Martingales and Stochastic Calculus

Alex Rutar*
University of Waterloo

Winter 2020†

*arutar@uwaterloo.ca
†Last updated: January 15, 2020

Contents

I. Stochastic Calculus

Definition. Given a measure space $(\Omega, \mathcal{F}, \mathbb{P})$, a measurable function $f : \Omega \to \mathbb{R}$ is called a **random variable**.

Definition. A **stochastic process** $X = \{X_t\}_{t\in T}$ is a collection of random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Typically $t \in \mathbb{Z}^+$ or $t \in \mathbb{R}^+$ (including 0); t is a discrete or continuous time parameter. Given some $\omega \in \Omega$ the map $t \mapsto X_t(\omega)$ is called a **realization** or **path** of this process. We will regard $\{X_t\}_{t\geq 0}$ as a random element in some path space, equipped with a proper σ−algebra and probability.

Consider $X_t(\omega)$ as a function $X : [0,\infty)\times\Omega \to \mathbb{R}$ equipped with the product σ−algebra.

Definition. The **distribution** of a stochastic process is the collection of all its finite-dimensional distributions.

Two processes X and Y can be "the same" in different senses:

Definition. Two process $X = \{X_t\}_{t\geq 0}$ and $Y = \{Y_t\}_{t\geq 0}$ are called **distinguishable** if almost all their sample paths agree; in other words, $\mathbb{P}(X_t = Y_t, 0 \leq t < \infty) = 1$. We say that Y is a **modification** of X if for each $t \geq 0$ we have $\mathbb{P}(X_t = Y_t) = 1$. Finally, X and Y are said to have the **same distribution** if all the finite dimensional distributions agree. In other words, if for all $n \in \mathbb{N}$ and $0 \leq t_1 < t_2 < \cdots < t_n < \infty$, we have $(X_{t_1}, \ldots, X_{t_n}) \overset{d}{=} (Y_{t_1}, \ldots, Y_{t_n})$.

Example. Let X be a continuous stochastic process and N a Poisson point process on $[0,\infty)$. Then define

$$Y_t := \begin{cases} X_t & : t \notin N \\ X_t + 1 & : t \in N \end{cases}$$

Thus $\mathbb{P}(X_t = Y_t) = 1$ for all t, so X is a modification of Y. However, $\mathbb{P}(X_t = Y_t, t \geq 0) = 0$, so that X and Y are not indistinguishable.

A filtration formalizes the idea of "information acquired over time".

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A **filtration** is a non-decreasing family $\{\mathcal{F}_t\}_{t\geq 0}$ of sub-σ-algebras of $\mathcal{F}$ so that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $0 \leq s < t < \infty$. We write $F_\infty = \sigma(\bigcup_{t\geq 0} \mathcal{F}_t)$.

Let $\{X_t\}_{t\geq 0}$ be a stochastic process. The filtration generated by $\{X_t\}_{t\geq 0}$ is $\{\sigma(X_s : 0 \leq s \leq t)\}_{t\geq 0}$, in other words $\mathcal{F}_t$ is the smallest σ−algbra which makes X_s measurable for all $s \in [0,t]$.

Definition. A stochastic process $\{X_t\}_{t\geq 0}$ is called **adapted** to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if X_t is $\mathcal{F}_t$−measurable for every $t \geq 0$.

The filtration generated by $\{X_t\}_{t\geq 0}$ is the smallest filtration which makes $(X_t)_{t\geq 0}$ adapted.

A filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is said to satisfy the "usual condition" if

1. It is right-continuous: $\lim_{s\to t^+} := \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$
2. $\mathcal{F}_0$ contains all the $\mathbb{P}$−null events in $\mathcal{F}$.

1 Martingale Theory

Consider a filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in S})$ where $S = \mathbb{N}$ or $S = \mathbb{R}^+$.

Definition. A **random time** T is called a stopping time if $\{T \leq t\} \in \mathcal{F}_t$ ("we know it happens when it happens").

Example. (i) Constants are trivial stopping times.
(ii) Last hit a constant before N is not a stopping time

1.1 Proposition. *If S, T are stopping times, $T \vee S$, $T \wedge S$, $T + S$ are stopping times.*

Proof That $T \wedge S$ and $T \vee S$ are stopping times are trivial. For $T+S$, $\{T+S > t\} = \{T = 0, S > t\} \cup \{0 < T \leq t, T+S > t\} \cup \{T > t\}$. It suffices to prove that

$$\{0 < T \leq t < T + S > t\} = \bigcup_{\substack{r\in\mathbb{Q}^+ \\ 0<r<t}} \{r < T \leq t, S > t - r\}.$$

If there exists r with $r < T \leq t$, then $S > t = r$ and $S + T > r + (t-r) = t$, so $\supseteq$ holds. Conversely, if $0 < T \leq t$ and $T + S \geq t$¡ then there exists $r \in \mathbb{Q}$ such that $r < T$ and $r + S > t$. Hence $r < T \leq t$ and $S > t - r$. ∎

Definition. The $\sigma-$algebra generated by a stopping time T is the collection of all the events A for which $A \cap \{T \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$. This is the "information you collect until the stopping time".

Exercise: show that the collection given in the definition above is actually a $\sigma-$algebra.

We write $X_{T\wedge t}$ is a random variable evaluated at time $T \wedge t$ (or T); in other words, $(X_{T\wedge t})(\omega) = X_{T\wedge t}(\omega)$. Then $\{T_{T\wedge t}\}_{t\geq 0}$, or X^T, is a stochastic process stopped at time t.

Definition. Consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, \mathbb{P})$. A $\{\mathcal{F}_t\}_{t\in T}-$adaped proces $\{X_t\}_{t\in T}$ is said to be a **submartingale** if
(i) For all $t \in T$, $X_t \in L^1(\Omega, \mathcal{F}, \mathcal{P})$
(ii) For all $s < t$ where $s, t \in T$,

$$\mathbb{E}(X_t|\mathcal{F}_s) \geq X_s \qquad \mathcal{P} \text{ a.s.}$$

and it is a **supermartingale** with the inequality in (ii) reversed. Then $\{X_t\}_{t\in T}$ is a **martingale** if it is both a submartingale and supermartingale.

Clearly if $0 \leq s < t$, then $\mathbb{E}(X_0) \leq \mathbb{E}(X_s) \leq \mathbb{E}(X_t)$ in the submartingale case, and other appropriate statements in the other case. One of the goals of martingale theory is to extend these results with respect to stopping times, rather than with respect to fixed time.

Let $X = \{X_n\}_{n=0}^{\infty}$ be a discrete time process, and fix levels $a < b$. The number $U_N[a,b]$ of **upcrossings** of $[a,b]$ by X by time N is the largest $k \in \mathbb{Z}^+$ such that there exists times $a \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_k < t_k \leq N$ such that for all i, $s_i \leq a$ and $t_i \geq b$.

Definition. A process $X = \{X_n\}_{n=0}^{\infty}$ is called **previsible** if $X_n \in \mathcal{F}_{n-1}$ for all $n \geq 1$.

Let $C = \{C_n\}_{n=1}^{\infty}$ be a previsible process. Then the martingale transform of X by C, denoted by $C \cdot X$, is defined as $(C \cdot X)_n = \sum_{k=1}^{n} C_k(X_k - X_{k-1})$ for $n > 0$ and $(C \cdot X)_0 = 0$.

1.2 Proposition. 1. *Let C be a bounded, non-negative previsible process and X a (super)matingale. Then $C \cdot X$ is a (super)martingale which is null at 0.*

2. *Let C be a bounded, previsible process and X a martingale. Then $C \cdot X$ is a martingale which is null at 0.*
3. *If C is an L^2 previsible process (resp. non-negative) and X an L^2 martingale (resp. supermartingale), then $C \cdot X$ is a martingale (resp. supermartingale).*

Proof 1. We have

$$\begin{aligned}\mathbb{E}[(C \cdot X)_n - (C \cdot X)_{n-1} | \mathcal{F}_{n-1}] &= \mathbb{E}[C_n(X_n - X_{n-1}) | \mathcal{F}_{n-1}] \\ &= C_n \cdot \mathbb{E}[X_n - X_{n-1} | \mathcal{F}_{n-1}] \\ &\leq 0\end{aligned}$$

where the second line follows since C is previsible. Since C is bounded, $C \cdot X$ is integrable.
2. Consider $C + k$ where k is a constant and $k \geq |C_n(w)|$ for all n and w.
3. Similar, but the integrability is now guaranteed by Hölder's inequality. ■

We now have the following result:

1.3 Proposition. (Doob's Upcrossing Inequality) *Let X be a supermartingale. Then $(b-a)\,\mathbb{E}(U_N[a,b]) \leq \mathbb{E}((X_N - a)^-)$.*

Proof Define a process $\{C_n\}_{n=1}^{\infty}$ by

$$\begin{aligned}C_1 &:= \chi_{\{X_0 < a\}} \\ C_k &:= \chi_{\{C_{n-1}=1\}}\chi_{\{X_{n-1} \leq b\}} + \chi_{\{C_{n-1}=0\}}\chi_{\{X_{n-1} < a\}}.\end{aligned}$$

Let $Y = C \cdot X$, i.e. $Y_n = \sum_{k=1}^{n} C_k(X_k - X_{k-1})$ almost surely. Note that each finished upcrossing increases the value of Y by at least $b - a$, we have

$$Y_N \geq (b-a)U_N[a,b] - (X_n - a)^-$$

where $(X_N - a)^-$ is the upper bound for the "loss" due to the last unfinished upcrossing. Since C is non-negative, bounded, and previsible, Y is a supermartingale. Moreover, $\mathbb{E}(Y_1) \leq 0$, so $\mathbb{E}(Y_N) \leq \mathbb{E}(Y_1) \leq 0$. Thus

$$(b-a)\,\mathbb{E}\,U_[a,b] \leq \mathbb{E}[(X_n - a)^i]$$

as required. ■

1.4 Theorem. *Suppose $\{X_t\}_{t \in T}$ is a martingale with $\mathbb{E}(X_t^2) \leq B < \infty$ for any $t \in T$ and B not depending on t. Then X_t converge in L^2 to a limit X_∞.*

Proof Note that we have the result for both discrete and continuous time martingales. By discretization, it is clear that it suffices to rove the result for discrete case.

We have the following orthogonality between the increments of a martingale $\{X_n\}_{n=0}^{\infty}$: if $n_1 < n_2 \leq n_2 < n_4$, then

$$\mathbb{E}[(X_{n_2} - X_{n_1})(X_{n_4} - X_{n_3})] = 0.$$

This result follows by conditioning on $\mathcal{F}_{n_3}$ and applying the law of total expectation.

Now the proof proceeds. Set $Y_n := X_n - X_{n-1}$, so

$$\|X_n\|_2^2 = \mathbb{E}(X_n^2) = \sum_{i=1}^{n} \|Y_i\|_2^2$$

and $\sum_{i=0}^{n} \|Y_n\|_2^2 \leq B$ for all n, so that $\sum_{i=0}^{\infty} \|Y_n\|_2^2 \leq B$. Thus $\{X_n\}_{n=0}^{\infty}$ is Cauchy in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. ■

1.5 Theorem. (Optimal Sampling for Bounded Stopping Times in Discrete Times) *Let $\{X_n\}_{n=0}^{\infty}$ be a $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, \mathbb{P})$ supermartingale and S, T be $\{\mathcal{F}_n\}$–stopping times such that $0 \leq S \leq T \leq N$ for some constant $N < \infty$. Then X_T is integrable and $\mathbb{E}(X_T|\mathcal{F}_s) \leq X_s$ almost surely.*

Proof Notice that

$$|X_T| \leq |X_0| + |X_1| + \cdots + |X_N|$$

so that $\mathbb{E}(|X_T|) < \infty$. To prove that $\mathbb{E}(X_t|\mathcal{F}_s) \leq X_s$ a.s., it suffices to prove that $\mathbb{E}(X_T; A) := \int_A X_T \, d\mathcal{P} \leq \int_A X_s \, d\mathcal{P} =: \mathbb{E}(X_s; A)$ for all $A \in \mathcal{F}_s$. Assuming this, then

$$\mathbb{E}(\mathbb{E}(X_T|\mathcal{F}_s) - X_s; A) \leq 0$$

for all $A \in \mathcal{F}_s$, so we may take $A = A_0 := \{\mathbb{E}(X_T|\mathcal{F}_s) - X_s > 0\}$, so that $\mathbb{P}(A_0) = 0$.

Let's prove the required statement. First note that

$$\sum_{n=1}^{N} \chi_{\{S<n\leq T\}}(X_n - X_{n-1}) = X_T - X_S$$

and taking expectation over A on both sides

$$\begin{aligned}\mathbb{E}(X_T - X_S; A) &= \sum_{n=1}^{N} \mathbb{E}(\chi_{\{S<n\leq T\}}(X_n - X_{n-1}); A) \\ &= \sum_{n=1}^{N} \mathbb{E}(X_n - X_{n-1}; A \cap \{S < n \leq T\})\end{aligned}$$

But $A \cap \{S < n \leq T\} = A \cap \{S \leq n-1\} \cap \{n-1 < T\} \in \mathcal{F}_{n-1}$. Thus

$$\mathbb{E}(X_n - X_{n-1}; A \cap \{S < n \leq T\}) = \mathbb{E}(\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}); A \cap \{S < n \leq T\}) \leq 0$$

so the required statement holds. ■

Definition. Let $\{X_n\}_{n=0}^{\infty}$ be a $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, \mathbb{P})$ a supermartingale. We say that $\{X_n\}_{n=0}^{\infty}$ is **closed** by a random variable X_∞ if X_∞ is $\mathcal{F}_\infty$–measurable and $X_n \geq \mathbb{E}(X_\infty|\mathcal{F}_n)$ almost surely for all $n = 0, 1, \ldots$.

Similar statements hold with $X_n \leq \mathbb{E}(X_\infty|\mathcal{F}_n)$ for a submartingale, or equality with a martingale.

1.6 Proposition. *Suppose that $\{X_n\}_{n=0}^{\infty}$ is a $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, \mathbb{P})$ a non-negative supermartingale, and $X_\infty = 0$. If $S, T : \Omega \to \overline{\mathbb{Z}^+}$ are $\{\mathcal{F}_n\}$–stopping times, $S \leq T$, then*

1. $\mathbb{E}(X_T) < \infty$
2. $\mathbb{E}(X_T|\mathcal{F}_s) \le X_s$

Proof 1. Note that $X_T \le \liminf_{n\to\infty} X_{T\wedge n}$ where $T \wedge n$ and 0 are two bounded stopping times. Thus $\mathbb{E}(X_{T\wedge n}) \le \mathbb{E}(X_0)$ for all $n = 0, 1, \ldots$. Thus by Fatou's lemma

$$\begin{aligned}\mathbb{E}(X_T) &\le \mathbb{E}(\liminf_{n\to\infty} X_{T\wedge n}) \le \liminf_{n\to\infty} \mathbb{E}(X_{T\wedge n}) \\ &\le \mathbb{E}(X_0) < \infty.\end{aligned}$$

2. Let $A \in \mathcal{F}_s$. For $n = 0, 1, \ldots,$,

$$\begin{aligned}\mathbb{E}(X_T; A \cap \{T \le n\}) &= \mathbb{E}(X_{T\wedge n}; A \cap \{T \le n\}) \\ &\le \mathbb{E}(X_{T\wedge n}; A \cap \{S \le n\}) \\ &\le \mathbb{E}(X_{S\wedge n}; A \cap \{S \le n\})\end{aligned}$$

Note that $S \wedge n$ and $T \wedge n$ are two bounded stopping times with $S \wedge n \le T \wedge n$. Also, $A \cap \{S \le n\} \in \mathcal{F}_{S\wedge n}$. Then apply the optimal sampling theorem for bounded stopping times. By the monotone convergence theorem,

$$\lim_{n\to\infty} \mathbb{E}(X_T; A \cap \{T \le n\})$$

and similarly for S. Thus

$$\mathbb{E}(X_T; A \cap \{T < \infty\}) \le \mathbb{E}(X_S : A \cap \{S < \infty\})$$

so that

$$\mathbb{E}(X_T; A \cap \{T = \infty\}) = \mathbb{E}(X_S : A \cap \{S = \infty\}) = 0$$

and $\mathbb{E}(X_T; A) \le \mathbb{E}(X_S; A)$. Since this holds for all $A \in \mathcal{F}_s$, $\mathbb{E}(X_T : \mathcal{F}_s) \le X_s$. ■

1.7 Lemma. *Let $\{M_n\}_{n=0}^{\infty}$ be a martingale closed by M_∞. Let $T : \Omega \to \overline{\mathbb{Z}^+}$ be a stopping time. Then $M_T = \mathbb{E}(M_\infty|\mathcal{F}_T)$.*

Proof First assume $M_\infty \ge 0$. For any $A \in \mathcal{F}_T$,

$$\begin{aligned}E(M_T; A) &= \sum_{n=0}^{\infty} \mathbb{E}(M_n; A \cap \{T = n\}) \\ &= \sum_{n=0}^{\infty} \mathbb{E}(M_\infty, A \cap \{T = n\}) \\ &= \mathbb{E}(M_\infty; A).\end{aligned}$$

For the general case, decompose M_∞ into positive and negative parts. ■

1.8 Theorem. (Optimal Sampling for Closed Supermartingales in Discrete Time) *Let $\{X_n\}_{n=0}^{\infty}$ be a $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, \mathbb{P})$ supermartingale closed by X_∞. Let $S, T : \Omega \to \overline{\mathbb{Z}^+}$ be two $\{\mathcal{F}_n\}_{n=0}^{\infty}$ stopping times with $S \le T$. Then*

1. $\mathbb{E}(|X_T|) < \infty$

2. $\mathbb{E}(X_T|\mathcal{F}_s) \leq X_s$ *almost surely*

Proof 1. Define $M_n = \mathbb{E}(X_\infty|\mathcal{F}_n)$ and $A_n = X_n - M_n$ for $n = 0, 1, \dots, \infty$. Since $\{X_n\}_{n=0}^\infty$ is a supermartingale closed by X_∞, $A_n \geq 0$, $A_\infty = 0$, and for $m \leq n$,

$$\begin{aligned}\mathbb{E}(A_n|\mathcal{F}_m) &= \mathbb{E}(X_n - \mathbb{E}(X_\infty|\mathcal{F}_n)|\mathcal{F}_m) \\ &\leq X_m - \mathbb{E}(X_\infty|\mathcal{F}_m) \\ &= A_m\end{aligned}$$

Thus $\{A_n\}_{n=0}^\infty$ is a non-negative supermartingale with $A_\infty = 0$.
One can prove that $\{M_n\}_{n=0}^\infty$ is a uniformly integrable martingale, i.e. $\lim_{c\to\infty} \sup_{n\in\mathbb{N}} \mathbb{E}(|X_n|; |X_n| > c) = 0$.
By the previous proposition, $\mathbb{E}(A_T) < \infty$. On the other hand, $\mathbb{E}(|M_T|) < \infty$ by definition of the $\{M_n\}$. Thus $\mathbb{E}(|X_T|) < \infty$.

2. Apply the previous lemma so $\mathbb{E}(M_T|\mathcal{F}_s) = \mathbb{E}(\mathbb{E}(M_\infty|\mathcal{F}_T)\mathcal{F}_S) = \mathbb{E}(M_\infty|\mathcal{F}_s) = M_s$ by the optimal sampling theorem for closed martingales. Meanwhile, as $\{A_n\}_{n=0}^\infty$ is a non-negative supermartingale with $A_\infty = 0$, $\mathbb{E}(A_T|\mathcal{F}_s) \leq A_s$ almost surely. Thus $\mathbb{E}(X_T|\mathcal{F}_s) \leq X_s$ almost surely. ∎

To prepare for the optimal sampling theorem for closed supermartingales in continuous time, we introduce one more discrete time notion:

Definition. A **negatively indexed supermartingale** $\{X_n\}_{n=0,-1,\dots}$ and $\{\mathcal{F}_n\}_{n=0,-1,\dots}$ where $\mathcal{F}_m \subseteq \mathcal{F}_n$ for $m \leq n$. Then $X_n \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, $\{X_n\}$ is adapted and $\mathbb{E}(X_n|\mathcal{F}_m) \leq X_m$.

1.9 Theorem. *Let* $\{X_n\}_{n=0,-1,\dots}$ *be a negatively indexed supermartingale such that* $\sup_{-\infty<n\leq 0} \mathbb{E}(X_n) < \infty$. *Then* $\{X_n\}_{n=0,-1,\dots}$ *is uniformly integrable.*