ECE511: Analysis of Random Signals	Fall 2015

Martingales

Dr. Salim El Rouayheb — *Scribe: Serge Kas Hanna*

1 Motivation: Fair Games

Example 1. *Suppose you play the following series of games. In game i, $i = 1, 2, \ldots,$ you bet \$1 and flip a fair coin. If the outcome is heads you win \$1, and if the outcome is tails you lose \$1. Let:*

$$X_i : \text{amount of money you win or lose } (X_i < 0) \text{ in bet } i, \text{and } Z_i = \sum_{i=1}^{n} X_i.$$

We say the game is fair since $E\left[X_i\right] = 0$ for all i.

Claim 1. $E\left[Z_n | X_1, \ldots, X_{n-1}\right] = Z_{n-1}$.

Proof.

$$\begin{aligned}
E\left[Z_n | X_1, \ldots, X_{n-1}\right] &= E\left[X_1 + \cdots + X_n | X_1, \ldots, X_{n-1}\right] \ , \\
&= E\left[X_1 | X_1, \ldots, X_{n-1}\right] + \cdots + E\left[X_{n-1} | X_1, \ldots, X_{n-1}\right] + \underbrace{E\left[X_n | X_1, \ldots, X_{n-1}\right]}_{=E[X_n]=0} \ , \\
&= X_1 + \cdots + X_{n-1} + 0 \ , \\
&= Z_{n-1}.
\end{aligned}$$

□

The property $E\left[Z_n | X_1, \ldots, X_{n-1}\right] = Z_{n-1}$ is a crucial property of the sequence $Z_0, Z_1, \ldots,$ that we will use later to define a martingale.

The betting amount in a fair game for each game does not have to be fixed. It can:

1. Depend on which game is played. For example, bet \$$2^i$ in game i.
2. Depend on the outcome of games in the past (betting strategy has memory). For example, bet \$$2^i$ if the previous bet was won, and bet \$$i$ if the previous game was lost.

Note: Regardless of the gambling strategy used, the game is fair as long as the coin used is fair.

2 Martingales

Definition 1. *A sequence of random variables $Z_0, Z_1, \ldots$ is a martingale with respect to the sequence $X_0, X_1, \ldots,$ if for all $n \geq 0$ all the following hold:*

1. $Z_n = f(X_0, X_1, \ldots, X_n)$*, i.e. f is a deterministic function of $X_0, \ldots, X_n$.*
2. $E[|Z_n|] < \infty$.
3. $E[Z_{n+1}|X_0, X_1, \ldots, X_n] = Z_n$.

Definition 2. *A sequence of random variables $Z_0, Z_1, \ldots$ is a martingale when it is a martingale with respect to itself, that is:*

1. $E[|Z_n|] < \infty$.
2. $E[Z_{n+1}|Z_0, Z_1, \ldots, Z_n] = Z_n$.

Lemma 1. *If $Z_0, Z_1, \ldots, Z_n$ is a martingale with respect to $X_0, X_1, \ldots, X_n$, then $E[Z_n] = E[Z_0]$.*

Proof. Since Z_i defines a martingale:

$$\begin{aligned} Z_i &= E[Z_{i+1}|X_0, X_1, \ldots, X_i]. \\ E[Z_i] &= E[E[Z_{i+1}|X_0, X_1, \ldots, X_i]] \ , \\ &= E[Z_{i+1}]. \end{aligned}$$

□

3 Stopping Time

Question: Is there a gambling strategy that can guarantee that one always wins on average in a fair game?

Example 2. *Possible Stopping Strategies*

1. *Stop after k games: Intuitively, if the game is fair then there is no such strategy that guarantees winning on average. In fact $E[Z_k] = E[Z_0]$.*
2. *Stop once you collect \$100: The problem with such a strategy is that you have to wait an infinite amount of time (and you may run out of money).*

Definition 3. *A non-negative, integer random variable T is a stopping time for the sequence $Z_0, Z_1, \ldots$ if the event "$T = n$" depends only on the value of random variables $Z_0, Z_1, \ldots, Z_n$.*

Intuition: T corresponds to a strategy for determining when to stop a sequence based only on values seen so far. In the gambling game:

1. First time I win 10 games in a row: is a stopping time.
2. The last time when I win: is not a stopping time (depends on the future).

Theorem 1. *(Martingale Stopping Theorem)*
If $Z_0, Z_1, \ldots$ is a martingale with respect to $X_1, X_2, \ldots$ and if T is a stopping time for $X_1, X_2, \ldots$ then:

$$E\left[Z_T\right] = E\left[Z_0\right],$$

whenever one *of the following holds:*

1. *There is a constant c such that, almost surely, $|Z_n| \leq c$ for all $n \leq T$.*
2. *T is bounded, almost surely.*
3. *$E\left[T\right] < \infty$, and there is a constant c such that $E\left[|Z_{i+1} - Z_i| \mid X_1, \ldots, X_i\right] < c$.*

Example 3. *The Gambler's Ruin*
The game consists of flipping a fair coin repeatedly, every time the outcome is heads the gambler wins \$1 and if the outcome is tails the gambler loses \$1. The gambler starts the game with \$$k$ and stops playing when he loses his all his money or when he reaches a total of \$$n$, where $n > k$. This problem can be seen as random as walk as shown in Figure 2, the gambler stops if he hits 0 or n.

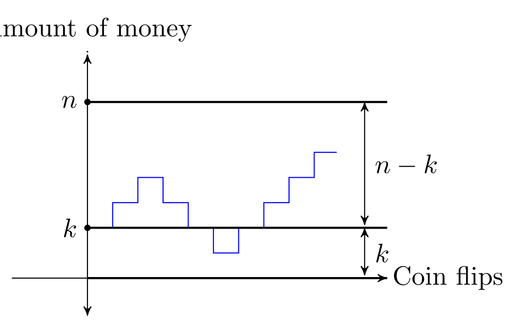

Figure 1: Gambler's Ruin Random Walk

Question: *What is the probability of winning?*
Answer: *Intuitively the probability of winning is proportional to k, $Pr(win) = \dfrac{k}{n}$.*

Now we calculate $Pr(win)$ again, this time using Theorem 2 (Martingale Stopping Theorem):
Let X_i be the amount of money the gambler gains after playing the i^{th} round, such that:

$$Pr(X_i = +\$1) = \underbrace{Pr(heads) = Pr(tails)}_{fair\ coin} = Pr(X_i = -\$1) = \frac{1}{2}.$$

Let $Z_0 = k$ *and* $Z_i = \sum_{k=1}^{i} X_k$.

$$\begin{aligned} E\left[Z_{i+1}|X_1,\dots,X_i\right] &= E\left[X_1+\cdots+X_{i+1}|X_1,\dots,X_i\right], \\ &= E\left[Z_i+X_{i+1}|X_1,\dots,X_i\right], \\ &= Z_i + \cancelto{0}{E\left[X_{i+1}\right]}, \\ &= Z_i. \end{aligned}$$

Therefore, $Z_0, Z_1, \dots,$ *is a martingale with respect to* $X_1, X_2, \dots$. *Let* T *be the time when the gambler hits* 0 *or* n, T *is a stopping time for* $X_1, \dots, X_n$. *Also* $|Z_i|$ *for* $i = 1, \dots, n$ *are bounded by* n. *By Martingale Stopping Theorem (theorem 2):*

$$\begin{aligned} E\left[Z_T\right] = E\left[Z_0\right] &= k. \\ n \times Pr(win) + 0 \times Pr(lose) &= k \\ Pr(win) &= \frac{k}{n}. \end{aligned}$$

Example 4. *A Ballot Theorem*
Two candidates, Hillary and Trump, run for election. Hillary gets a votes and Trump gets b votes, such that $a > b$. *Votes are counted in random order: chosen from all permutations on* $n = a + b$ *votes. What is the probability that Hillary is always ahead in the count?*
Let S_i *be the number of votes Hillary is leading by after i votes (if Hillary is trailing:* $S_i < 0$*),* $S_n = a - b$. *The cases we're interested in are the ones similar to example shown in table 1, where Hillary is always in the lead (*$S_i > 0$*).*

Hillary	Trump	S_i
1	0	1
2	0	2
3	0	3
4	0	4
4	1	3
4	2	2

Table 1: Count Example.

Let p be the probability that Hillary is always leading. Let us start by a toy example with $a = 2$ *and* $b = 1$ *in table 2 and will be proven later using martingales.*

Hillary	Trump	S_i
1	0	1
2	0	2
2	1	1

(a) case 1

Hillary	Trump	S_i
0	1	-1
1	1	0
2	1	1

(b) case 2

Hillary	Trump	S_i
1	0	1
1	1	0
2	1	1

(c) case 3

Table 2: $a = 2$, $b = 1$ Example.

Out of the 3 possible cases shown in table 2, case 1 is the only case where Hillary is always in the lead. Therefore:

$$p = \frac{1}{3}.$$

Now we determine p using martingales. Define:

$$X_k = \frac{S_{n-k}}{n-k},\ 0 \leq k \leq n-1.$$

Claim 2. *$X_0, X_1, \ldots, X_n$ is a martingale.*

Let us assume that the claim is true for now. Let:

$$T = \begin{cases} min\{k : X_k = 0\} & \text{if such } k \text{ exists,} \\ n-1 & \text{otherwise.} \end{cases}.$$

The first case (case 1, $T = min\{k : X_k = 0\}$), is the first time that the score of Hillary is equal to the score of Trump. This case happening means that Hillary does not lead throughout the whole count. The second case (case 2, $T = n-1$) is when $n = a + b$ is reached and Hillary was always in the lead.
T is bounded and depends only on all the votes up to T (past). Therefore, T is a stopping time and by Theorem 2 (Martingale Stopping Theorem):

$$E[X_T] = E[X_0] = \frac{E[S_n]}{n} = \frac{a-b}{a+b}.$$

Case 1: For some $k : S_k = 0$, then $X_k = 0$. $T = k < n-1$ and $X_T = 0$.
Case 2: For $0 \leq k \leq n-1 : S_{n-k} > 0$, then $X_k > 0$. $T = n-1$ and $X_T = X_{n-1} = S_1 = 1$ (definitely leads after first vote).

$$E[X_T] = 0 \times Pr(case\ 1) + 1 \times Pr(case\ 2) = \frac{a-b}{a+b}.$$

Therefore,

$$Pr(case\ 2) = p = \frac{a-b}{a+b}.$$

Now let's prove Claim 2, $E\left[X_{i+1}|X_0,\ldots,X_i\right] \stackrel{?}{=} X_i$.
Conditioning on $X_0, X_1, \ldots, X_{k-1}$ *is the same as conditioning on* $S_n, S_{n-1}, \ldots, S_{n-k+1}$.
Let a_i *be the number of votes for Hillary after first* i *votes are counted.*

$$S_{n-k} = \begin{cases} S_{n-k+1}+1 & \text{if } n-k+1th \text{ vote is for Trump,} \\ S_{n-k+1}-1 & \text{if } n-k+1th \text{ vote is for Hillary.} \end{cases}$$

$$S_{n-k} = \begin{cases} S_{n-k+1}+1 & \text{with prob. } \dfrac{n-k+1-a_{n-k+1}}{n-k+1}, \\ S_{n-k+1}-1 & \text{with prob. } \dfrac{a_{n-k+1}}{n-k+1}. \end{cases}$$

$$E\left[S_{n-k}|S_{n-k+1}\right] = \left(S_{n-k+1}+1\right)\frac{n-k+1-a_{n-k+1}}{n-k+1} + \left(S_{n-k+1}-1\right)\frac{a_{n-k+1}}{n-k+1}. \tag{1}$$

$$\begin{aligned} \text{number of votes for Hillary} + \text{number of votes for Trump} &= n-k+1, \\ a_{n-k+1} + \left(a_{n-k+1} - S_{n-k+1}\right) &= n-k+1, \\ a_{n-k+1} &= \frac{n-k+1+S_{n-k+1}}{2}. \end{aligned} \tag{2}$$

Replacing (2) *in* (1) *we get:*

$$E\left[S_{n-k}|S_{n-k+1}\right] = S_{n-k+1}\frac{n-k}{n-k+1}. \tag{3}$$

$$\begin{aligned} E\left[X_k|X_0, X_1, \ldots, X_{k-1}\right] &= E\left[\frac{Sn-k}{n-k}|S_n, \ldots, S_{n-k+1}\right], \\ &= \frac{S_{n-k+1}}{n-k+1}, \\ &= X_{k-1}. \end{aligned}$$

$$\Rightarrow X_0, X_1, \ldots, X_n \text{ is a martingale, which proves Claim 2.}$$

4 Doob Martingale

Theorem 2. *(Martingale Convergence Theorem)*
If $Z_0, Z_1, \ldots$ *is a martingale and* $E\left[Z_n^2\right] < c$ *for all* $n \geq 0$ *and for some constant* c*, then* $Z_n \xrightarrow{a.s.} Z$ *when* $n \to \infty$.

Proof. Refer to textbook. □

Let $X_1, X_2, \ldots, X_n$ be a sequence of random variables.

Definition 4 (Doob Martingale)**.**
Let $Y = f(X_1, \ldots, X_n)$ be a random variable with $E[|Y|] < \infty$ that is a function of $X_1, \ldots, X_n$. Let $Z_0 = E[Y]$, and $Z_i = E[Y|X_1, X_2, \ldots, X_i], i = 0, 1, \ldots, n$.

Lemma 2. *The sequence $Z_0, Z_1, \ldots, Z_n$ defined above is a martingale with respect to $X_0, X_1, \ldots, X_n$, called Doob martingale.*

Proof. The proof uses the fact that $E[E[V|U,W]|W] = E[V|W]$.

$$Z_i = E[Y|X_0, X_1, \ldots, X_i], i = 0, 1, \ldots, n.$$

$$\begin{aligned} E[Z_{i+1}|X_0, X_1, \ldots, X_i] &= E[E[Y|X_0, X_1, \ldots, X_{i+1}]|X_0, X_1, \ldots, X_i] \ , \\ &= E[Y|X_0, X_1, \ldots, X_i] \ , \\ &= Z_i. \end{aligned}$$

□

Example 5. *(Edge Exposure Martingale)*
Consider the random graph $G_{n,p}$, i.e. a graph with n vertices such that the probability that two vertices are connected is p. Consider the $m = \binom{n}{2}$ possible edges in arbitrary order such that:

$$X_i = \begin{cases} 1 & \textit{if edge } i \textit{ is present,} \\ 0 & \textit{otherwise.} \end{cases} .$$

The clique number (maximum number of vertices that are pairwise connected) is denoted by $w(G)$. We can think of $w(G)$ as a function of the edges, i.e. $w(G) = f(X_1, X_2, \ldots, X_m)$. The Doob martingale is defined as follows:
$Z_0 = E[w(G)]$.
$Z_i = E[w(G)|X_1, X_2, \ldots, X_i], i = 1, \ldots, m$.
$Z_0, Z_1, \ldots, Z_m$ *is a Doob martingale.*

Let us take the example of a random graph on $n = 3$ vertices ($G_{3,\frac{1}{2}}$). The complete graph of G is shown in Figure 1, where $X_1 = X_2 = X_3 = 1$ and $w(G) = 3$.

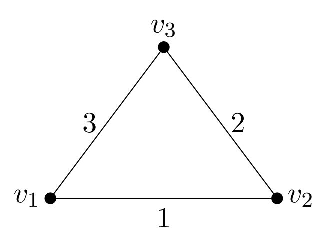

Figure 2: Complete graph for $n = 3$.

Now for $G_{3,\frac{1}{2}}$ we determine Z_0 and Z_1, the first two terms of the sequence, and verify the martingale property. The other terms of the sequence can also be verified to satisfy the martingale property by the same procedure.

1. $Z_0 = E[w(G)]$ *and* $w(G) = \{1, 2, 3\}$.

(a) $Pr(w(G)=1) = Pr(G \textit{ has no edges}) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$.

(b) $Pr(w(G)=2) = Pr(G \textit{ has 1 or 2 edges}) = 3\left(\frac{1}{2}\right)^3 + 3\left(\frac{1}{2}\right)^3 = \frac{3}{4}$.

(c) $Pr(w(G)=3) = Pr(G \textit{ has 3 edges}) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$. *(This is the case of Figure 1).*

$Z_0 = E\left[w(G)\right] = 1 \times \frac{1}{8} + 2 \times \frac{3}{4} + 3 \times \frac{1}{8} = 2.$

2. $Z_1 = E\left[w(G)|X_1\right]$ *which is a random variable.*

   (a) *If* $X_1 = 1$ *the possible values of* $w(G)$ *are* 2 *and* 3.

   $Pr\left(w(G)=3|X_1=1\right) = Pr(\textit{edges 2 and 3 exist}) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$.

   $Pr\left(w(G)=2|X_1=1\right) = Pr(\textit{at most one edge other than edge 1 exits}) = 3\left(\frac{1}{2}\right)^2 = \frac{3}{4}$.

   $E\left[w(G)|X_1=1\right] = 3 \times \frac{1}{4} + 2 \times \frac{3}{4} = \frac{9}{4}$.

   (b) *If* $X_1 = 0$, *the possible values of* $w(G)$ *are* 1 *and* 2.

   $Pr\left(w(G)=2|X_1=0\right) = Pr(\textit{at least one of the edges 2 and 3 exist}) = 3\left(\frac{1}{2}\right)^2 = \frac{3}{4}$.

   $Pr\left(w(G)=1|X_1=0\right) = Pr(\textit{edges 2 and 3 don't exist}) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$.

   $E\left[w(G)|X_1=0\right] = 2 \times \frac{3}{4} + 1 \times \frac{1}{4} = \frac{7}{4}$.

$Z_1 \in \left\{\frac{7}{4}, \frac{9}{4}\right\}$ *such that* $Pr\left(Z_1 = \frac{9}{4}\right) = Pr(X_1=1) = Pr(X_1=0) = Pr\left(Z_1 = \frac{7}{4}\right) = \frac{1}{2}$.

Now having determined Z_0 *and* Z_1, *we verify that* $E\left[Z_1\right]$ *(which can be thought of as equal to* $E\left[Z_1|X_0\right]$ *for some constant* X_0*) is equal to* Z_0. *In fact:*

$$\begin{aligned} E\left[Z_1\right] &= \frac{1}{2} \times \frac{7}{4} + \frac{1}{2} \times \frac{9}{4}, \\ &= 2 \\ &= Z_0. \end{aligned}$$

5 Azuma-Hoeffding Inequality

Motivation: If the martingale is stopped at Z_k for a fixed k ($E[Z_k] = E[Z_0]$), how far can Z_k be from $E[Z_k]$?

Example 6.
Let $X_1, \ldots, X_n$ *be a sequence of iid random variables such that* $Pr(X_i = +1) = Pr(X_i = -1) = \frac{1}{2}$. *Let* $Z_0 = 0$ *and* $Z_i = \sum_{k=1}^{i} X_k$. $Z_0, Z_1, \ldots$ *is a martingale.*

$$E[Z_i] = 0 = E[X_1] + \cdots + E[X_i] = E[Z_0] = 0 \textit{ for all } i = 1, 2, \ldots.$$

The CLT,

$$P\left(|Z_i| \geq a\right) \xrightarrow{i \to \infty} 2\int_a^{+\infty} \frac{e^{-\frac{x^2}{2i}}}{\sqrt{2\pi i}} dx \leq 2e^{-\frac{a^2}{2i}} \text{ (Chernoff Bound).}$$

$\Rightarrow$ *There is a high concentration around the average. And the probability that Z_i is far from the average decreases exponentially.*

The next theorem proves that the same phenomenon happens even if the increments $X_i' s$ are not independent (so we can't apply the CLT), as long as their partial sums form a Martingale,

Theorem 3 (Azuma-Hoeffding Inequality)**.**
Let $Z_0, Z_1, \ldots, Z_n$ be a martingale with respect to $X_1, X_2, \ldots$ such that $|Z_k - Z_{k-1}| \leq c_k$ for given constants c_k. Then, for all $t \geq 0$ and any $\lambda > 0$:

$$Pr\left(|Z_t - Z_0| \geq \lambda\right) \leq 2e^{-\frac{\lambda^2}{2\sum_{k=1}^{t} c_k^2}}.$$

Corollary 1. *Let $X_0, X_1, \ldots$ be a martingale such that for all $k \geq 1$,*

$$|X_k - X_{k-1}| \leq c.$$

Then, for all $t \geq 1$ and $\lambda > 0$,

$$Pr\left(|X_t - X_0| \geq \lambda c\sqrt{t}\right) \leq 2e^{-\frac{\lambda^2}{2}}.$$

Example 7. *Pattern Matching*
Given a long string A and a short pattern B, we ask the following questions:

1. *Does this pattern appear more often than is expected in a random string?*
2. *Is the number of occurrences of the pattern concentrated around the expectation?*

Let $A = (a_1, a_2, \ldots, a_n)$ be a string of characters, each chosen independently and uniformly at random from Σ, with $|\Sigma| = m$ (size of the alphabet), and $B = (b_1, \ldots, b_k)$ a fixed string representing the pattern such that $k < n$ and $b_i \in \Sigma$. Let F be the number of occurrences of B in random string A.

$$E[F] = (n - k + 1)\left(\frac{1}{m}\right)^k.$$

A possible application for such a pattern matching is looking for a certain pattern in a patient's DNA sequence, that is not present in a healthy human's DNA, in order to diagnose a disease. Where $\Sigma = \{A, C, G, T\} \Rightarrow |\Sigma| = m = 4$.

Can we bound the deviation of F from its expectation? The idea is illustrated in figure 3 which shows the histogram, frequency of appearing, of all substrings of length k in the big sequence A.

Let us construct a Doob martingale:

$$Z_0 = E[F],$$
$$Z_i = E[F|a_1, \ldots, a_i] \text{ for } i = 1, \ldots, n.$$

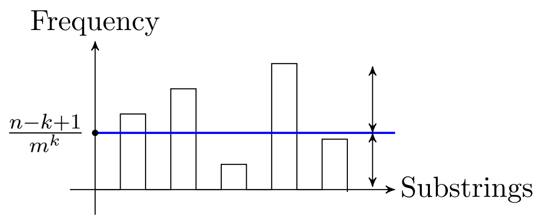

Figure 3: Histogram of all substrings of length k in the big sequence A.

$$\Rightarrow Z_0, Z_1, \dots, \text{ is a Doob martingale such that } E[Z_n] = F.$$

Each character in A can participate in no more than k occurrences of B:

$$|Z_i - Z_{i+1}| \leq k.$$

By Azuma-Hoeffding inequality:

$$Pr\left(|F - E[F]| \geq \lambda\right) \leq e^{-\frac{\lambda^2}{2nk^2}}.$$

Example 8. *Bins and Balls*
We are throwing m balls independently and uniformly at random into n bins. We are interested in F the number of empty bins at the end.

After the balls are thrown, let:

$$X_i = \begin{cases} 1 & \text{if bin } i \text{ is empty,} \\ 0 & \text{otherwise.} \end{cases}$$

Let F be the number of empty bins. $F = f\left(X_1, X_2, \dots, X_m\right)$, in particular we can write:

$$F = X_1 + \cdots + X_n.$$

Note that the $X_i's$ are not independent and F is a not a Binomial random variable. For example, not all $X_i's$ can be zero at the same time. Also, if $m < n$ and it happens that $X_1 = X_2 = \cdots = X_m = 0$ (first m bins not empty) then definitely we have $X_{m+1} = \cdots = X_n = 1$ (the remaining bins must be empty because we are out of balls).
A bin i is empty if all the m balls are thrown in the $n-1$ bins other than bin i. Therefore,

$$Pr(\text{bin } i \text{ is empty}) = \left(\frac{n-1}{n}\right)^m.$$

$$E[F] = E(X_1 + \cdots + X_n) = nE(X_1) = n\left(\frac{n-1}{n}\right)^m.$$

We are interested in bounding the deviation of F from its expectation. Let us construct a Doob martingale,

$$Z_0 = E(F), \text{ and, } Z_{i+1} = E\left[F|X_1, \dots, X_i\right], i = 1, \dots n.$$

After revealing the status of bin $i+1$, the number of empty bins can at most increase by one. Therefore,

$$|Z_{i+1} - Z_i| \leq 1.$$

Therefore, we can apply the Azuma-Hoeffding inequality to obtain:

$$Pr\left(|F - E[F]| \geq \lambda\right) \leq e^{-\frac{\lambda^2}{2n}}.$$