Martingales in finite probability space

Lecturer: Dr. Hong-Gwa Yeh
Department of Mathematics
National Central University
hgyeh@math.ncu.edu.tw
Finite Probability Space

\( \Omega \) = the set of all possible outcomes of a random experiment (\( \Omega \) is called a sample space, and \( \omega \in \Omega \) is called a sample point) note: we only consider \( |\Omega| < \infty \).

\( \mathcal{F} \subseteq 2^\Omega \) is called a \( \subseteq \)-algebra if

1. \( \Omega \in \mathcal{F} \)
2. \( A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F} \)
3. \( A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \)

Ex: Toss a coin 3 times. \( \Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\} \)

Let \( \mathcal{F}_0 = \{\emptyset, \Omega\} \),

\( \mathcal{F}_1 = \{\emptyset, \Omega, A_H, A_T\} \), where \( A_H = \{HHH, HHT, HTH, HTT\} \)
EX (continued)  After the random experiment, you are not told the outcome, but you are told, for every set in $\mathcal{F}$, whether or not the outcome is in that set. For example, you would be told that the outcome is not in $\emptyset$, and is in $\Omega$. Moreover, you might be told that outcome is not in $A_1$, but is in $A_T$. In effect, you have been told that the first toss was a $T$.

We interpret the $\sigma$-algebra $\mathcal{F}_1$ as a record of the "information of the first toss."
\((\Omega, \mathcal{F}, \mathbb{P})\)

**Probability measure:** Let \(\mathbb{P}\) be a \(\sigma\)-algebra on \(\Omega\).
\[\mathbb{P}: \mathcal{F} \rightarrow [0,1]\]

1. \(\mathbb{P}(\Omega) = 1\)
2. If \(A_1, A_2, \ldots\) is a sequence of disjoint sets in \(\mathcal{F}\), then \(\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)\)

**Probability space:** \((\Omega, \mathcal{F}, \mathbb{P})\) is called a probability space.

**Remark:** Here we only consider the case that \(|\Omega| < \infty\).
\[ \alpha(D) \]

**decomposition of** \( \Omega \): i.e. a collection \( \mathcal{D} = \{ D_1, D_2, \ldots, D_n \} \) of sets in \( \Omega \)

s.t. ① \( D_i \neq \emptyset \) for all \( i \) ② \( D_1 + D_2 + \cdots + D_n = \Omega \)

**Fact:** Let \( \mathcal{F} = \{ S : S \text{ is a union of sets in } \mathcal{D} \} \cup \{ \emptyset \} \).

Then \( \mathcal{F} \) is a \( \sigma \)-algebra, and is called the \( \sigma \)-algebra generated by \( \mathcal{D} \), and is denoted by \( \alpha(\mathcal{D}) \).

**Fact:** Let \( B \) be a \( \sigma \)-algebra of subsets of a finite space \( \Omega \).

Then \( \exists! \) decomposition \( \mathcal{D} \) of \( \Omega \) s.t. \( B = \alpha(\mathcal{D}) \)

**Hint:** see Shiryaev p.13. Let \( \mathcal{D} = \{ D \in B : D \neq \emptyset, D \cap B = D \text{ or } \emptyset \text{ for any } B \in B \} \).

This \( \mathcal{D} \) will meet our need.
**G-measurable**

**Note**: There is a one-to-one correspondence between σ-algebras and decomposition of a finite space Ω.

- Let $\mathcal{D} = \{D_1, \ldots, D_k\}$ be a decomposition of Ω. Let $Y = Y(\omega)$ be a function on Ω.

**Def**: $Y$ is said to be $\mathcal{D}$-measurable if $Y$ has the form

$$Y(\omega) = \sum_{i=1}^{k} y_i I_{D_i}(\omega)$$

i.e., $Y$ takes constant values on the blocks of $\mathcal{D}$.

**Remark**: In above def., we also said that $Y$ is $G$-measurable where $G$ is the σ-algebra $\sigma(\mathcal{D})$. 
Random Variable $X$, $\mathcal{D}_X$

**Def:** For a finite probability space $(\Omega, \mathcal{F}, P)$, we say that $X$ is a random variable on it if $X$ is an $\mathcal{F}$-measurable real-valued function defined on $\Omega$.

**Def:** Let $X$ be a r.v having the values $x_1, x_2, \ldots, x_k$ with positive probabilities i.e. $X = \sum_{i=1}^{k} x_i I_{D_i}(w)$, where $D_i = \{ \omega \in \Omega : X(\omega) = x_i \}$

We define the decomposition $\mathcal{D}_X = \{ D_1, D_2, \ldots, D_k \}$.

If $X_1, X_2, \ldots X_m$ are r.vs then decomposition $\mathcal{D}_{X_1, X_2, \ldots, X_m}$ is defined in the same way.

**Note:** $\alpha(\mathcal{D}_X)$ is the smallest $\sigma$-algebra over which $X$ is measurable.
\( P(A|\mathcal{D}) \)

- \((\Omega, \mathcal{F}, P)\) is a finite prob. space, \( A \in \mathcal{F}, \ \{D_i \in \mathcal{F} \ \forall i \)
- \( \mathcal{D} = \{ D_1, \ldots, D_k \} \) is a decomposition of \( \Omega \) with \( P(D_i) > 0 \ \forall i \).

**Def:** Define the \( P(A|\mathcal{D}) : \Omega \rightarrow \mathbb{R} \) as follows

\[
P(A|\mathcal{D})(\omega) = \sum_{i=1}^{k} P(A|D_i) I_{D_i}(\omega)
\]

**Facts:**

1. \( ANB = \emptyset \Rightarrow P(A \cup B|\mathcal{D}) = P(A|\mathcal{D}) + P(B|\mathcal{D}) \)
2. \( P(A|\mathcal{D}(\Omega)) = P(A) \) constant rv
3. \( \mathbb{E} P(A|\mathcal{D}) = P(A) \)
4. \( P(A|\mathcal{D}) \) is \( \mathcal{D} \)-measurable and hence is \( \mathcal{F} \)-measurable

i.e. \( P(A|\mathcal{D}) \) is a rv.
\[ E(X|D) \]

**Def:** Define the function \( E(X|D) : \Omega \to \mathbb{R} \) as follows:

\[
E(X|D)(\omega) = \sum_{i=1}^{n} x_i P(X=x_i|D)(\omega)
\]

**Fact:** (2) \( E(X|D) \) is \( D \)-measurable and hence \( \mathcal{F} \)-measurable, i.e., \( E(X|D) \) is a rv.

\[
E(X|D) = \sum_{j=1}^{k} \left( \sum_{i=1}^{n} x_i P(X=x_i|D_j) \right) I_{D_j}
\]  

\[ = \sum_{j=1}^{k} E(X|D_j) I_{D_j}. \]  

the average value of \( X \) on the block \( D_j \)

\[
E\left( E(X|D) \right) = EX
\]
Remark: If $\mathcal{F} = \sigma(\mathcal{C})$ then

$\mathcal{P}(A1\mathcal{F})$ is also denoted by $\mathcal{P}(A1\mathcal{I})$,

and $\mathcal{E}(A1\mathcal{F})$ is also denoted by $\mathcal{E}(A1\mathcal{I})$.

Warning: we only use this notation for finite probability space.
**Tower Property**

**Thm** If two $\sigma$-algebras $\mathcal{F}, \mathcal{G}$ have $\mathcal{F} \subseteq \mathcal{G}$ then
\[ E (E (X | \mathcal{G}) | \mathcal{F}) = E (X | \mathcal{F}) \]

**Fact (1)** If $X$ is $\mathcal{D}$-measurable for a decomposition $\mathcal{D}$ of $\mathcal{S}$, then
\[ E (X | \mathcal{D}) = X \quad \text{and} \quad E (XY | \mathcal{D}) = X E (Y | \mathcal{D}) \]

"taking out what is known"

**Fact (2)** If $X$ is independent of decomposition $\mathcal{D}$ (i.e. $\forall D_i \in \mathcal{D}$, $X$ and $I_{D_i}$ are independent) then
\[ E (X | \mathcal{D}) = EX \]

**bt of (2):** Say $\mathcal{D} = \{ D_1, \ldots, D_k \}$. Then
\[ E (X | \mathcal{D}) = \sum_{i=1}^k E (X | D_i) I_{D_i} = \sum_{i=1}^k EX I_{D_i} = EX. \]
\[ P(A|X_1, X_2, ..., X_M) \]

- \((\Omega, \mathcal{F}, \mathbb{P})\) is a finite prob space, \(A \in \mathcal{F}\)
- \(\forall X: \Omega \rightarrow \{x_i, ..., x_k\} \text{ with } P(X=x_i) > 0 \ \forall i\)

\textbf{Def:} \[ P(A|X) \overset{\text{def}}{=} P(A|D_X) \]

\[ P(A|X_1, X_2, ..., X_m) \overset{\text{def}}{=} P(A|D_{X_1, X_2, ..., X_m}) \]

\textbf{Facts:}

1. \[ P(A|X)(\omega) = \sum_{i=1}^{k} P(A|X=x_i) I_{\{X=x_i\}}(\omega) \]
2. \[ P(A|X_1, X_2, ..., X_m)(\omega) \]
   \[ = \sum_{y_1, ..., y_m} P(A|X_1=x_1, ..., X_m=y_m) I_{\{X_1=x_1, ..., X_m=y_m\}}(\omega) \]
\[ E(X|Y_1, Y_2, \ldots, Y_M) \]

**Def:** In a finite probability space, we define a random variable \( E(X|Y) \) as follows:

\[ E(X|Y) \stackrel{\text{def}}{=} E(X|\mathcal{D}_Y) \]

\[ E(X|Y_1, Y_2, \ldots, Y_m) \stackrel{\text{def}}{=} E(X|\mathcal{D}_{Y_1, Y_2, \ldots, Y_m}) \]

**Fact:** The random variable \( E(X|Y) \) is the random variable \( f(Y) \) such that \( f(y) = E(X|Y=y) \).

**Pf:** Let \( Y: \Omega \rightarrow \{y_1, \ldots, y_k\} \) with \( P(Y=y_i) > 0 \) \( \forall i \).

\[ E(X|Y) = E(X|\mathcal{D}_Y) = \sum_{j=1}^{k} E(X|D_j) I_{D_j}, \text{ where } D_j = \{Y=y_j\} \]

**Remark:** We can generalize the above fact to the random variable \( E(X|Y_1, \ldots, Y_m) \).
Example for $\mathbb{E}(X_1 | Y_1, Y_2, \ldots, Y_m)$

**Example** Consider independent throws of an unbiased 6-sided die. For $1 \leq i \leq 6$, let $X_i$ denote the number of times the value $i$ appears in $n$ throws of the die. Then

$\mathbb{E}(X_1 | X_2) = \frac{n - X_2}{5}, \quad \mathbb{E}(X_1 | X_2, X_3) = \frac{n - X_2 - X_3}{4}.$

**Proof:**

$\mathbb{E}(X_1 + X_2 + \ldots + X_6 | X_2 = \alpha, X_3 = \beta) = n$

$\Rightarrow 4 \mathbb{E}(X_1 | X_2 = \alpha, X_3 = \beta) = n - \alpha - \beta \Rightarrow \mathbb{E}(X_1 | X_2 = \alpha, X_3 = \beta) = \frac{n - \alpha - \beta}{4}.$

$\Rightarrow \mathbb{E}(X_1 | X_2, X_3) = \frac{n - X_2 - X_3}{4}.$

QED
A filter in finite prob. spaces

- Given a finite probability space \((\Omega, 2^\omega, P)\).

**Def:** A filter is a nested sequence of \(\sigma\)-algebras in \(-2\)

\[\emptyset \subseteq \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n = 2^\omega\]

**Remark:** In above definition, if we have decompositions \(\mathcal{D}_0 \cdots \mathcal{D}_n\)

s.t. \(\mathcal{D}_0 = \{\Omega\}\), \(\mathcal{D}_i(\mathcal{D}_i) = \mathcal{F}_i\), \(\forall i\), and \(\mathcal{D}_n = \{\{\omega\} : \omega \in \Omega\}\)

Then sometimes we write the above filter as

\[\{\Omega\} = \mathcal{D}_0 \subseteq \mathcal{D}_1 \subseteq \cdots \subseteq \mathcal{D}_n = \{\{\omega\} : \omega \in \Omega\}\]
Martingale (I) general setting

**Def:** Given a finite r.s. \((\Omega, 2^\Omega, P)\) with a filter \(\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n = 2^\Omega\), a sequence of r.v.s \(X_0, X_1, \ldots, X_n\) is called a **martingale** w.r.t. the filter \(\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_n\) if \(\mathbb{E}(X_k | \mathcal{F}_k) = X_k\) for each \(k = 0, 1, 2, \ldots, n-1\).

**Notation:** Sometimes we use \((X_k, \mathcal{F}_k)_{k=0}^n\) to denote the above martingale.

**Fact:** (1) \(X_k\) is \(\mathcal{F}_k\)-measurable, \(k = 0, 1, 2, \ldots, n\).
(2) \(\mathbb{E}X_k = \mathbb{E}X_0\) for each \(k = 1, 2, \ldots, n\).
**Martingale (II) special case**

**Def:** A sequence of rvs $X_0, X_1, \ldots, X_m$ is called a martingale if for $0 \leq i \leq m$, $\mathbb{E}(X_{i+1} | X_0, X_1, \ldots, X_i) = X_i$.

**Recall:** $\mathbb{E}(X_{i+1} | X_0, X_1, \ldots, X_i) \overset{\text{def}}{=} \mathbb{E}(X_{i+1} | \mathcal{D}_{X_0, X_1, \ldots, X_i})$

**Note:** In the above definition, $X_0, X_1, \ldots, X_m$ is a martingale w.r.t. the filter $\alpha(\mathcal{D}_{X_0}), \alpha(\mathcal{D}_{X_0, X_1}), \alpha(\mathcal{D}_{X_0, X_1, X_2}), \ldots, \alpha(\mathcal{D}_{X_0, X_1, \ldots, X_m}), \mathbb{P}$ in the general setting.
Recall

Suppose $\mathcal{Z}_k = \alpha(\{D_{k1}, D_{k2}, \ldots, D_{kt}\})$. Then $E(X_{k+1} \mid \mathcal{Z}_k) = X_k$ implies

$$X_k = \sum_{i=1}^{t} E(X_{k+1} \mid D_{ki}) I_{D_{ki}}$$
An illustration of a martingale

- We assume that $\mathcal{I}_0 = \mathcal{F}_0$, $\mathcal{I}(\mathcal{D}_1) = \mathcal{F}_1$, $\mathcal{I}(\mathcal{D}_2) = \mathcal{F}_2$.
- The values of $X_i$ are indicated by the red lines.

\begin{align*}
X_0 &= \mathbb{E}X_1 I_\Omega \\
X_1 &= \mathbb{E}(X_2 | D_1) I_{D_1} + \mathbb{E}(X_2 | D_2) I_{D_2} \\
X_2 &= \mathbb{E}(X_3 | D_1) I_{D_1} + \mathbb{E}(X_3 | D_2) I_{D_2} + \mathbb{E}(X_3 | D_3) I_{D_3} \\
&\quad + \mathbb{E}(X_3 | D_\alpha) I_{D_\alpha} + \mathbb{E}(X_3 | D_\beta) I_{D_\beta}
\end{align*}

Note $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2$. 
Doob Martingales

- Let \((\Omega, 2^\Omega, P)\) be a finite ps. with a filter \(\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n\).

**Thm** Let \(X\) be a rv on \((\Omega, 2^\Omega, P)\). Define \(X_i = \mathbb{E}(X_1 | \mathcal{F}_i)\) for \(i = 0, 1, 2, \ldots, n\). Then \((X_i, \mathcal{F}_i)_{i=0}^n\) is a martingale.

**Prf:** For \(k = 0, 1, 2, \ldots, n-1\), \(\mathbb{E}(X_{k+1} | \mathcal{F}_k) = \mathbb{E}(\mathbb{E}(X_1 | \mathcal{F}_{km+1}) | \mathcal{F}_k) = \mathbb{E}(X_1 | \mathcal{F}_k)\) by Tower Thm.

**Ex:** Toss a fair coin three times. Let \(X_i = 1\) if \(i\)th toss is head, \(X_i = 0\) otherwise, \(i = 1, 2, 3\). Let \(f(X_1, X_2, X_3) = \sum_{i=1}^3 X_i\).

The Doob process. \(Y_0 = \mathbb{E}(f(X) | \mathcal{F}_0) = \mathbb{E}f(X) = \frac{3}{2}\), where \(X = (X_1, X_2, X_3)\)

\(Y_1 = \mathbb{E}(f(X) | \mathcal{D}_1) = \mathbb{E}(\sum_{i=1}^3 X_i | X_1) = X_1 + 1\)

\(Y_2 = \mathbb{E}(f(X) | \mathcal{D}_2) = \mathbb{E}(f(X) | X_1, X_2) = X_1 + X_2 + \frac{1}{2}\)

\(Y_3 = \mathbb{E}(f(X) | \mathcal{D}_3) = X_1 + X_2 + X_3\)

\(\{\Omega\} = \mathcal{D}_0 \subseteq \mathcal{D}_1 \subseteq \mathcal{D}_2 \subseteq \mathcal{D}_3 = 2^\Omega, \text{ where }\)

\(\mathcal{D}_0 = \{ \{X_1=1\}, \{X_1=0\} \}, \mathcal{D}_1 = \{ \{X_1=1, X_2=0\}\{X_1=1, X_2=1\}, \{X_1=0, X_2=1\}\}, \mathcal{D}_2 = \{ \{X_1=1, X_2=0\}\{X_1=1, X_2=1\}, \{X_1=0, X_2=1\}\}, \mathcal{D}_3 = \{ \{X_1=1, X_2=0\}\{X_1=1, X_2=1\}, \{X_1=0, X_2=0\}\}\)
**Edge Exposure Martingale**

- Consider random graph space $G_{n,p}$. Label the $\binom{n}{2}$ possible edges with the sequence $1, 2, 3, \ldots, m$.

Define the rvs $I_j(\omega) = \begin{cases} 1 & \text{if edge } j \text{ appears in } \omega \\ 0 & \text{o.w.} \end{cases}$

Consider any real-valued function $F$ over $G_{n,p}$, e.g., the clique number.

The **edge exposure martingale** if defined to be the sequence of rvs $X_0, X_1, X_2, \ldots, X_m$ s.t.

- $X_0 = \mathbb{E}(F)$
- $X_1 = \mathbb{E}(F | I_.)$
- $\vdots$
- $X_{m-1} = \mathbb{E}(F | I_1, I_2, \ldots, I_{m-1})$
- $X_m = F(G_{n,p})$

**Note:** $X_0, X_1, \ldots, X_m$ is a Doob martingale.
**Vertex Exposure Martingale**

- In the same setting as in the edge exposure martingale.

Let $I_{xy}(w) = \begin{cases} 1 & \text{if edge } xy \text{ appears in } w \\ 0 & \text{o.w.} \end{cases}$

The vertex exposure martingale is defined to be the sequence of rvs $Y_1, Y_2, \ldots, Y_n$ s.t.

$Y_1 = \mathcal{E}(F)$

$Y_2 = \mathcal{E}(F \mid I_{xy}, \{x,y\} \in \binom{[2]}{2})$

$Y_3 = \mathcal{E}(F \mid I_{xy}, \{x,y\} \in \binom{[3]}{2})$

$\vdots$

$Y_n = \mathcal{E}(F \mid I_{xy}, \{x,y\} \in \binom{[n-1]}{2})$

$Y_n = \mathcal{F}(G_{n,p})$

**Note** $Y_1, Y_2, \ldots, Y_n$ is a Doob Martingale.

By ordering the edge appropriately, the vertex exposure martingale is a subsequence of the edge exposure martingale.
References