Martingales

Samy Tindel

Purdue University

Probability Theory 1 - MA 538

Taken from
Probability: Theory and examples – Sections 5.2 $\rightarrow$ 5.4
by R. Durrett

Outline

1. Definitions and first properties
2. Strategies and stopped martingales
3. Convergence
4. Convergence in L^p
5. Optional stopping theorems

Outline

1. Definitions and first properties
2. Strategies and stopped martingales
3. Convergence
4. Convergence in L^p
5. Optional stopping theorems

Adaptation

Context: We are given

- A probability space $(\Omega, \mathcal{F}, \mathbf{P})$
- A filtration $\{\mathcal{F}_n;\ n \geq 0\}$
  $\hookrightarrow$ Sequence of σ-algebras such that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$.

Definition 1.

A sequence of random variables $\{X_n;\ n \geq 0\}$ is adapted if:

$$X_n \in \mathcal{F}_n.$$

Martingales, Supermartingales, Submartingales

Definition 2.

We consider a sequence of random variables $X = \{X_n;\ n \geq 0\}$ such that

1. $\{X_n;\ n \geq 0\}$ is adapted.
2. $X_n \in L^1(\Omega)$ for all $n \geq 0$.

Then

- X is a martingale if $X_n = \mathbf{E}[X_{n+1} | \mathcal{F}_n]$.
- X is a supermartingale if $X_n \geq \mathbf{E}[X_{n+1} | \mathcal{F}_n]$.
- X is a submartingale if $X_n \leq \mathbf{E}[X_{n+1} | \mathcal{F}_n]$.

Interpretation

Adaptation: The data X_n only depends on information until instant n.

Martingale: $n \mapsto X_n$ constant plus fluctuations.

Submartingale: $n \mapsto X_n$ increasing plus fluctuations.

Supermartingale: $n \mapsto X_n$ decreasing plus fluctuations.

Random walk

Definition: Let

- $\{Z_i;\ i \geq 1\}$ independent Rademacher r.v
  $\hookrightarrow \mathbf{P}(Z_i = -1) = \mathbf{P}(Z_i = 1) = 1/2$
- We set $X_0 = 0$, and for $n \geq 1$,

$$X_n = \sum_{i=1}^{n} Z_i.$$

X is called random walk in $\mathbb{Z}$.

Property: X is a martingale.

Conditional expectation in the past

Proposition 3.

Let X be a $\mathcal{F}_n$-martingale and $m \geq 0$.
For all $n \geq m$ we have

$$\mathbf{E}\left[X_n \middle| \mathcal{F}_m\right] = X_m.$$

Proof: Recursive procedure.

Important corollary: Let X be a $\mathcal{F}_n$-martingale and $m \geq 0$.
For all $n \geq m$ we have

$$\mathbf{E}\left[X_n\right] = \mathbf{E}\left[X_m\right] = \mathbf{E}\left[X_0\right]. \tag{1}$$

Composition with a convex function

Proposition 4.

Let

- X a $\mathcal{F}_n$-martingale.
- $\varphi : \mathbb{R} \to \mathbb{R}$ a convex function
  $\hookrightarrow$ such that $\varphi(X_n) \in L^1(\Omega)$ for all $n \geq 0$.
- $Y_n = \varphi(X_n)$

Then Y is a submartingale.

Proof: application of Jensen for conditional expectation.

Example: If X_n is a random walk, X_n^2 is a submartingale
$\hookrightarrow$ Fluctuations increase with time.

Outline

1. Definitions and first properties
2. Strategies and stopped martingales
3. Convergence
4. Convergence in L^p
5. Optional stopping theorems

Martingale transformation

Definition 5.

Let $(\mathcal{F}_n)_{n\geq 0}$ be a filtration and X, H $\mathcal{F}_n$-adapted-processes.

- We say that H is predictable if $H_n \in \mathcal{F}_{n-1}$.
- The transform of X by H is

$$[H \cdot X]_n = \sum_{j=1}^{n} H_j \, \Delta X_j, \quad \text{where} \quad \Delta X_j = X_j - X_{j-1}$$

Interpretation:

1. $H \equiv$ game strategy
   $\hookrightarrow$ Today's decision depends on the information until yesterday
2. $H \cdot X \equiv$ value if strategy H is used

D'Alembert

Some facts about d'Alembert:

- Abandoned after birth
- Mathematician
- Contribution in fluid dynamics
- Philosopher
- Participation in 1st Encyclopedia

D'Alembert's Martingale

Example: Let $X_n = \sum_{i=1}^{n} \xi_i$ be a random walk.
We interpret ξ_i as a gain ou a loss at ith iteration of the game.
The filtration is $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$

Strategy: We define H in the following way:

- $H_1 = 1$, thus $H_1 \in \mathcal{F}_0$.
- $H_n = 2\, H_{n-1}\, \mathbf{1}_{(\xi_{n-1}=-1)}$

Let $N = \inf\{j \geq 1;\ \xi_j = 1\}$. Then

$$[H \cdot X]_N = \sum_{j=1}^{N} H_j\, \Delta X_j = \sum_{j=1}^{N} H_j\, \xi_j = -\sum_{j=1}^{N-1} 2^{j-1} + 2^{N-1} = 1$$

We get an almost sure gain!

Strategies and martingales

Theorem 6.

Let

- X a martingale.
- H a predictable process such that $H_j \, \Delta X_j \in L^1$ for all j.

Then $H \cdot X$ is a martingale.

Interpretation: One cannot win in a fair game context
$\hookrightarrow$ Compare with d'Alembert's martingale

Proof

Main ingredients: We write

$$[H \cdot X]_{n+1} = [H \cdot X]_n + H_{n+1} \left(X_{n+1} - X_n \right).$$

Then we use the fact that

1. H is predictable
2. X is a martingale

Stopping time

Definition 7.

Let

- $T : \Omega \to \bar{N}$ random time
- $\mathcal{F}_n$ a filtration

We say that T is a stopping time for $\mathcal{F}_n$ if

- For all $n \in \mathbb{N}$, the set $\{\omega;\ T(\omega) = n\}$ is $\mathcal{F}_n$-measurable.

Note: basic examples are hitting times.

Stopped martingales

Theorem 8.

Let

- X martingale
- N stopping time

We set $Y_n = X_{n \wedge N}$. Then Y is a martingale.

Proof

Decomposition of Y: We have

$$Y_j - Y_{j-1} = (X_j - X_{j-1}) \, \mathbf{1}_{(j-1<N)}.$$

Expression as transformed martingale: Set $H_j = \mathbf{1}_{(j-1<N)}$. Then

$$\begin{aligned} Y_n &= Y_0 + \sum_{j=1}^{n} (Y_j - Y_{j-1}) \\ &= Y_0 + \sum_{j=1}^{n} (X_j - X_{j-1}) \, \mathbf{1}_{(j-1<N)} \\ &= Y_0 + \sum_{j=1}^{n} H_j \, \Delta X_j \end{aligned}$$

In addition H is predictable. Thus Y is a martingale.

Outline

1. Definitions and first properties
2. Strategies and stopped martingales
3. Convergence
4. Convergence in L^p
5. Optional stopping theorems

Convergence in L^2

Theorem 9.

Let X such that

- $\{X_n;\, n \geq 1\}$ is a martingale.
- For all n we have $X_n \in L^2(\Omega)$ and

$$\sup\left\{\mathbf{E}[X_n^2];\, n \geq 0\right\} \equiv M < \infty. \qquad (2)$$

Then

1. $L^2 - \lim_{n\to\infty} X_n = X_\infty$.
2. For all $n \geq 0$, we have $X_n = \mathbf{E}[X_\infty | \mathcal{F}_n]$.

Proof

Step 1: We set $a_n = \mathbf{E}[X_n^2]$. We will show that if $n \geq m$, then

$$\mathbf{E}\left[(X_n - X_m)^2\right] = a_n - a_m.$$

Indeed,

$$\mathbf{E}[X_m X_n] = \mathbf{E}\left\{X_m \, \mathbf{E}[X_n | \mathcal{F}_m]\right\} = \mathbf{E}\left[X_m^2\right].$$

Therefore

$$\begin{aligned} \mathbf{E}\left[(X_n - X_m)^2\right] &= \mathbf{E}\left[X_n^2\right] + \mathbf{E}\left[X_m^2\right] - 2\,\mathbf{E}[X_m X_n] \\ &= \mathbf{E}\left[X_n^2\right] - \mathbf{E}\left[X_m^2\right] \\ &= a_n - a_m. \end{aligned}$$

Proof (2)

Step 2: Convergence in L^2.

- $a_{n+1} - a_n = \mathbf{E}[(X_{n+1} - X_n)^2] \Longrightarrow n \mapsto a_n$ increasing.
- Inequality (2) $\Longrightarrow (a_n)_{n\geq 0}$ bounded $\Longrightarrow (a_n)_{n\geq 0}$ convergent.
- $\mathbf{E}\left[(X_n - X_m)^2\right] = a_n - a_m \Longrightarrow (X_n)_{n\geq 0}$ Cauchy in $L^2(\Omega)$

Conclusion: $(X_n)_{n\geq 0}$ converges in $L^2(\Omega)$ towards X_∞.

Proof (3)

Step 3: We have $X_n = \mathbf{E}[X_\infty | \mathcal{F}_n]$.

Set

$$V = |\mathbf{E}[X_\infty | \mathcal{F}_n] - X_n| .$$

We are reduced to show that $\mathbf{E}[V] = 0$.

Computation: For $n, k \geq 0$,

$$\begin{aligned} V &= |\mathbf{E}[X_\infty | \mathcal{F}_n] - \mathbf{E}[X_{n+k} | \mathcal{F}_n]| \\ &= |\mathbf{E}[X_\infty - X_{n+k} | \mathcal{F}_n]| \leq \mathbf{E}\left[|X_\infty - X_{n+k}| \,|\, \mathcal{F}_n\right] \end{aligned}$$

Hence

$$\mathbf{E}[V] \leq \mathbf{E}\left[|X_\infty - X_{n+k}|\right] \leq \mathbf{E}^{1/2}\left[(X_\infty - X_{n+k})^2\right]$$

We get $\mathbf{E}[V] = 0$ whenever $k \to \infty$ above.

Almost sure convergence

Theorem 10.

Let X satisfying

- $\{X_n;\ n \geq 0\}$ is a martingale or a submartingale.
- We have

$$\sup\left\{\mathbf{E}[X_n^+];\ n \geq 0\right\} \equiv M < \infty. \qquad (3)$$

Then

1. $\mathrm{a.s} - \lim_{n\to\infty} X_n = X_\infty$.
2. We have $\mathbf{E}[|X_\infty|] < \infty$.

Particular cases

Particular case 1:
$(X_n)_{n\geq 0}$ positive martingale $\Longrightarrow \mathrm{a.s} - \lim_{n\to\infty} X_n = X_\infty$.

Particular case 2:
$\sup\{\mathbf{E}[X_n^2];\, n \geq 0\} \equiv M < \infty \Longrightarrow \mathrm{a.s} - \lim_{n\to\infty} X_n = X_\infty$.
$\hookrightarrow$ We have both a.s and L^2 convergence.

Convergence counterexample

Example 11.

Let

- $\{\xi_n;\ n \geq 1\}$ i.i.d Rademacher sequence
- $\{S_n;\ n \geq 0\}$ defined by
  - $S_0 = 1$
  - $S_n = S_{n-1} + \xi_n$ for $n \geq 1$
- $N = \inf\{n \geq 1;\ S_n = 0\}$
- $X_n = S_{n \wedge N}$

Then the following holds true:

1. X_n converges almost surely to 0
2. X_n does not converge in $L^1(\Omega)$

Proof

Almost sure convergence: We have

- Theorem 8 $\Longrightarrow X$ is a martingale
- $X_n \geq 0$

Thus X_n converges almost surely to $X_\infty \geq 0$

Identification of the limit: Assume $\mathbf{P}(\Omega_k) > 0$ with $k > 0$ and

$$\Omega_k = \left\{\omega;\ \lim_{n\to\infty} X_n(\omega) = k\right\}.$$

For $\omega \in \Omega_k$, we have the following:

- Set $n_0(\omega) = \inf\{n \geq 0;\ X_m(\omega) = k \text{ for } m \geq n\}$.
- For $m \geq n_0$ we have $X_{m+1} = X_m \pm 1$

This yields a contradiction. Hence $\mathbf{P}(\Omega_k) = 0$ and $X_\infty = 0$ a.s

Proof (2)

Convergence in $L^1(\Omega)$: According to (1) we have

$$\mathbf{E}[X_n] = \mathbf{E}[X_0] = 1$$

Thus we cannot have $L^1(\Omega) - \lim_{n\to\infty} X_n = 0$

Doob's decomposition

Theorem 12.

Let

- X submartingale

Then X can be decomposed uniquely as:

$$X_n = M_n + A_n,$$

where

- M is a martingale
- A is a predictable increasing process such that $A_0 = 0$

Proof

Expression for M and A: We wish

- $\mathbf{E}[M_n|\mathcal{F}_{n-1}] = M_{n-1}$
- $A_n \in \mathcal{F}_{n-1}$

Therefore if $X_n = M_n + A_n$ we have

$$\begin{aligned}\mathbf{E}\left[X_n|\,\mathcal{F}_{n-1}\right] &= \mathbf{E}\left[M_n|\,\mathcal{F}_{n-1}\right] + \mathbf{E}\left[A_n|\,\mathcal{F}_{n-1}\right] \\ &= M_{n-1} + A_n \\ &= X_{n-1} + A_n - A_{n-1}\end{aligned}$$

We thus take:

$$A_n - A_{n-1} = \mathbf{E}\left[X_n|\,\mathcal{F}_{n-1}\right] - X_{n-1}, \quad \text{and} \quad M_n = X_n - A_n$$

Proof (2)

Expression for A and M: recall that

$$A_n - A_{n-1} = \mathbf{E}\left[X_n \middle| \mathcal{F}_{n-1}\right] - X_{n-1}, \quad \text{and} \quad M_n = X_n - A_n$$

Proof of Doob's properties: We have

1. A_n is increasing since $\mathbf{E}\left[X_n \middle| \mathcal{F}_{n-1}\right] \geq X_{n-1}$
2. $A_n \in \mathcal{F}_{n-1}$ by induction
3. The martingale property for M is obtained as follows:

$$\begin{aligned} \mathbf{E}\left[M_n \middle| \mathcal{F}_{n-1}\right] &= \mathbf{E}\left[X_n - A_n \middle| \mathcal{F}_{n-1}\right] \\ &= \mathbf{E}\left[X_n \middle| \mathcal{F}_{n-1}\right] - A_n \\ &= A_n - A_{n-1} + X_n - A_n \\ &= M_{n-1} \end{aligned}$$

Outline

1 Definitions and first properties

2 Strategies and stopped martingales

3 Convergence

4 Convergence in L^p

5 Optional stopping theorems

Doob's inequality

Theorem 13.

We consider

- A submartingale X
- $\bar{X}_n \equiv \max_{m \le n} X_m^+$
- A real number $\lambda > 0$
- The set $A = \{\bar{X}_n \ge \lambda\}$

Then we have

$$\lambda \, \mathbf{P}(A) \le \mathbf{E}\left[X_n \mathbf{1}_A\right] \le \mathbf{E}\left[X_n^+\right]$$

Otherwise stated:

$$\mathbf{P}\left(\bar{X}_n \ge \lambda\right) \le \frac{\mathbf{E}\left[X_n^+\right]}{\lambda}$$

L^p maximum inequality

Theorem 14.

We consider

- A submartingale X
- $\bar{X}_n \equiv \max_{m \le n} X_m^+$
- $p \in (1, \infty)$

Then we have

$$\mathbf{E}\left[\bar{X}_n^p\right] \le c_p \, \mathbf{E}\left[\left(X_n^+\right)^p\right], \quad \text{with} \quad c_p = \left(\frac{p}{p-1}\right)^p$$

L^p bound for $|Y|$

Theorem 15.

We consider

- A martingale Y
- $Y_n^* \equiv \max_{m \le n} |Y_m|$
- $p \in (1, \infty)$

Then we have

$$\mathbf{E}\left[|Y_n^*|^p\right] \le c_p \, \mathbf{E}\left[|Y_n|^p\right], \quad \text{with} \quad c_p = \left(\frac{p}{p-1}\right)^p$$

Counterexample in $L^1(\Omega)$

Example 16.

As in Example 11, set

- $\{\xi_n;\ n \geq 1\}$ i.i.d Rademacher sequence
- $\{S_n;\ n \geq 0\}$ defined by
  - $S_0 = 1$
  - $S_n = S_{n-1} + \xi_n$ for $n \geq 1$
- $N = \inf\{n \geq 1;\ S_n = 0\}$
- $X_n = S_{n \wedge N}$

Then Theorem 14 is not satisfied for $p = 1$ and X:

1. $\lim_{n\to\infty} \mathbf{E}[X_n] = 1$
2. $\mathbf{E}[\bar{X}_\infty] = \infty$

Proof

Item 1: We have already seen in Example 11 that

$$\mathbf{E}[X_n] = 1, \quad \text{for all} \quad n \geq 0.$$

Hence we trivially have

$$\lim_{n\to\infty} \mathbf{E}[X_n] = 1.$$

Proof (2)

Hitting times: For $x \in \mathbb{Z}$ set

$$T_x = \inf\{n \geq 0;\ S_n = x\}.$$

Then for $a < 1 < b$ we have (see Section 5):

$$\mathbf{P}\left(T_b < T_a\right) = \frac{1-a}{b-a}. \qquad (4)$$

Item 2: Thanks to (4) we have, for all $M > 1$

$$\mathbf{P}\left(\bar{X}_\infty \geq M\right) = \mathbf{P}\left(T_M < T_0\right) = \frac{1}{M}.$$

Therefore

$$\mathbf{E}\left[\bar{X}_\infty\right] = \sum_{M=1}^{\infty} \frac{1}{M} = \infty.$$

Convergence in L^p

Theorem 17.

Let X and p such that

- $\{X_n;\, n \geq 1\}$ is a martingale.
- $p > 1$.
- For all n we have $X_n \in L^p(\Omega)$ and

$$\sup \{\mathbf{E}[|X_n|^p];\, n \geq 0\} \equiv M < \infty.$$

Then

1. $L^p - \lim_{n\to\infty} X_n = X_\infty$.
2. $\text{a.s} - \lim_{n\to\infty} X_n = X_\infty$.
3. For all $n \geq 0$, we have $X_n = \mathbf{E}[X_\infty | \mathcal{F}_n]$.

Outline

1 Definitions and first properties

2 Strategies and stopped martingales

3 Convergence

4 Convergence in L^p

5 Optional stopping theorems

Simple optional stopping theorem

Theorem 18.

Let

- $\{X_n;\, n \geq 0\}$ a martingale.
- T stopping time.

We assume that **ONE** of those two assumptions is satisfied:

1. T is a.s bounded by a constant M_1.
2. The sequence of random variables $\{X_{n\wedge T};\, n \geq 0\}$ is a.s bounded by a constant M_2.

Then

$$\mathbf{E}[X_T] = \mathbf{E}[X_0].$$

Proof

Under Hypothesis 1: Let $\kappa \in \mathbb{N}$ such that $T \leq \kappa$ a.s.

Computation: We use the fact that $\{X_{n\wedge T};\, n \leq \kappa\}$ is a martingale.

$$\mathbf{E}[X_0] = \mathbf{E}[X_{0\wedge T}] = \mathbf{E}[X_{n\wedge T}] = \mathbf{E}[X_{\kappa\wedge T}] = \mathbf{E}[X_T]$$

Under Hypothesis 2: We set $Y_n = X_{n\wedge T}$. Then

- $(Y_n)_{n\geq 0}$ bounded martingale in L^2
$\Longrightarrow Y_n \to Y_\infty$ in L^2 and a.s. Hence $\mathbf{E}[Y_\infty] = \mathbf{E}[Y_0]$.
- We have $Y_\infty = X_T$ and $Y_0 = X_0$. Therefore

$$\mathbf{E}[Y_\infty] = \mathbf{E}[Y_0] \quad \Longrightarrow \quad \mathbf{E}[X_T] = \mathbf{E}[X_0].$$

Case of a submartingale

Proposition 19.

Let

- $\{X_n;\, n \geq 0\}$ a submartingale.
- T stopping time.

We assume that:

- T is a.s bounded by a constant M.

Then

$$\mathbf{E}[X_0] \leq \mathbf{E}[X_T] \leq \mathbf{E}[X_M].$$

$L^p(\Omega)$ bound for stopped martingales

Proposition 20.

Let X and p such that

- $\{X_n;\, n \geq 1\}$ is a submartingale.
- $p > 1$.
- For all n we have $X_n \in L^p(\Omega)$ and

$$\sup \{\mathbf{E}[|X_n|^p];\, n \geq 0\} \equiv M < \infty.$$

For a stopping time N we set

$$Y_n = X_{n \wedge N}.$$

Then

$$\sup \{\mathbf{E}[|Y_n|^p];\, n \geq 0\} \leq M.$$

Proof

Definition of a submartingale: If X_n is a submartingale then

$$|X_n|^p \quad \text{is a submartingale}$$

Application of Proposition 19:
Since $N \wedge n$ is a stopping time bounded by n, we have

$$\mathbf{E}\left[|X_{N\wedge n}|^p\right] \leq \mathbf{E}\left[|X_n|^p\right],$$

and hence

$$\sup_{n\geq 0} \mathbf{E}\left[|X_{N\wedge n}|^p\right] \leq \sup_{n\geq 0} \mathbf{E}\left[|X_n|^p\right] = M$$

Optional stopping in $L^p(\Omega)$

Theorem 21.

Let X and $p > 1$ such that

- $\{X_n;\ n \geq 1\}$ is a submartingale.
- For all n we have $X_n \in L^p(\Omega)$ and

$$\sup\left\{\mathbf{E}[|X_n|^p];\ n \geq 0\right\} \equiv M < \infty.$$

- $X_\infty \equiv \text{a.s} - \lim_{n\to\infty} X_n$

Then for any stopping time N we have

$$\mathbf{E}\left[X_0\right] \leq \mathbf{E}\left[X_N\right] \leq \mathbf{E}\left[X_\infty\right]$$

Proof

Application of Proposition 19: for $n \geq 1$ we have

$$\mathbf{E}[X_0] \leq \mathbf{E}[X_{N \wedge n}] \leq \mathbf{E}[X_n] \tag{5}$$

Application of Proposition 20:
$n \mapsto X_{N \wedge n}$ and $n \mapsto X_n$ are bounded submartingales in $L^p(\Omega)$.
Thus:

$$X_{n \wedge N} \xrightarrow{\text{a.s, } L^p} X_N, \quad \text{and} \quad X_n \xrightarrow{\text{a.s, } L^p} X_\infty.$$

Therefore taking limits in (5) we get:

$$\mathbf{E}[X_0] \leq \mathbf{E}[X_N] \leq \mathbf{E}[X_\infty]$$

Optional sampling: the general form

Theorem 22.

Let X, $p > 1$ and two stopping times M, N such that

- $\{X_n;\, n \geq 1\}$ is a submartingale.
- For all n we have $X_{n\wedge N} \in L^p(\Omega)$ and

$$\sup\{\mathbf{E}[|X_{n\wedge N}|^p];\, n \geq 0\} \equiv A < \infty.$$

- $M \leq N$ almost surely

Then we have

$$\mathbf{E}[X_M] \leq \mathbf{E}[X_N], \quad \text{and} \quad X_M \leq \mathbf{E}[X_N|\,\mathcal{F}_M].$$

Proof

Proof of $\mathbf{E}[X_M] \leq \mathbf{E}[X_N]$: Set $Y_n = X_{n \wedge N}$. Then

- Y is a submartingale satisfying the assumptions of Theorem 21
- $Y_\infty = X_N$

Invoking Theorem 21 we thus get

$$\mathbf{E}[Y_M] \leq \mathbf{E}[Y_\infty] \quad \Longleftrightarrow \quad \mathbf{E}[X_M] \leq \mathbf{E}[X_N].$$

Proof (2)

Definition of a stopping time: For $A \in \mathcal{F}_M$ we set

$$T = M\,\mathbf{1}_A + N\,\mathbf{1}_{A^c}.$$

Then T is a stopping time. Indeed:

$$\begin{aligned}
\{T \le n\} &= (\{M \le n\} \cap A) \bigcup (\{N \le n\} \cap A^c) \\
&= (\{M \le n\} \cap A) \bigcup (\{N \le n\} \cap \{M \le n\} \cap A^c),
\end{aligned}$$

and hence:

$$\{T \le n\} \in \mathcal{F}_n$$

Proof (3)

Inequality involving A:

For A, T as before, applying $\mathbf{E}[X_T] \le \mathbf{E}[X_N]$ we get

$$\begin{aligned} & \mathbf{E}\left[X_T\right] \le \mathbf{E}\left[X_N\right] \\ \Longleftrightarrow\ & \mathbf{E}\left[X_M\, \mathbf{1}_A\right] + \mathbf{E}\left[X_N\, \mathbf{1}_{A^c}\right] \le \mathbf{E}\left[X_N\, \mathbf{1}_A\right] + \mathbf{E}\left[X_N\, \mathbf{1}_{A^c}\right] \\ \Longleftrightarrow\ & \mathbf{E}\left[X_M\, \mathbf{1}_A\right] \le \mathbf{E}\left[X_N\, \mathbf{1}_A\right] \end{aligned}$$

Therefore, by definition of the conditional expectation we get:

$$\mathbf{E}\left[X_M\, \mathbf{1}_A\right] \le \mathbf{E}\left\{\mathbf{E}\left[X_N \middle| \mathcal{F}_M\right] \mathbf{1}_A\right\} \tag{6}$$

Proof (4)

Conclusion: For $k \geq 1$ we set

$$A_k = \left\{ X_M - \mathbf{E}\left[X_N \middle| \mathcal{F}_M \right] \geq \frac{1}{k} \right\}.$$

Then $A_k \in \mathcal{F}_M$ and according to (6) we have

$$\mathbf{P}(A_k) = 0.$$

Hence:

$$\mathbf{P}\left(X_M - \mathbf{E}\left[X_N \middle| \mathcal{F}_M \right] > 0 \right) = \mathbf{P}\left(\cup_{k \geq 1} A_k \right) = 0$$

and thus:

$$X_M \leq \mathbf{E}\left[X_N \middle| \mathcal{F}_M \right], \quad \text{almost surely.}$$