Discrete time martingales

Aureli Alabert

Departament de Matemàtiques
Universitat Autònoma de Barcelona

Draft Version: 2004/4/15. A. Alabert

Table of contents

1. Definitions and examples

$(\Omega, \mathfrak{F}, P)$ probability space.

$\{\mathfrak{F}_n,\ n \geqslant 0\}$ increasing sequence of sub-σ-fields of $\mathfrak{F}$ =: *Filtration.*

A filtration models the evolution of the information: $A \in \mathfrak{F}_n$ means that at time n we can tell if the event A has ocurred or not.

1.0.1 Def. $X = \{X_n,\ n \geqslant 0\}$ is a *martingale w.r.t.* $\{\mathfrak{F}_n,\ n \geqslant 0\}$ if, $\forall n$,

1) $X_n \in L^1$.
2) X_n is $\mathfrak{F}_n$-measurable.
3) $\mathrm{E}\left[{}^{X_{n+1}}/{}_{\mathfrak{F}_n}\right] = X_n$. □

The martingale property is a property of the law of the process, just like the Markov dependence, the independence of the variables of the process, etc.

Notation: $\{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$.

1.0.2 Def. $\big\{\sigma\{X_1, \dots, X_n\},\ n \geqslant 0\big\}$ is the *natural filtration* of the process X.

$\{X_n,\ n \geqslant 0\}$ is a *martingale* if satisfies 1,2 of the previous definition and $\mathrm{E}\left[{}^{X_{n+1}}/{}_{X_0, \dots, X_n}\right] = X_n$. □

1.0.3 Def. $X = \{X_n,\ n \geqslant 0\}$ is a *submartingale (resp. supermartingale) w.r.t.* $\{\mathfrak{F}_n,\ n \geqslant 0\}$ if, $\forall n$,

1) $X_n \in L^1$.
2) X_n is $\mathfrak{F}_n$-measurable.
3) $\mathrm{E}\left[{}^{X_{n+1}}/{}_{\mathfrak{F}_n}\right] \geqslant X_n$ (resp. $\mathrm{E}\left[{}^{X_{n+1}}/{}_{\mathfrak{F}_n}\right] \geqslant X_n$). □

1.0.4 First (obvious) property.

1) X submartingale $\Rightarrow \mathrm{E}[X_{n+1}] \geqslant \mathrm{E}[X_n]$.
2) X supermartingale $\Rightarrow \mathrm{E}[X_{n+1}] \leqslant \mathrm{E}[X_n]$.
3) X martingale $\Rightarrow \mathrm{E}[X_{n+1}] = \mathrm{E}[X_n] = \mathrm{E}[X_0]$. □

1.0.5 Example 1.

a) $X_n \equiv X_0 \Rightarrow \{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a martingale.
b) $X_{n+1} \geqslant X_n \Rightarrow \{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a submartingale.
c) $X_{n+1} \leqslant X_n \Rightarrow \{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a supermartingale.

(assuming properties 1,2 satisfied).

The proof is very easy. For instance,

$$X_{n+1} \geqslant X_n \Rightarrow \mathrm{E}\left[{}^{X_{n+1}}/_{\mathfrak{F}_n}\right] \geqslant \mathrm{E}\left[{}^{X_n}/_{\mathfrak{F}_n}\right] = X_n$$

1.0.6 Example 2. If

a) $\{Y_n,\ n \geqslant 0\}$ indep. r.v. in L^1, centred.

b) $X_n = \sum_{k=0}^n Y_k$.

c) $\mathfrak{F}_n = \sigma\{Y_0, \dots, Y_n\}$.

then $\{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a martingale.

Proof:

$$\mathrm{E}\left[{}^{X_{n+1}}/_{\mathfrak{F}_n}\right] = \mathrm{E}\left[{}^{\sum_{k=0}^{n+1} Y_k}/_{\mathfrak{F}_n}\right] = \sum_{k=0}^{n} Y_k + \mathrm{E}\left[{}^{Y_{n+1}}/_{\mathfrak{F}_n}\right] = X_n + \mathrm{E}[Y_{n+1}].$$

Note that if $\mathrm{E}[Y_n] \geqslant 0$, $\forall n$, we obtain a submartingale; if $\mathrm{E}[Y_n] \leqslant 0$, $\forall n$, we obtain a supermartingale.

1.0.7 Example 3. This is a particular case of Example 2, with additional comments.

Take

$$Y_n = \begin{cases} +1, & \text{with prob. } 1/2 \\ -1, & \text{with prob. } 1/2 \end{cases}$$

and define $X_n = x_0 + \sum_{k=1}^n Y_k$. This process models the evolution of your fortune when you are playing a game "heads/tails" or roulette "black/red", with initial fortune x_0.

By Example 2, this is a martingale. The property

$$\mathrm{E}\left[{}^{X_{n+1}}/_{Y_1, \dots, Y_n}\right] = X_n$$

says that, in each play, we don't expect to win or lose anything.

Let us elaborate on this example. What about intelligent strategies for winning? For example:

- $n = 1$: We bet 1 euro.
- $n = 2$: $\begin{cases} \text{If we won at time } n = 1, \text{ we bet 1 euro.} \\ \text{If we lost at time } n = 1, \text{ we bet 2 euros.} \end{cases}$
- ...
- At step n, $\begin{cases} \text{If we won at time } n-1, \text{ we bet 1 euro.} \\ \text{If we lost at time } n-1, \text{ we double the previous stake.} \end{cases}$

Our "strategy" is a process ϕ, defined by $\phi_1 = 1$ and

$$\phi_n = \begin{cases} 1, & \text{if } Y_{n-1} = 1 \\ 2\phi_{n-1}, & \text{if } Y_{n-1} = -1 \end{cases}$$

Note that, after losing n times in a row, and then winning once, our net profit will always be

$$-2^0 - 2^1 - 2^2 - \cdots - 2^{n-1} + 2^n = 1\ ,$$

but if we abandon the game in the middle of a losing run, we can lose a lot of money.

Notice also that ϕ_n is a $\mathfrak{F}_{n-1}$ measurable r.v. A process $\phi = \{\phi_n,\ n \geqslant 1\}$ with this property is said to be *predictable w.r.t.* $\{\mathfrak{F}_n,\ n \geqslant 0\}$.

In general, for any predictable strategy ϕ, and assuming that the profit in each play per unit bet is +1 or -1, then the quantity $\phi_n Y_n$ is the net profit in game n, and the total profit after n games is

$$\sum_{k=1}^{n} \phi_k Y_k = \sum_{k=1}^{n} \phi_k (X_k - X_{k-1}) =: (\phi \bullet X)_n.$$

This process is called the *stochastic integral* of ϕ w.r.t. X. (The name comes from an analogous concept in continuous time, which is not so easy to define).

In turns out that the stochastic integral $\phi \bullet X$ of a predictable process w.r.t. a $\mathfrak{F}_n$-martingale is again a $\mathfrak{F}_n$-martingale, provided that its variables belong to L^1. Indeed:

$$\begin{aligned} \mathrm{E}\left[(\phi \bullet X)_{n+1}/_{\mathfrak{F}_n}\right] &= \sum_{k=1}^{n} \phi_k (X_k - X_{k-1}) + \mathrm{E}\left[\phi_{n+1}(X_{n+1} - X_n)/_{\mathfrak{F}_n}\right] \\ &= (\phi \bullet X)_n + \phi_{n+1} \cdot \mathrm{E}\left[X_{n+1} - X_n/_{\mathfrak{F}_n}\right] = (\phi \bullet X)_n \ . \end{aligned}$$

Notice that a similar statement can be stated for sub- and supermartingales: If X is a sub- (resp. super-) martingale, ϕ is predictable and $\phi \geqslant 0$, then $\phi \bullet X$ is a sub- (resp. super-) martingale. (We will use this fact in Chapter 2.)

1.0.8 Example 4. If $Y \in L^1$ and $\{\mathfrak{F}_n,\ n \geqslant 0\}$ is a filtration, then $X_n := \mathrm{E}\left[Y/_{\mathfrak{F}_n}\right]$ is an $\mathfrak{F}_n$-martingale.

Interpretation: This is the evolution of the best prediction of Y with the information known up to time n.

Proof:

$$\mathrm{E}\left[X_{n+1}/_{\mathfrak{F}_n}\right] = \mathrm{E}\left[\mathrm{E}\left[Y/_{\mathfrak{F}_{n+1}}\right]/_{\mathfrak{F}_n}\right] = \mathrm{E}\left[Y/_{\mathfrak{F}_n}\right] = X_n \ .$$

1.0.9 Example 5. (Dyadic martingales).

It's a specific situation of Example 4, which helps visualizing what we are doing.

$\Omega = [0,1]$, $\mathfrak{F} = \mathfrak{B}([0,1])$, $P = \mathrm{Unif}([0,1])$.

$\mathfrak{F}_0 = \{\emptyset, \Omega\}$.

$\mathfrak{F}_1 = \sigma\{[0,1/2],\ [1/2,1]\}$.

$\mathfrak{F}_2 == \sigma\{[0,1/4],\ [1/4,2/4],\ [2/4,3/4],\ [3/4,1]\}$.

$\mathfrak{F}_n == \sigma\{[0,1/2^n],\dots,[(2^n-1)/2^n,1]\}$.

Take any $Y\colon [0,1] \to \mathbb{R}$ in L^1.

Define

$$X_n(\omega) = \frac{1}{P(I_k)} \int_{I_k} Y(\omega)\, d\omega = \mathrm{E}\left[Y/_{\mathfrak{F}_n}\right](\omega) \ ,$$

if $\omega \in I_k$, with I_k a dyadic interval of $\mathfrak{F}_n$. So, by Example 4, $\{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a martingale.

[Picture 1]

We will prove later that $X_n \to Y$ as $n \to \infty$ almost surely and in L^1.

1.0.10 Properties.

0) $(X_n, \mathfrak{F}_n)$ is a (sub-, super-) martingale $\Rightarrow X_n$ is a (sub-, super-) martingale w.r.t. its natural filtration.
1) $\{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a (sub-, super-) martingale $\Leftrightarrow \mathrm{E}\left[X_{n+k}/_{\mathfrak{F}_n}\right]$ is $(\geqslant, \leqslant) = X_n$.
2) A sub- or supermartingale with $\mathrm{E}[X_n] \equiv \mathrm{E}[X_0]$ is a martingale.
3) $\{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a submartingale $\Leftrightarrow \{(-X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a supermartingale.
4) X, Y submartingales w.r.t. $\mathfrak{F}_n$, and $a, b \geqslant 0 \Rightarrow aX + bY$ is a submartingale w.r.t. $\mathfrak{F}_n$.
5) X, Y submartingales w.r.t. $\mathfrak{F}_n \Rightarrow \max(X, Y)$ submartingale.
6) X submartingale; $f\colon \mathbb{R} \to \mathbb{R}$ convex, increasing; $f(X) \in L^1 \Rightarrow f(X)$ is a submartingale. (For instance, X submartingale $\Rightarrow X^+$ submartingale.)
7) X martingale; $f\colon \mathbb{R} \to \mathbb{R}$ convex; $f(X) \in L^1 \Rightarrow f(X)$ is a submartingale. (For instance, if X is a martingale in L^p $(p \geqslant 1)$, then X^p is a submartingale.)

Proof of 6:

$$\mathrm{E}\left[f(X_{n+1})/\mathfrak{F}_n\right] \geqslant f\Big(\mathrm{E}\left[X_{n+1}/\mathfrak{F}_n\right]\Big) \geqslant f(X_n) \; .$$

The first inequality is Jensen's, and the second is due to the submartingale property and the increasing nature of f.

Proof of 7:

$$\mathrm{E}\left[f(X_{n+1})/\mathfrak{F}_n\right] \geqslant f\Big(\mathrm{E}\left[X_{n+1}/\mathfrak{F}_n\right]\Big) = f(X_n) \; .$$

The first inequality is Jensen's as before, and the equality is the martingale property.

2. Convergence theorems

Convergence of sequences of random variables is obviously an important issue in Statistics. We have for instance the law of large numbers, which is a very satisfactory result in that it shows that the (theoretic) mean of a random variable arises as a limit of averaging when the number of observed values tends to infinity. Some statistical schools take this fact as a definition of expectation, or even as a definition of probability, when the random variables are indicators.

2.1 The basic convergence theorem

2.1.1 Theorem

$\{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ sub- or supermartingale such that

$$\sup_n \mathrm{E}[|X_n|] < \infty \qquad (\text{“bounded” in } L^1).$$

Then,

$$X_n \xrightarrow[n\to\infty]{\text{a.s.}} \ell\ , \quad \text{with } \ell \in L^1.$$

Proof: Let $\{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ be a supermartingale. We want to see that the set

$$\{\omega \in \Omega : \ \overline{\lim_{n\to\infty}}\, X_n(\omega) \neq \underline{\lim}_{n\to\infty}\, X_n(\omega)\}$$

has probability zero.

If for some fixed ω, it holds that $\overline{\lim}_{n\to\infty} X_n(\omega) \neq \underline{\lim}_{n\to\infty} X_n(\omega)$, then clearly we can find two rational numbers a, b with

$$\overline{\lim_{n\to\infty}}\, X_n(\omega) \leqslant a < b \leqslant \underline{\lim}_{n\to\infty}\, X_n(\omega)\ .$$

This amounts to say that $M_a^b(\omega) = \infty$, where $M_a^b(\omega)$ is the number of upcrossings of the interval $[a, b]$. We want to prove that

$$P\{\omega \in \omega : \ M_a^b(\omega) = \infty\} = 0$$

for all $a, b \in \mathbb{Q}$, and the countable union of these sets will have probability zero as well.

The following inequalities prove this fact (the first one will be proved later):

$$\mathrm{E}[M_a^b] \leqslant \frac{1}{b-a} \sup_n \mathrm{E}[(X_n - a)^-] = \frac{1}{b-a} \sup_n \Big[\int_{\{X_n \leqslant a\}} (a - X_n)\, dP \Big] \leqslant \frac{1}{b-a} \Big(|a| + \sup_n \mathrm{E}[|X_n|] \Big) < \infty\ .$$

And we can easily see that the limit ℓ is integrable:

$$\mathrm{E}[|\ell|] = \mathrm{E}[\lim_{n\to\infty} |X_n|] \leqslant \varliminf_{n\to\infty} \mathrm{E}[|X_n|] \leqslant \sup_n \mathrm{E}[|X_n|] < \infty \ ,$$

using Fatou lemma. □

Notice that we do not claim that the convergence takes place in L^1. There are counterexamples.

2.1.2 Lemma

$$\mathrm{E}[M_a^b] \leqslant \frac{1}{b-a} \sup_n \mathrm{E}[(X_n - a)^-] \ .$$

Proof: Suppose that $X_n - X_{n-1}$ are your winnings per euro bet on game n. Consider the following predictable strategy:

1. Wait until X gets below a.
2. Bet 1 euro until X gets above b.
3. Go to 1.

Formally,

$\phi_1 := \mathbf{1}_{\{X_0<a\}}$.

For $n > 1$, $\phi_n := \mathbf{1}_{\{\phi_{n-1}=1\}} \cdot \mathbf{1}_{\{X_{n-1}\leqslant b\}} + \mathbf{1}_{\{\phi_{n-1}=0\}} \cdot \mathbf{1}_{\{X_{n-1}<a\}}$

Define $Y = \phi \bullet X$ (your total winnings) $= \sum_{k=1}^n \phi_k(X_k - X_{k-1})$.

[Picture 2]

Then,

$$Y_n(\omega) \geqslant (b-a)\big[M_a^b(n)\big](\omega) - (X_n(\omega) - a)^- \ . \tag{2.1.1}$$

Y is a supermartingale (see Example 3 of Chapter 1, and notice that ϕ is bounded). Moreover $Y_0 \equiv 0$, so that $\mathrm{E}[Y_n] \leqslant 0$.

Taking expectations in (2.1.1),

$$\begin{aligned} 0 &\geqslant (b-a)\,\mathrm{E}\left[M_a^b(n)\right] - \mathrm{E}\left[(X_n-a)^-\right] \ , \\ &\Rightarrow \mathrm{E}\left[M_a^b(n)\right] \leqslant \frac{1}{b-a}\,\mathrm{E}\left[(X_n-a)^-\right] \ , \\ &\Rightarrow \mathrm{E}\left[M_a^b\right] \leqslant \frac{1}{b-a} \sup_n \mathrm{E}\left[(X_n-a)^-\right] \ . \end{aligned}$$

2.2 Uniformly integrable sequences

2.2.1 Def. A family $\{X_i\}_{i\in I}$ of r.v. is *uniformly integrable* if

$$\lim_{a\to\infty} \sup_{i\in I} \int_{\{|X_i|\geqslant a\}} |X_i|\, dP = 0 \ .$$

2.2.2 Remarks

1) If one has only one variable, then

$$\lim_{a\to\infty} \int_{\{|X|\geqslant a\}} |X|\, dP = 0 \tag{2.2.1}$$

is a necessary and sufficient condition for integrability. The necessity is obvious. It is sufficient because, taking a such that $\int_{\{|X|\geqslant a\}} |X|\, dP \leqslant 1$, then

$$\mathrm{E}[|X|] = \int_{\{|X|\geqslant a\}} |X|\, dP + \int_{\{|X|<a\}} |X|\, dP \leqslant 1 + a \ .$$

In the definition of u.i., we are saying that

$$\forall \varepsilon > 0,\ \exists a : \int_{\{|X_i|\geqslant a\}} |X_i|\, dP < \varepsilon \ ,$$

and that a can be taken to be the same for all variables in the family.

2) If $X \in L^1$, then a stronger condition holds:

$$\forall \varepsilon > 0,\ \exists \delta > 0 :\ \forall A \in \mathfrak{F},\ P(A) < \delta \Rightarrow \int_A |X|\, dP < \varepsilon \ .$$

(2.2.1) can be deduced from this using Chebyshev inequality: since $P\{|X| > a\} \leqslant \frac{1}{a}\mathrm{E}[|X|] < \infty$, we can choose a such that $P\{|X| > a\} < \delta$, and we apply this statement to $A = P\{|X| > a\} < \delta$.

2.2.3 Prop.

1. $\{X_i\}_{i\in I}$ u.i. $\Rightarrow \sup_{i\in I} \mathrm{E}[|X_i|] < \infty$ (bounded in L^1).
2. $\sup_{i\in I} \mathrm{E}[|X_i|^p] < \infty$ (for some $p > 1$) $\Rightarrow \{X_i\}_{i\in I}$ u.i.
3. $|X_i| \leqslant Y \in L^1 \Rightarrow \{X_i\}_{i\in I}$ u.i.

The converse statements are false.

Proof:

1. Same argument as in Remark 2.2.2 (1) above.
2. $$\int_{\{|X_i|\geqslant a\}} |X_i|\, dP \leqslant a^{1-p} \int_{\{|X_i|\geqslant a\}} |X_i|^p\, dP \leqslant a^{1-p} \sup_{i\in I} \mathrm{E}[|X_i|^p] \ .$$
   The first inequality comes from
   $$x \geqslant a > 0 \ \Rightarrow\ x^{p-1} \geqslant a^{p-1} \ \Rightarrow\ x \leqslant x^p a^{1-p} \ .$$
3. $$\int_{\{|X_i|\geqslant a\}} |X_i|\, dP \leqslant \int_{\{Y\geqslant a\}} Y\, dP \xrightarrow[a\to\infty]{} 0 \ ,$$
   by Remark 2.2.2.

2.2.4 Theorem (Extension of the Dominated Convergence Theorem)

$$\left.\begin{array}{r} X_n \xrightarrow[n\to\infty]{\Pr} X \\ \{X_n,\ n\geqslant 0\} \text{ u.i.} \end{array}\right\} \Leftrightarrow X_n \xrightarrow[n\to\infty]{L^1} X \ \Rightarrow\ \int \underline{\lim}\, X_n \leqslant \underline{\lim} \int X_n \leqslant \overline{\lim} \int X_n \leqslant \int \overline{\lim}\, X_n \ .$$

Additional remarks about the basic convergence Theorem 2.1.1:

- The basic convergence theorem looks surprising. Two conditions that look like very weak give a.s. convergence. Think of it as an stochastic analogue of
  $$\left.\begin{array}{r} \{x_n\}_n \text{ bounded} \\ \{x_n\}_n \text{ monotonous} \end{array}\right\} \Rightarrow \{x_n\}_n \text{ convergent}$$
- In fact, we know that for real sequences
  $$\left.\begin{array}{c} \{x_n\}_n \text{ bounded from above} \\ \{x_n\}_n \text{ increasing} \end{array}\right\} \Rightarrow \{x_n\}_n \text{ convergent}$$

so for submartingales we may expect that $\sup_n \mathrm{E}[X_n^+] < \infty$ must be enough. Indeed,

$$|X_n| = X_n^+ + X_n^- = 2X_n^+ + X_n^- - X_n^+ = 2X_n^+ - X_n \ , \ \Rightarrow$$
$$\mathrm{E}[|X_n|] = 2\,\mathrm{E}[X_n^+] - \mathrm{E}[X_n] \leqslant 2\,\mathrm{E}[X_n^+] - \mathrm{E}[X_0] \ .$$

A similar statement holds for supermartingales.

2.2.5 Example of a uniformly integrable martingale Recall Example 4:

$$\left.\begin{array}{c}\{\mathfrak{F}_n,\ n \geqslant 0\}\text{ filtration}\\ Y \in L^1\end{array}\right\} \ \Rightarrow\ \{(\mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_n}\right])\}\text{ is a martingale}$$

Let us prove that it is uniformly integrable:

Fix $\varepsilon > 0$.

Take $\delta > 0$ such that

$$\forall A \in \mathfrak{F},\ P(A) < \delta \Rightarrow \int_A |Y|\,dP < \varepsilon \ .$$

Such a δ exists because Y is u.i. (see Remark 2.2.2 (2)).

Take $a > 0$ such that

$$\frac{1}{a}\,\mathrm{E}[|Y|] < \delta \ .$$

Applying Chebyshev and Jensen inequalities,

$$P\Big\{\big|\,\mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_n}\right]\big| > a\Big\} \leqslant \frac{1}{a}\,\mathrm{E}\Big[\big|\,\mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_n}\right]\big|\Big] \leqslant \frac{1}{a}\,\mathrm{E}\Big[\,\mathrm{E}\left[{}^{|Y|}/_{\mathfrak{F}_n}\right]\Big] = \frac{1}{a}\,\mathrm{E}[|Y|] < \delta \ .$$

Now, applying Jensen inequality and the definition of conditional expectation,

$$\int_{\{|\,\mathrm{E}[{}^{Y}/_{\mathfrak{F}_n}]|>a\}} \big|\,\mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_n}\right]\big|\,dP \leqslant \int_{\{|\,\mathrm{E}[{}^{Y}/_{\mathfrak{F}_n}]|>a\}} \mathrm{E}\left[{}^{|Y|}/_{\mathfrak{F}_n}\right]dP = \int_{\{|\,\mathrm{E}[{}^{Y}/_{\mathfrak{F}_n}]|>a\}} |Y|\,dP < \varepsilon \ .$$

Finally, take $\sup_n$ and that's it.

2.2.6 Remark

It $Y \in L^1$, and $\{F_i\}_{i\in I}$ is an arbitrary family of sub-σ-fields of $\mathfrak{F}$, then $\Big\{\mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_i}\right],\ i \in I\Big\}$ is an u.i. family of r.v. (the proof is exactly the same).

2.2.7 Remark

Not every martingale is u.i.!

2.3 Convergence of uniformly integrable martingales

2.3.1 Theorem

$$\left.\begin{array}{c}\{\mathfrak{F}_n,\ n \geqslant 0\}\text{ filtration, }\mathfrak{F}_\infty = \sigma\{\bigcup\limits_n \mathfrak{F}_n\}\\ Y \in L^1\end{array}\right\} \ \Rightarrow\ X_n := \mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_n}\right] \xrightarrow[n\to\infty]{\text{a.s., } L^1} \mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_\infty}\right] \ .$$

(We know from 2.2.5 that X is a u.i. martingale).

Proof: We know from Proposition 2.2.3 (1) that X u.i. $\Rightarrow \sup_{n\in\mathbb{N}} \mathrm{E}[|X_n|] < \infty$.

We can apply Theorem 2.1.1 (the "basic theorem"):

$$X_n \xrightarrow[n\to\infty]{\text{a.s.}} \ell \ .$$

By Theorem 2.2.4, the convergence is also in L^1. We therefore have to show that $\ell = \mathrm{E}\left[{}^{Y}\!/_{\mathfrak{F}_\infty}\right]$: Take $A \in \mathfrak{F}_n$. Then:

$$\int_A Y\,dP = \int_A \mathrm{E}\left[{}^{Y}\!/_{\mathfrak{F}_n}\right] = \int_A X_n\,dP \xrightarrow[n\to\infty]{} \int_A \ell\,dP \ ,$$

because of L^1-convergence. That means

$$\int_A Y\,dP = \int_A \ell\,dP \ ,$$

for all $A \in \mathfrak{F}_n$, $\forall n$. By a monotone class argument, it is true also for $A \in \mathfrak{F}_\infty$. On the other hand, X_n is $\mathfrak{F}_\infty$-measurable $\Rightarrow$ ℓ is $\mathfrak{F}_\infty$-measurable.

Finally,

$$\left.\begin{array}{c}\displaystyle\int_A Y\,dP = \int_A \ell\,dP \\ \ell \text{ is } \mathfrak{F}_\infty\text{-measurable.}\end{array}\right\} \ \Rightarrow\ \ell = \mathrm{E}\left[{}^{Y}\!/_{\mathfrak{F}_\infty}\right] \ .$$

2.3.2 Theorem

$$\left.\begin{array}{c}\{\mathfrak{F}_n,\ n \geqslant 0\} \text{ decreasing sequence of } \sigma\text{-fields, } \mathfrak{F}_\infty = \bigcap_n \mathfrak{F}_n \\ Y \in L^1\end{array}\right\} \ \Rightarrow\ X_n := \mathrm{E}\left[{}^{Y}\!/_{\mathfrak{F}_n}\right] \xrightarrow[n\to\infty]{\text{a.s., } L^1} \mathrm{E}\left[{}^{Y}\!/_{\mathfrak{F}_\infty}\right] \ .$$

2.3.3 Theorem $\{(X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ u.i. (sub-, super-) martingale. Then:

1. $$X_n \xrightarrow[n\to\infty]{\text{a.s., } L^1} \ell$$

2. Denoting $\mathfrak{F}_\infty = \sigma\{\bigcup_n \mathfrak{F}_n\}$ and $X_\infty = \ell$, then

$$\{(X_n, \mathfrak{F}_n),\ n = 0, \dots, \infty\} \text{ is a (sub-, super-) martingale.}$$

Proof: The proof of 1. is identical to the corresponding part of Theorem 2.3.1. Let's prove 2. in the submartingale case:

We know from property 1 in 1.0.10 that $\mathrm{E}\left[{}^{X_{n+k}}\!/_{\mathfrak{F}_n}\right] \geqslant X_n$. That means, if $A \in \mathfrak{F}_n$,

$$\int_A X_{n+k} \geqslant \int_A X_n \ .$$

Letting $k \to \infty$, and by L^1-convergence,

$$\int_A \mathrm{E}\left[{}^{X_\infty}\!/_{\mathfrak{F}_n}\right] = \int_A X_\infty \geqslant \int_A X_n \ .$$

Since $A \in \mathfrak{F}_n$ is arbitrary, this implies that

$$\mathrm{E}\left[{}^{X_\infty}\!/_{\mathfrak{F}_n}\right] \geqslant X_n \ ,$$

which gives that $\{(X_n, \mathfrak{F}_n),\ n = 0, \ldots, \infty\}$ is a submartingale.

2.3.4 Def.

If $\mathfrak{F}_\infty \supset \bigcup_n \mathfrak{F}_n$ and $\{(X_n, \mathfrak{F}_n),\ n = 0, \ldots, \infty\}$ is a (sub-, super-) martingale, then the (sub-, super-) martingale is said to have a *last element*, or that *is closed*. □

A consequence of Theorem 2.3.3 in the martingale case:

$X = \{X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ martingale.

X u.i. $\Rightarrow X = \{X_n, \mathfrak{F}_n),\ n = 0, \ldots, X_\infty\}$ is a martingale (where $X_\infty = (\text{a.s.}, L^1) - \lim X_n)) \Rightarrow$ $\mathrm{E}\left[X_\infty / \mathfrak{F}_n\right] = X_n \Rightarrow X$ is u.i. (see Example 2.2.5).

So, we obtain the following

2.3.5 Corollary

$X = \{X_n, \mathfrak{F}_n),\ n \geqslant 0\}$ is a u.i. martingale iff $\exists Y \in L^1$ such that $X_n = \mathrm{E}\left[Y / \mathfrak{F}_n\right]$. And in that case,

$$X_n \xrightarrow[n\to\infty]{\text{a.s., } L^1} \mathrm{E}\left[Y / \mathfrak{F}_\infty\right] .$$

Moreover, if we require Y to be $\mathfrak{F}_\infty$-measurable, then it is unique (coincides with X_∞ (the limit) a.s.)

2.3.6 Remark

Closed $\not\Rightarrow$ u.i. for sub- or supermartingales. □

3. Applications

3.1 Kolmogorov 0-1 law

$\{X_n,\ n \geqslant 0\}$ independent r.v.

$\mathfrak{G}_n := \sigma\{X_{n+1}, X_{n+2}, \dots\}$, $\mathfrak{G} = \bigcap_n \mathfrak{G}_n$.

Then,

$$\forall A \in \mathfrak{G},\ P(A) = 0 \text{ or } 1\ .$$

Proof:

$\mathfrak{F}_n := \sigma\{X_1, \dots, X_n\}$, $\mathfrak{F}_\infty := \sigma\{\bigcup_n \mathfrak{F}_n\}$.

$Y = \mathbf{1}_A$.

$Y \in L^1$, obviously, so

$$Y = \mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_\infty}\right] = \lim_{n\to\infty} \mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_n}\right] \text{ a.s.}$$

But Y is $\mathfrak{G}_n$-measurable, and $\mathfrak{G}_n$ and $\mathfrak{F}_n$ are independent. Therefore,

$$\mathrm{E}\left[{}^{Y}/_{\mathfrak{F}_n}\right] = \mathrm{E}[Y] = P(A)$$

and we obtain

$$Y = P(A) \text{ a.s. } \Rightarrow\ Y = 0 \text{ or } 1 \text{ a.s.}$$

3.2 Strong law of large numbers

$$\left.\begin{array}{c}\{X_n,\ n \geqslant 1\} \text{ i.i.d.r.v. in } L^1, \text{ with expectation } m \in \mathbb{R} \\ S_n := X_1 + \cdots X_n\end{array}\right\} \Rightarrow\ \frac{S_n}{n} \xrightarrow[n\to\infty]{\text{a.s., } L^1} m\ .$$

(Note: Not only a.s. convergence, but also in L^1).

Proof:

$\mathfrak{H}_n := \sigma\{S_n, S_{n+1}, S_{n+2}, \dots\}$, $\mathfrak{H}_\infty = \bigcap_n \mathfrak{H}_n$.

$$\frac{S_n}{n} = \mathrm{E}\left[X_1/\mathfrak{H}_n\right] \xrightarrow[n\to\infty]{a.s.,\ L^1} \mathrm{E}\left[X_1/\mathfrak{H}_\infty\right]$$

The first equality is left as exercise, but it is intuitively obvious. The convergence comes from Theorem 2.3.2.

This gives the existence of a limit. We want to see that this limit is the constant m.

$$\lim_{n\to\infty}\frac{S_n}{n} = \lim_{n\to\infty}\left(\frac{X_1+\cdots+X_{k-1}}{n} + \frac{X_k+\cdots+X_n}{n}\right) = \lim_{n\to\infty}\frac{X_k+\cdots+X_n}{n}\ ,$$

which is measurable w.r.t. $\mathfrak{G}_k := \sigma\{X_k, X_{k+1}, \dots\}$, $\forall k$. Therefore,

$$\lim_{n\to\infty}\frac{S_n}{n} \text{ is } \bigcap_k \mathfrak{G}_k\text{-measurable.}$$

By Kolmogorov 0-1 law, this σ-field has only events of probability 0 or 1, and we conclude that the limit is a constant. Which constant? That's easy: From the L^1-convergence,

$$\mathrm{E}\left[\lim_n \frac{S_n}{n}\right] = \lim_n \mathrm{E}\left[\frac{S_n}{n}\right] = \lim_n \frac{1}{n}\,\mathrm{E}\left[X_1+\cdots+X_n\right] = \frac{nm}{n} = m\ .$$

3.3 Extinction of family names

Assume that family names are transmitted by men. We are interested in the evolution of the number of men with a given family name. Assume the time evolves in a discrete fashion (generations), so we are considering a process $\{X_n,\ n\geqslant 0\}$.

Assume:

- $X_0 = 1$.
- If $X_n = k$, then $X_{n+1} = Y_1+\cdots+Y_k$, where Y's are i.i.d.r.v. with some law $P\{Y_i = r\} = p_r$, $r\in\mathbb{N}$, and represent the number of male offspring of men $i = 1,\dots,k$.

This process is clearly a Markov chain, by construction.

Set $m := \mathrm{E}[Y_i]$. Assume $0 < m < \infty$. Then:

$$\left\{\frac{X_n}{m^n},\ n\geqslant 0\right\} \text{ is a martingale (w.r.t. its natural filtration).}$$

Indeed:

$$\begin{aligned}
&\mathrm{E}\left[\frac{X_{n+1}}{m^{n+1}}/X_0 = i_0,\dots,X_n = i_n\right] = \frac{1}{m^{n+1}}\,\mathrm{E}\left[X_{n+1}/X_n = i_n\right]\\
&= \frac{1}{m^{n+1}}\,\mathrm{E}\left[Y_1+\cdots+Y_{i_n}\right] = \frac{i_n\cdot m}{m^{n+1}} = \frac{i_n}{m^n}\ ,\\
&\Rightarrow \mathrm{E}\left[\frac{X_{n+1}}{m^{n+1}}/X_0,\dots,X_n\right] = \frac{X_n}{m^n}\ . \quad \square
\end{aligned}$$

Now we will consider different cases and subcases:

- *Case 1: $m < 1$*

$$\begin{aligned}
&\mathrm{E}\left[X_{n+1}/X_n = k\right] = k\cdot \mathrm{E}[Y_i] = k\cdot m\ ,\ \Rightarrow\ \mathrm{E}\left[X_{n+1}/X_n\right] = m\cdot X_n\ ,\\
&\Rightarrow\ \mathrm{E}[X_{n+1}] = m\cdot\mathrm{E}[X_n] = \cdots = m^n\ ,\\
&\Rightarrow\ \mathrm{E}\left[\sum_{n=0}^{\infty} X_n\right] = \sum_{n=0}^{\infty} m^n < \infty\ ,\\
&\Rightarrow\ X_n \xrightarrow[n\to\infty]{\text{a.s.}} 0\ .
\end{aligned}$$

But X_n are integer valued. So, $X_n(\omega) = 0$ for some n onwards. Extinction!

• *Case 2: $m > 1$*

Define

$$g(s) = \sum_{k=0}^{\infty} p_k s^k \ , \quad 0 \leqslant s \leqslant 1 \ .$$

Then:

$$\begin{aligned} g(0) &= p_0 \\ g(1) &= 1 \\ g'(0) &= p_1 < 1 \\ g'(1) &= m > 1 \end{aligned}$$

Let r be the root of $g(s) = s$ in $[0, 1)$.

- *Case 2.1: $r > 0$*

The process $\{r^{X_n},\ n \geqslant 0\}$ is a martingale (exercise; it's very similar to X_n/m^n).

It is a *positive* martingale. But any positive (super-)martingale converges almost surely, because $\sup_n \mathrm{E}[X_n^-] = 0 < \infty$, which is equivalent (for supermartingales) to $\sup_n \mathrm{E}[|X_n|] < \infty$.

This implies that the process in the exponent $\{X_n\}$ converges a.s., and, since it takes values on $\mathbb{N}$, we must have, either

$$\begin{aligned} &X_n(\omega) = K \text{ from some } n_0 \text{ onwards, for some } K \in \mathbb{N} \\ &\text{or} \\ &X_n(\omega) \to \infty \ . \end{aligned}$$

Let us prove that K can only be zero: Assume $K \geqslant 1$. Then, using the Markov property,

$$P\{X_n = K \ \text{ for all } n \geqslant n_0\} = P\{X_{n_0} = K\} \cdot \lim_{j\to\infty} P\big\{X_{n+1} = K / X_n = K\big\}^j = 0 \ ,$$

the last equality coming from the fact that the conditional probability is less than 1, because we are assuming that $p_0 > 0$.

Therefore, we conclude that

$$X_n \xrightarrow[n\to\infty]{\text{a.s.}} X_\infty \equiv 0 \text{ or } \infty \ .$$

Now, using that X_n is a martingale and the Dominated Convergence Theorem,

$$\mathrm{E}[r^{X_0}] = \mathrm{E}[r^{X_n}] \to \mathrm{E}[r^{X_\infty}] = P\{X_\infty = 0\} \ .$$

Therefore, the probability of extinction is exactly r. With probability $1 - r$ we have *explosion* of the family name.

- *Case 2.2: $r = 0$ ($\Leftrightarrow p_0 = 0$)*

$$p_0 = 0 \Rightarrow X_{n+1} \geqslant X_n \geqslant 1 \Rightarrow \ X_n \text{ increases to a limit } X_\infty \text{ which is never 0.}$$

For any $K \in \mathbb{N}$, $K \geqslant 1$,

$$\begin{aligned} &P\{X_n = K \ \text{ for all } n \geqslant n_0\} = P\{X_{n_0} = K\} \cdot \lim_{j\to\infty} P\big\{X_{n+1} = K / X_n = K\big\}^j \\ &= \begin{cases} 1, & \text{if } p_1 = 1 \ (\Rightarrow K = 1) \ . \\ 0, & \text{if } p_1 < 1 \text{ (reasoning as before)} \ . \end{cases} \end{aligned}$$

The first is the trivial case in which each father has exactly one son (no extinction, no explosion). In the second case, $P\{X_\infty = \infty\} = 1$, we have explosion.

• *Case 3: $m = 1$*

We know that $\left\{\dfrac{X_n}{m^n},\ n \geqslant 0\right\}$ is a positive martingale. Therefore it converges a.s. to a limit $X_\infty \in L^1$.

With the same analysis as in case 2, we find that X_n cannot converge to any constant $\geqslant 1$. Hence, $X_n \to 0$, a.s. Extinction.