e^{-rT} \int_{\mathbb{R}^d} f(\sqrt{T}y) \frac{1}{(2\pi)^{d/2}} e^{-|y|^2/2} dy.$$

The computational cost for such integrals grows exponentially in d. Second, it may be of interest to compute not just the time-0 price for a single initial asset price, but the pricing surface

$$V(t,s) = e^{-r(T-t)} \mathbb{E}^*(C|S_t = s), \quad 0 \leqslant t \leqslant T, \quad s \in (0,\infty).$$

As we will now show, this can often be done using computational methods for the heat equation.

Consider the case of a terminal-value option $C = g(S_T)$. Then

$$V(t,s) = e^{-r(T-t)} \mathbb{E}\left(g(se^{\sigma B_{T-t}+\mu(T-t)})\right) = e^{-r(T-t)} u(T-t, (\log s + \mu(T-t))/\sigma)$$

where

$$u(t,x) = \mathbb{E}_x(f(B_t)), \quad f(x) = g(e^{\sigma x}).$$

Assume that g is continuous on $(0,\infty)$ and of no more than linear growth. Then, by continuity of Brownian motion and dominated convergence, u is continuous on $[0,T] \times \mathbb{R}$. Also

$$u(t,x) = \int_{\mathbb{R}} p_t(x,y) f(y) dy, \quad p_t(x,y) = \frac{1}{\sqrt{2\pi t}} e^{-(y-x)^2/(2t)}$$

and

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}$$

so, by differentiation under the integral sign, u is $C^{1,2}$ on $(0,T] \times \mathbb{R}$ with

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2}.$$

Hence we can attempt to compute u by any standard numerical scheme for the heat equation, such as we will now describe.

Fix some L sufficiently large, chosen so that we can approximate accurately the values $u(t, \pm L)$ for all t, by some other means, assuming this to be possible. Then use a grid

$$\{(ik, jh) : i = 0,1,\ldots,N \text{ and } j = -M,\ldots,M\} \subseteq [0,T] \times [-L,L]$$

where $k = T/N$ and $h = L/M$. We compute values U_j^i at the grid points with the aim that $U_j^i \approx u(ik, jh)$. The values at the inital grid points $(0, jh)$ and the upper and lower boundary grid points are assumed given.

The *FTCS* (*forward-in-time, central-in-space*) method uses

$$\frac{U_j^{i+1} - U_j^i}{k} = \frac{U_{j-1}^i - 2U_j^i + U_{j+1}^i}{2h^2}, \quad i = 0,1,\ldots N-1, \quad |j| \leqslant M-1$$

while the *BTCS* (*backward-in-time, central-in-space*) method uses

$$\frac{U_j^{i+1}-U_j^i}{k}=\frac{U_{j-1}^{i+1}-2U_j^{i+1}+U_{j+1}^{i+1}}{2h^2},\quad i=0,1,\dots N-1,\quad |j|\leqslant M-1.$$

The *Crank–Nicolson* method uses the average of these two equations

$$\frac{U_j^{i+1}-U_j^i}{k}=\frac{1}{2}\left(\frac{U_{j-1}^{i}-2U_j^{i}+U_{j+1}^{i}}{2h^2}+\frac{U_{j-1}^{i+1}-2U_j^{i+1}+U_{j+1}^{i+1}}{2h^2}\right)$$

Note that the FTCS method is explicit, while BTCS and Crank–Nicolson require to solve a $2M\times 2M$ matrix inversion in each time-step. The second two methods have better stability properties than FTCS, while Crank-Nicolson has the merit of being second-order in time, both FTCS and BTCS methods being first order in time.

A second type of computational approach to the pricing formula is Monte Carlo. Having chosen a time-step $k=T/N$, the restriction of $(B_t)_{0\leqslant t\leqslant T}$ to $\{ik:i=0,1,\dots,N\}$ is a random walk with step distribution $N(0,k)$. Write $(B_t^{(N)})_{0\leqslant t\leqslant T}$ for the linear interpolation of $(B_t:t=ik,i=0,1,\dots,N)$. By Theorem 5.6, for any continuous function F on $C[0,T]$, as $N\to\infty$,

$$\mathbb{E}\left(F(B^{(N)})\right)\to\mathbb{E}\left(F(B)\right)$$

as $N\to\infty$. On the other hand, we can generate a random sample $(B^{(N),m}:m=1,\dots,n)$ of random walk paths. Then, provided $F(B^{(N)})$ is integrable, by the law of large numbers, as $n\to\infty$,

$$\frac{1}{n}\sum_{m=1}^{n}F(B^{(N),m})\to\mathbb{E}\left(F(B^{(N)})\right)\quad\text{almost surely.}$$

This is the basis of a computational method for general path-dependent options, and without the need to chose a finite interval $[-L,L]$ and approximate boundary values. We have not addressed how the (approximate) evaluation of $F(B^{(N)})$ could be done numerically.

The same argument applies if we use, in place of $(B_t:t=ik,i=0,1,\dots,N)$, simple symmetric random walk paths $(X_t:t=ik,i=0,1,\dots,N)$ of step-size $h=\sqrt{k}$. For the variance of each step is still k, so Theorem 5.6 still applies. Then, for suitable functions f, we can compute $\mathbb{E}(f(X_T))$ by Monte-Carlo as above. On the other hand, if we set

$$U_j^i=\mathbb{E}_{jh}(f(X_{ik})),\quad i=0,1,\dots,N,\quad j\in\mathbb{Z}$$

then, by conditioning on the first step, we have

$$U_j^{i+1}=\frac{U_{j-1}^i+U_{j+1}^i}{2},\quad i=0,1,\dots N-1,\quad j\in\mathbb{Z}.$$

This system of linear equations provides another way to compute $\mathbb{E}(f(X_T))$. In fact these equations can be rewritten as

$$\frac{U_j^{i+1}-U_j^i}{k}=\frac{U_{j-1}^{i}-2U_j^{i}+U_{j+1}^{i}}{2h^2}.$$

Hence, in fact this simply the FTCS method for the heat equation in the case $h=\sqrt{k}$.

6.5 Black–Scholes formula for the price of a European call

The *European call* of *maturity* T and *strike* $K \geqslant 0$ is the contingent claim $(S_T - K)^+$. This may be thought of as conferring the right but not the obligation to buy one unit of stock at time T at price K. We will compute a formula for the fair value of this claim at time 0, in the Black–Scholes model. The fair value is given by

$$EC(x, K, \sigma, r, T) = \mathbb{E}^* \left(e^{-rT}(S_T - K)^+ \right).$$

Recall that we write Φ for the standard normal distribution function, given by

$$\Phi(a) = \int_{-\infty}^{a} \phi(y)dy, \quad \phi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

and that

$$\bar{\Phi}(a) = 1 - \Phi(a) = \Phi(-a).$$

Proposition 6.3 (Black–Scholes formula)**.** *We have*

$$EC(x, K, \sigma, r, T) = x\Phi(d^+) - e^{-rT} K \Phi(d^-)$$

where

$$d^{\pm} = \frac{\log(x/K) + rT}{\sigma\sqrt{T}} \pm \frac{\sigma\sqrt{T}}{2}.$$

Proof. It suffices to consider the case $\mu = \mu^*$, when $\mathbb{P}^* = \mathbb{P}$. Then $B_T \sim \sqrt{T} B_1$ and $B_1 \sim N(0, 1)$. We compute

$$\begin{aligned} EC(x, K, \sigma, r, T) &= \mathbb{E}\left(e^{-rT}(S_T - K)^+\right) = \mathbb{E}\left((xe^{\sigma B_T - \sigma^2 T/2} - e^{-rT}K)^+\right) \\ &= \int_a^\infty (xe^{\sigma\sqrt{T}y - \sigma^2 T/2} - e^{-rT}K) \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \end{aligned}$$

where a is given by $xe^{\sigma\sqrt{T}a - \sigma^2 T/2} = e^{-rT}K$ so

$$\log(x/K) + \sigma\sqrt{T}a + rT = \sigma^2 T/2$$

and so $a = -d^-$. Hence

$$\begin{aligned} EC(x, T, \sigma, r, T) =& x \int_a^\infty \frac{1}{\sqrt{2\pi}} e^{-(y - \sigma\sqrt{T})^2/2} dy - e^{-rT} K \int_a^\infty \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\ =& x\bar{\Phi}(a - \sigma\sqrt{T}) - e^{-rT} K \bar{\Phi}(a) \\ =& x\Phi(-a + \sigma\sqrt{T}) - e^{-rT} K \Phi(-a) \\ =& x\Phi(d^+) - e^{-rT} K \Phi(d^-). \end{aligned}$$

□

The *European put* of *maturity* $T > 0$ and *strike* $K \geqslant 0$ on $(S_t)_{t\geqslant 0}$ is the contingent claim $(S_T - K)^-$. This may be thought of as conferring the right but not the obligation to sell one unit of stock at time T at price K. In the Black–Scholes model $(S_t^0, S_t)_{t\geqslant 0}$ the fair value of this claim at time 0 is given by

$$EP(x, K, \sigma, r, T) = \mathbb{E}^* \left(e^{-rT}(S_T - K)^- \right).$$

Note the *put-call parity equation*

$$(S_T - K)^+ - (S_T - K)^- = S_T - K.$$

Thus a portfolio long in the call and short in the put is equivalent to the forward contract $S_T - K$ of maturity T and strike K. The fair value of the forward contract is

$$\mathbb{E}^* \left(e^{-rT}(S_T - K) \right) = x - e^{-rT} K.$$

Hence we obtain

$$EP(x, K, \sigma, r, T) = EC(x, K, \sigma, r, T) - x + e^{-rT} K = e^{-rT} K \bar{\Phi}(d_-) - x\bar{\Phi}(d_+).$$

6.6 Sensitivities for contingent claims in the Black–Scholes model

Recall that the value of a contingent claim C at time T in the Black–Scholes model $(S_t^0, S_t)_{t\geqslant 0}$ is given by

$$v(x) = v(x, C, \sigma, r, T) = \mathbb{E}^*(e^{-rT} C).$$

where $\mathbb{E}^*$ denotes expectation with respect to the equivalent martingale measure $\mathbb{P}^*$. The derivatives of v with respect to parameters of the model are known as *sensitivities*. The following terminology is widely used

$$\text{Delta} = \Delta = \frac{\partial v}{\partial x}, \quad \text{Gamma} = \Gamma = \frac{\partial^2 v}{\partial x^2}, \quad \text{Vega} = \mathcal{V} = \frac{\partial v}{\partial \sigma}, \quad \text{Rho} = \rho = \frac{\partial v}{\partial r}.$$

These and other sensitivities are also known as *Greeks*.

In the case of a European call of maturity T and strike K, the Black–Scholes formula can be differentiated explicitly to obtain

$$\Delta = \Phi(d^+), \quad \mathcal{V} = x\phi(d_+)\sqrt{T}.$$

6.7 Implied volatility

Proposition 6.4. *The map* $\sigma \mapsto EC(x, K, \sigma, r, T)$ *is an increasing bijection* $(0, \infty) \to ((x - e^{-rT} K)^+, x)$.

Proof. It is straightforward to check that, for all values of x, K, r and T, we have

$$\lim_{\sigma\to 0} EC(x,K,\sigma,r,T) = (x-e^{-rT})^+, \quad \lim_{\sigma\to\infty} EC(x,K,\sigma,r,T) = x.$$

Since

$$\frac{\partial}{\partial\sigma} EC(x,K,\sigma,r,T) = \mathcal{V} = x\phi(d_+)\sqrt{T} > 0$$

this implies the claim. $\square$

European calls are widely traded in financial markets, so the actual market price may be observed, for a range of maturities T and strikes K. Since the current stock price S_0 and interest rate r are also known, if we accept the Black–Scholes model, this determines a unique *implied volatility* $\sigma_{\text{implied}}(K,T)$ such that

$$EC(S_0, K, \sigma_{\text{implied}}(K,T), r, T) = EC_{\text{market}}(K,T).$$

In principle, $\sigma_{\text{implied}}(K,T)$ should be independent of K and T. In practice, the implied volatility surface is not flat, indicating that the Black–Scholes model is not an exact fit. However, implied volatility, acting as an encoding of the call price $EC_{\text{market}}(K,T)$, is a convenient and widely used means to quote these market prices.

6.8 Pricing of exotic options in the Black–Scholes model

The reflection principle allows us to obtain pricing formulas for certain barrier options in the Black–Scholes model. We illustrate this by the case of the up-and-out call

$$C = h(S_T)1_{\{\sup_{0\leqslant t\leqslant T} S_t < A\}}, \quad h(s) = (s-K)^+$$

where $A \geqslant \max\{S_0, K\}$. The fair price at time 0 for C is given by the Black–Scholes pricing formula

$$V_0 = e^{-rT}\mathbb{E}^*(C)$$

where $\mathbb{P}^*$ is the equivalent martingale measure, corresponding to drift $\mu^* = r - \sigma^2/2$. In calculating V_0, we will assume that $\mu = \mu^*$ and drop the stars. Then

$$S_t = S_0 e^{\sigma B_t + rt - \sigma^2 t/2} = S_0 e^{\sigma \tilde{B}_t}, \quad \tilde{B}_t = B_t + ct, \quad c = (r-\sigma^2/2)/\sigma$$

so

$$\left\{\sup_{0\leqslant t\leqslant T} S_t < A\right\} = \left\{\sup_{0\leqslant t\leqslant T} \tilde{B}_t < a\right\} = \{\tilde{T}_a > t\}, \quad \tilde{T}_a = \inf\{t \geqslant 0 : \tilde{B}_t = a\}$$

where $a > 0$ is determined by $S_0 e^{\sigma a} = A$. We showed in Section 5.6 that

$$\mathbb{E}\left(f(\tilde{B}_t)1_{\{\tilde{T}_a > t\}}\right) = \int_{-\infty}^{a} f(y)e^{cy - c^2 t/2} p_t^a(0,y)dy$$

where

$$p_t^a(0,y) = p_t(0,y) - p_t(0, 2a-y).$$

Hence

$$V_0 = e^{-rT} \int_{-\infty}^{a} h(S_0 e^{\sigma y}) e^{cy - c^2 T/2} p_T^a(0, y) dy.$$

In the special case $h(s) = (s - K)^+$, the right-hand side can be expressed in terms of the normal distribution function Φ by using the identity

$$\int_a^b e^{\lambda y} p_t(0, y) dy = e^{\lambda^2 t/2} \left(\Phi\left(\frac{b - \lambda t}{\sqrt{t}} \right) - \Phi\left(\frac{a - \lambda t}{\sqrt{t}} \right) \right).$$

7 Appendix

Recall the statement of Theorem 3.4 for an asset price model $(\bar{S}_n)_{0\leqslant n\leqslant T}$. *The following are equivalent:*

(a) $(\bar{S}_n)_{0\leqslant n\leqslant T}$ *has no arbitrage,*

(b) $(\bar{S}_n)_{0\leqslant n\leqslant T}$ *has an equivalent martingale measure.*

This is the discrete-time version of the fundamental theorem of asset pricing. We will now prove this statement – the proof is not examinable. The main step is to show the following conditional version of Proposition 3.3.

Proposition 7.1. *Let Y be a random variable in $\mathbb{R}^d$ and let $\mathcal{F}_0$ be a sub-σ-algebra of $\mathcal{F}$. Then the following are equivalent:*

(a) *there exists no $\mathcal{F}_0$-measurable random variable Θ in $\mathbb{R}^d$ such that $\Theta.Y \geqslant 0$ almost surely and $\Theta.Y > 0$ with positive probability,*

(b) *there exists an equivalent probability measure $\tilde{\mathbb{P}}$ such that $\tilde{\mathbb{P}} = \mathbb{P}$ on $\mathcal{F}_0$ and, almost surely,*

$$\tilde{\mathbb{E}}(|Y||\mathcal{F}_0) < \infty, \quad \tilde{\mathbb{E}}(Y|\mathcal{F}_0) = 0.$$

Morever, given any finite non-negative random variable U, we can choose $\tilde{\mathbb{P}}$ in (b) *so that $\tilde{\mathbb{E}}(U|\mathcal{F}_0) < \infty$ almost surely.*

Proof. Suppose that (a) holds. It will suffice to consider the case where, almost surely, for all $\theta \in \mathbb{R}^d$,

$$\mathbb{E}(e^{\theta.Y}|\mathcal{F}_0) < \infty, \quad \mathbb{E}(Ue^{\theta.Y}|\mathcal{F}_0) < \infty.$$

For the general case, we can replace $\mathbb{P}$ by the equivalent probability measure $\tilde{\mathbb{P}}$ given by $d\tilde{\mathbb{P}}/d\mathbb{P} = W/\mathbb{E}(W|\mathcal{F}_0)$ where $W = e^{-|Y|^2}e^{-U}$ for which $\tilde{\mathbb{P}} = \mathbb{P}$ on $\mathcal{F}_0$ and, almost surely,

$$\tilde{\mathbb{E}}(e^{\theta.Y}|\mathcal{F}_0) < \infty, \quad \tilde{\mathbb{E}}(Ue^{\theta.Y}|\mathcal{F}_0) < \infty.$$

We will assume that this has been done and drop the tildes.

There are $\mathcal{F}_0$-measurable kernels M and N on $\Omega \times \mathcal{B}(\mathbb{R}^d)$ such that, for all $B \in \mathcal{B}(\mathbb{R}^d)$, almost surely,

$$M(B) = \mathbb{P}(Y \in B|\mathcal{F}_0), \quad N(B) = \mathbb{E}(U1_{\{Y\in B\}}|\mathcal{F}_0).$$

We write here $M(B)$ for the random variable $M(.,B)$ and will use similar notation elsewhere. Define $\mathcal{F}_0$-measurable fields of subspaces of $\mathbb{R}^d$ by

$$E_0(\omega) = \{\theta \in \mathbb{R}^d : M(\omega, \{\theta.y = 0\}) = 1\}, \quad E_1(\omega) = \operatorname{span}(\operatorname{supp} M(\omega, .)) = E_0(\omega)^\perp.$$

Write $P(\omega)$ for the orthogonal projection $\mathbb{R}^d \to E_1(\omega)$. Fix $\theta \in \mathbb{R}^d$ and consider the event

$$B(\theta) = \{P\theta \neq 0\} = \{\theta \notin E_0\}.$$

Fix $B \in \mathcal{F}_0$ with $B \subseteq B(\theta)$ and $\mathbb{P}(B) > 0$. Consider the $\mathcal{F}_0$-measurable random variable $\Theta = 1_B\theta$. Define

$$W = M(\{y \in \mathbb{R}^d : \theta.y \neq 0\}), \quad W' = M(\{y \in \mathbb{R}^d : \theta.y > 0\}).$$

Then, almost surely,

$$W = \mathbb{P}(\theta.Y \neq 0|\mathcal{F}_0), \quad W' = \mathbb{P}(\theta.Y > 0|\mathcal{F}_0).$$

Now $W > 0$ on $B(\theta)$, so $\mathbb{P}(\Theta.Y \neq 0) = \mathbb{E}(1_B W) > 0$. Condition (a) then implies that $\mathbb{E}(1_B W') = \mathbb{P}(\Theta.Y > 0) > 0$ so, since B is arbitrary, $W' > 0$ almost surely on $B(\theta)$.

Set $\psi(t) = 0 \vee t \wedge 1$ and define $F : \Omega \times \mathbb{R}^d \to [0, \infty)$ by

$$F(\omega, \theta) = \int_{\mathbb{R}^d} \psi(\theta.y) M(\omega, dy).$$

Then $F(\omega, .)$ is continuous on $\mathbb{R}^d$ for all ω, $F(\theta) = \mathbb{E}(\psi(\theta.Y)|\mathcal{F}_0)$ almost surely, and $F(\theta) > 0$ almost surely on $B(\theta)$ for all θ. Hence, by Fubini's theorem, $F(\theta) > 0$ for Lebesgue almost all $\theta \in \mathbb{R}^d \setminus E_0$, almost surely. Since F is continuous on $\mathbb{R}^d$ and $\mathbb{R}^d \setminus E_0$ is open, this implies that $F(\theta) > 0$ for all $\theta \in \mathbb{R}^d \setminus E_0$ almost surely. Set $S(\omega) = \{\theta \in E_1(\omega) : |\theta| = 1\}$. Then $S(\omega)$ is compact. Define

$$\varepsilon(\omega) = \frac{1}{2} \inf\{F(\omega, \theta) : \theta \in S(\omega)\}, \quad \Omega_0 = \{\omega \in \Omega : \varepsilon(\omega) > 0\}.$$

Then ε is $\mathcal{F}_0$-measurable, $\mathbb{P}(\Omega_0) = 1$ and, for all $\omega \in \Omega_0$ and $\theta \in S(\omega)$,

$$\begin{aligned}\int_{\mathbb{R}^d} 1_{\{\theta.y \geqslant \varepsilon(\omega)\}} M(\omega, dy) &\geqslant \int_{\mathbb{R}^d} \psi(\theta.y - \varepsilon(\omega)) M(\omega, dy) \\ &\geqslant \int_{\mathbb{R}^d} \psi(\theta.y) M(\omega, dy) - \varepsilon(\omega) = F(\omega, \theta) - \varepsilon(\omega) \geqslant \varepsilon(\omega).\end{aligned}$$

Define functions $\Phi : \Omega \times \mathbb{R}^d \to (0, \infty)$ and $\Psi : \Omega \times \mathbb{R}^d \to [0, \infty)$ by

$$\Phi(\omega, \theta) = \int_{\mathbb{R}^d} e^{\theta.y} M(\omega, dy), \quad \Psi(\omega, \theta) = \int_{\mathbb{R}^d} e^{\theta.y} N(\omega, dy).$$

Then Φ is $\mathcal{F}_0$-measurable on Ω, differentiable on $\mathbb{R}^d$, and strictly convex on E_1, with

$$\Phi'(\omega, \theta) = \int_{\mathbb{R}^d} y e^{\theta.y} M(\omega, dy)$$

and, almost surely,

$$\Phi(\theta) = \mathbb{E}(e^{\theta.Y}|\mathcal{F}_0), \quad \Phi'(\theta) = \mathbb{E}(Y e^{\theta.Y}|\mathcal{F}_0).$$

Similarly, Ψ is $\mathcal{F}_0$-measurable on Ω, continuous on $\mathbb{R}^d$ and, almost surely,

$$\Psi(\theta) = \mathbb{E}(U e^{\theta.Y}|\mathcal{F}_0).$$

Then, for $\omega \in \Omega_0$, $\theta_0 \in E_0(\omega)$, $\theta_1 \in S(\omega)$ and $t \geqslant (1/\varepsilon(\omega))\log(1/\varepsilon(\omega))$,

$$\Phi(\omega, \theta_0 + t\theta_1) = \Phi(\omega, t\theta_1) \geqslant \varepsilon e^{t\varepsilon} \geqslant 1 = \Phi(\omega, 0).$$

Hence, for all $\omega \in \Omega_0$, there exists a unique $\Theta^*(\omega) \in E_1(\omega)$ such that $\Phi(\omega, \theta) \geqslant \Phi(\omega, \Theta^*(\omega))$ for all $\theta \in \mathbb{R}^d$. Then $\Phi'(\omega, \Theta^*(\omega)) = 0$. Set $\Phi^*(\omega) = 0$ for $\omega \notin \Omega_0$. Define recursively a sequence of $\mathcal{F}_0$-measurable random variables $(\Theta_n : n \geqslant 0)$ in $\mathbb{R}^d$ by

$$\Theta_0 = 0, \quad \Theta_{n+1} = \Theta_n - \gamma\Phi'(\Theta_n).$$

Then $\Theta_n \in E_1$ for all n almost surely, so by the properties of gradient descent, for γ sufficiently small, we have $\Theta_n \to \Theta^*$ almost surely. Hence Θ^* is an $\mathcal{F}_0$-measurable random variable.

Define an equivalent probability measure $\tilde{\mathbb{P}}$ by $d\tilde{\mathbb{P}}/d\mathbb{P} \propto e^{\Theta^*.Y}$. Then, almost surely,

$$\tilde{\mathbb{E}}(U|\mathcal{F}_0) = \mathbb{E}(Ue^{\Theta^*.Y}|\mathcal{F}_0) = \Psi(\Theta^*) < \infty$$

and similarly $\tilde{\mathbb{E}}(|Y||\mathcal{F}_0) < \infty$ almost surely. For $\omega \in \Omega_0$, since $\Phi(\omega, .)$ is minimized at $\Theta^*(\omega)$, we have $\Phi'(\omega, \Theta^*(\omega)) = 0$. So, for all $B \in \mathcal{F}_0$,

$$\mathbb{E}(Y1_B e^{\Theta^*.Y}|\mathcal{F}_0) = 1_B\mathbb{E}(Ye^{\Theta^*.Y}|\mathcal{F}_0) = 1_B\Phi'(\Theta^*) = 0$$

and so

$$\tilde{\mathbb{E}}(Y1_B) = \frac{\mathbb{E}(Y1_B e^{\Theta^*.Y})}{\mathbb{E}(e^{\Theta^*.Y})} = 0.$$

Hence $\tilde{\mathbb{E}}(Y|\mathcal{F}_0) = 0$ almost surely and (b) holds.

Suppose on the other hand that (b) holds and let Θ be an $\mathcal{F}_0$-measurable random variable in $\mathbb{R}^d$. Then, almost surely,

$$\tilde{\mathbb{E}}(\Theta.Y|\mathcal{F}_0) = \Theta.\tilde{\mathbb{E}}(Y|\mathcal{F}_0) = 0$$

and so $\tilde{\mathbb{E}}(\Theta.Y) = 0$. Hence, if $\Theta.Y \geqslant 0$ almost surely, then $\Theta.Y = 0$ almost surely. Hence (a) holds. □

Proof of Theorem 3.4. Suppose that (a) holds. Write $Y_n = X_n - X_{n-1}$. Consider the following hypothesis: there is a random variable ρ_n such that, almost surely, $\rho_n > 0$ and $\mathbb{E}(\rho_n|\mathcal{F}_{n-1}) = 1$ and $U_n < \infty$ and $\mathbb{E}(Y_k\rho_n|\mathcal{F}_{k-1}) = 0$ for $k = n, \dots, T$, where we write U_n for a version of

$$\sum_{k=n}^{T} \mathbb{E}(|Y_k|\rho_n|\mathcal{F}_{n-1}).$$

The hypothesis holds for $n = T+1$ with $\rho_{T+1} = 1$, the third and fourth conditions holding vacuously. Suppose, for a reverse induction, that the hypothesis holds for $n+1$. Given an $\mathcal{F}_{n-1}$-measurable random variable Θ in $\mathbb{R}^d$ such that $\Theta.Y_n \geqslant 0$ almost surely, we obtain a previsible process $(\theta_1, \dots, \theta_T)$ in $\mathbb{R}^d$ by setting $\theta_k = \Theta 1_{\{k=n\}}$. The associated value process satisfies

$$V_T = \sum_{k=1}^{T} \theta_k.Y_k = \Theta.Y_n \geqslant 0.$$

The no-arbitrage condition (a) then forces $\Theta.Y_n = 0$ almost surely. Hence, condition (a) of Proposition 7.1 holds for the $\mathcal{F}_n$-measurable random variable Y_n and the sub-σ-algebra $\mathcal{F}_{n-1}$. Then, by Proposition 7.1, there is an $\mathcal{F}_n$-measurable random variable α_n such that, almost surely, $\alpha_n > 0$ and $\mathbb{E}(\alpha_n|\mathcal{F}_{n-1}) = 1$ and

$$\mathbb{E}(|Y_n|\alpha_n|\mathcal{F}_{n-1}) < \infty, \quad \mathbb{E}(Y_n\alpha_n|\mathcal{F}_{n-1}) = 0$$

and moreover we can choose α_n so that $\mathbb{E}(U_n\alpha_n|\mathcal{F}_{n-1}) < \infty$ amost surely. Define

$$\rho_n = \alpha_n \rho_{n+1}.$$

Then, almost surely, $\rho_n > 0$ and

$$\mathbb{E}(\rho_n|\mathcal{F}_{n-1}) = \mathbb{E}(\alpha_n\mathbb{E}(\rho_{n+1}|\mathcal{F}_n)|\mathcal{F}_{n-1}) = \mathbb{E}(\alpha_n|\mathcal{F}_{n-1}) = 1$$

and

$$U_n = \mathbb{E}(U_{n+1}\alpha_n|\mathcal{F}_{n-1}) + \mathbb{E}(|Y_n|\alpha_n|\mathcal{F}_{n-1}) < \infty$$

and, for $k = n+1, \dots, T$,

$$\mathbb{E}(Y_k\rho_n|\mathcal{F}_{k-1}) = \rho_n\mathbb{E}(Y_k|\mathcal{F}_{k-1}) = 0$$

while

$$\mathbb{E}(Y_n\rho_n|\mathcal{F}_{n-1}) = \mathbb{E}(Y_n\alpha_n\mathbb{E}(\rho_{n+1}|\mathcal{F}_n)|\mathcal{F}_{n-1}) = \mathbb{E}(Y_n\alpha_n|\mathcal{F}_{n-1}) = 0.$$

Hence the hypothesis holds for n and the induction proceeds.

Since the random variables $\mathbb{E}(|Y_k|\rho_1|\mathcal{F}_0)$ are all finite-valued and $\mathcal{F}_0$-measurable, and so is X_0, there is an $\mathcal{F}_0$-measurable random variable α_0 such that $\alpha_0 > 0$ almost surely, $\mathbb{E}(\alpha_0) = 1$ and

$$\alpha_0 \left(|X_0| + \sum_{k=1}^{T} \mathbb{E}(|Y_k|\rho_1|\mathcal{F}_0) \right)$$

is integrable. We can define an equivalent probability measure $\tilde{\mathbb{P}}$ by $d\tilde{\mathbb{P}}/d\mathbb{P} = \alpha_0\rho_1$. Then, under $\tilde{\mathbb{P}}$, the random variable $|X_0| + |Y_1| + \cdots + |Y_T|$ is integrable and and $\tilde{\mathbb{E}}(Y_k|\mathcal{F}_{k-1}) = 0$ for $k = 1, \dots, T$. Hence $(X_n)_{0 \leqslant n \leqslant T}$ is a martingale under $\tilde{\mathbb{P}}$. Hence (b) holds. The reverse implication is the content of Proposition 3.2. $\square$