Mathematical Foundations for Finance

Chapter I: Financial Markets in Finite Discrete Time

Prof. Dr. Beatrice Acciaio
beatrice.acciaio@math.ethz.ch

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Table of contents

Basic Probabilistic Concepts

General Setting

We start with a *probability space* $(\Omega, \mathcal{F}, P)$, which consists of

- Ω, arbitrary *state space*
- $\mathcal{F}$, *σ-algebra* on Ω
- $P : \mathcal{F} \to [0, 1]$, *probability measure*

Time evolves in discrete steps over a finite horizon. We label *trading dates* by $k = 0, 1, \ldots, T$.

We start with a *probability space* $(\Omega, \mathcal{F}, P)$, which consists of

- Ω, arbitrary *state space*
- $\mathcal{F}$, *σ-algebra* on Ω
- $P : \mathcal{F} \to [0, 1]$, *probability measure*

Time evolves in discrete steps over a finite horizon. We label *trading dates* by $k = 0, 1, \ldots, T$.

The flow of information over time is described by a *filtration* $\mathbb{F} = (\mathcal{F}_k)_{k=0,\ldots,T}$. This is a family of σ-algebras $\mathcal{F}_k \subseteq \mathcal{F}$ which is increasing, that is, $\mathcal{F}_k \subseteq \mathcal{F}_\ell$ for $k \leq \ell$.

Intuition: $\mathcal{F}_k$ contains all events that are known/observable up to and including time k.

We will use the concept of discrete ($\mathbb{R}^d$-valued) *stochastic process* $X = (X_k)_{k=0,\dots,T}$, which consists of a family of random variables X_k's defined on the same probability space $(\Omega, \mathcal{F}, P)$.

Intuition: This describes the random evolution over time of $d \in \mathbb{N}$ quantities, for example bank accounts, asset prices, or the holdings in a portfolio of assets.

We will use the concept of discrete ($\mathbb{R}^d$-valued) *stochastic process* $X = (X_k)_{k=0,\dots,T}$, which consists of a family of random variables X_k's defined on the same probability space $(\Omega, \mathcal{F}, P)$.

Intuition: This describes the random evolution over time of $d \in \mathbb{N}$ quantities, for example bank accounts, asset prices, or the holdings in a portfolio of assets.

A stochastic process X is called

- ($\mathbb{F}$-)*adapted* if each X_k is $\mathcal{F}_k$-measurable
- ($\mathbb{F}$-)*predictable* if each X_k is $\mathcal{F}_{k-1}$-measurable.

Exercise. If you are not already, you should get familiar with the notion of measurability (see Chapter 8 in the lecture notes).

Example - Multiplicative Model

Consider random variables $r_1, \dots, r_T$ s.t. $r_k > -1$ P-a.s. for all k, and define the process $\widetilde{S}^0$ by

$$\widetilde{S}^0_k := \prod_{j=1}^{k}(1 + r_j), \qquad k = 0, 1, \dots, T.$$

(Here and throughout the course we use the convention that an empty product equals 1, and an empty sum equals 0.)

Then we have

$$\widetilde{S}^0_0 = 1 \quad \text{and} \quad \frac{\widetilde{S}^0_k}{\widetilde{S}^0_{k-1}} = 1 + r_k \;\; \forall\; k = 1, \dots T.$$

Interpretation: the random variable r_k describes the *(simple) interest rate* for the period $(k-1, k]$. So $\widetilde{S}^0$ models a *bank account* with that interest rate evolution, and $r_k > -1$ ensures that $\widetilde{S}^0 > 0$, in the sense that $\widetilde{S}^0_k > 0$ P-a.s. for $k = 0, \dots, T$.

Example - Multiplicative Model

Let now consider random variables $Y_1, \dots, Y_T$ s.t. $Y_k > 0$ P-a.s. for all k. Take a constant $S_0^1 > 0$ and define the process $\widetilde{S}^1$ by

$$\widetilde{S}_k^1 := S_0^1 \prod_{j=1}^{k} Y_j, \qquad k = 0, 1, \dots, T.$$

Then we have

$$\widetilde{S}_0^1 = S_0^1 \quad \text{and} \quad \frac{\widetilde{S}_k^1}{\widetilde{S}_{k-1}^1} = Y_k \;\; \forall \; k = 1, \dots T.$$

Interpretation: $\widetilde{S}^1$ models a *stock*, and Y_k is the *growth factor* for the time period $(k-1, k]$. (Of course, we could strengthen the analogy by writing $Y_k = 1 + R_k$, then $R_k > -1$ would describe the (simple) return on the stock for the period $(k-1, k]$.)

Consider a market consisting of the above defined

- bank account $\widetilde{S}^0$
- and stock $\widetilde{S}^1$

Which *filtration* should we use? The most usual choice for $\mathbb{F}$ is the filtration generated by Y, that is

$$\mathcal{F}_k = \sigma(Y_1, \dots, Y_k) = \sigma(\widetilde{S}^1_0, \widetilde{S}^1_1, \dots, \widetilde{S}^1_k),$$

the smallest σ-field that makes all stock prices up to time k observable.

Somme comments on this choice will follow.

With this choice of filtration, the process $\widetilde{S}^1$ is adapted to $\mathbb{F}$.

The bank account is naturally *less risky* than a stock. In particular, the interest rate for the period $(k-1, k]$ is usually already known at time $k-1$. Therefore, r_k needs to be $\mathcal{F}_{k-1}$-measurable, that means the process $r = (r_k)_{k=1,\dots,T}$ should be predictable, eqv. $\widetilde{S}^0$ should be predictable.

Then, with the above choice for $\mathbb{F}$, the interest rate r_k for the period $(k-1, k]$ only depends on $Y_1, \dots, Y_{k-1}$ or equivalently on the stock prices $\widetilde{S}^1_0, \widetilde{S}^1_1, \dots, \widetilde{S}^1_{k-1}$, but not on other factors.

Suppose all r_k are equal to a constant $r > -1$. Then the bank account evolves as

$$\widetilde{S}_k^0 = (1+r)^k \quad \text{for } k = 0, \dots, T.$$

Assume that $Y_1, \dots, Y_T$ are independent and only take two values, $1+u$ with probability p, and $1+d$ with probability $1-p$, where $d < 0 < u$. Then the stock price moves either up or down at each step k:

$$\frac{\widetilde{S}_k^1}{\widetilde{S}_{k-1}^1} = Y_k = \begin{cases} 1+u & \text{with probability } p, \\ 1+d & \text{with probability } 1-p. \end{cases} \tag{1}$$

This is known as the *Cox–Ross–Rubinstein (CRR) binomial model*.

In the general multiplicative model, one could also start with the filtration

$$\mathcal{F}'_k := \sigma(Y_1, \ldots, Y_k, r_1, \ldots, r_k) = \sigma\big(\widetilde{S}^1_0, \widetilde{S}^1_1, \ldots, \widetilde{S}^1_k, \widetilde{S}^0_0, \widetilde{S}^0_1, \ldots, \widetilde{S}^0_k\big)$$

generated by both Y and r, or equivalently by both assets $\widetilde{S}^0$ and $\widetilde{S}^1$. But since we assume that the process r (or $\widetilde{S}^0$) is predictable, one can show by induction (exercise for you) that

$$\mathcal{F}'_k = \sigma(Y_1, \ldots, Y_k) = \mathcal{F}_k \quad \text{for all } k.$$

This explains why we have started directly with $\mathbb{F}$ generated by Y.

Financial Markets and Trading (FMT)

As before, we work on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,\dots,T}$ for some $T \in \mathbb{N}$. We will only be more specific when exploiting special properties of a particular model $(\Omega, \mathcal{F}, \mathbb{F}, P)$.

We sometimes assume that $\mathcal{F}_0$ is *(P-)trivial*, that means that $P[A] \in \{0, 1\}$ for all $A \in \mathcal{F}_0$. This also means that any $\mathcal{F}_0$-measurable random variable is P-a.s. constant, hence, there is only trivial information at time 0.

For notational convenience, we sometimes also assume that $\mathcal{F} = \mathcal{F}_T$, this means that any event is observable by time T.

The basic *asset prices* in our financial market are specified by a strictly positive adapted process $\widetilde{S}^0 = (\widetilde{S}^0_k)_{k=0,1,\dots,T}$ and an $\mathbb{R}^d$-valued adapted process $\widetilde{S} = (\widetilde{S}_k)_{k=0,1,\dots,T}$ (so we have $d+1$ tradable assets).

Interpretation:

- $\widetilde{S}^0$ models a *reference asset* or *numéraire*. This explains why we assume that $\widetilde{S}^0_0 = 1$ and $\widetilde{S}^0$ is strictly positive, that is $\widetilde{S}^0_k > 0$ P-a.s. for all k. In many cases, we think of $\widetilde{S}^0$ as a *bank account* and then in addition also assume that $\widetilde{S}^0$ is predictable.

- $\widetilde{S} = \{\widetilde{S}^i\}_{i=1,\dots,d}$ describes the prices of d genuinely *risky assets* (often called *stocks*). So $\widetilde{S}^i_k$ is the price of asset i at time k, and because this becomes known at time k and not earlier, each $\widetilde{S}^i$ and hence also the vector process $\widetilde{S}$ is adapted. For financial reasons one might want $\widetilde{S}^i_k \geq 0$ P-a.s. for all i and k, but mathematically this is not needed.

We will express everything in units of the reference asset $\widetilde{S}^0$. We call this procedure "*discounting with* $\widetilde{S}^0$" or "*using* $\widetilde{S}^0$ *as numéraire*". Mathematically, this amounts to dividing at each time k every traded quantity by $\widetilde{S}^0_k$. So the *discounted price of the reference asset* is

$$S^0_k := \widetilde{S}^0_k / \widetilde{S}^0_k = 1 \quad \text{at all times}$$

and the *discounted asset prices* $S = (S_k)_{k=0,1,\dots,T}$ are given by

$$S_k := \widetilde{S}_k / \widetilde{S}^0_k \,.$$

If $\widetilde{S}^0$ is viewed as a bank account, then, in terms of interest rates, using discounted prices is equivalent to working with *zero interest*. Hence, our basic (discounted) model has $S^0 \equiv 1$, and we call asset 0 the bank account.

We assume that we have a *frictionless financial market*, that means:

- There are *no transaction costs* so that assets can be bought or sold at the same price,
- Money can be borrowed or lent at the same (zero) interest rate,
- Assets are available in arbitrarily small or large quantities,
- There are *no constraints* on the numbers of assets one holds, and in particular, one may decide to own a negative number of shares (known as 'short selling'),
- *Investors* are *small* so that their trading activities have no effect on asset prices (which means that S is an exogenously and a priori given and fixed stochastic process).

All this is of course unrealistic, but for explaining and understanding basic concepts, one has to start with the simplest case, and this is in many cases at least a reasonable first approximation.

Definition. A *trading strategy* is an $\mathbb{R}^{d+1}$-valued stochastic process $\varphi = (\varphi^0, \vartheta)$, where $\varphi^0 = (\varphi^0_k)_{k=0,1,\dots,T}$ is real-valued and adapted, and $\vartheta = (\vartheta_k)_{k=0,1,\dots,T}$ with $\vartheta_0 = 0$ is $\mathbb{R}^d$-valued and predictable. The *(discounted) value process* of a strategy φ is the real-valued adapted process $V(\varphi) = (V_k(\varphi))_{k=0,1,\dots,T}$ given by

$$V_k(\varphi) := \varphi^0_k S^0_k + \vartheta^{\mathrm{tr}}_k S_k = \varphi^0_k + \sum_{i=1}^{d} \vartheta^i_k S^i_k \qquad \text{for } k = 0, 1, \dots, T.$$

(For the scalar product we write $\vartheta^{\mathrm{tr}}_k S_k$ rather than $\vartheta_k \cdot S_k$ for convenience.)

Interpretation: A trading strategy describes a *dynamically evolving portfolio* in the $d+1$ basic assets available for trade. At time k, we have φ^0_k units of the bank account and ϑ^i_k units (shares) of asset (stock) i, so that straightforward financial book-keeping gives the above time-k value, in units of the bank account, of the portfolio holdings.

Trading Strategy - Intuition

- The vector φ_k is the portfolio with which we *arrive* at time k.
- ϑ_k are the asset holdings on $[k-1,k)$, so ϑ_k is $\mathcal{F}_{k-1}$-measurable, hence ϑ is predictable.
- φ_k^0 are the bank account holdings on $[k-1,k)$; but as the bank account is riskless (at least locally for each time step, by predictability), one can allow φ^0 to be adapted without giving investors any extra advantages. So φ_k^0 can be $\mathcal{F}_k$-measurable.

Trading Strategy - Intuition

With the above interpretation, we arrive at time k with the portfolio $\varphi_k = (\varphi_k^0, \vartheta_k)$ and change this at time k to a new portfolio $\varphi_{k+1} = (\varphi_{k+1}^0, \vartheta_{k+1})$ with which we then keep till date $k+1$. Hence $V_k(\varphi)$ is more precisely the *pre-trade value* of the strategy φ at time k.

As there are no activities before time 0, we demand via $\vartheta_0 = 0$ that investors start out without any shares. At time 0 they are endowed with an initial amount $V_0(\varphi) = \varphi_0^0$ in the reference asset or bank account.

Remark. If the numéraire $\widetilde{S}^0$ is just strictly positive and adapted, but not necessarily predictable, then also φ^0 must be predictable. We shall see later that this is automatically satisfied if the strategy φ is *self-financing*.

With the above interpretation, we arrive at time k with the portfolio $\varphi_k = (\varphi_k^0, \vartheta_k)$ and change this at time k to a new portfolio $\varphi_{k+1} = (\varphi_{k+1}^0, \vartheta_{k+1})$ with which we then keep till date $k+1$. Hence $V_k(\varphi)$ is more precisely the *pre-trade value* of the strategy φ at time k.

As there are no activities before time 0, we demand via $\vartheta_0 = 0$ that investors start out without any shares. At time 0 they are endowed with an initial amount $V_0(\varphi) = \varphi_0^0$ in the reference asset or bank account.

Remark. If the numéraire $\widetilde{S}^0$ is just strictly positive and adapted, but not necessarily predictable, then also φ^0 must be predictable. We shall see later that this is automatically satisfied if the strategy φ is *self-financing*.

Notation. For any stochastic process $X = (X_k)_{k=0,1,\dots,T}$, we denote the *increment* from $k-1$ to k of X by

$$\Delta X_k := X_k - X_{k-1}.$$

We work out the *costs* associated with a trading strategy $\varphi = (\varphi^0, \vartheta)$.

The *initial cost* for φ comes from putting φ_0^0 into the bank account, so

$$C_0(\varphi) = \varphi_0^0 = V_0(\varphi). \tag{2}$$

Then, transactions occur at each date $k = 0, 1, \dots, T-1$, when φ_k is changed to φ_{k+1}. So the *incremental cost* for φ over the time interval $(k, k+1]$ occurs at time k (when we change from φ_k to φ_{k+1}) at the time-k prices S_k, and it is given by (post-trade - pre-trade value)

$$\begin{aligned} \Delta C_{k+1}(\varphi) &:= (\varphi_{k+1}^0 - \varphi_k^0) S_k^0 + (\vartheta_{k+1} - \vartheta_k)^{\mathrm{tr}} S_k \\ &= \varphi_{k+1}^0 - \varphi_k^0 + \sum_{i=1}^{d} (\vartheta_{k+1}^i - \vartheta_k^i) S_k^i. \end{aligned} \tag{3}$$

Thus

$$C_k(\varphi) = \varphi_k^0 + \sum_{j=1}^{k} \sum_{i=1}^{d} (\vartheta_j^i - \vartheta_{j-1}^i) S_{j-1}^i \; .$$

Note that this is all in units of the bank account.

Elementary rewriting of (3), by adding and subtracting $\vartheta_{k+1}^{\mathrm{tr}} S_{k+1}$, leads to

$$\begin{aligned} \Delta C_{k+1}(\varphi) &= \varphi_{k+1}^0 - \varphi_k^0 + (\vartheta_{k+1} - \vartheta_k)^{\mathrm{tr}} S_k \\ &= \varphi_{k+1}^0 + \vartheta_{k+1}^{\mathrm{tr}} S_{k+1} - \varphi_k^0 - \vartheta_k^{\mathrm{tr}} S_k - \vartheta_{k+1}^{\mathrm{tr}} (S_{k+1} - S_k) \\ &= V_{k+1}(\varphi) - V_k(\varphi) - \vartheta_{k+1}^{\mathrm{tr}} \Delta S_{k+1} \\ &= \Delta V_{k+1}(\varphi) - \vartheta_{k+1}^{\mathrm{tr}} \Delta S_{k+1}, \end{aligned} \tag{4}$$

which gives

$$C_k(\varphi) = V_k(\varphi) - \sum_{j=1}^{k} \vartheta_j^{\mathrm{tr}} \Delta S_j \qquad \text{for } k = 0, 1, \dots, T.$$

Note: $\vartheta^{\mathrm{tr}}_{k+1} \Delta S_{k+1}$ is the *(discounted) incremental gain or loss* arising over $(k, k+1]$ from our trading strategy due to the price fluctuations of S. There is no such gain or loss from the bank account, since $S^0 \equiv 1$.

Note: $\vartheta_{k+1}^{\mathrm{tr}}\Delta S_{k+1}$ is the *(discounted) incremental gain or loss* arising over $(k, k+1]$ from our trading strategy due to the price fluctuations of S. There is no such gain or loss from the bank account, since $S^0 \equiv 1$.

Definition. Let $\varphi = (\varphi^0, \vartheta)$ be a trading strategy. The *(discounted) gains process* associated to φ or to ϑ is the real-valued adapted process $G(\vartheta)$ given by

$$G_k(\vartheta) := \sum_{j=1}^{k} \vartheta_j^{\mathrm{tr}}\Delta S_j \qquad \text{for } k = 0, 1, \dots, T\,, \tag{5}$$

where $G_0(\vartheta) = 0$ by convention. The *(discounted) cost process* of φ can be then defined by

$$C_k(\varphi) := V_k(\varphi) - G_k(\varphi) \qquad \text{for } k = 0, 1, \dots, T, \tag{6}$$

as justified by (2) and (4).

If we think of a continuous-time model with infinitely close successive trading dates, then the increment ΔS in (5) becomes a differential $\mathrm{d}S$ and the sum becomes an integral. This is why the *stochastic integral* $G(\vartheta) = \int \vartheta \, \mathrm{d}S$ provides the natural description of gains from trade in a continuous-time financial market model.

By construction, $C_k(\varphi)$ describes the *cumulative (total) costs* for the strategy φ on the time interval $[0, k]$. If we do not want to worry how to pay these costs, we ideally try to make sure they never occur, by imposing this as a condition on φ. This motivates the next definition.

Definition. A trading strategy $\varphi = (\varphi^0, \vartheta)$ is called *self-financing* if its cost process $C(\varphi)$ is constant over time (and hence equal to the initial cost $C_0(\varphi) = V_0(\varphi) = \varphi_0^0$).

Due to (3), a strategy is self-financing if and only if it satisfies for each k

$$\varphi_{k+1}^0 - \varphi_k^0 + (\vartheta_{k+1} - \vartheta_k)^{\mathrm{tr}} S_k = \Delta C_{k+1}(\varphi) = 0 \qquad P\text{-a.s.} \qquad (7)$$

From economic intuition, this means that changing the portfolio from φ_k to φ_{k+1} at time k is done cost-neutrally, that means with zero gains or losses at that time ($\varphi_k^0 + \vartheta_k^{\mathrm{tr}} S_k = \varphi_{k+1}^0 + \vartheta_{k+1}^{\mathrm{tr}} S_k$).

In particular, all losses from the portfolio due to stock price changes must be fully compensated by gains from the bank account holdings and vice versa, without adding or withdrawing money.

Due to (6), and since $G_0(\vartheta) = 0$, another *equivalent description* of a self-financing strategy $\varphi = (\varphi^0, \vartheta)$ is that it satisfies

$$V(\varphi) = V_0(\varphi) + G(\vartheta) = \varphi_0^0 + G(\vartheta) \tag{8}$$

(in the sense that $V_k(\varphi) = V_0(\varphi) + G_k(\vartheta)$ P-a.s. for each k). This gives the following very useful result.

Proposition. Any self-financing trading strategy $\varphi = (\varphi^0, \vartheta)$ is uniquely determined by its initial wealth $V_0(\varphi)$ and its "risky asset component" ϑ. In particular, any pair (V_0, ϑ), where V_0 is an $\mathcal{F}_0$-measurable random variable and ϑ is an $\mathbb{R}^d$-valued predictable process with $\vartheta_0 = 0$, specifies in a unique way a self-financing strategy. We sometimes write $\varphi \mathrel{\widehat{=}} (V_0, \vartheta)$ for the resulting strategy φ.

Moreover, if $\varphi = (\varphi^0, \vartheta)$ is self-financing, then $(\varphi_k^0)_{k=1,\dots,T}$ is automatically predictable.

Proof of the Proposition. By (8) (or directly from the definitions of self-financing and of $C(\varphi)$ in (6)), a strategy φ is self-financing if and only if, for each k,

$$V_k(\varphi) = V_0(\varphi) + G_k(\vartheta) \qquad P\text{-a.s..}$$

Because $V_k(\varphi) = \varphi_k^0 + \vartheta_k^{\mathrm{tr}} S_k$ by definition, we can solve the above equation for φ_k^0 to get

$$\varphi_k^0 = V_0(\varphi) + G_k(\vartheta) - \vartheta_k^{\mathrm{tr}} S_k,$$

which shows that φ^0 is determined by $V_0(\varphi)$ and ϑ under the the self-financing condition.

To see that φ^0 is predictable, note that

$$G_k(\vartheta) - G_{k-1}(\vartheta) = \Delta G_k(\vartheta) = \vartheta_k^{\mathrm{tr}} \Delta S_k = \vartheta_k^{\mathrm{tr}}(S_k - S_{k-1}).$$

Therefore

$$\begin{aligned} \varphi_k^0 &= V_0(\varphi) + G_{k-1}(\vartheta) + \Delta G_k(\vartheta) - \vartheta_k^{\mathrm{tr}} S_k \\ &= V_0(\varphi) + G_{k-1}(\vartheta) - \vartheta_k^{\mathrm{tr}} S_{k-1} \end{aligned}$$

is directly seen to be $\mathcal{F}_{k-1}$-measurable, because $G(\vartheta)$ and S are adapted and ϑ is predictable. **q.e.d.**

Another way to see that φ^0 is predictable (exercise for you) is to use (7).

(1) The notion of a strategy being self-financing is a kind of *economic budget constraint*. Exactly like the cost process, this is formulated via basic *financial book-keeping* requirements, and hence there cannot be any alternative (different) definitions that make sense financially.

(2) We have expressed all prices and values in units of the bank account. However, as basic intuition suggests, this has no effect on whether or not a strategy is self-financing. Indeed, as $\widetilde{S}_k^0 > 0$, (7) is equivalent to

$$(\varphi_{k+1}^0 - \varphi_k^0)\widetilde{S}_k^0 + (\vartheta_{k+1} - \vartheta_k)^{\mathrm{tr}}\widetilde{S}_k = 0, \tag{9}$$

by recalling that $S = \widetilde{S}/\widetilde{S}^0$. But (9) is clearly the self-financing condition expressed in terms of the original units.

Let $\tau : \Omega \to \{0, 1, \dots, T\}$ be some mapping to be thought of as a *random time*. One specific example might be the first time that stock i's price exceeds that of stock j.

Now we want to use the strategy to "buy and then hold until time τ", because we believe for some reason that this might be a good idea. For ease of notation, we take $d = 1$ so that there is just one risky asset.

Formally, let us take $V_0 := S_0$ and

$$\vartheta_k(\omega) := I_{\{k \le \tau(\omega)\}} = \begin{cases} 1 & \text{for } k = 1, \dots, \tau(\omega) \\ 0 & \text{for } k = \tau(\omega) + 1, \dots, T. \end{cases}$$

This means exactly that we hold one unit of S up to and including time $\tau(\omega)$. The value process of the corresponding self-financing strategy $\varphi \mathrel{\widehat{=}} (V_0, \vartheta)$ is then by (8) and (5) given by

$$
\begin{aligned}
V_k(\varphi) = V_0 + G_k(\vartheta) = S_0 + \sum_{j=1}^{k} \vartheta_j \Delta S_j
&= S_0 + \sum_{j=1}^{k} I_{\{j \le \tau\}} (S_j - S_{j-1}) \\
&= S_0 + \sum_{j=1}^{k\wedge\tau} (S_j - S_{j-1}) \\
&= S_{k\wedge\tau} = \begin{cases} S_k & \text{if } \tau > k \\ S_\tau & \text{if } \tau \le k \end{cases},
\end{aligned}
$$

where we use the standard notation $a \wedge b := \min(a, b)$.

The stochastic process $S^\tau = (S^\tau_k)_{k=0,1,\dots,T}$ defined by

$$S^\tau_k(\omega) := S_{k\wedge\tau}(\omega) := S_{k\wedge\tau(\omega)}(\omega) \tag{10}$$

is called the *process S stopped at τ*, because it coincides with S up to time τ, and remains constant from time τ on. Of course, for every $\omega \in \Omega$, this operation and notation per se make sense for any stochastic process S and any "random time" τ as above.

The stochastic process $S^\tau = (S_k^\tau)_{k=0,1,\dots,T}$ defined by

$$S_k^\tau(\omega) := S_{k\wedge\tau}(\omega) := S_{k\wedge\tau(\omega)}(\omega) \tag{10}$$

is called the *process S stopped at τ*, because it coincides with S up to time τ, and remains constant from time τ on. Of course, for every $\omega \in \Omega$, this operation and notation per se make sense for any stochastic process S and any "random time" τ as above.

However, S^τ could fail to be a stochastic process because $S_k^\tau = S_{k\wedge\tau}$ could fail to be a random variable (i.e. fail to be measurable). But in discrete time this is not a problem if we assume that τ is *measurable*, which is mild and reasonable enough.

Moreover, for φ to be a strategy, we need ϑ to be predictable, and this translates into a stronger requirement for τ.

Definition. A random time $\tau : \Omega \to \{0, 1, \ldots, T\}$ is called a ($\mathbb{F}$-)*stopping time* if it satisfies

$$\{\tau \leq j\} \in \mathcal{F}_j \qquad \text{for all } j. \tag{11}$$

Intuition: (11) says that at each time j, we can observe from the then available information $\mathcal{F}_j$ whether τ has already past or not.

Definition. A random time $\tau : \Omega \to \{0, 1, \dots, T\}$ is called a ($\mathbb{F}$-)*stopping time* if it satisfies

$$\{\tau \le j\} \in \mathcal{F}_j \qquad \text{for all } j. \tag{11}$$

<u>**Intuition:**</u> (11) says that at each time j, we can observe from the then available information $\mathcal{F}_j$ whether τ has already past or not.

Note that ϑ defined by $\vartheta_k = I_{\{k \le \tau\}}$ is predictable if and only if τ is a stopping time w.r.t. $\mathbb{F}$. Indeed, ϑ_k is $\mathcal{F}_{k-1}$-measurable if and only if $\{\tau \ge k\} \in \mathcal{F}_{k-1}$, and to have this for all k is equivalent to (11) by passing to complements (this uses $\{\tau \ge k\} = \{\tau < k\}^c = \{\tau \le k-1\}^c$).

Typical examples of stopping times are the first (or the n-th) time that an adapted process does something that only involves current information, for example

$$\tau(\omega) := \inf \left\{ k \,\middle|\, S^i_k(\omega) > S^j_k(\omega) \right\} \wedge T$$

(the first time that stock i's price exceeds that of stock j); or

$$\tau'(\omega) := \inf \left\{ k \,\middle|\, S^1_k(\omega) \geq 10 \max_{j=0,1,\dots,k-1} S^1_j(\omega) \right\} \wedge T$$

(the first time that stock 1's price goes above ten times its past maximum value).

Typical examples of stopping times are the first (or the *n*-th) time that an adapted process does something that only involves current information, for example

$$\tau(\omega) := \inf \left\{ k \,\middle|\, S^i_k(\omega) > S^j_k(\omega) \right\} \wedge T$$

(the first time that stock i's price exceeds that of stock j); or

$$\tau'(\omega) := \inf \left\{ k \,\middle|\, S^1_k(\omega) \geq 10 \max_{j=0,1,\dots,k-1} S^1_j(\omega) \right\} \wedge T$$

(the first time that stock 1's price goes above ten times its past maximum value).

On the other hand, times looking at the future like

$$\tau''(\omega) := \sup \left\{ k \,\middle|\, S^\ell_k(\omega) > 5 \right\} \vee 0$$

(the *last* time that stock ℓ's price exceeds 5) are typically *not* stopping times and cannot be used for constructing such buy-and-hold strategies.

Suppose we have a model where the stock price can in each step only go up or down. A well-known idea for a strategy to force winnings is then to bet on a rise and keep on betting, doubling the stakes at each date, until the rise occurs (this is the so-called *doubling strategy*).

More formally, consider the *binomial model* with parameters $d < 0 = r < u$. The stock price S_k is either $(1 + u)S_{k-1}$ or $(1 + d)S_{k-1}$. To simplify computations, suppose $d = -u$ so that the growth factors $Y_k = S_k/S_{k-1}$ are symmetric around 1. Note that

$$\Delta S_k = S_k - S_{k-1} = S_{k-1}(Y_k - 1). \tag{12}$$

Suppose we have a model where the stock price can in each step only go up or down. A well-known idea for a strategy to force winnings is then to bet on a rise and keep on betting, doubling the stakes at each date, until the rise occurs (this is the so-called *doubling strategy*).

More formally, consider the *binomial model* with parameters $d < 0 = r < u$. The stock price S_k is either $(1+u)S_{k-1}$ or $(1+d)S_{k-1}$. To simplify computations, suppose $d = -u$ so that the growth factors $Y_k = S_k/S_{k-1}$ are symmetric around 1. Note that

$$\Delta S_k = S_k - S_{k-1} = S_{k-1}(Y_k - 1). \tag{12}$$

We are now going to define a doubling strategy in this market model.

Now denote by

$$\tau := \inf\{k \mid Y_k = 1 + u\} \wedge T \tag{13}$$

the (random) time of the first stock price rise and define

$$\vartheta_k := \frac{1}{S_{k-1}} 2^{k-1} I_{\{k \leq \tau\}}. \tag{14}$$

Now denote by

$$\tau := \inf\{k \mid Y_k = 1 + u\} \wedge T \tag{13}$$

the (random) time of the first stock price rise and define

$$\vartheta_k := \frac{1}{S_{k-1}} 2^{k-1} I_{\{k \leq \tau\}}. \tag{14}$$

Then τ is a stopping time, because

$$\{\tau \leq j\} = \{\max(Y_1, \ldots, Y_j) \geq 1 + u\} \in \mathcal{F}_j$$

for each j. This in turn implies that ϑ is predictable, because each ϑ_k is $\mathcal{F}_{k-1}$-measurable (exercise for you).

For $V_0 := 0$, we now take the self-financing strategy φ corresponding to (V_0, ϑ). Its value process is by (8) and (5) given by

$$V_k(\varphi) = \sum_{j=1}^{k} \vartheta_j \Delta S_j = \sum_{j=1}^{k} 2^{j-1} I_{\{j \le \tau\}}(Y_j - 1),$$

using (12) and (14). By the definition (13) of τ, we have $Y_j - 1 = d$ for $j < \tau$ and $Y_j - 1 = u$ for $j = \tau$. Therefore, for all $k < T$,

$$\begin{aligned} V_k(\varphi) &= I_{\{\tau > k\}} \sum_{j=1}^{k} 2^{j-1} d + I_{\{\tau \le k\}} \left(\sum_{j=1}^{\tau-1} 2^{j-1} d + 2^{\tau-1} u \right) \\ &= (2^k - 1) d \, I_{\{\tau > k\}} + \big((2^{\tau-1} - 1) d + 2^{\tau-1} u \big) I_{\{\tau \le k\}}. \end{aligned}$$

Because $d = -u$ and $u > 0$, we can rewrite this as

$$V_k(\varphi) = uI_{\{\tau \leq k\}} - u(2^k - 1)I_{\{\tau > k\}},$$

which says that we obtain a value, and hence net gain, of u in all the (usually many) cases that S goes up at least once up to time k, and make a (big) loss of $u(2^k - 1)$ in the (hopefully unlikely) event that S always goes down up to time k.

Admissibility - Definition

One *problem* with the doubling strategy in the above example is that its value process goes very far below 0 in those cases where “things go badly”. In continuous time or over an infinite time horizon, one obtains quite pathological effects if one does not forbid such strategies in some way. The next definition aims at that.

One *problem* with the doubling strategy in the above example is that its value process goes very far below 0 in those cases where "things go badly". In continuous time or over an infinite time horizon, one obtains quite pathological effects if one does not forbid such strategies in some way. The next definition aims at that.

Definition. For $a \geq 0$, a trading strategy φ is called *a-admissible* if its value process $V(\varphi)$ is *uniformly bounded from below by $-a$*, that means $V(\varphi) \geq -a$ (i.e. $V_k(\varphi) \geq -a$ P-a.s. for all k).

A trading strategy is *admissible* if it is a-admissible for some $a \geq 0$.

Interpretation. An admissible strategy has some credit line which imposes a lower bound on the associated value process; so one can have debts, but only within clearly defined limits.

Remark. If Ω (or more generally $\mathcal{F}$) is finite, any random variable can only take finitely many values; for any model with finite discrete time, every trading strategy is then admissible. But if $\mathcal{F}$ (or the time horizon) is infinite or time is continuous, imposing admissibility is usually a genuine and important restriction.

Some Important Martingale Results

Let $(\Omega, \mathcal{F}, Q)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$.

Definition. A (real-valued) stochastic process $X = (X_k)_{k=0,1,\dots,T}$ is called a *martingale* (with respect to Q and $\mathbb{F}$) if:

- X is adapted to $\mathbb{F}$
- X is Q-integrable in the sense that $X_k \in \mathcal{L}^1(Q)$ for each k
- X satisfies the *martingale property*

$$E_Q[X_\ell \,|\, \mathcal{F}_k] = X_k \qquad Q\text{-a.s. for } k \leq \ell. \tag{15}$$

Intuition: This means that the best prediction for the later value X_ℓ given the information $\mathcal{F}_k$ is just the current value X_k.

Let $(\Omega, \mathcal{F}, Q)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$.

Definition. A (real-valued) stochastic process $X = (X_k)_{k=0,1,\dots,T}$ is called a *martingale* (with respect to Q and $\mathbb{F}$) if:

- X is adapted to $\mathbb{F}$
- X is Q-integrable in the sense that $X_k \in \mathcal{L}^1(Q)$ for each k
- X satisfies the *martingale property*

$$E_Q[X_\ell \,|\, \mathcal{F}_k] = X_k \qquad Q\text{-a.s. for } k \leq \ell. \tag{15}$$

Intuition: This means that the best prediction for the later value X_ℓ given the information $\mathcal{F}_k$ is just the current value X_k.

If we have "$\leq$" in (15) (a tendency to go down), X is called a *supermartingale*; if we have "$\geq$" (a tendency to go up), then X is a *submartingale*. An $\mathbb{R}^d$-valued process X is a (super/sub)martingale if each coordinate X^i is a (super/sub)martingale.

In the binomial model with parameters r, u, d, the discounted stock price $S^1 = \widetilde{S}^1/\widetilde{S}^0$ is a P-martingale if and only if $r = pu + (1-p)d$.

Indeed, by induction, one easily sees that it is enough to check (the one-step martingale property) that

$$E_P\left[\frac{\widetilde{S}^1_{k+1}}{\widetilde{S}^0_{k+1}}\,\middle|\,\mathcal{F}_k\right] = \frac{\widetilde{S}^1_k}{\widetilde{S}^0_k} \qquad \text{for each } k$$

or equivalently that

$$1 = E_P\left[\frac{\widetilde{S}^1_{k+1}}{\widetilde{S}^0_{k+1}}\middle/\frac{\widetilde{S}^1_k}{\widetilde{S}^0_k}\,\middle|\,\mathcal{F}_k\right] = E_P\left[\frac{Y_{k+1}}{1+r}\,\middle|\,\mathcal{F}_k\right].$$

Recall that Y_{k+1} is independent of $\mathcal{F}_k$ and takes the values $1+u, 1+d$ with probabilities $p, 1-p$. This gives

$$\begin{aligned} E_P\left[\frac{Y_{k+1}}{1+r}\,\middle|\,\mathcal{F}_k\right] &= \frac{1}{1+r}E_P[Y_{k+1}] \\ &= \frac{1}{1+r}\big(p(1+u)+(1-p)(1+d)\big) \\ &= \frac{1+pu+(1-p)d}{1+r}. \end{aligned}$$

This equals 1 if and only if $r = pu + (1-p)d$, which proves the assertion.

Local Martingale - Definition

Let $(\Omega, \mathcal{F}, Q)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$.

Definition. An adapted process $X = (X_k)_{k=0,1,\dots,T}$ is called a *local martingale* (with respect to Q and $\mathbb{F}$) if there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ increasing to T such that for each $n \in \mathbb{N}$, the stopped process $X^{\tau_n} = (X_{k\wedge\tau_n})_{k=0,1,\dots,T}$ is a $(Q, \mathbb{F})$-martingale. We then call $(\tau_n)_{n\in\mathbb{N}}$ a *localising sequence*.

Let $(\Omega, \mathcal{F}, Q)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$.

Definition. An adapted process $X = (X_k)_{k=0,1,\dots,T}$ is called a *local martingale* (with respect to Q and $\mathbb{F}$) if there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ increasing to T such that for each $n \in \mathbb{N}$, the stopped process $X^{\tau_n} = (X_{k\wedge\tau_n})_{k=0,1,\dots,T}$ is a $(Q, \mathbb{F})$-martingale. We then call $(\tau_n)_{n\in\mathbb{N}}$ a *localising sequence*.

Remark.

(1) Especially in continuous time, local martingales can be substantially different from (true) martingales; the concept is rather subtle.

(2) In parts of the recent finance literature, local martingales have come up in studies of price bubbles.

Theorem. Suppose $X = (X_k)_{k=0,1,\dots,T}$ is an $\mathbb{R}^d$-valued martingale or local martingale null at 0 (i.e. $X_0 = 0$). Then, for any $\mathbb{R}^d$-valued predictable process ϑ, the stochastic integral process $\vartheta \cdot X$ defined by

$$\vartheta \cdot X_k := \sum_{j=1}^{k} \vartheta_j^{\mathrm{tr}} \Delta X_j \qquad \text{for } k = 0, 1, \dots, T$$

is a (real-valued) *local martingale* null at 0.

Moreover, if X is a martingale and ϑ is bounded, then $\vartheta \cdot X$ is even a *martingale*.

Proof. We start by proving the second statement, thus considering a martingale X and a bounded predictable process ϑ. Then $\vartheta \cdot X$ is also integrable, it is always adapted, and

$$E_Q\big[\vartheta \cdot X_{k+1} - \vartheta \cdot X_k \,\big|\, \mathcal{F}_k\big] = E_Q\big[\vartheta^{\mathrm{tr}}_{k+1} \Delta X_{k+1} \,\big|\, \mathcal{F}_k\big] = \sum_{i=1}^{d} E_Q\big[\vartheta^i_{k+1} \Delta X^i_{k+1} \,\big|\, \mathcal{F}_k\big].$$

Now, ϑ^i_{k+1} is bounded and $\mathcal{F}_k$-measurable (since ϑ is predictable). Thus,

$$E_Q\big[\vartheta^i_{k+1} \Delta X^i_{k+1} \,\big|\, \mathcal{F}_k\big] = \vartheta^i_{k+1} E_Q\big[\Delta X^i_{k+1} \,\big|\, \mathcal{F}_k\big] = 0\,,$$

because X^i is a martingale. So $\vartheta \cdot X$ also has the martingale property.

We will now show the first statement.

Since ϑ is predictable,

$$\sigma_n := \inf\left\{k \,\middle|\, |\vartheta_{k+1}| > n\right\}$$

is a stopping time, and $|\vartheta_k| \le n$ for $k \le \sigma_n$ by definition. So if $(\tau_n)_{n\in\mathbb{N}}$ is a localising sequence for X, one can easily check with the above argument that $\tau_n' := \tau_n \wedge \sigma_n$ yields a localising sequence for $\vartheta \cdot X$. This gives the general result. **q.e.d.**

We have seen earlier that if τ is any stopping time, then $\vartheta_k := I_{\{k \leq \tau\}}$ is predictable, and of course bounded. So an immediate consequence of the last theorem is the following corollary.

Corollary. *For any martingale X and any stopping time τ, the stopped process X^τ is also a martingale. In particular, $E_Q[X_{k \wedge \tau}] = E_Q[X_0]$ for all k.*

We have seen earlier that if τ is any stopping time, then $\vartheta_k := I_{\{k \leq \tau\}}$ is predictable, and of course bounded. So an immediate consequence of the last theorem is the following corollary.

Corollary. *For any martingale X and any stopping time τ, the stopped process X^τ is also a martingale. In particular, $E_Q[X_{k \wedge \tau}] = E_Q[X_0]$ for all k.*

Interpretation. A martingale describes a *fair game* in the sense that the current value is the average of the future values. The above corollary says that one cannot change this fundamental character by cleverly stopping the game, while the previous Theorem says that as long as one can only use information from the past, not even complicated clever trading will help.

In general, the stochastic integral with respect to a local martingale is only a local martingale. But there is one situation where things are very nice in discrete time.

Theorem. *Suppose that X is an $\mathbb{R}^d$-valued local Q-martingale null at 0 and ϑ is an $\mathbb{R}^d$-valued predictable process. If the stochastic integral process $\vartheta \cdot X$ is uniformly bounded below (i.e. $\vartheta \cdot X_k \geq -b$ Q-a.s. for all k, for a constant $b \geq 0$), then $\vartheta \cdot X$ is a Q-martingale.*

Proof. See Föllmer/Schied (2011), Theorem 5.15. and lecture notes for details. **q.e.d.**

An Example: The Multinomial Model

We now take a closer look at the multinomial model introduced in the first section. This is the multiplicative model with i.i.d. returns given by

$$\frac{\widetilde{S}_k^0}{\widetilde{S}_{k-1}^0} = 1 + r > 0 \qquad \text{for all } k,$$
$$\frac{\widetilde{S}_k^1}{\widetilde{S}_{k-1}^1} = Y_k \qquad \text{for all } k,$$

where $\widetilde{S}_0^0 = 1$, $\widetilde{S}_0^1 = S_0^1 > 0$ is a constant, and $Y_1, \dots, Y_T$ are i.i.d. and take the finitely many values $1 + y_1, \dots, 1 + y_m$ with respective probabilities $p_1, \dots, p_m$.

We now take a closer look at the multinomial model introduced in the first section. This is the multiplicative model with i.i.d. returns given by

$$\frac{\widetilde{S}^0_k}{\widetilde{S}^0_{k-1}} = 1 + r > 0 \qquad \text{for all } k,$$
$$\frac{\widetilde{S}^1_k}{\widetilde{S}^1_{k-1}} = Y_k \qquad \text{for all } k,$$

where $\widetilde{S}^0_0 = 1$, $\widetilde{S}^1_0 = S^1_0 > 0$ is a constant, and $Y_1, \dots, Y_T$ are i.i.d. and take the finitely many values $1 + y_1, \dots, 1 + y_m$ with respective probabilities $p_1, \dots, p_m$.

To avoid degeneracies and fix the notation, we assume that $p_j > 0$ for all j and that $y_m > y_{m-1} > \cdots > y_1 > -1$. This also ensures that $\widetilde{S}^1$ remains strictly positive.

Interpretation: At each step, the bank account changes by a factor of $1 + r$, while the stock changes by a random factor that can only take the m different values $1 + y_j$, $j = 1, \ldots, m$. The choice of these factors happens randomly, with the same mechanism (identically distributed) at each date, and independently across dates.

Interpretation: At each step, the bank account changes by a factor of $1 + r$, while the stock changes by a random factor that can only take the m different values $1 + y_j$, $j = 1, \ldots, m$. The choice of these factors happens randomly, with the same mechanism (identically distributed) at each date, and independently across dates.

Intuition suggests that, for a reasonable market model, the sure factor $1 + r$ should lie between the minimal and maximal values $1 + y_1$ and $1 + y_m$ of the (uncertain) random factor; we shall come back to this issue in the next chapter.

The simplest and in fact canonical model for this setup is a *path space*. Let

$$\Omega = \{1, \dots, m\}^T = \left\{\omega = (x_1, \dots, x_T) \,\middle|\, x_k \in \{1, \dots, m\} \text{ for } k = 1, \dots, T\right\}$$

be the set of all sequences of length T formed by elements of $\{1, \dots, m\}$.

The simplest and in fact canonical model for this setup is a *path space*. Let

$$\Omega = \{1, \dots, m\}^T = \Big\{\omega = (x_1, \dots, x_T) \,\Big|\, x_k \in \{1, \dots, m\} \text{ for } k = 1, \dots, T\Big\}$$

be the set of all sequences of length T formed by elements of $\{1, \dots, m\}$.

Take $\mathcal{F} = 2^\Omega$, the family of all subsets of Ω, and define P by setting

$$P[\{\omega\}] = p_{x_1} p_{x_2} \cdots p_{x_T} = \prod_{k=1}^{T} p_{x_k}. \tag{16}$$

The simplest and in fact canonical model for this setup is a *path space*. Let

$$\Omega = \{1, \dots, m\}^T = \Big\{\omega = (x_1, \dots, x_T) \,\Big|\, x_k \in \{1, \dots, m\} \text{ for } k = 1, \dots, T\Big\}$$

be the set of all sequences of length T formed by elements of $\{1, \dots, m\}$.

Take $\mathcal{F} = 2^\Omega$, the family of all subsets of Ω, and define P by setting

$$P[\{\omega\}] = p_{x_1} p_{x_2} \cdots p_{x_T} = \prod_{k=1}^{T} p_{x_k}. \tag{16}$$

Define $Y_1, \dots, Y_T$ by

$$Y_k(\omega) := 1 + y_{x_k} \tag{17}$$

so that $Y_k(\omega) = 1 + y_j$ if and only if $x_k = j$.

The simplest and in fact canonical model for this setup is a *path space*. Let

$$\Omega = \{1, \dots, m\}^T = \big\{\omega = (x_1, \dots, x_T) \,\big|\, x_k \in \{1, \dots, m\} \text{ for } k = 1, \dots, T\big\}$$

be the set of all sequences of length T formed by elements of $\{1, \dots, m\}$.

Take $\mathcal{F} = 2^\Omega$, the family of all subsets of Ω, and define P by setting

$$P[\{\omega\}] = p_{x_1} p_{x_2} \cdots p_{x_T} = \prod_{k=1}^{T} p_{x_k}. \tag{16}$$

Define $Y_1, \dots, Y_T$ by

$$Y_k(\omega) := 1 + y_{x_k} \tag{17}$$

so that $Y_k(\omega) = 1 + y_j$ if and only if $x_k = j$.

This mathematically formalises the idea that at each step k, we choose the value $1 + y_j$ for Y_k with probability p_j, and we do this independently over time (k) because P is obtained by multiplication.

As usual, we take as *filtration* the one generated by $\widetilde{S}^1$ (equivalently, by Y):

$$\mathcal{F}_k = \sigma(Y_1, \dots, Y_k) \qquad \text{for } k = 0, 1, \dots, T.$$

Intuition: Up to time k, we can observe the values of $Y_1, \dots, Y_k$ and hence the first k "bits" of the trajectory or sequence ω.

Formally, this translates as follows.
A set $A \subseteq \Omega$ is an *atom of $\mathcal{F}_k$* if and only if there exists a sequence $(\bar{x}_1, \dots, \bar{x}_k)$ of length k with elements $\bar{x}_i \in \{1, \dots, m\}$ such that A consists of all those $\omega \in \Omega$ that start with the substring $(\bar{x}_1, \dots, \bar{x}_k)$, i.e.

$$A = A_{\bar{x}_1, \dots, \bar{x}_k} := \Big\{ \omega = (x_1, \dots, x_T) \in \{1, \dots, m\}^T \,\Big|\, x_i = \bar{x}_i \text{ for } i = 1, \dots, k \Big\}.$$

With this, we have the following consequences:

- ▷ Each $\mathcal{F}_k$ is parametrised by substrings of length k and therefore contains precisely m^k atoms.

With this, we have the following consequences:

- Each $\mathcal{F}_k$ is parametrised by substrings of length k and therefore contains precisely m^k atoms.
- When going from time k to time $k+1$, each atom $A = A_{\bar{x}_1,\dots,\bar{x}_k}$ from $\mathcal{F}_k$ splits into precisely m subsets $A_1 = A_{\bar{x}_1,\dots,\bar{x}_k,1}, \dots, A_m = A_{\bar{x}_1,\dots,\bar{x}_k,m}$ that are atoms of $\mathcal{F}_{k+1}$. So we can see very precisely and graphically how information about the past, that is the initial part of trajectories ω, is growing and refining over time.

With this, we have the following consequences:

- Each $\mathcal{F}_k$ is parametrised by substrings of length k and therefore contains precisely m^k atoms.
- When going from time k to time $k+1$, each atom $A = A_{\bar{x}_1,\ldots,\bar{x}_k}$ from $\mathcal{F}_k$ splits into precisely m subsets $A_1 = A_{\bar{x}_1,\ldots,\bar{x}_k,1}, \ldots, A_m = A_{\bar{x}_1,\ldots,\bar{x}_k,m}$ that are atoms of $\mathcal{F}_{k+1}$. So we can see very precisely and graphically how information about the past, that is the initial part of trajectories ω, is growing and refining over time.

It is clear from the above description that for any k, the atoms of $\mathcal{F}_k$ are pairwise disjoint and their union is Ω. Finally, each set B in $\mathcal{F}_k$ is a union of atoms of $\mathcal{F}_k$; so the family $\mathcal{F}_k$ of events observable up to time k consists of 2^{m^k} sets.

With the help of the atoms introduced above, we can also give a very precise and intuitive description of *all probability measures* Q on $\mathcal{F}_T$:

- First of all, we identify each atom in $\mathcal{F}_k$ with a node at time k of the non-recombining tree, namely that node which is reached via the substring $(\bar{x}_1, \dots, \bar{x}_k)$ that parametrises the atom.

With the help of the atoms introduced above, we can also give a very precise and intuitive description of *all probability measures* Q on $\mathcal{F}_T$:

- First of all, we identify each atom in $\mathcal{F}_k$ with a node at time k of the non-recombining tree, namely that node which is reached via the substring $(\bar{x}_1, \ldots, \bar{x}_k)$ that parametrises the atom.
- For any atom $A = A_{\bar{x}_1,\ldots,\bar{x}_k}$ of $\mathcal{F}_k$, we then look at its m successor atoms $A_1 = A_{\bar{x}_1,\ldots,\bar{x}_k,1}, \ldots, A_m = A_{\bar{x}_1,\ldots,\bar{x}_k,m}$ of $\mathcal{F}_{k+1}$, and we define the *one-step transition probabilities* for Q at the node corresponding to A by the conditional probabilities

$$Q[A_j \mid A] = \frac{Q[A_j]}{Q[A]} \qquad \text{for } j = 1, \ldots, m. \tag{18}$$

Because A is the disjoint union of $A_1, \dots, A_m$, we have $0 \leq Q[A_j \mid A] \leq 1$ for $j = 1, \dots, m$ and $\sum_{j=1}^{m} Q[A_j \mid A] = 1$. (If $Q[A]$ is zero, then so are all the $Q[A_j]$, because $A_j \subseteq A$.)

Because A is the disjoint union of $A_1, \ldots, A_m$, we have $0 \leq Q[A_j \,|\, A] \leq 1$ for $j = 1, \ldots, m$ and $\sum_{j=1}^{m} Q[A_j \,|\, A] = 1$. (If $Q[A]$ is zero, then so are all the $Q[A_j]$, because $A_j \subseteq A$.)

By attaching these one-step transition probabilities to each branch from each node, we then have by construction a decomposition or factorisation of Q in such a way that for every trajectory $\omega \in \Omega$, its probability $Q[\{\omega\}]$ is the product of the successive one-step transition probabilities along ω.

Multinomial Model

Because A is the disjoint union of $A_1, \ldots, A_m$, we have $0 \leq Q[A_j \,|\, A] \leq 1$ for $j = 1, \ldots, m$ and $\sum_{j=1}^{m} Q[A_j \,|\, A] = 1$. (If $Q[A]$ is zero, then so are all the $Q[A_j]$, because $A_j \subseteq A$.)

By attaching these one-step transition probabilities to each branch from each node, we then have by construction a decomposition or factorisation of Q in such a way that for every trajectory $\omega \in \Omega$, its probability $Q[\{\omega\}]$ is the product of the successive one-step transition probabilities along ω.

This follows in an elementary way from the definition of conditional probabilities, $Q[C \cap D] = Q[C]\, Q[D \,|\, C]$, and by iteration.