An Analytical Introduction to Probability Theory

12. Conditional Expectations and Martingales

DASN, NTU https://personal.ntu.edu.sg/wptay/

12.1 Definition

Given a probability space $(\Omega, \mathscr{A}, \mathbb{P})$, let us start off with a simple example that we are familiar with from undergraduate probability courses. Suppose we have a r.v. X and a discrete r.v. $Y \in \{1, 2, \ldots\}$. We can partition $\Omega = \bigcup_{y \geq 1} \Omega_y$, where $\Omega_y = \{\omega : Y(\omega) = y\}$. Then the conditional expectation of X given Y can be defined as

$$\mathbb{E}[X \mid Y = y] = \frac{\mathbb{E}\left[X \mathbf{1}_{\Omega_y}\right]}{\mathbb{P}(\Omega_y)}$$

for each value of y. Note that the expectation is conditioned on a set $\Omega_y = \{Y = y\}$. It depends on the value $Y(\omega) = y$, and is hence a function of ω. Since Ω_y is measurable, this is a measurable function and is hence a *random variable*!

We wish to generalize this definition to all sets in a sub-σ-algebra in $\mathscr{A}$. Furthermore, when we average out the conditional expectation over a set B of feasible Y values, we should get back the expectation over this set:

$$\sum_{y \in B} \mathbb{E}[X \mid Y = y] \mathbb{P}(Y = y) = \sum_{y \in B} \mathbb{E}\left[X \mathbf{1}_{\Omega_y}\right] = \mathbb{E}[X \mathbf{1}_{Y \in B}],$$

where the last inequality follows from Fubini's theorem.

Definition 12.1. *Suppose $\mathbb{E}|X| < \infty$ and the σ-algebra $\mathscr{F} \subset \mathscr{A}$. A random variable $Y : \Omega \mapsto \mathbb{R}$ is a conditional expectation of X given $\mathscr{F}$ if*

(i) $Y^{-1}(B) \in \mathscr{F}$, $\forall\, B \in \mathscr{B}(\mathbb{R})$, i.e., Y is $\mathscr{F}$-measurable (we denote it as $Y \in \mathscr{F}$).

(ii) $\forall\, A \in \mathscr{F}$, $\mathbb{E}[Y \mathbf{1}_A] = \int_A Y \,\mathrm{d}\mathbb{P} = \int_A X \,\mathrm{d}\mathbb{P} = \mathbb{E}[X \mathbf{1}_A]$.

If Y is a conditional expectation of X given $\mathscr{F}$, we write $Y = \mathbb{E}[X \mid \mathscr{F}]$. We also write $\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid \sigma(Y)]$, where $\sigma(Y)$ is the σ-algebra generated by Y. The notion that expectation is an operator comes from here: $\mathbb{E}[\cdot \mid \mathscr{F}] : L^1(\Omega, \mathscr{A}, \mathbb{P}) \mapsto L^1(\Omega, \mathscr{F}, \mathbb{P})$ is a linear (which will be shown later) transformation.

The **existence** of conditional expectations is given by Radon–Nikodym Theorem (Theorem 4.4). Suppose $X \geq 0$. We can define a measure

$$\mu(A) = \int_A X \,\mathrm{d}\mathbb{P}, \text{ where } A \in \mathscr{F} \text{ and } \mu \ll \mathbb{P}.$$

Since X is integrable, μ is a finite measure. Then there exists $Y = \dfrac{\mathrm{d}\mu}{\mathrm{d}\mathbb{P}} \in \mathscr{F}$ such that $\int_A X \,\mathrm{d}\mathbb{P} = \mu(A) = \int_A Y \,\mathrm{d}\mathbb{P}$. The existence of conditional expectations for general $X = X^+ - X^-$ now follows.

We next show that conditional expectations are **unique** almost surely. Suppose that Y and Y' are both versions of $\mathbb{E}[X \,|\, \mathscr{F}]$ and $\mathbb{P}(Y \neq Y') > 0$, i.e., $\mathbb{P}(Y > Y') > 0$ or $\mathbb{P}(Y < Y') > 0$. Let $A = \{Y > Y'\} \in \mathscr{F}$ and suppose that $\mathbb{P}(A) > 0$. Then, we have

$$\begin{aligned} 0 < \mathbb{E}[(Y - Y')\mathbf{1}_A] &= \mathbb{E}Y\mathbf{1}_A - \mathbb{E}Y'\mathbf{1}_A \\ &= \mathbb{E}X\mathbf{1}_A - \mathbb{E}X\mathbf{1}_A \\ &= 0, \end{aligned}$$

which is a contradiction. A similar argument holds for the case $\mathbb{P}(Y < Y') > 0$. Therefore, $\mathbb{P}(Y \neq Y') = 0$, and $Y = Y'$ a. s.

Example 12.1. *Suppose that the joint pdf of (X, Y) is $f(x, y)$. Let*

$$h(y) = \int g(x) f(x|y) \,\mathrm{d}x = \frac{\int g(x) f(x, y) \,\mathrm{d}x}{\int f(x, y) \,\mathrm{d}x}.$$

We show that $h(Y) = \mathbb{E}[g(X) \,|\, Y]$. Let $A \in \sigma(Y)$. Then $A = \{\omega : Y(\omega) \in B\}$ for some $B \in \mathscr{B}(\mathbb{R})$. We check that

$$\begin{aligned} \mathbb{E}[h(Y)\mathbf{1}_A] &= \int_B \int h(y) f(x, y) \,\mathrm{d}x \,\mathrm{d}y \\ &= \int_B h(y) \int f(x, y) \,\mathrm{d}x \,\mathrm{d}y \\ &= \int_B \int g(x) f(x, y) \,\mathrm{d}x \,\mathrm{d}y \\ &= \mathbb{E}[g(X)\mathbf{1}_B(Y)] \\ &= \mathbb{E}[g(X)\mathbf{1}_A], \end{aligned}$$

and the claim is proved.

Example 12.2. *A sensor makes an observation $X \in \mathbb{R}$ and sends a summary $Z = \gamma(X) \in \mathbb{R}$ to a fusion center, where γ is a randomized function. The fusion center uses Z to perform hypothesis testing for*

$$H = \begin{cases} H_0 : & X \sim \mathbb{P}_0, \\ H_1 : & X \sim \mathbb{P}_1. \end{cases}$$

We assume that $\mathbb{P}_0$ and $\mathbb{P}_1$ are absolutely continuous w.r.t. each other. Note that these are the laws of X under different hypotheses.

For $i = 0, 1$, let $\mathbb{P}_{i,Z}$ be the restriction of $\mathbb{P}_i$ on $\sigma(Z)$. Suppose $A \in \sigma(Z)$. We have

$$\begin{aligned}
\int_A \frac{\mathrm{d}\mathbb{P}_{1,Z}}{\mathrm{d}\mathbb{P}_{0,Z}} \, \mathrm{d}\mathbb{P}_0 &= \int_A \frac{\mathrm{d}\mathbb{P}_{1,Z}}{\mathrm{d}\mathbb{P}_{0,Z}} \, \mathrm{d}\mathbb{P}_{0,Z} \\
&= \int_A \mathrm{d}\mathbb{P}_{1,Z} \\
&= \int_A \mathrm{d}\mathbb{P}_1 \\
&= \int_A \frac{\mathrm{d}\mathbb{P}_1}{\mathrm{d}\mathbb{P}_0} \mathrm{d}\mathbb{P}_0.
\end{aligned}$$

Therefore,

$$\frac{\mathrm{d}\mathbb{P}_{1,Z}}{\mathrm{d}\mathbb{P}_{0,Z}} = \mathbb{E}\left[\frac{\mathrm{d}\mathbb{P}_1}{\mathrm{d}\mathbb{P}_0} \,\middle|\, \sigma(Z) \right].$$

As a special case, suppose X and Z are continuous r.v.s. Then γ corresponds to a conditional pdf $p(z|x)$ and letting f_i be the pdf of $\mathbb{P}_i$, $i = 0, 1$, we have from Example 12.1,

$$\begin{aligned}
\mathbb{E}_0\left[\frac{f_1(x)}{f_0(x)} \,\middle|\, Z = z \right] &= \frac{\int \frac{f_1(x)}{f_0(x)} f_0(x, z) \, \mathrm{d}z}{\int f_0(x, z) \, \mathrm{d}x} \\
&= \frac{\int f_1(x) p(z|x) \, \mathrm{d}x}{\int f_{0,Z}(z)} \\
&= \frac{f_{1,Z}(z)}{f_{0,Z}(z)}.
\end{aligned}$$

12.2 Properties

In the following, we list some fundamental properties of conditional expectations, many without proofs. The proofs are left for your exercise.

1. If $\sigma(X) \subset \mathscr{F}$, then $X \in \mathscr{F}$ and by Definition 12.1, $X = \mathbb{E}[X \,|\, \mathscr{F}]$ a.s. As a special case, we have $\mathbb{E}[X \,|\, \mathscr{A}] = X$. If c is a constant, viewed as a r.v., its σ-algebra is the trivial one and so $\mathbb{E}[c \,|\, \mathscr{F}] = c$.

2. If $\sigma(X) \perp\!\!\!\perp \mathscr{F}$, then for $A \in \mathscr{F}$, $\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}X\mathbb{E}\mathbf{1}_A = \int_A \mathbb{E}X \, \mathrm{d}\mathbb{P} \Rightarrow \mathbb{E}[X \,|\, \mathscr{F}] = \mathbb{E}X$ a.s. As a special case, if $\mathscr{F} = \{\emptyset, \Omega\}$, we have $\mathbb{E}[X \,|\, \mathscr{F}] = \mathbb{E}X$.

3. Suppose that c is a constant, then $\mathbb{E}[cX + Y \,|\, \mathscr{F}] = c\mathbb{E}[X \,|\, \mathscr{F}] + \mathbb{E}[Y \,|\, \mathscr{F}]$.

4. Suppose $\mathscr{G}$ and $\mathscr{F}$ are σ-algebras with $\mathscr{G} \subset \mathscr{F}$, then $\mathbb{E}[X \,|\, \mathscr{G}] = \mathbb{E}[\mathbb{E}[X \,|\, \mathscr{G}] \,|\, \mathscr{F}] = \mathbb{E}[\mathbb{E}[X \,|\, \mathscr{F}] \,|\, \mathscr{G}]$.

Proof. Let $A \in \mathscr{G} \subset \mathscr{F}$, then we have

$$\begin{aligned} \mathbb{E}[\mathbb{E}[\mathbb{E}[X \mid \mathscr{F}] \mid \mathscr{G}]\mathbf{1}_A] &= \mathbb{E}[\mathbb{E}[X \mid \mathscr{F}]\mathbf{1}_A] \text{ by definition of } \mathbb{E}[\cdot \mid \mathscr{G}] \\ &= \mathbb{E}X\mathbf{1}_A \text{ by definition of } \mathbb{E}[\cdot \mid \mathscr{F}]. \end{aligned}$$

□

For example, setting $\mathscr{G} = \{\emptyset, \Omega\}$, we obtain $\mathbb{E}[\mathbb{E}[X \mid \mathscr{F}]] = \mathbb{E}X$ for all $\mathscr{F}$.

5. If $X \leq Y$ a.s., then $\mathbb{E}[X \mid \mathscr{F}] \leq \mathbb{E}[Y \mid \mathscr{F}]$.

   Lemma 12.1. *$\mathbb{E}[X \mid \mathscr{F}] \leq \mathbb{E}[Y \mid \mathscr{F}]$ a.s. iff $\mathbb{E}X\mathbf{1}_A \leq \mathbb{E}Y\mathbf{1}_A$ for all $A \in \mathscr{F}$.*

   Proof. Similar to the uniqueness proof. □

6. Monotone Convergence Theorem. Suppose $\mathbb{E}|X_n| < \infty$, $\forall n \geq 1$, $\mathbb{E}|X| < \infty$ and $X_n \uparrow X$ a.s. as $n \to \infty$, then $\mathbb{E}[X_n \mid \mathscr{F}] \uparrow \mathbb{E}[X \mid \mathscr{F}]$ a.s.

   Proof. Since $X_n \uparrow X$, we have

   $$\mathbb{E}[X_n \mid \mathscr{F}] \leq \mathbb{E}[X_{n+1} \mid \mathscr{F}] \leq \mathbb{E}[X \mid \mathscr{F}]$$

   and there exists

   $$Y \triangleq \lim_{n\to\infty} \mathbb{E}[X_n \mid \mathscr{F}] \leq \mathbb{E}[X \mid \mathscr{F}].$$

   From Lemma 4.2, Y is $\mathscr{F}$-measurable. Moreover, for each $A \in \mathscr{F}$, we have

   $$\mathbb{E}[X_n \mid \mathscr{F}]\mathbf{1}_A \uparrow Y\mathbf{1}_A$$

   since $\mathbb{E}[X_n \mid \mathscr{F}] \uparrow Y$. From the MCT, we obtain

   $$\mathbb{E}[\mathbb{E}[X_n \mid \mathscr{F}]\mathbf{1}_A] \to \mathbb{E}[Y\mathbf{1}_A].$$

   But $\mathbb{E}[\mathbb{E}[X_n \mid \mathscr{F}]\mathbf{1}_A] = \mathbb{E}[X_n\mathbf{1}_A] \xrightarrow{\text{MCT}} \mathbb{E}[X\mathbf{1}_A]$. Therefore, $\mathbb{E}[Y\mathbf{1}_A] = \mathbb{E}[X\mathbf{1}_A]$ and $Y = \mathbb{E}[X \mid \mathscr{F}]$ a.s. □

7. Dominated Convergence Theorem. If $|X_n| \leq Y$ a.s., $\mathbb{E}Y < \infty$ and $X_n \to X$ a.s., then $\lim_{n\to\infty} \mathbb{E}[X_n \mid \mathscr{F}] = \mathbb{E}[X \mid \mathscr{F}]$.

8. Suppose $X \perp\!\!\!\perp Y$ and $\mathbb{E}|\phi(X,Y)| < \infty$. Let $g(y) = \mathbb{E}[\phi(X,y)]$, then $\mathbb{E}[\phi(X,Y) \mid Y] = g(Y)$.

9. If $\mathbb{E}|X| < \infty$, $\mathbb{E}|XY| < \infty$, and $Y \in \mathscr{F}$, then $\mathbb{E}[XY \mid \mathscr{F}] = Y\mathbb{E}[X \mid \mathscr{F}]$.

10. The usual inequalities apply. We define $\mathbb{P}(A \mid \mathscr{F}) = \mathbb{E}[\mathbf{1}_A \mid \mathscr{F}]$.

    - Markov's inequality: For $a > 0$, $\mathbb{P}(X > a \mid \mathscr{F}) \leq a\mathbb{E}[X \mid \mathscr{F}]$.
    - Chebyshev's inequality: For $a > 0$, $\mathbb{P}(|X| \geq a \mid \mathscr{F}) \leq a^2\mathbb{E}[X^2 \mid \mathscr{F}]$.
    - Cauchy-Schwarz inequality: $\mathbb{E}[XY \mid \mathscr{F}]^2 \leq \mathbb{E}[X^2 \mid \mathscr{F}]\mathbb{E}[Y^2 \mid \mathscr{F}]$.
    - Jensen's inequality: Given a convex function ϕ with $\mathbb{E}|\phi(X)| < \infty$, we have

      $$\phi(\mathbb{E}[X \mid \mathscr{F}]) \leq \mathbb{E}[\phi(X) \mid \mathscr{F}].$$

12.3 L^2 Interpretation

Proposition 12.1. *Consider $L^2(\Omega, \mathscr{F}, \mathbb{P}) = \{X : X \in \mathscr{F},\ \mathbb{E}X^2 < \infty\}$, which is a subspace of $L^2(\Omega, \mathscr{A}, \mathbb{P})$. If $X \in L^2(\Omega, \mathscr{A}, \mathbb{P})$, then*

$$\mathbb{E}[X \mid \mathscr{F}] = \operatorname*{arg\,min}_{Y \in L^2(\Omega, \mathscr{F}, \mathbb{P})} \mathbb{E}(X - Y)^2.$$

Proof. For $Y \in L^2(\Omega, \mathscr{F}, \mathbb{P})$, let $Z = \mathbb{E}[X \mid \mathscr{F}] - Y \in \mathscr{F}$. Then, we have

$$\begin{aligned}
\mathbb{E}(X - Y)^2 &= \mathbb{E}\Big[(X - \mathbb{E}[X \mid \mathscr{F}] + \mathbb{E}[X \mid \mathscr{F}] - Y)^2\Big] \\
&= \mathbb{E}(X - \mathbb{E}[X \mid \mathscr{F}])^2 + \mathbb{E}Z^2 + 2\mathbb{E}[Z(X - \mathbb{E}[X \mid \mathscr{F}])]
\end{aligned}$$

and since $Z \in \mathscr{F}$,

$$\begin{aligned}
\mathbb{E}[Z(X - \mathbb{E}[X \mid \mathscr{F}])] &= \mathbb{E}[ZX] - \mathbb{E}[\mathbb{E}[ZX \mid \mathscr{F}]] \\
&= \mathbb{E}[ZX] - \mathbb{E}[ZX] \\
&= 0.
\end{aligned}$$

Therefore,

$$\mathbb{E}(X - Y)^2 \geq \mathbb{E}(X - \mathbb{E}[X \mid \mathscr{F}])^2,$$

and the proposition is proved. ∎

We can define $\mathrm{var}(X\mathscr{F}) = \mathbb{E}[(X - \mathbb{E}[X \mid \mathscr{F}])^2 \mid \mathscr{F}] = \mathbb{E}[X^2 \mid \mathscr{F}] - \mathbb{E}[X \mid \mathscr{F}]^2$. One can show (exercise) that

$$\mathrm{var}(X) = \mathbb{E}[\mathrm{var}(X \mid \mathscr{F})] + \mathrm{var}(\mathbb{E}[X \mid \mathscr{F}]).$$

12.4 Martingales

Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. A sequence of sub-$\sigma$-algebras $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots \subset \mathscr{F}$ is called a *filtration*. Let the r.v. $M_n \in \mathscr{F}_n$ (i.e., M_n is $\mathscr{F}_n$-measurable). We say that M_n is adapted to $\mathscr{F}_n$.

Definition 12.2 (Martingale)**.** *We say that $(M_n, \mathscr{F}_n)_{n \geq 0}$ is a martingale if $\mathbb{E}|M_n| < \infty$ for all $n \geq 0$ and*

$$\mathbb{E}[M_m \mid \mathscr{F}_n] = M_n,\ \forall\, m \geq n. \tag{1}$$

By induction, the condition (1) is equivalent to $\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] = M_{n-1}$. We give examples of martingales below.

Example 12.3. *Suppose that* $(X_n)_{n\geq 1}$ *are independent r.v.s,* $\mathscr{F}_n = \sigma(X_1,\dots,X_n)$ *and* $\mathbb{E}X_n = 0$. *Let* $S_n = \sum_{i=1}^n X_i$. *Then,*

$$\begin{aligned}\mathbb{E}[S_n \mid \mathscr{F}_{n-1}] &= \mathbb{E}[S_{n-1} + X_n \mid \mathscr{F}_{n-1}] \\ &= S_{n-1} + \mathbb{E}[X_n \mid \mathscr{F}_{n-1}]\end{aligned}$$

Since $\mathbb{E}[X_n \mid \mathscr{F}_{n-1}] = \mathbb{E}X_n = 0$, *we have*

$$\mathbb{E}[S_n \mid \mathscr{F}_{n-1}] = S_{n-1}.$$

Therefore, $(S_n, \mathscr{F}_n)_{n\geq 1}$ *is a martingale.*

Example 12.4. *Suppose that* $(X_n)_{n\geq 1}$ *are independent r.v.s,* $\mathrm{var}(X_n) = \sigma^2$ *and* $\mathbb{E}X_n = 0$. *Let* $M_0 = 0$, *and* $M_n = S_n^2 - n\sigma^2$. *We have*

$$\begin{aligned}\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] &= \mathbb{E}\left[S_{n-1}^2 + 2S_{n-1}X_n + X_n^2 - n\sigma^2 \,\middle|\, \mathscr{F}_{n-1}\right] \\ &= S_{n-1}^2 - (n-1)\sigma^2 \\ &= M_{n-1},\end{aligned}$$

and $(M_n, \mathscr{F}_n)_{n\geq 1}$ *is a martingale.*

Example 12.5. *Suppose that* $(X_i)_{i\geq 1}$ *are independent r.v.s.,* $X_i \geq 0$ *and* $\mathbb{E}X_i = 1$. *Let* $M_0 = 1$, *and* $M_n = \prod_{i=1}^n X_i$. *We have*

$$\begin{aligned}\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] &= \mathbb{E}[M_{n-1} \cdot X_n \mid \mathscr{F}_{n-1}] \\ &= M_{n-1}\mathbb{E}[X_n \mid \mathscr{F}_{n-1}]\end{aligned}$$

Since $\mathbb{E}[X_n \mid \mathscr{F}_{n-1}] = \mathbb{E}X_i = 1$, *we have*

$$\mathbb{E}[S_n \mid \mathscr{F}_{n-1}] = M_{n-1},$$

and $(M_n, \mathscr{F}_n)_{n\geq 1}$ *is a martingale.*

Example 12.6. *Suppose that* $(Y_n)_{n\geq 1}$ *are i.i.d. and* $\phi(\lambda) = \mathbb{E}e^{\lambda Y_1} < \infty$. *Let* $X_n = \dfrac{e^{\lambda Y_n}}{\phi(\lambda)}$. *Then* $\mathbb{E}X_n = 1$. *Let*

$$M_n = \frac{\exp\left(\lambda \sum_{i=1}^n Y_n\right)}{\phi(\lambda)^n}$$

and from Example 12.5, we have $(M_n, \mathscr{F}_n)_{n\geq 1}$ *is a martingale.*

Example 12.7. *Given* $\mathbb{E}|X| < \infty$ *and a filtration* $(\mathscr{F}_n)_{n\geq 1}$, *let* $M_n = \mathbb{E}[X \mid \mathscr{F}_n]$. *Then,*

$$\begin{aligned}\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] &= \mathbb{E}[\mathbb{E}[X \mid \mathscr{F}_n] \mid \mathscr{F}_{n-1}] \\ &= \mathbb{E}[X \mid \mathscr{F}_{n-1}] \\ &= M_{n-1}.\end{aligned}$$

Therefore, $(M_n, \mathscr{F}_n)_{n\geq 1}$ *is a martingale.*

Definition 12.3. *$(A_n)_{n\geq 1}$ is predictable w.r.t. $(\mathscr{F}_n)_{n\geq 0}$ if $A_n \in \mathscr{F}_{n-1}, \forall\, n \geq 1$.*

We call $(\widetilde{M}_n)_{n\geq 0}$ the martingale transform of $(M_n)_{n\geq 0}$ by $(A_n)_{n\geq 1}$ if

(i) $\widetilde{M}_0 = M_0$. (In general, we can choose any integrable $\widetilde{M}_0 \in \mathscr{F}_0$.)

(ii) $\widetilde{M}_n = \widetilde{M}_0 + \sum_{k=1}^{n} A_k(M_k - M_{k-1})$.

Theorem 12.1 (MTT)**.** *If $(A_n)_{n\geq 1}$ is predictable w.r.t. $(\mathscr{F}_n)_{n\geq 0}$ and bounded, and $(M_n, \mathscr{F}_n)_{n\geq 0}$ is a martingale, then $(\widetilde{M}_n, \mathscr{F}_n)_{n\geq 0}$ is a martingale.*

Proof. It is obvious that $\widetilde{M}_n \in \mathscr{F}_n$ and $\mathbb{E}|\widetilde{M}_n| < \infty$. Furthermore, we have

$$\begin{aligned}\mathbb{E}\Big[\widetilde{M}_n - \widetilde{M}_{n-1} \,\Big|\, \mathscr{F}_{n-1}\Big] &= \mathbb{E}[A_n(M_n - M_{n-1}) \,|\, \mathscr{F}_{n-1}] \\ &= A_n \mathbb{E}[M_n - M_{n-1} \,|\, \mathscr{F}_{n-1}] = 0.\end{aligned}$$

□

Example 12.8. *Suppose we divide time into discrete equal intervals (e.g., days). Let M_n be the price of a stock at time instant $n \geq 1$ and $\mathscr{F}_n$ be the available information up to time n. From the efficient-market hypothesis (assuming no dividends and discount factor), M_n is a "fair" price that has priced in all expected future gains or losses, i.e., $\mathbb{E}[M_m \,|\, \mathscr{F}_n] = M_n$ for $m \geq n$ and $(M_n, \mathscr{F}_n)$ is a martingale. Let $(A_n)_{n\geq 1}$ be a strategy that decides to hold A_n units of the stock in the time period $[n-1, n)$. Clearly, (A_n) has to be predictable w.r.t. $(\mathscr{F}_n)$. Then $\widetilde{M}_n = \sum_{k\leq n} A_k(M_k - M_{k-1})$ is the change in wealth of this strategy up to time n. The MTT says that the expected wealth of* any *strategy is the same and what you start with, i.e., you cannot beat the market! We will see a much stronger version of this result in Doob's Optional Stopping Theorem in the next session.*