Submartingale Convergence

A Measure-Theoretic Approach

Robbie Nohra

A paper submitted for *Topics in Ergodic Theory: Introduction to Random Walks on Groups* under the guidance of Dr Tiozzo

Department of Mathematics
The University of Toronto
12 December 2016

Submartingale Convergence

A Measure-Theoretic Approach

Robbie Nohra

Abstract

A measure-theoretic proof of submartingale convergence is provided with a historical overview.

Contents

Chapter 1

Introduction

1.1 History

> An old man, who had spent his life looking for a winning formula (martingale), spent the last days of his life putting it into practice, and his last pennies to see it fail. The martingale is as elusive as the soul.
>
> —Alexandre Dumas Père
> *Les mille et un fantômes, 1849*

For centuries, gamblers have been preoccupied with probability and its potential uses for developing a "winning" strategy. Indeed, the term *martingale* — which as mathematicians we understand as a sequence of random variables satisfying certain conditional expectation properties — actually has equestrian origins.

The dictionary definition of a martingale is:

> "...a strap, or set of straps, attached at one end to the noseband (standing martingale) or reins (running martingale) of a horse and at the other end to the girth. It is used to prevent the horse from raising its head too high."

(Although one would guess that the word's relation to both horses and gambling likely has to do with betting on horses, this has not been corroborated.)

The first citation of the term "martingale" is found in the thesis of Jean Ville (a French mathematician from the early 19th century). Specifically, the word is introduced in chapter IV, third paragraph, in the expression "game system or martingal". Jean Ville discarded the term "game system" in subsequent chapters and exclusively used the term "martingale". He also later mentioned that the name was borrowed directly from the vocabulary of gamblers.

The Académie Française (the French Academy) — the authority on the French language at the time— introduced the word into the French dictionary in 1762; they defined it as such:

> "To play the martingale is to always bet all that was lost."

Earlier, in 1750, Abbé Prévost (a French author and novelist) described a strategy in which the gambler doubles his bet at each loss "in order to quit with a sure profit, provided that he wins once". The strategy is often called"d'Alembert's martingale", after the French mathematician Jean-Baptiste le Rond d'Alembert. [1]

It is in this spirit that the term *martingale* arose. If the outcomes of a series of bets are known, is there an optimal strategy for the next bet? Ironically, while martingales (as gambling strategies) were used in the hope that knowledge of past outcomes might inform one's next bet, the probabilistic definition as it is understood today describes a scenario in which knowledge of past outcomes is useless.

Just over a century later, in 1933, Kolmogorov published his *Foundations of the Theory of Probability*, establishing the modern axiomatic foundations of probability theory. One year prior, Joseph Doob graduated from Harvard. The Great Depression was still in effect, and Doob is quoted as saying:

> "B. O. Koopman at Columbia told me that I should approach Harold Hotelling, professor of statistics there, that there was money in probability and statistics. Hotelling said he could get me a Carnegie Corporation grant to work with him, and thus the force of economic circumstances got me into probability." [2]

Kolmogorov, Doob and others, equipped with a more rigorous measure-theoretic basis of probability, proceeded to explore martingales in more detail. As a result, the martingale convergence theorem was born and is today used in a variety of fields, especially in stochastic processes and finance.

1.2 Preliminary Definitions

Definition 1. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $(\mathcal{A}_j)$ be a sequence of σ-algebras such that:

$$\mathcal{A}_0 \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_j \subset \cdots \subset \mathcal{A} \ .$$

Such a sequence of is called a *filtration*.

Recall that a σ-algebra defines a collection of measurable subsets of the underlying space. In a probability context these measurable subsets are referred to as "events". Therefore, a filtration can be thought to represent a discrete time-change of the underlying σ-algebra under consideration, as a function of a gain or loss of information.

For example, in mathematical finance the underlying sigma algebra often changes throughout the time-evolution of a given stock price.

Definition 2. Let $(X, \mathcal{A}, \mu)$ be a measure space and let (A_j) be a respective filtration. If $(X, \mathcal{A}_0, \mu)$ is σ-finite then we call $(X, \mathcal{A}, \mathcal{A}_j, \mu)$ a *σ-finite filtered measure space .*

The σ-algebra generated by the union $\cup \mathcal{A}_k$ is denoted as $\mathcal{A}_\infty := \sigma(\cup \mathcal{A}_j)$.

*The σ-finite condition is standard when dealing with probability measures (which is how we wish to interpret μ in the context of this paper) .

Definition 3. Let $(X, \mathcal{A}, \mathcal{A}_j, \mu)$ be a σ-finite filtered measure space and let (u_j) be a sequence of $\mathcal{A}$- measurable functions.

If

1. $u_j \in L^1(\mathcal{A}_j)$ for all j
2. $\displaystyle\int_A u_{j+1} = \int_A u_j$ for all $A \in \mathcal{A}_j$

then (u_j) is called a *martingale* with respect to the filtration $(\mathcal{A}_j)$.

If $\displaystyle\int_A u_{j+1} \geq \int_A u_j$ for all $A \in \mathcal{A}_j$ then (u_j) is called a *submartingale* .

Table 1.1 illustrates the probabilistic representation of the conditions defining a martingale as compared with the above measure-theoretic representation.

Measure-theoretic	Probablistic
Sequence (u_j) of $\mathcal{A}$-measurable functions	Sequence of random variables (X_j)
$u_j \in L^1(\mathcal{A}_j)$ for all j	$\mathbb{E}(X_j) < \infty$ for all j
$\displaystyle\int_A u_{j+1} = \int_A u_j$ for all $A \in \mathcal{A}_j$	$\mathbb{E}(X_{j+1} \vert X_1, \ldots, X_j) = X_j$

Table 1.1

Chapter 2

Submartingale Convergence

2.1 The Martingale Convergence Theorem

Theorem 1. If (u_j) be a submartingale on the the σ-finite filtered measure space $(X, \mathcal{A}, \mathcal{A}_j, \mu)$ such that $\sup_j \int u_j^+ < \infty$ for all j (where $u_j^+ = \max(u, 0) := u \vee 0$) then

$$u_\infty(x) := \lim_{j \to \infty} u_j(x)$$

exists for almost all $x \in \mathbb{R}$ and defines an $\mathcal{A}_\infty$-measurable function.

2.2 Upcrossings

To motivate a sketch of the proof of Theorem 1 suppose first that (u_j) was a sequence of real numbers such that $u_j \longrightarrow \ell < \infty$ where $\ell \in (a, b)$.

Then it is necessary that $u_j \in \mathbb{R} \backslash (a, b)$ for *finitely* many j.

That is, if $u_j > b$ and $u_j < a$ for *infinitely* many j for some $a < b$ then $u_j \not\to \ell$ for any $\ell \in \mathbb{R}$.

Definition 4. If $u_j \leq a$ and $u_{j+k} \geq b$ for some k then we call such an occurrence an *upcrossing* of $[a, b]$.

The prior statement can then be translated as saying that if u_j has an infinite number of upcrossings then $\lim\limits_{j\to\infty} u_j$ does not exist. $\circledast$

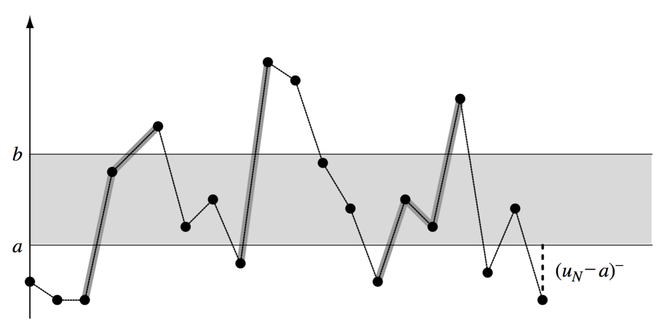
Lemma 1. If $\lim\limits_{j\to\infty} u_j$ d.n.e. then $u_j > b$ and $u_j < a$ for infinitely many j, for some $a, b \in \mathbb{R}$ (i.e. the converse of $\circledast$ is also true—if $\lim_{j\to\infty} u_j$ d.n.e. then u_j has an infinite number of upcrossings of $[a, b]$ for some $a, b \in \mathbb{R}$) .

Proof. $\lim\limits_{j\to\infty} u_j$ dne $\Rightarrow \limsup\limits_j u_j > \liminf\limits_j u_j$. If $\limsup u_j = \infty$ or $\liminf u_j = -\infty$ then the result follows trivially. Suppose that $|\limsup u_j|, |\liminf u_j| < \infty$. Let $a, b \in (\liminf u_j, \limsup u_j)$ such that $a < b$. Choose $\epsilon > 0$ small enough such that $b < \limsup u_j - \epsilon < \limsup u_j$ and $\liminf u_j < \liminf u_j + \epsilon < a$.

By definition of $\limsup u_j$ and $\liminf u_j$ (in the case that both are finite as they are by hypothesis) $u_j > \limsup u_j - \epsilon > b$ and $u_j < \liminf u_j + \epsilon < a$ for infinitely many j .

□

Figure 2.1 Upcrossings of u_j



2.3 Stopping Times

To count upcrossings of a martingale (u_j) we must be able to admit indices which may depend on x .

i.e. we must consider martingales of the form:

$$u_j(x) := u_{\tau(x)}(x)$$

Definition 5. Let $(X, \mathcal{A}, \mathcal{A}_j, \mu)$ be a σ-finite filtered measure space. Let $\tau : X \to \mathbb{N} \cup \{\infty\}$ be a map such that:

$$\{\tau \leq j\} := \{x \in X : \tau(x) \leq j\} \in \mathcal{A}_j \text{ for all } j \ .$$

Such a map τ is referred to as a *stopping time*.

The σ-algebra generated by τ is denoted as:

$$\mathcal{A}_\tau := \{A \in \mathcal{A} : A \cap \{\tau \leq j\} \in \mathcal{A}_j \text{ for all } j\} \ .$$

Lemma 2. Let $(X, \mathcal{A}, \mathcal{A}_j, \mu)$ be a σ-finite filtered measure space and (u_j) a submartingale such that $u_j \in L^1(\mathcal{A}_j)$ for each j . If $\sigma \leq \tau \leq N$ are bounded stopping times then:

$$\int_A u_\sigma \leq \int_A u_\tau$$

for all $A \in A_\sigma$.

For a proof of Lemma 2 see Appendix.

Our goal is to use stopping times to establish a maximal inequality.

2.4 Doob's Upcrossing Estimate

Theorem 2. Let (u_j) be a submartingale and $(U([a,b];N;x)$ the number of upcrossings of $(u_j(x))$ across $[a,b]$ for $1 \leq j \leq N$ for some $a,b \in \mathbb{R}$ where $a < b$.

Then
$$\int_A (U([a,b];N;x) \leq \frac{1}{b-a} \int_A (u_N - a)^+$$

for all $A \in \mathcal{A}_0$ $\Big($where $a = u_0(x)\Big)$.

Proof. Introduce the following stopping times:

- $\sigma_k = \min(\inf\{j > \tau_{k-1} : u_j \leq a\}, N)$
- $\tau_k = \min(\inf\{j > \sigma_k : u_j \geq b\}, N)$

Where $\tau_0 = 0$.

Then σ_k denotes the k^{th} occurrence of the event that $u_j \leq a$ and τ_k denotes the k^{th} occurrence of the event that $u_j \geq b$.

It follows from the definition that

$$\tau_0 = 0 < \sigma_1 \leq \tau_1 \leq \sigma_2 \leq \cdots \leq \sigma_N = \tau_N = N \tag{2.1}$$

In particular,

$$u_{\tau_k} \geq b, u_{\sigma_k} \leq a \implies u_{\tau_k} - u_{\sigma_k} \geq b - a \tag{2.2}$$

for each k .

(For the sake of brevity we use the shorthand $u_\tau(x) = u_\tau$)

Then

$$\underbrace{(u_{\tau_1} - a)}_{\geq b-a} + \underbrace{(u_{\tau_2} - u_{\sigma_2})}_{\geq b-a} + \cdots + \underbrace{(u_{\tau_N} - u_{\sigma_N})}_{\geq b-a} \geq U\big([a,b];N\big) \cdot (b-a) \tag{2.3}$$

$\Big($by (2.2) and the definition of $U([a,b];N)$ $\Big)$.

Rearranging (2.3) gives

$-a \quad + \quad (u_{\tau_1} \quad - \quad u_{\sigma_2}) \quad + \quad \cdots \quad + \quad (u_{\tau_{N-1}} \quad - \quad u_{\sigma_N}) \quad + \quad u_{\tau_N}$

Let $A \in \mathcal{A}_0$.

Then

$$(b-a)\int_A U([a,b];N) \tag{2.4}$$

$$\leq -\int_A a + \underbrace{\int_A (u_{\tau_1} - u_{\sigma_2})}_{\leq 0} + \cdots + \underbrace{\int_A (u_{\tau_{N-1}} - u_{\sigma_N})}_{\leq 0} + \int_A u_{\tau_N} \tag{2.5}$$

$$\leq \int_A (u_{\tau_N} - a) \leq \int_A (u_{\tau_N} - a)^+. \tag{2.6}$$

Where:

- The non-positivity of the integrals in (2.4) follows by (2.1) and Theorem 2
- (2.6) follows by the definition of u^+

Dividing by $b-a$ proves Theorem 3.

□

Note that Doob's estimate depends only on the the endpoints u_{τ_N} and $u_0 = a$ — which is what we would expect given the defining submartingale property.

2.5 The Proof

We are ready to prove Theorem 1 — *The Martingale Convergence Theorem* .

Theorem 1. If (u_j) be a submartingale on the the σ-finite filtered measure space $(X, \mathcal{A}, \mathcal{A}_j, \mu)$ such that $\sup_j \int u_j^+ < \infty$ for all j (where $u_j^+ = \max(u, 0) := u \vee 0$) then

$$u_\infty(x) := \lim_{j\to\infty} u_j(x)$$

exists for almost all $x \in \mathbb{R}$ and defines an $\mathcal{A}_\infty$-measurable function.

Proof. For the first part of the theorem it is enough to show that

$$\left\{x : \lim_{j\to\infty} u_j(x) \text{ dne}\right\}$$

is a null set.

$$\left\{x : \lim_{j\to\infty} u_j(x) \text{ dne}\right\} = \left\{x : \limsup_{j\to\infty} u_j(x) > \liminf_{j\to\infty} u_j(x)\right\} \tag{2.7}$$

$$= \bigcup_{\substack{a<b \\ a,b\in\mathbb{Q}}} \left\{x : \sup_{n\in\mathbb{N}} U([a,b]; N; x) = \infty\right\}. \tag{2.8}$$

Where (2.7) $\longrightarrow$ (2.8) holds by Lemma 1.

Claim: the above set is null.

Recall that $\mu|_{\mathcal{A}_0}$ is σ-finite by hypothesis. Then there exists an exhaustive sequence $(A_k) \subset \mathcal{A}_0$ such that $A_k \uparrow X$ and $\mu(A_k) < \infty$.

Also note that, in general, the inequality $(\beta - \alpha)^+ \leq \beta^+ + |\alpha|$ holds.

Let A_k be a member of the exhaustive sequence mentioned above.

Then

$$\int_{A_k} \sup_{N\in\mathbb{N}} U([a,b];N) = \sup_{N\in\mathbb{N}} \int_{A_k} U([a,b];N) \tag{2.9}$$
$$\leq \frac{1}{b-a} \sup_{N\in\mathbb{N}} \int_{A_k} (u_N - a)^+ \qquad \text{(Doob)} \tag{2.10}$$
$$\leq \frac{1}{b-a} \left(\sup_{N\in\mathbb{N}} \int_{A_k} u_N^+ + |a|\mu(A_k) \right) \tag{2.11}$$

By writing

$$\left\{ x : \sup_{N\in\mathbb{N}} U([a,b];N;x) = \infty \right\} = \bigcap_j \left\{ x : \sup_{N\in\mathbb{N}} U([a,b];N;x) \geq j \right\} ,$$

a routine application of the Markov inequality, combined with the above finiteness result, gives

$$\mu \left(\left\{ x : \sup_{N\in\mathbb{N}} U([a,b];N;x) = \infty \right\} \cap A_k \right) = 0 \ .$$

Therefore,

$$\bigcup_{\substack{a<b \\ a,b\in\mathbb{Q}}} \left\{ x : \sup_{n\in\mathbb{N}} U([a,b];N;x) = \infty \right\} \tag{2.12}$$
$$= \bigcup_{\substack{a<b \\ a,b\in\mathbb{Q}}} \left\{ x : \sup_{n\in\mathbb{N}} U([a,b];N;x) \infty \right\} \cap X \tag{2.13}$$
$$= \bigcup_{\substack{a<b \\ a,b\in\mathbb{Q}}} \left\{ x : \sup_{n\in\mathbb{N}} U([a,b];N;x) \infty \right\} \bigcap (\cup A_k) \tag{2.14}$$
$$= \bigcup_{k\in\mathbb{N}} \bigcup_{\substack{a<b \\ a,b\in\mathbb{Q}}} \left(\left\{ x : \sup_{N\in\mathbb{N}} U([a,b];N;x) = \infty \right\} \cap A_k \right) \tag{2.15}$$

is a countable union of null sets and therefore itself null.

$\therefore u_\infty(x) := \lim_{j\to\infty} u_j(x)$ exists for almost all $x \in \mathbb{R}$.

Finally we want to show that u_∞ is $\mathcal{A}_\infty$-measurable .

But

$$\int |u_\infty| = \int \liminf_{j\to\infty} |u_j| \overset{\text{Fatou}}{\leq} \liminf_{j\to\infty} \int |u_j| \leq \sup_{j\in\mathbb{N}} \int |u_j|$$

and $|u_j| = u_j^+ + u_j^-$, $u_j = u_j^+ - u_j^- \implies |u_j| = 2u_j^+ - u_j$.

Therefore,

$$\sup_{j\in\mathbb{N}} \int |u_j| = \sup \left(2 \int u_j^+ - \int u_j\right) \tag{2.16}$$

$$\leq 2 \sup_{j\in\mathbb{N}} \int u_j^+ \, d\mu - \int u_1 \tag{2.17}$$

$$< \infty \tag{2.18}$$

Where

- (2.16) $\longrightarrow$ (2.17) follows from the fact that (u_j) is a martingale
- (2.18) follows from the fact that $\sup_{j\in\mathbb{N}} \int u_j^+ < \infty$ for all j by hypothesis

This concludes the proof.

□

Chapter 3

Appendix

Lemma 2. Let $(X, \mathcal{A}, \mathcal{A}_j, \mu)$ be a σ-finite filtered measure space and (u_j) a submartingale such that $u_j \in L^1(\mathcal{A}_j)$ for each j . If $\sigma \leq \tau \leq N$ are bounded stopping times then:

$$\int_A u_\sigma \leq \int_A u_\tau$$

for all $A \in A_\sigma$.

Proof. For the sake of brevity we will omit the proofs that if σ, τ are stopping times such that $\sigma \leq \tau$ then $\sigma \wedge \tau$ and $\sigma + k$ are also stopping times for any $k \in \mathbb{N}$ and $\mathcal{A}_\sigma \subset \mathcal{A}_\tau$.

Firstly, note that $u_\tau, u_\sigma \in L^1(X, \mathcal{A}, \mu)$ since

$$\int |u_\tau| = \sum_{j=1}^{N} \int_{\{\tau=j\}} |u_j| \leq \sum_{j=1}^{N} \int |u_j| < \infty$$

where the last step follows by the $\mathcal{A}$-measurability of u_j.

Next, introduce the intermediate stopping times

$$\rho_j := (\sigma + j) \wedge \tau$$

where $j = 0, 1, 2, \ldots, k \leq N$.

It follows from this definition that

1. $\sigma = \rho_0 \leq \rho_1 \leq \cdots \leq \rho_k = \tau$ for some $k \leq N$
2. $\rho_{j+1} - \rho_j \leq 1$

In general, suppose that $\sigma \leq \tau$ and $\tau - \sigma \leq 1$.

Then $0 \leq \tau - \sigma \leq 1 \implies \tau = \sigma$ or $\tau = \sigma + 1$ ⊛

and

$$\int u_\sigma = \int_{\{\tau=\sigma\}} u_\sigma + \sum_{j=1}^{N-1} \int_{\{\sigma<\tau\}\cap\{\sigma=j\}} u_j \tag{3.1}$$

$$\leq \int_{\{\tau=\sigma\}} u_\tau + \sum_{j=1}^{N-1} \int_{\{\sigma<\tau\}\cap\{\sigma=j\}} u_{j+1} \tag{3.2}$$

$$= \int_{\{\tau=\sigma\}} u_\tau + \int_{\{\sigma<\tau\}} u_\tau \tag{3.3}$$

$$= \int u_\tau \tag{3.4}$$

where

- (3.1) follows by splitting the integral into the case $\sigma = \tau$ and $\sigma < \tau$
- (3.1) $\longrightarrow$ (3.2) because (u_j) is a submartingale
- (3.2) $\longrightarrow$ (3.3) because $\sigma = j$ and $\sigma < \tau \implies \tau = j + 1$ (by ⊛)

By the two properties of ρ_k listed above we can iterate this result to obtain

$$\int u_\sigma = \int u_{\rho_0} \leq \int u_{\rho_1} \leq \cdots \leq \int u_{\rho_k} = \int u_\tau$$

Note that $\{\sigma < \tau\} \cap \{\sigma = j\} = \{\tau > j\} \cap \{\sigma = j\} = \{\tau \leq j\}^c \cap \{\sigma = j\} \in \mathcal{A}_j$ so that the integrals in (3.1) and (3.2) are well-defined.

Finally, let $\rho := \sigma \cdot \chi_A + \tau \cdot \chi_{A^c}$.

$\{\rho \leq j\} = (\{\sigma \leq j\} \cap A) \cup (\{\tau \leq j\} \cap A^c) \in \mathcal{A}_j \implies \rho$ is also a bounded stopping time

(where we use the fact that $\sigma \leq \tau \implies \mathcal{A}_\sigma \subset \mathcal{A}_\tau$)

Therefore, from the previous result,

$$\int (u_\sigma \cdot \chi_A + u_\tau \cdot \chi_{A^c}) = \int u_\rho \leq \int u_\tau$$
$$\implies \int_A u_\sigma = \int u_\sigma \cdot \chi_A \leq \int u_\tau \cdot \chi_{A^c} = \int_A u_\tau$$

□

Bibliography

[1] Mansuy, R. (2009). *The Origins of the Word "Martingale"*. Electronic Journal for History of Probability and Statistics, 5(1).

[2] Snell, J. (1997). *A conversation with Joe Doob*. Statistical Science, 12(4).

[3] Schilling, R. (2005). *Measures, Integrals and Martingales*. Cambridge University Press.