6. Review of Probability

S. Ortiz-Latorre

STK-MAT 3700/4700 An Introduction to Mathematical Finance
October 21, 2021

Department of Mathematics
University of Oslo

Outline

Information and Measurability

Conditional Expectation

Information and Measurability

Information and measurability

- Our standing assumption is that $\#\Omega = K < \infty$.

Definition 1

Outcomes of an experiment $\omega_1, ..., \omega_K$ are called **elementary events** or **sample points** and the finite set $\Omega = \{\omega_1, ..., \omega_K\}$ is called the **space of of elementary events** or the **sample space**.

Definition 2

Events are all subsets $A \subseteq \Omega$ for which, under the conditions of the experiment, one can conclude that either "the outcome $\omega \in A$" or "the outcome $\omega \notin A$".

Information and measurability

Example 3

- The random experiment consists in tossing a coin three times.
- Then, $\#\Omega = 8$ and

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

- Event = " 2 heads in all " $= \{HHT, HTH, THH\} \subset \Omega$.

Definition 4

A collection $\mathcal{F}$ of subsets of Ω is called an **algebra** on Ω if

1. $\Omega \in \mathcal{F}$.
2. $A \in \mathcal{F} \Rightarrow A^c := \Omega \setminus A \in \mathcal{F}$.
3. $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$.

Remark 5

- Note that $\emptyset = \Omega^c \in \mathcal{F}$ and

$$A, B \in \mathcal{F} \Rightarrow A \cap B = (A^c \cup B^c)^c \in \mathcal{F}.$$

  Hence, an algebra $\mathcal{F}$ is a family of subsets of Ω which is closed under complementation and finitely many set operations (intersection and union).
- If $\#\Omega = \infty$, we need the closedness property to hold for infinitely many set operations.
- In this case, we say that a collection $\mathcal{F}$ of subsets of Ω is a σ**-algebra** on Ω if 1., 2. and

  $3'$. $\{A_n\}_{n\geq 1} \subseteq \mathcal{F} \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{F}$.
- For Ω with $\#\Omega < \infty$ both concepts coincide.

Example 6

Consider the following examples

1. $\mathcal{F}_1 = \{\emptyset, \Omega\}$ trivial algebra. (contains no information)
2. $\mathcal{F}_2 = \mathcal{P}(\Omega)$ collection of all subsets of Ω. (contains all the information)
3. $\mathcal{F}_3 = \{\emptyset, \Omega, A, A^c\}$ algebra generated by the event A. (contains the minimal information needed to decide if A has occurred or not)

Definition 7

Let S be a class of subsets of Ω. Then $\mathfrak{a}(S)$, the **algebra generated by** S, is the smallest algebra on Ω containing S. That is,

1. $S \subseteq \mathfrak{a}(S)$,
2. If $S \subseteq \mathcal{F}$, where $\mathcal{F}$ is an algebra, then $S \subseteq \mathfrak{a}(S) \subseteq \mathcal{F}$.

Note that

- If $S_1 \subseteq S_2$ then $\mathfrak{a}(S_1) \subseteq \mathfrak{a}(S_2)$.
- The intersection of an arbitrary number of algebras is an algebra.
- $\mathfrak{a}(S)$ is the intersection of all the algebras on Ω containing S.

Example 8

Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$.

1. $S_1 = \{\{\omega_1\}\}$, then

$$\mathfrak{a}(S_1) = \{\Omega, \emptyset, \{\omega_1\}, \{\omega_2, \omega_3, \omega_4\}\}.$$

2. $S_2 = \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4\}\}$, then

$$\mathfrak{a}(S_2) = \{\Omega, \emptyset, \{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4\}, \{\omega_2, \omega_3, \omega_4\}, \{\omega_1, \omega_4\}, \{\omega_1, \omega_2, \omega_3\}\}.$$

3. $S_3 = \{\{\omega_1\}, \{\omega_1, \omega_4\}\}$, then

$$\mathfrak{a}(S_3) = \{\Omega, \emptyset, \{\omega_1\}, \{\omega_1, \omega_4\}, \{\omega_2, \omega_3, \omega_4\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_3\}, \{\omega_4\}\}.$$

Example 8

- Since $S_1 \subseteq S_2$, we have that $\mathfrak{a}(S_1) \subseteq \mathfrak{a}(S_2)$.
- The algebra $\mathfrak{a}(S_2)$ contains the events in $\mathfrak{a}(S_1)$ and more.
- Hence, $\mathfrak{a}(S_2)$ is more informative than $\mathfrak{a}(S_1)$.
- Note that, $S_2 \not\subseteq S_3$ and $S_3 \not\subseteq S_2$, but $\mathfrak{a}(S_2) = \mathfrak{a}(S_3)$ and, therefore, $\mathfrak{a}(S_2)$ and $\mathfrak{a}(S_3)$ contain the same information.

An interesting class of subsets of Ω are those which form a partition of Ω.

Definition 9

A class of subsets $\pi = \{A_1, \ldots, A_m\}$ of Ω is a **partition** of Ω if

1. $A_i \cap A_j = \emptyset, \qquad i \neq j,$
2. $\cup_{i=1}^{m} A_i = \Omega.$

Definition 10

Given two partitions π_1, π_2 of Ω, we say that π_2 **is finer than (or refines)** π_1, if for any $A \in \pi_2$ there exists $B \in \pi_1$ such that $A \subseteq B$ and we will denote it by $\pi_1 \subseteq \pi_2$.

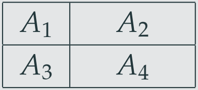
Definition 11

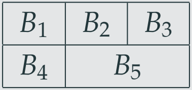
Given two partitions π_1, π_2 of Ω, we may define its **intersection** $\pi_1 \cap \pi_2$ to be the following partition

$$\pi_1 \cap \pi_2 = \{A \cap B : A \in \pi_1 \text{ and } B \in \pi_2\}.$$

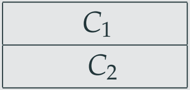
Note that, in general, neither $\pi_1 \subseteq \pi_2$ nor $\pi_2 \subseteq \pi_1$, but $\pi_1 \subseteq \pi_1 \cap \pi_2$ and $\pi_2 \subseteq \pi_1 \cap \pi_2$.

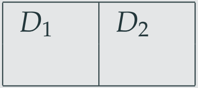
Example 12

- 

π_1
$\subseteq$


π_2

- 

π_3
neither $\subseteq$ nor $\supseteq$


π_4

But $\pi_3 \cap \pi_4 = \pi_1$ and $\pi_3 \subseteq \pi_1$, $\pi_4 \subseteq \pi_1$.

Remark 13

Why are partitions interesting?

- For any algebra $\mathcal{F}$ on Ω, there exists a partition π such that $\mathcal{F} = \mathfrak{a}(\pi)$ (**bijection**).
- The elements of $\mathfrak{a}(\pi)$ are all possible unions of the elements in π. (**easy structure**)
- Let $X : \Omega \to \{x_1, \ldots, x_M\}$, where $M \leq K = \#\Omega$, represent a measurement in a random experiment. Then, the following class of subsets of Ω is a partition

$$\pi_X = \left\{X^{-1}(x_i) = \{\omega \in \Omega : X(\omega) = x_i\}, i = 1, \ldots, M\right\}.$$

(**easy to interpret**)

Definition 14

Let $\mathcal{F}$ be an algebra on Ω. We say that function $X : \Omega \to \{x_1, \ldots, x_M\}$ is $\mathcal{F}$**-measurable** (measurable with respect to $\mathcal{F}$) if

$$X^{-1}(x_i) = \{\omega \in \Omega : X(\omega) = x_i\} \in \mathcal{F}, \quad i = 1, ..., M.$$

X is a random variables if and only if X is $\mathcal{P}(\Omega)$-measurable.

Definition 15

The **algebra generated by a finite number of r.v.** $X_1, X_2, \ldots, X_n$, denoted by $\mathfrak{a}(X_1, X_2, \ldots, X_n)$, is defined as $\mathfrak{a}\left(\bigcap_{i=1}^{n} \pi_{X_i}\right)$.

Remark 16

- $\mathfrak{a}(X) = \mathfrak{a}(\pi_X)$ is the smallest algebra $\mathcal{F}$ such that X is $\mathcal{F}$-measurable.
- Let $\mathcal{F} = \mathfrak{a}(\pi)$ where π is a partition of Ω. Then, X is $\mathcal{F}$-measurable if and only if X is constant on each element of the partition π.
- Usually, $\mathcal{P}(\Omega)$ is strictly finer than $\mathfrak{a}(X)$, that is, by observing X we cannot get all the information available in the sample space Ω.
- $\mathfrak{a}(X) = \mathcal{P}(\Omega)$ if and only if X takes $K = \#\Omega$ different values.

Information and measurability

Example 17

- Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$.
- Consider the random variables

$$X(\omega) = \begin{cases} 2 & \text{if} \quad \omega = \omega_1, \omega_2 \\ 4 & \text{if} \quad \omega = \omega_3, \omega_4 \end{cases}$$

$$Y(\omega) = \begin{cases} 1 & \text{if} \quad \omega = \omega_1 \\ 2 & \text{if} \quad \omega = \omega_2 \\ 3 & \text{if} \quad \omega = \omega_3 \\ 4 & \text{if} \quad \omega = \omega_4 \end{cases} .$$

- Then,

$$\pi_X = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\},$$

$$\mathfrak{a}(X) = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\},$$

Example 17

$$\pi_Y = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\},$$
$$\mathfrak{a}(Y) = \mathfrak{a}(\pi_Y) = \mathcal{P}(\Omega).$$

- Let Z be the “random variable” $Z \equiv 1$.
- Then, $\pi_Z = \{\Omega\}$ and $\mathfrak{a}(Z) = \mathfrak{a}(\pi_Z) = \{\emptyset, \Omega\}$.
- Note that Z (in fact any constant random variable) is measurable with respect to any algebra on Ω.

Information and measurability

Definition 18

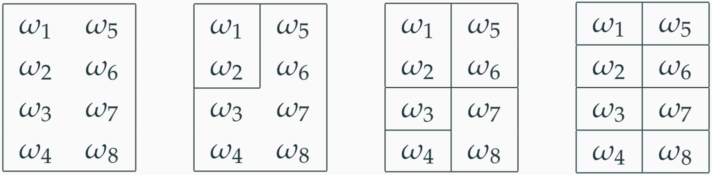
A **filtration** $\mathbb{F} = \{\mathcal{F}_t\}_{t=0,\dots,T}$ on Ω is a sequence of algebras on Ω such that $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$, $t = 0, \dots, T$.

- We will always assume that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and usually $\mathcal{F}_T = \mathcal{P}(\Omega)$.
- A filtration models the evolution of the information at our disposal through time.
- At time $t = 0$ we have no information and at time T, if $\mathcal{F}_T = \mathcal{P}(\Omega)$, we have full information.

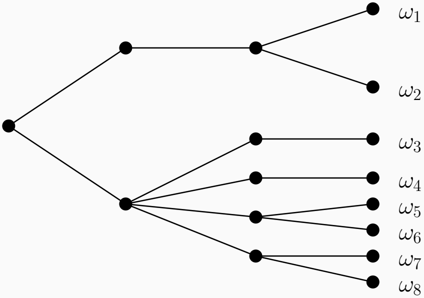
Information and measurability

Two graphical ways to represent the flow of information:

- Partitions



- Trees



Definition 19

A **stochastic process** $X = \{X(t)\}_{t=0,\dots,T}$ is a collection of random variables indexed by $t = 0, \dots, T$. You can see it as a function $X : \Omega \times \{0, \dots, T\} \to \mathbb{R}$ or as random variable $X : \Omega \to \mathbb{R}^{\{0,\dots,T\}}$, where $\mathbb{R}^{\{0,\dots,T\}}$ denotes the set of all real-valued functions with domain of definition $\{0, \dots, T\}$.

Definition 20

We say that a stochastic process X is **adapted to the filtration** $\mathbb{F}$ or $\mathbb{F}$**-adapted** if X_t is $\mathcal{F}_t$-measurable, $t = 0, \dots, T$.

Information and measurability

Definition 21

The **natural filtration generated by a stochastic process** X, denoted by $\mathbb{F}^X$, is defined by

$$\mathbb{F}^X = \left\{ \mathcal{F}_t^X = \mathfrak{a}\left(X\left(0\right), X\left(1\right), \ldots, X\left(t\right)\right) \right\}_{t=0,\ldots,T}.$$

- $\mathbb{F}^X$ is the minimal filtration to which X is adapted to. It contains the information that you can get by observing the process X.

Definition 22

We say that a process $X = \{X\left(t\right)\}_{t=1,\ldots,T}$ is **predictable with respect to a filtration** $\mathbb{F}$ or $\mathbb{F}$**-predictable** if X_t is $\mathcal{F}_{t-1}$-measurable, $t = 1, \ldots, T$.

Information and measurability

Example 23

- Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and $X = \{X(t)\}_{t=0,1,2}$ with $X(0) = 3$,

$$X(1,\omega) = \begin{cases} 5 & \text{if} \quad \omega = \omega_1, \omega_2 \\ 2 & \text{if} \quad \omega = \omega_3, \omega_4 \end{cases},$$

$$X(2,\omega) = \begin{cases} 6 & \text{if} \quad \omega = \omega_1, \omega_2 \\ 3 & \text{if} \quad \omega = \omega_3 \\ 2 & \text{if} \quad \omega = \omega_4 \end{cases}.$$

- Then,

$$\mathcal{F}_0^X = \mathfrak{a}(X(0)) = \mathfrak{a}\left(\pi_{X(0)}\right) = \{\emptyset, \Omega\},$$

Example 23

$$
\begin{aligned}
\mathcal{F}_1^X &= \mathfrak{a}\left(X\left(0\right), X\left(1\right)\right) = \mathfrak{a}\left(\pi_{X(0)} \cap \pi_{X(1)}\right) = \mathfrak{a}\left(\pi_{X(1)}\right) \\
&= \mathfrak{a}\left(\left\{\left\{\omega_1, \omega_2\right\}, \left\{\omega_3, \omega_4\right\}\right\}\right) = \left\{\emptyset, \Omega, \left\{\omega_1, \omega_2\right\}, \left\{\omega_3, \omega_4\right\}\right\}, \\
\mathcal{F}_2^X &= \mathfrak{a}\left(X\left(0\right), X\left(1\right), X\left(2\right)\right) = \mathfrak{a}\left(\pi_{X(0)} \cap \pi_{X(1)} \cap \pi_{X(2)}\right) \\
&= \mathfrak{a}\left(\pi_{X(2)}\right) = \mathfrak{a}\left(\left\{\left\{\omega_1, \omega_2\right\}, \left\{\omega_3\right\}, \left\{\omega_4\right\}\right\}\right) \\
&= \left\{\emptyset, \Omega, \left\{\omega_1, \omega_2\right\}, \left\{\omega_3\right\}, \left\{\omega_4\right\}, \left\{\omega_1, \omega_2, \omega_3\right\}, \left\{\omega_1, \omega_2, \omega_4\right\}, \right. \\
&\qquad \left. \left\{\omega_3, \omega_4\right\}\right\}.
\end{aligned}
$$

- In this case $\mathcal{F}_2^X \neq \mathcal{P}\left(\Omega\right)$.
- Check what happens if $X\left(2, \omega_2\right) = 3$.

Remark 24

- The systematic way to compute $\mathfrak{a}(S)$, where $S \subseteq \mathcal{P}(\Omega)$, is to identify the finest partition of Ω that you can obtain by basic set operations on all elements of S, denoted by π_S.
- Then, the elements of $\mathfrak{a}(S)$ will be all possible unions of elements in π_S.

Conditional Expectation

Conditional expectation

- Recall that a **probability measure** P on a finite sample space $\Omega = \{\omega_1, \ldots, \omega_K\}$ is a function $P : \Omega \to [0, 1]$ such that $\sum_{i=1}^{K} P(\omega_i) = 1$.
- The triple $(\Omega, \mathcal{P}(\Omega), P)$ is a **probability space**.
- In addition, we will assume that $P(\omega_i) > 0$, $i = 1, \ldots, K$. This assumption is not essential but implies that all sets in $\mathcal{P}(\Omega)$ have strictly positive probability, which simplifies the statements about conditional probabilities and conditional expectations.
- Given an event $A \in \mathcal{P}(\Omega)$ the **probability of** A happening is given by

$$P(A) = \sum_{\omega \in A} P(\omega).$$

- We say that **two events** $A, B \in \mathcal{P}(\Omega)$ **are independent** if

$$P(A \cap B) = P(A)\, P(B)\,.$$

- Given two events $A, B \in \mathcal{P}(\Omega)$, the **probability of** A **given** B, denoted by

$$P(A|\, B) = P(A \cap B)\, / P(B)\,.$$

Remark 25

In general, we would need to assume that $P(B) > 0$ for this probability to be well defined. However, thanks to the assumption on the strict positivity of P, this probability is always well defined in our setup.

Conditional expectation

Definition 26

Given **two algebras** $\mathcal{F}_1, \mathcal{F}_2$ on Ω we say that they **are independent** if for all $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$ we have that A and B are independent.

Definition 27

Given a random variable X we define its **expectation** by

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega) P(\omega).$$

Conditional expectation

Definition 28

Given an algebra $\mathcal{F}$ and a random variable X we define the **conditional expectation of X given $\mathcal{F}$** as the unique random variable Z, denoted by $\mathbb{E}\left[X \middle| \mathcal{F}\right]$, satisfying

1. Z is $\mathcal{F}$-measurable.
2. $\mathbb{E}\left[\mathbf{1}_A X\right] = \mathbb{E}\left[\mathbf{1}_A Z\right]$, $A \in \mathcal{F}$.

- Note that since $\mathbb{E}\left[X \middle| \mathcal{F}\right]$ is $\mathcal{F}$-measurable, it is constant on the partition that generates $\mathcal{F}$.
- How we compute $\mathbb{E}\left[X \middle| \mathcal{F}\right]$?

Conditional expectation

Definition 29

Let $A \in \mathcal{P}(\Omega)$ and X be a random variable. Then, the **conditional expectation of** X **given** A is the quantity

$$\mathbb{E}[X|A] = \sum_x xP(X = x|A),$$

where x are the values taken by X and

$$P(X = x|A) = \frac{P(\{\omega : X(\omega) = x\} \cap A)}{P(A)}.$$

- A remark analogous to Remark 25 applies to the previous definition.

Conditional expectation

Proposition 30

Let $\mathcal{F}$ be an algebra on Ω, X be a random variable and let $\pi = \{A_1, \ldots, A_m\}$ be the partition of Ω such that $\mathcal{F} = \mathfrak{a}(\pi)$. Then,

$$\mathbb{E}[X|\mathcal{F}](\omega) = \sum_{i=1}^{m} \mathbb{E}[X|A_i] \mathbf{1}_{A_i}(\omega).$$

Proof.

Smartboard. □

Remark 31

- Usually we are given (or we guess) a candidate Z to be $\mathbb{E}\left[X|\mathcal{F}\right]$, then we need to check conditions $1)$ and $2)$ in Definition 28.
- When $\mathcal{F} = \sigma\left(\pi\right)$, π a partition it suffices to check that the candidate Z is constant over the elements of π ($\mathcal{F}$-measurable) and check condition $2)$ in Definition 28 only for $A_i \in \pi$.

Example 32

- Let $\Omega = \{\omega_1, \ldots, \omega_4\}$ and $P(\omega_i) = 1/4, i = 1, ..., 4$.
- Consider the algebra $\mathcal{F} = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$ and the random variable X given by

$$X(\omega) = \begin{cases} 9 & \text{if} \quad \omega = \omega_1 \\ 6 & \text{if} \quad \omega = \omega_2, \omega_3 \\ 3 & \text{if} \quad \omega = \omega_4 \end{cases}$$
$$= 9\mathbf{1}_{\{\omega_1\}}(\omega) + 6\mathbf{1}_{\{\omega_2,\omega_3\}}(\omega) + 3\mathbf{1}_{\{\omega_4\}}(\omega).$$

- We will compute $\mathbb{E}[X|\mathcal{F}]$ on the smartboard.

Conditional expectation

Theorem 33

Suppose X and Y are random variables on $(\Omega, \mathcal{P}(\Omega), P)$, $\mathcal{G}$ is an algebra on Ω, $a, b \in \mathbb{R}$. Then,

1. ***Linearity***: $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$.
2. ***Law of total expectation***: $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.
3. ***Independence***: *If X is independent of $\mathcal{G}$ then* $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.
4. ***Measurability***: *If Y is $\mathcal{G}$-measurable then* $\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$.
5. ***Tower property***: *If $\mathcal{H}$ is an algebra on Ω such that $\mathcal{H} \subseteq \mathcal{G}$, then* $\mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.

Proof.

Smartboard. □

Conditional expectation

Theorem 34

Let X be a random variable on $(\Omega, \mathcal{P}(\Omega), P)$ and $\mathcal{G}$ an algebra on Ω. Then,

$$\mathbb{E}[X|\mathcal{G}] = \arg\min\left\{\mathbb{E}\left[(X-Y)^2\right] : \ Y \textit{ being } \mathcal{G}\textit{-measurable}\right\}.$$

Proof.

Smartboard. □

Remark 35

The conditional expectation is the best prediction of X based on the information contained in $\mathcal{G}$, in the sense of minimizing the L^2 error (variance).

Conditional expectation

Definition 36

Let $\mathbb{F} = \{\mathcal{F}_t\}_{t=0,\dots,T}$ be a filtration on $(\Omega, \mathcal{P}(\Omega), P)$. A stochastic process $X = \{X(t)\}_{t=0,\dots,T}$ is a **($\mathbb{F}$-) martingale** if

1. X is $\mathbb{F}$-adapted.
2. For $t \in \{0, \dots, T\}$,$s \geq 0$,$t + s \in \{0, \dots, T\}$ we have

$$\mathbb{E}\left[X(t+s) \mid \mathcal{F}_t\right] = X(t).$$

- Intuitively, the best forecast of the process at some future time $t + s$ given today's information $\mathcal{F}_t$ is the value of the process today.

Remark 37

- An $\mathbb{F}$-adapted process X is called a **(sub) supermartingale** if
$$\mathbb{E}\left[X(t+s)|\,\mathcal{F}_t\right](\geq) \leq X(t).$$
- If $\#\Omega = +\infty$ then we need to impose that $\mathbb{E}\left[|X(t)|\right] < \infty$ for all $t = 0, \ldots, T$.
- In the previous definitions we can change $X(t+s)$ by $X(t+1)$.

Proposition 38 (Martingale transform or stochastic integral)

Let $\mathbb{F} = \{\mathcal{F}_t\}_{t=0,\dots,T}$ be a filtration on $(\Omega, \mathcal{P}(\Omega), P)$. Let H be an $\mathbb{F}$-predictable process and M an $\mathbb{F}$-martingale. Then, the process Y defined by $Y_0 = c$ (a constant) and

$$Y(t) = \sum_{s=1}^{t} H(s)\left(M(s) - M(s-1)\right)$$
$$= \sum_{s=1}^{t} H(s)\,\Delta M(s), \qquad t = 1, \dots, T,$$

is an $\mathbb{F}$-martingale with $\mathbb{E}\left[Y(t)\right] = c$.

Proof.

Smartboard. □