Stochastic Analysis I Processes and Martingales

Chapter · May 2017

1 author:

Paolo Vanini
University of Basel

Stochastic Analysis I

Martingales

Paolo Vanini
Version 1

May 5, 2017

Contents

0.1 Overview

We introduce to stochastic processes, focussing on discrete time processes, and the probabilistic in particular the use of filtrations.

Martingales are key concepts in general equilibrium models (necessary for the existence of equilibria) and they are basic in derivative pricing (the absence of arbitrage is equivalent to the existence of a probability such that discounted price process are a martingale).

These notes consider the discrete time setup.

0.2 Stochastic Process in Discrete Time - Investment Case

The goal is to consider the stochastic model for prices of securities in a finite, discrete time set-up. We assume that prices S_t can move up or down at each time date in a recombining binomial tree; that is the so-called Cox-Ross-Rubinstein (CRR) model is considered. That is, the dynamics reads

$$S_{t+1} = \begin{cases} S_t(1+u) \\ S_t(1+d) \quad t = 0, 1, \ldots, T-1. \end{cases}$$

The price at time k is then

$$S_k = S_0(1+u)^{N_k}(1+d)^{k-N_k}$$

with $N_k := \sum_{i=1}^{k} \omega_k$ the number of upwards moves. For second asset is a risk-less asset B_t with solves a deterministic first order difference equation with solution $B_t = (1+r)^t$, i.e. $B_0 = 1$.

The underlying model structure is a stochastic base, i.e. filtered probability space. It consists of four objects $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbf{Z}_+}, P)$ with Ω the state space, $\mathcal{A}$ the event space, $(\mathcal{F}_t)_{t\in\mathbf{Z}_+}$ the filtration and P a probability distribution.

The set Ω has only a finite number of observable states ω_k since we consider a discrete and finite time model. Each ω_k represents a possible price paths. Say, $\omega_1 = (u, u, u)$ is the observable path if the asset prices moves three times up. After T periods, 2^T price paths realizations are possible.

We consider next the information flow dynamics in the model if prices are realized in the CRR model, see Figure 1.

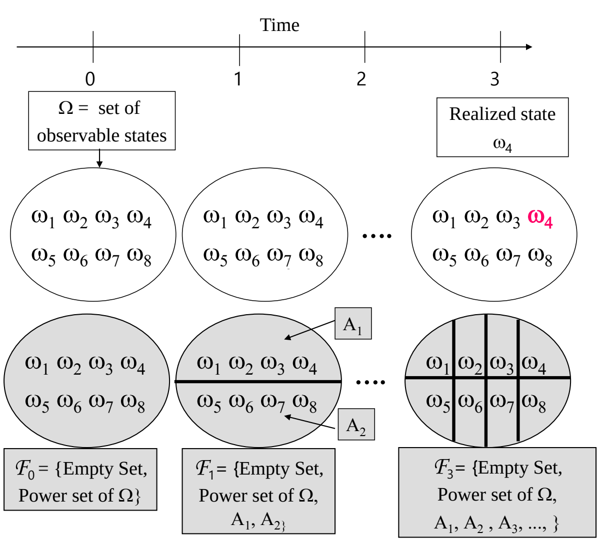

Figure 1: Illustration of the information and filtration structure for the three period CRR.

Suppose that the first move was up. Then several paths are still possible but other paths are no longer possible. If in the second step there is a down move, some more paths will become impossible and so on. This shows that besides observable elements also **possible** events are important. If we have 8 observable events, the power set $\mathcal{F} = 2^8$ defines all possible events. $\mathcal{F}_t \in \mathcal{F}$ represents the possible information up to time t. Filtrations $(\mathcal{F}_t)$ are used to describe the information flow as follows:

$$\mathcal{F}_t \subset \mathcal{F}_{t+1} \ , \mathcal{F}_t \in \mathcal{F}, \ \forall t$$

and element F_t is a σ-algebra (see below). The first property means that information uncertainty is decreasing over time - we know more and more bout the possible price path as times goes by. The second property means that each set of the filtrations is measurable, i.e. we can assign a probability to each set.

The event set $\mathcal{F}$ is by definition a σ-algebra, which is set equal to the power set of Ω in our simple example. A σ-algebra is a system of sets that includes the empty set, each complement of a set and which is closed under countable unions and intersections. The events are the sets which can be **assigned** probabilities.

Finally, a stochastic processes $S = (S_t)_{t \in \mathbf{Z}_+}$ is a function

$$S_t(\omega) : \Omega \times \mathbf{Z}_+ \to \mathbf{R} \ .$$

Hence, it is a set of random variables. If we fix the state ω, then the function X_t is called a path (i.e. an observable element in the CRR model). The information evolution will be generated by the stochastic process since all uncertainty in the models will be due to stochastic prices of some assets. More precisely, at time t we know the realization of S_t and S_s for all times $s < t$. Hence, the information set is the smallest set $\mathcal{F}_t$ generated by all events the stochastic process could attain up to time t, i.e. the following σ-algebra

$$\mathcal{F}_t = \sigma(S_s \ , \ s = 0, \ldots, t) \ .$$

We call the filtration, which is given by minimal σ-algebras generated by the observation of prices, the *natural* filtration. In this example the information $\mathcal{F}_t$ at time t is generated by the price process S history - this is the usual case in applications but one could consider information structures generated by other sources.

Consider the three periods model $t = 0, 1, 2, 3$ with price moves up 'u' or down 'd' with probabilities p and $1 - p$ respectively. The eight observable states are:

$$\begin{aligned}\omega_1 &= (u,u,u)\ ,\ \omega_2 = (u,u,d)\ ,\ \omega_3 = (u,d,u)\ ,\ \omega_4 = (u,d,d)\\ \omega_5 &= (d,u,d)\ ,\ \omega_6 = (d,d,u)\ ,\ \omega_7 = (d,u,d)\ ,\ \omega_8 = (d,d,d)\ .\end{aligned}$$

The possible event

$$A = \{\text{at least 2 upside moves}\} \in \mathcal{F}$$

consists of the elements $A = \{\omega_1, \omega_2, \omega_3\}$ which is a non-observable event. We define a filtration. At t_0 we know S_0 with certainty and $\mathcal{F}_0 = \{\emptyset, \mathcal{F}\}$. This is a σ-algebra since it is closed under countable intersection and complement formation. The definition of $\mathcal{F}_0$ means to an investor that anything can happen ($\mathcal{F}$) since all information is still random. The e inclusion of $\emptyset$ guarantees that $\mathcal{F}_0$ is a σ-algebra. At $t = 1$, either the price S_0 increased or decreased. Therefore, we define the sets

$$A_1 = \{\omega_1, \omega_2, \omega_3, \omega_4\}\ ,\ A_2 = \{\omega_5, \omega_6, \omega_7, \omega_8\}\ .$$

Hence, A_1 (A_2) is the set of all events where the first price move is 'up' ('down'). Therefore, we can set

$$\mathcal{F}_1 = \{\emptyset, \mathcal{A}, A_1, A_2\}\ .$$

This assures that $\mathcal{F}_0 \subset \mathcal{F}_1$ and that $\mathcal{F}_1$ is a σ-algebra. It is clear, how $\mathcal{F}_2$ can be defined and so on.

If S_t is $\mathcal{F}_t$-measurable for all $t \in \mathbf{Z}_+$, we call the process S *adapted to the filtration* $\mathbf{F}$ and we write $S_t \in \mathcal{F}_t$. This means that the realization of the process at time S_t is known at time t. This makes sense for risky assets since adaptedness means the impossibility to look in into the future.

If $X_t \in \mathcal{F}_{t-1}$ for all $t \in \mathbf{Z}_+$, we call the process X *predictable*[1]. The value of X_t is then known at time $t-1$. To give an interpretation of predictable process consider a portfolio strategy ϕ_t again in discrete time. This strategy tells us at each date t how many stocks with price S_t we possess. Hence, $V_t = \phi_t S_t$ is the portfolio value at time t. From this vista time, we don't know the price of S_{t+1} until time $t+1$ is reached - hence S is an adapted process. What can be said about the other process ϕ_t? If we do not want to consider any inflows and outflows of our portfolio between $[t, t+1)$, that is

[1]By definition $\mathcal{F}_{-1} = \mathcal{F}_0$.

the change in portfolio value is entirely given by the change in asset prices, then ϕ_t should remain constant in the period and the strategy is changed only after the new prices are realized at time $t+1$. Such a portfolio strategy is called self-financing. Therefore, ϕ_t is a predictable process.

There is a second, purely mathematical interpretation of predictable processes. Consider the definition of the Riemann integral of a continuous real-valued function f. The integral is constructed by approximating the function by a finite number of stepwise function and then the integral definition is extended by taking a limit where the stepwise function approximate the function f as close as possible. The limit, if it exists is called the Riemann integral. The same method is used for stochastic integrals, i.e. integrals where both the integrand and integrator are stochastic process. Predictable processes then are the generalization of step-wise ordinary function - they turn out to be the good integrands for stochastic integration theory. We encounter this in the Notes Stochastic Calculus, Semi Martingales and Local Martingales.

0.3 Stochastic Processes in Continuous Time

When the index t is continuous the definition of filtrations and predictability become more subtle and the definition of the properties a stochastic process should have in order to be useful for integration theory and applications for example is not a trivial task at all. From a practitioners view where all processes uses are in discrete time one might wonder why anybody which is not a pure mathematician should care about continuous time processes. Continuous time models, despite the difficulties to set and understand the fundamentals, are more elegant than discrete time ones and they are widely used in formulation of theories. To understand option pricing models or investment models without relying on continuous time approaches is therefore simply not meaningful.

We discuss some of the subtleties without going into the mathematical details.

First, in stochastic processes one is not interested in sets with zero probability. One therefore call two processes indistinguishable if they have with probability one the same sample paths. Since probability theory concerns countable operations by definition one often can construct counter example of statements where un-countability is used. For example $X_t = Y_t$ with probability one does not imply that they are indistinguishable. One therefore defines that a process X to be a version of Y if they are equal with probability one at each date. The most often used versions are right-continuous ones. This then rules out the above mentioned case - if for all time $X_t = Y_t$ with probability holds, then they are indistinguishable. In order to control for the set of measure zero under time evolution one requires that all zero probability set are yet contained in $\mathcal{F}_0$. The filtered probability space is then complete.

The concept of limits is also needed for filtrations. Given a filtration, the right limits at any time are defined as:

$$\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$$

and similar for the left limit. The filtration is right-continuous if $\mathcal{F}_t = \mathcal{F}_{t+}$. In this case, no additional information is gained by taking any infinitesimal step forward in time.

The workhorse in stochastic process theory are the so-called usual conditions assumptions: the filtered probability space is complete and the filtration is right-continuous. Since adding the zero probability sets generated completeness and by replacing $\mathcal{F}_t$ by $\mathcal{F}_{t+}$ any filtration becomes right-continuous, any filtered probability space can be enlarged in a minimal way to one satisfying the usual conditions. The usual conditions allows to assume that the processes have cadlag (right continuous with only simple jump discontinuities) paths. The cadlag property eliminates specific path properties of the processes due to uncountable operations and allow us to work with countable operations. Hence it suffices to understand properties of stochastic process only to the level where they are indistinguishable. The regularity of the paths of cadlag processes then also simplify to prove properties of processes which are needed for example in optimal stopping or integration theory.

0.4 Martingales

We first recall basic properties of conditional expectation.

Definition 0.4.1. *Consider a probability space $(\Omega, \mathcal{A}, P)$ and assume that the random variable X, Y are integrable. Further, H, G are sub σ-algebras of $\mathcal{A}$.*

1. *(Linearity) $E[aX + bY|H] = aE[X|H] + bE[Y|H]$ for constants a, b.*
2. *If $X \in H$ then $E[X|F] = X$.*
3. *(Iterated expectation law) $E[E[X|H]] = E[X]$.*
4. *(Projection property) Let $H \subset G$. Then : $E[E[X|G]|H] = E[X|H]$ (P-a.s).*
5. *If the random variable Z is bounded and H-measurable, then $E[ZX|H] = ZE[X|H]$.*
6. *If X is independent of H, then $E[X|H] = E[X]$.*

Properties (2) and (4)-(6) hold only P-almost surely (P-a.s.). That is, two random variables X and Y are equal P-almost surely, if the the set of states where they may differ has probability zero. This technical point is of no importance if Ω is a set with a finite number of elements. It follows from general theory that conditional expectation is the best guess of a random variable given an information set. The projection property then means that if F has more information than G, you gain nothing in forecasting a random variable by first projecting on G and then on F compared to directly projecting on F.

Definition 0.4.2. *Let $(X_t)_{t=0,1,\ldots,}$ be an adapted to the filtration* **F** *process with integrable random variables X_t, i.e. $E[|X_t|] < \infty$ for all t. The process $(X_t)_t$ is a* martingale *iff:*

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

$(X_t)_t$ is a sub-martingale*, iff:*

$$E[X_{t+1}|\mathcal{F}_t] \geq X_t \quad \forall t$$

$(X_t)_t$ is a super-martingale*, iff:*

$$E[X_{t+1}|\mathcal{F}_t] \leq X_t \quad \forall t$$

Intuitive, for a martingale, the expected value tomorrow given today's information equals today value. Therefore, martingales are often called fair processes or fair games. We consider some examples next. Straightforward proofs are left to the readers.

Lemma 0.4.3. *1. Let X be a martingale and g a convex function. Then $Y = g(X)$ is a submartingale.*

2. Let X be a martingale. Then $E[X_t] = E[X_0]$ $\forall t$.

Proof. Exercise. □

Let (X_t) be a sequence of i.i.d. Bernoulli random variables with

$$P(X_t = 1) = p \ , \ P(X_t = -1) = 1 - p \ .$$

The natural filtration is then given by $\mathcal{F}_t = \sigma(X_0, \ldots, X_t)$ Interpreting the realization of the random variables as "success" or "failure" in the respective games we define the $\mathcal{F}_t$-predictable doubling-bet strategy

$$W_t = W_t(X_0, \ldots, X_{t-1}) = \frac{1}{2^{t-1}} \chi_{\{X_1=-1,\ldots,X_{t-1}=-1\}} \ .$$

The indicator (characteristic) function $\chi_{\{A\}}$ is defined by

$$\chi_{\{A\}}(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise.} \end{cases}$$

The total payoff in t games is

$$V_t = V_{t-1} + W_t X_t = \sum_{s=0}^{t} W_s X_s \ .$$

Lemma 0.4.4. *The process $V = (V_t)$ is a martingale if and only if $p = 0.5$.*

Proof. Exercise. □

Let X_t be $\mathcal{F}_t$-measurable. $X = (X_t)$ is called a *martingale-difference* if

$$E[X_t | \mathcal{F}_{t-1}] = 0$$

for all t.

Lemma 0.4.5.

$$Y_t = \sum_{s=0}^{t} X_s \text{ is a martingale} \iff (X_t) \text{ is a martingale-difference} \ .$$

Proof. Exercise. □

In the next example we consider the famous Doob decomposition of submartingales.

Theorem 0.4.6. *Let X be a submartingale. Then there exists a martingale M and a predictable, increasing process A such that for any t*

$$X_t = M_t + A_t \ , \ P - a.s. \tag{1}$$

holds. This decomposition is unique.

Proof. Let $X_0 = M_0, A_0 = 0$ and consider

$$M_t = M_0 + \sum_{s=0}^{t-1} (X_{t+1} - E[X_{t+1}|\mathcal{F}_t]) \tag{2}$$

$$A_t = \sum_{s=0}^{t-1} (E[X_{t+1}|\mathcal{F}_t] - X_{t+1}) \geq 0 \text{ increasing} \ . \tag{3}$$

The first claim then follows. To prove uniqueness, consider that a second decomposition with M' and A' exists. Then

$$\Delta A'_{t+1} = A'_{t+1} - A'_t = \Delta A_{t+1} + \Delta M_{t+1} - \Delta M'_{t+1}$$

follows. Since A and A' are predictable and M, M' are martingales,

$$\Delta A'_{t+1} = E[\Delta A'_{t+1}|\mathcal{F}_t] = E[\Delta A_{t+1}|\mathcal{F}_t] = \Delta A_{t+1}$$

follows. Hence, $A'_t = A_t$ and finally $M'_t = M_t$ for all t. □

Let M be a square integrable martingale ($E[M_t^2] < \infty$ for all $t \geq 1$). Lemma 0.4.3 implies that M^2 is a submartingale and from (2) we obtain from Doob's theorem the decomposition

$$M_t^2 = \underbrace{\sum_{1 \leq s \leq t} E[\Delta M_s^2|\mathcal{F}_{t-1}]}_{A_t =: \langle M \rangle_t} + \underbrace{\sum_{1 \leq s \leq t} \left(\Delta M_s^2 - E[\Delta M_s^2|\mathcal{F}_{t-1}]\right)}_{=: m_t} \tag{4}$$

with m_t a martingale and $\langle M \rangle_t$ a predictable, increasing sequence called the *quadratic characteristic*. The following properties then hold:

Lemma 0.4.7. *Consider the square integrable M defined above. Then*

$$E[M_t^2] = E[\langle M \rangle_t] \ , \ t \in \mathbf{Z}_+ \tag{5}$$

$$E[M_t - M_s|\mathcal{F}_s] = 0 \ , \ s \leq t \ . \tag{6}$$