Martingales
February 19-21, 2001

Definitions. Let $(\Omega, \mathcal{A}, P)$ be a probability space. Then an increasing sequence of sub
sigma-algebras $\mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \mathcal{A}$ is said to be a filtration. A sequence $Y_0, Y_1, Y_2, \cdots$
of random variables is said to be adapted to a filtration $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots$ if

\[ Y_n \text{ is } \mathcal{A}_n \text{—measurable} \quad (1) \]

for every $n \geq 1$, and $Y_0, Y_1, Y_2, \cdots$ is said to be a martingale with respect to a filtration $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots$ iff for every $n \geq 1$:

\[ E|Y_n| < \infty \quad (2) \]

and

\[ E(Y_{n+1}|\mathcal{A}_n) = Y_n \text{ w.p.1.} \quad (3) \]

Examples: From Independent Random Variables. Let $X_1, X_2, \cdots$ be independent random
variables, and let $\mathcal{A}_n = \sigma\{X_1, \cdots, X_n\}$ for $n \geq 1$ and $\mathcal{A}_0 = \{\emptyset, \Omega\}$.

a): Sums. If $E(X_k) = 0$ for all $k$, then $S_n = X_1 + \cdots + X_n$, $n \geq 0$, defines a martingale
with respect to $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots$. For (1) and (2) are clear and

\[ E(S_{n+1}|\mathcal{A}_n) = E(S_n|\mathcal{A}_n) + E(X_{n+1}|\mathcal{A}_n) = S_n + 0 = S_n \]

w.p.1 for all $n \geq 0$.

b): Products. If $E(X_k) = 1$ for all $k$, then $Y_n = X_1 \times \cdots \times X_n$, $n \geq 1$, defines a martingale
with respect to $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots$. Again, (1) and (2) are clear and

\[ E(Y_{n+1}|\mathcal{A}_n) = Y_n E(X_n) = Y_n \]

w.p.1 for each $n$. \quad \diamond

Example: Successive Conditional Expectation. Let $\mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \mathcal{A}$, and let $Y$ be
an integrable random variable. Then

\[ Y_n = E(Y|\mathcal{A}_n), \quad n \geq 0, \]

defines a martingale with respect to $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots$. Again, (1) and (2) are clear, and

\[ E(Y_{n+1}|\mathcal{A}_n) = E[E(Y|\mathcal{A}_{n+1})|\mathcal{A}_n] = E(Y|\mathcal{A}_n) = Y_n \]

w.p.1 for each $n$. \quad \diamond

A sequence $Y_0, Y_1, Y_2, \cdots$ of random variables is said to be a sub (super) martingale
with respect to $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots$ iff (1), (2), and

\[ E(Y_{n+1}|\mathcal{A}_n) \geq (\leq) Y_n \text{ w.p.1} \quad (4) \]
for each \(n\). Thus, \(Y_0, Y_1, Y_2, \cdots\) is a martingale with respect to \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\) iff it is both a sub and super martingale.

**Remarks a.** Let \(Y_0, Y_1, Y_2, \cdots\) be a (sub) martingale with respect to some sequence \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\) and let \(\mathcal{B}_n = \sigma\{Y_0, \cdots, Y_n\}\) for \(n \geq 1\). Then \(\mathcal{B}_n \subseteq \mathcal{A}_n\), by (2) and, therefore,

\[
E(Y_{n+1}|\mathcal{B}_n) = E[E(Y_{n+1}|\mathcal{A}_n)|\mathcal{B}_n] = (\geq) E(Y_n|\mathcal{B}_n) = Y_n
\]

w.p.1. So, if \(Y_0, Y_1, Y_2, \cdots\) is a (sub) martingale with respect to any sequence \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\), then it is a (sub) martingale with respect to \(\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2, \cdots\). Below, the unqualified term (sub) martingale means a (sub) martingale with respect to \(\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2, \cdots\).

b). If \(Y_0, Y_1, Y_2, \cdots\) is a (sub) martingale with respect to \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\) then

\[
\int_A Y_{n+1} dP = \int_A E(Y_{n+1}|\mathcal{A}_n) dP = (\geq) \int_A Y_n dP
\]

for every \(A \in \mathcal{A}_n\) and every \(n\). Conversely, if (5) holds, then \(Y_0, Y_1, Y_2, \cdots\) is a (sub) martingale with respect to \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\).

c). In particular, if \(Y_0, Y_1, \cdots\) is a (sub) martingale, then \(E(Y_0) = (\leq) E(Y_1) = (\leq) \cdots\).

**Convex Functions.** We need Jensen’s Inequality for conditional expectations.

**Jensen’s Inequality.** Let \(\mathcal{B} \subseteq \mathcal{A}\) be a sigma-algebra; let \(Y\) be an integrable random variable; and let \(\psi: \mathbb{R} \to \mathbb{R}\) be a convex function for which \(Z = \psi(Y)\) is integrable. Then

\[
E(Z|\mathcal{B}) \geq \psi(E(Y|\mathcal{B})).
\]

**Proof.** Apply Jensen’s Inequality to the conditional distribution of \(Y\) given \(\mathcal{B}\). \(\diamondsuit\)

**Corollary.** Let \(Y_0, Y_1, Y_2, \cdots\) be a (sub) martingale with respect to \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\), and let \(\psi: \mathbb{R} \to \mathbb{R}\) be a (non-decreasing) convex function for which \(Z_n = \psi(Y_n)\) are integrable. Then let \(Z_0, Z_1, Z_2, \cdots\) is a (sub) martingale with respect to \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\).

**Proof.** Using Jensen’s Inequality and the monotonicity of \(\psi\) in the submartingale case,

\[
E(Z_{n+1}|\mathcal{A}_n) \geq \psi[E(Y_{n+1}|\mathcal{A}_n)] \geq \psi(Y_n)
\]

w.p.1. \(\diamondsuit\)

**Examples:** If \(Y_n\) is a square integrable martingale, then \(Y_n^2\) is a sub-martingale; If \(Y_n\) is submartingale, then so it \(Y_n^+\).

**Exercise:** If the random variables \(X_k\) have finite variances \(\sigma_k^2\) in a) of the first example, then \(Y_n = S_n^2 - \sum_{k=1}^{n} \sigma_k^2\) is a martingale with respect to \(\mathcal{A}_n\).

**Martingale Transforms.** Let \(Y_0, Y_1, Y_2, \cdots\) be a martingale with respect to a sequence \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\), and let

\[
\Delta_n = Y_n - Y_{n-1}
\]

for \(n \geq 1\). Then \(\Delta_n\) satisfy (1), (2), and

\[
E(\Delta_{n+1}|\mathcal{A}_n) = 0
\]

w.p.1 for each \(n\). Any such sequence \(\Delta_n\) are called *martingale differences with respect to* \(\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots\).
Theorem. Martingale Transforms. Let $Y_0, Y_1, \Delta_2, \cdots$ be a (sub) martingale with respect to $A_0, A_1, A_2, \cdots$; define $\Delta_n$ by (6); and let $W_n$ be (non-negative) random variables for which $W_n$ is $A_{n-1}$-measurable and $E|W_n \Delta_n| < \infty$ for each $n$; and let $Z_0$ be an integrable $A_0$-measurable random variable. Then

$$Z_n = Z_0 + \sum_{k=1}^{n} W_k \Delta_k, \ n \geq 0,$$

defines a (sub) martingale with respect to $A_0, A_1, A_2, \cdots$.

Proof. It is clear that $Y_n$ is $A_n$-measurable for each $n$ and that

$$E|Z_n| \leq E|Z_0| + \sum_{k=1}^{n} E|W_k \Delta_k| < \infty$$

For the (sub) martingale property, observe that

$$E(Z_{n+1} - Z_n | A_n) = W_{n+1} E(\Delta_{n+1} | A_n) = (\geq 0 \ w.p.1$$

for each $n$. $

\diamond$

Remark: Gambling Systems. If $\Delta_1, \Delta_2 \cdots$ are i.i.d. taking the values $\pm 1$ with probability $1/2$ each, and $A_n = \sigma\{\Delta_1, \cdots, \Delta_n\}$, then $\Delta_1, \Delta_2 \cdots$ are martingale differences with respect to $A_0, A_1, A_2, \cdots$. So, if $W_k$ is $A_{k-1}$-measurable and $E|W_k| < \infty$ for all $k$, then $Y_n$ defines a martingale. $

\diamond$

Stopping Times. An integer or $\infty$ valued random variable $\tau$ is said to be a stopping time with respect to a filtration $A_0, A_1, A_2, \cdots$ if

$$\{\tau \leq n\} \in A_n.$$ (7)

Since $\{\tau = n\} = \{\tau \leq n\} \cap \{\tau \leq n - 1\}'$ and $\{\tau \leq n\} = \{\tau = 0\} \cup \cdots \cup \{\tau = n\}$ for all $n$, it is easily seen that $\{\tau \leq n\}$ may be replaced by $\{\tau = n\}$ in (7).

Constants may be regarded as stopping times. For if $\tau(\omega) = m$ for all $\omega \in \Omega$, then $\{\tau \leq n\} = \emptyset$ or $\Omega$ for $n < m$ or $n \geq m$, and (7) is satisfied. For less trivial examples, let $Y_0, Y_1, \cdots$ be adapted to a filtration $A_0, A_1, A_2, \cdots$ and $B_n \subseteq \mathbb{R}^{n+1}$ be Borel sets. Then

$$\tau = \inf\{(Y_0, \cdots, Y_n) \in B_n\}$$

defines a stopping time. For,

$$\{\tau \leq n\} = \bigcup_{k=0}^{n}\{(Y_0, \cdots, Y_n) \in B_n\}$$

for all $n$. If $\tau$ is a stopping time, with respect to $A_0, A_1, A_2, \cdots$, let

$$A_\tau = \{B \in A : B \cap \{\tau \leq n\} \in A_n, \ \text{for all} \ n\}.$$

As above "$\tau \leq n$" may be replaced by "$\tau = n$" in the definition.
Lemma 1. If $\tau$ is a stopping time, then $A_\tau$ is a sigma-algebra.

Proof. $\Omega \in A_\tau$ by (7); if $A \in A_\tau$, then $A' \cap \{\tau \leq n\} = \{\tau \leq n\} - A \cap \{\tau \leq n\} \in A_n$ for all $n$, so that $A' \in A_\tau$; and if $A_1, A_2, \cdots \in A_\tau$, then

$$(\cup_{k=1}^{\infty} A_k) \cap \{\tau \leq n\} = \cup_{k=1}^{\infty} A_k \cap \{\tau \leq n\} \in A_n$$

for all $n$, so that $A_\tau$ is closed under countable unions. \hfill \Box

Lemma 2. If $\eta$ and $\tau$ are stopping times, then so are $\tau \wedge \eta$ and $\tau \vee \eta$. If $\eta$ and $\tau$ are stopping times for which $\eta \leq \tau$, then $A_\eta \subseteq A_\tau$.

Proof. If $\eta$ and $\tau$ are stopping times and $n \geq 1$, then

$$\{\eta \wedge \tau \leq n\} = \{\eta \leq n\} \cup \{\tau \leq n\} \in A_n,$$

so that $\eta \wedge \tau$ is a stopping time. The maximum may be handled similarly. Suppose now that $\eta \leq \tau$ and let $A \in A_\eta$. If $n \geq 1$, then

$$A \cap \{\tau \leq n\} = A \cap \{\eta \leq n\} \cap \{\tau \leq n\} \in A_n,$$

since $A \cap \{\eta \leq n\} \in A_n$ and $\{\tau \leq n\} \in A_n$. \hfill \Box

The Stopped Process. Let $\tau$ be a stopping time with respect to $A_0, A_1, A_2, \cdots$ for which $\tau < \infty$ w.p.1., and let $Y_0, Y_1, Y_2 \cdots$ be adapted to the same sequence. Let

$$Y_\tau(\omega) = \begin{cases} Y_\tau(\omega) & \text{if } \tau(\omega) < \infty \\ \liminf_{n \to \infty} Y_n(\omega) & \text{if otherwise} \end{cases}$$

Lemma 3. If $\tau$ is a stopping time, then $Y_\tau$ is $A_\tau$-measurable.

Proof. If $c \in \mathbb{R}$, then

$$\{Y_\tau \leq c\} \cap \{\tau = n\} = \{Y_n \leq c\} \cap \{\tau = n\} \in A_n$$

for all $n$, so that $\{Y_\tau \leq c\} \in A_\tau$. \hfill \Box

Lemma 4. Let $\tau$ be a stopping time for which $\tau < \infty$ w.p.1. If

$$E|Y_\tau| < \infty,$$

$$\lim_{n \to \infty} \int_{\{\tau > n\}} |Y_n| dP = 0,$$

then

$$\lim_{n \to \infty} \int |Y_\tau - Y_{\tau \wedge n}| dP = 0.$$

Proof. \hfill \Box

$$\int |Y_\tau - Y_{\tau \wedge n}| dP \leq \int_{\{\tau > n\}} |Y_\tau| dP + \int_{\{\tau > n\}} |Y_n| dP,$$

which approaches zero under the assumptions of the lemma.
The Optional Stopping Theorem. Let $Y_0, Y_1, Y_2, \cdots$ be a (sub) martingale and $\eta$ and $\tau$ be stopping times with respect to the same filtration $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \cdots$. If

$$\eta \leq \tau < \infty \text{ w.p.1,}$$

$$E|Y_\eta| + E|Y_\tau| < \infty,$$

and

$$\lim_{n \to \infty} \int_{\{\tau > n\}} |Y_n| dP = 0,$$

Then

$$E(Y_\tau | \mathcal{A}_\eta) = (\geq) Y_\eta \text{ w.p.1.}$$

Proof. It suffices to do the case of a submartingale, and there is no loss of generality in supposing the $\eta \leq \tau$ everywhere, since $\eta$ and $\tau$ can be replaced by $\eta \wedge \tau$ and $\eta \vee \tau$. Then it suffices to show that

$$\int_A Y_\eta dP \leq \int_A Y_\tau dP$$  \hspace{1cm} (8)

for each $A \in \mathcal{A}_\eta$; and for this, it suffices to show that

$$\int_{A \cap \{\eta = m\}} Y_\eta dP \leq \int_{A \cap \{\eta = m\}} Y_\tau dP$$  \hspace{1cm} (9)

for each integer $m$. If $N \geq m$ and $\tau \geq m$, then

$$Y_\tau - Y_m = \sum_{n=m+1}^{\infty} (Y_n - Y_{n-1}) \cdot 1_{\{\tau \geq n\}}$$

$$= \sum_{n=m+1}^{N} (Y_n - Y_{n-1}) \cdot 1_{\{\tau \geq n\}} + (Y_\tau - Y_{\tau \wedge N}).$$

So, if $A \in \mathcal{A}_\eta$, then

$$\int_{A \cap \{\eta = m\}} (Y_\tau - Y_\eta) dP = \sum_{n=m+1}^{N} \int_{A \cap \{\eta = m, \tau \geq n\}} (Y_n - Y_{n-1}) dP$$

$$+ \int_{A \cap \{\eta = m\}} (Y_\tau - Y_{\tau \wedge N}) dP.$$

Here $A \cap \{\eta = m\} \cap \{\tau \geq n\} \in \mathcal{A}_{n-1}$ for each $n > m$, since $A \cap \{\eta = m\} \in \mathcal{A}_m$ and $\{\tau \geq n\} = \{\tau \leq n-1\}' \in \mathcal{A}_{n-1}$. Thus,

$$\sum_{n=m+1}^{N} \int_{A \cap \{\eta = m, \tau \geq n\}} (Y_n - Y_{n-1}) dP = \int_{A \cap \{\eta = m, \tau \geq n\}} E(Y_n - Y_{n-1} | \mathcal{A}_{n-1}) dP \geq 0$$

for each $N$; and

$$| \int_{A \cap \{\eta = m\}} (Y_\tau - Y_{\tau \wedge N}) dP | \leq \int_{A \cap \{\eta = m\}} |Y_\tau - Y_{\tau \wedge N}| dP,$$

which tends to 0 as $N \to \infty$ by Lemma 4. \(\diamond\)
Corollary. If \( \tau \) satisfies the conditions with \( \eta = 0 \), then \( E(Y_\tau) \geq E(Y_0) \), and there is equality in the martingale case.

Proof. This is clear.

Problem. Let \( Y_0, Y_1, Y_2, \ldots \) be a super martingale and \( \tau \) a stopping time with respect to a filtration \( \mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \ldots \). If there is a constant \( c \) for which \( Y_n \geq -c \) w.p.1 for all \( n \), then \( E(Y_\tau) \leq E(Y_0) \).

The Gambler’s Ruin Problem. Let \( X_1, X_2, \ldots \) be i.i.d. \( \pm 1 \) with probabilities \( p \) and \( q = 1 - p \); let \( S_n = X_1 + \cdots + X_n, \ n \geq 0; \) and let

\[
\tau = \inf \{ n \geq 1 : S_n = -a \text{ or } S_n = b \},
\]

where \( a \) and \( b \) are positive integers. The problem is to compute \( P\{S_n = b\} \), and the answer is

\[
P\{S_n = b\} = \frac{b}{a + b}
\]
or

\[
P\{S_n = b\} = \frac{(q/p)^a - 1}{(q/p)^{a+b} - 1}
\]

accordingly as \( p = 1/2 \) or \( p \neq 1/2 \). This formula will be derived for the case \( p \neq 1/2 \). In this case, it is clear from the Law of Large Numbers that \( \tau < \infty \) w.p.1. Let

\[
Y_n = \left(\frac{q}{p}\right)^S_n = \prod_{k=1}^n \left(\frac{q}{p}\right)^{X_k}
\]

for \( n \geq 0 \). Then \( Y_n \) is a martingale with respect to \( \mathcal{A}_n = \sigma\{X_1, \ldots, X_n\} \), because

\[
E\left(\left(\frac{q}{p}\right)^{X_k}\right) = p\left(\frac{q}{p}\right) + q\left(\frac{p}{q}\right) = 1
\]

for all \( k \). Moreover,

\[
E(Y_\tau) = \alpha\left(\frac{q}{p}\right)^b + (1 - \alpha)\left(\frac{q}{p}\right)^{-a} < \infty
\]

and

\[
\int_{\{\tau > n\}} Y_n dP \leq \left[ (\frac{q}{p})^b + (\frac{q}{p})^{-a} \right] P\{\tau > n\} \to 0
\]
as \( n \to \infty \). So,

\[
1 = E(Y_\tau) = \alpha\left(\frac{q}{p}\right)^b + (1 - \alpha)\left(\frac{q}{p}\right)^{-a}
\]

and, therefore,

\[
\alpha = \frac{1 - (q/p)^{-a}}{(q/p)^b - (q/p)^{-a}} = \frac{(q/p)^a - 1}{(q/p)^{a+b} - 1},
\]
as asserted. \( \diamond \)

Problem. When \( p = 1/2 \), show that \( P\{S_\tau = b\} = a/(a + b) \) and \( E(\tau) = ab \).

Random variables \( X_1, X_2, \ldots \) are said to be independently adapted to \( \mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \) if \( X_n \) is \( \mathcal{A}_n \)-measurable and if \( X_{n+1} \) and \( \mathcal{A}_n \) are independent for each \( n \).
Lemma 5. If \( X_1, X_2, \cdots \) are said to be independently adapted to \( \mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \), then \( X_1, X_2, \cdots \) are independent. Conversely, if \( X_1, X_2, \cdots \) are independent, then they are independently adapted to \( \sigma\{X_1, \cdots, X_n\}, n \geq 1 \).

Proof. Exercise

Theorem. Wald’s Lemma. Let \( X_1, X_2, \cdots \) be i.i.d. and independently adapted to a filtration \( \mathcal{A}_n, n \geq 0 \), and suppose that \( \mu = E(X_k) \) is finite. If \( \tau \) is any stopping time with respect to the same filtration for which \( E(\tau) < \infty \), then

\[
E(S_\tau) = \mu E(\tau).
\]

Proof. Suppose first that the random variables are non-negative. Then

\[
S_\tau = \sum_{k=1}^{\infty} 1_{\{\tau \geq k\}} X_k
\]

and, therefore,

\[
E(S_\tau) = \sum_{k=1}^{\infty} \int_{\{\tau \geq k\}} X_k dP = \sum_{k=1}^{\infty} P\{\tau \geq k\} E(X_k) = \mu E(\tau).
\]

The general case follows by applying the special case to \( X_k^+ \) and \( X_k^- \).

Problem: In the gambler’s ruin problem with \( p = 1/2 \), let \( \tau = \inf\{n \geq 1 : S_n \geq 1\} \). Show that \( \tau < \infty \) w.p.1 and that \( E(S_\tau) = 1 \). Conclude that \( E(\tau) = \infty \).