MARTINGALES AND GAMBLING

Louis H. Y. Chen
Department of Mathematics
National University of Singapore

1. Introduction

The word *martingale* refers to a strap fastened between the girth and the noseband of a horse to keep its head down. It also means the policy of doubling the stake on losing in gambling. In probability theory, *martingale* is used as a name for a class of stochastic processes which has application to gambling.

Notable early results on martingales were due to P. Lévy (1937) and J. Ville (1939). J. L. Doob (1940) was the first person to explore martingales fully and later develop them systematically in his book "Stochastic Processes" (1953). Since then further development of martingales and their widespread use in probability theory have occurred. Today, martingale theory has become a branch of probability theory in its own right. Not only does it have applications in many branches of probability theory and statistics, but it also plays a central role in the theory of stochastic integration and has a deep relation with harmonic analysis.

A proper treatment of martingales requires a sophisticated mathematical background. However, some understanding of martingales at an elementary level can be achieved by considering some special cases and relating them to gambling.

In this article we attempt to discuss some of the basic ideas and properties of martingales in an elementary fashion and apply them to gambling. In order to avoid the use of advanced tools from analysis, we confine ourselves to discrete-time discrete-value martingales.

2. Definitions and Examples

A sequence of discrete random variables $f = (f_1, f_2, \ldots)$ is called a martingale if for $n \geqslant 1$, $E|f_n| < \infty$ and for $n \geqslant 2$,

$$(1) \quad E(f_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) = x_{n-1}$$

for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. Here the expression on the left hand side of (1) is the conditional expectation of f_n given $f_1 = x_1, \ldots, f_{n-1} = x_{n-1}$.

It is worth noting that a consequence of (1) is $Ef_1 = Ef_2 = \ldots$. This can be proved by multiplying both sides of (1) by $P(f_1 = x_1, \ldots, f_{n-1} = x_{n-1})$ and then summing over all possible values $x_1, \ldots, x_{n-1}$.

There is some similarity between the definition of a martingale and that of a Markov chain in the sense that in both cases "the presence dominates the past in influencing the future". There also exist examples of martingales which are Markov chains. However, a martingale is distinctively different from a Markov chain, and one can easily find an example of a martingale which is not a Markov chain and vice versa.

How do martingales occur? This question is perhaps best answered by looking at some examples.

Example 1

Let $Z_0, Z_1, Z_2, \ldots$ be a sequence of independent discrete random variables such that $E|Z_0| < \infty$ and for $n \geq 1$, $EZ_n = 0$. Define $f_n = Z_0 + Z_1 + \ldots + Z_n$. Then for $n \geq 1$, $E|f_n| \leq E|Z_0| + E|Z_1| + \ldots + E|Z_n| < \infty$ and for $n \geq 2$ and any set of values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$,

$$
\begin{aligned}
(2)\quad & E(f_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \\
&= E(f_{n-1} + Z_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \\
&= E(f_{n-1} \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) + E(Z_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \\
&= x_{n-1} + EZ_n = x_{n-1}
\end{aligned}
$$

where the penultimate equality follows from the independence between Z_n and $f_1, \ldots, f_{n-1}$. So, $f = (f_1, f_2, \ldots)$ is a martingale.

If $Z_0 = a$, where a is an integer, and for $n \geq 1$, $Z_n = 1$ or -1 with probability ½, then f in Example 1 is a symmetric simple random walk. In this case, f_n represents the position, after n steps, of a particle which starts from position a and moves on the set of integers according to the following rule: From positive i it moves to $i + 1$ or $i - 1$ with probability ½.

Example 2

Suppose a gambler plays a series of games each of which has two outcomes each occurring with probability ½. Assume that each game is fair so that the gambler's payoff equals his bet. Suppose further that the gambler adopts a strategy by which he decides on the amount of his bet in each game on the basis of the outcomes of the previous games and his fortunes throughout these games. Let f_0 be the initial fortune of the gambler, b_n his bet in the nth game and f_n his fortune after the nth game. Then the games may be represented by a sequence of independent random variables $Z_1, Z_2, \ldots$ where for $n \geq 1$, $P(Z_n = 1) = P(Z_n = -1) = ½$ and the gambler wins in the nth game if $Z_n = 1$ and he loses if $Z_n = -1$. Furthermore, for $n \geq 1$, b_n is a function of $(Z_1, \ldots, Z_{n-1}, f_0, f_1, \ldots, f_{n-1})$ such that

$$(3)\qquad 0 \leq b_n(Z_1, \ldots, Z_{n-1}, f_0, f_1, \ldots, f_{n-1}) \leq f_{n-1}$$

and

$$(4)\qquad f_n = f_{n-1} + b_n(Z_1, \ldots, Z_{n-1}, f_0, f_1, \ldots, f_{n-1})Z_n.$$

Equation (3) simply means that at any time the gambler's bet will not exceed his current fortune. Both equations (3) and (4) inductively imply that for $n \geq 1$, f_n depends on $Z_1, \ldots, Z_n$ and consequently b_n depends only on $Z_1, \ldots, Z_{n-1}$. Hence without loss of generality we may rewrite (3) and (4) as

$$(5)\qquad 0 \leq b_n(Z_1, \ldots, Z_{n-1}) \leq f_{n-1}$$

and

$$(6)\qquad f_n = f_{n-1} + b_n(Z_1, \ldots, Z_{n-1})Z_n$$

where for n = 1, $b_n(z_1, \ldots, Z_{n-1})$ equals a constant $b_1 \leqslant f_0$. Also by induction, we have for $n \geqslant 1$, $E|f_n| < \infty$.

The expression $b_n(Z_1, \ldots, Z_{n-1})Z_n$ is the gain of the gambler in the nth game. Since the games are fair, the gambler's expected gain in each game is zero regardless of the outcomes of the previous games. This can be checked as follows: For n = 1

$$\begin{aligned} E(b_n(Z_1, \ldots, Z_{n-1})Z_n) &= E(b_1 Z_n) \\ &= b_1 E Z_n = 0; \end{aligned}$$

and for $n \geqslant 2$ and any set of values $z_1, \ldots, z_{n-1}$ of $Z_1, \ldots, Z_{n-1}$,

$$\begin{aligned} (7) \quad & E(b_n(Z_1, \ldots, Z_{n-1})Z_n \mid Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}) \\ &= b_n(z_1, \ldots, z_{n-1})\, E(Z_n \mid Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}) \\ &= b_n(z_1, \ldots, z_{n-1})\, E Z_n = 0, \end{aligned}$$

where the penultimate equality follows from the independence of $Z_1, \ldots, Z_n$.

Now let $d_n = b_n(Z_1, \ldots, Z_{n-1})Z_n$. Multiplying the first and the last terms of (7) by $P(Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1} \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1})$ and summing over $(z_1, \ldots, z_{n-1}) \in \Lambda(x_1, \ldots, x_{n-1})$, where $\Lambda(x_1, \ldots, x_{n-1}) = \{(z_1, \ldots, z_{n-1}) : f_1(z_1) = x_1, \ldots, f_{n-1}(z_1, \ldots, z_{n-1}) = x_{n-1}\}$, we obtain

$$\begin{aligned} (8) \quad & \sum_{(z_1, \ldots, z_{n-1}) \in \Lambda(x_1, \ldots, x_{n-1})} \cdots \sum E(d_n \mid Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}) \\ & \quad \times P(Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1} \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \\ & = 0 \end{aligned}$$

Since $(z_1, \ldots, z_{n-1}) \in \Lambda(x_1, \ldots, x_{n-1})$, it follows that $\{Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}\} \subset \{f_1 = x_1, \ldots, f_{n-1} = x_{n-1}\}$ and so $P(Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1} \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) = P(Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}) / P(f_1 = x_1, \ldots, f_{n-1} = x_{n-1})$. By this and the definition of conditional expectation, the left hand side of (8) equals

$$\sum_{(z_1, \ldots, z_{n-1}) \in \Lambda(x_1, \ldots, x_{n-1})} \cdots \sum \frac{E d_n I(Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1})}{P(Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1})} \cdot \frac{P(Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1})}{P(f_1 = x_1, \ldots, f_{n-1} = x_{n-1})}$$

$$= \sum_{(z_1, \ldots, z_{n-1}) \in \Lambda(x_1, \ldots, x_{n-1})} \cdots \sum \frac{E d_n\, I(Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1})}{P(f_1 = x_1, \ldots, f_{n-1} = x_{n-1})}$$

$$= \frac{E d_n\, I(f_1 = x_1, \ldots, f_{n-1} = x_{n-1})}{P(f_1 = x_1, \ldots, f_{n-1} = x_{n-1})} = E(d_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1})$$

where I(A) denotes the indicator function of the set A. Hence we have .

$$(9) \qquad E(d_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) = 0$$

and this holds for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. Using this and taking conditional expectation of both sides of (6) given $f_1 = x_1, \ldots, f_{n-1} = x_{n-1}$, we have, for $n \geqslant 2$,

$$
\begin{aligned}
E(f_n \mid f_1 &= x_1, \ldots, f_{n-1} = x_{n-1}) \\
&= E(f_{n-1} \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) + E(d_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \\
&= x_{n-1}
\end{aligned}
$$

for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. Hence the sequence of fortunes $f = (f_1, f_2, \ldots)$ of the gambler is a martingale. By (5), it is in fact a non-negative martingale, that is, $f_n \geqslant 0$ for $n \geqslant 1$.

In Example 2, we have shown that for $n \geqslant 2$, $E(d_n \mid Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}) = 0$ for all values $z_1, \ldots, z_{n-1}$ of $Z_1, \ldots, Z_{n-1}$ implies that $E(d_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) = 0$ for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. This is a consequence of the fact that $f_1, \ldots, f_{n-1}$ are functions of $Z_1, \ldots, Z_{n-1}$. This property holds in general.

Example 3

In real life, when a gambler plays a series of games, he may not play the same type of game throughout and the games may not give rise to only two outcomes. In such situations, the games may be represented by a sequence of independent random variables $Z_1, Z_2, \ldots$ which need not have a common distribution. The rules of the games may allow the gambler to place more than one bet, either all at once or sequentially, or to increase his bet progressively. Sometimes bets are not even required. In any case, his sequence of fortunes $f = (f_1, f_2, \ldots)$ can always be expressed as

$$(10) \qquad f_n = f_{n-1} + d_n, \; n \geqslant 1,$$

where f_0 is his initial fortune and d_n his gain in the nth game. As in Example 2, for $n \geqslant 1$, f_n and d_n depend on $Z_1, \ldots, Z_n$ and $E \mid f_n \mid < \infty$. If the games are fair, then $Ed_1 = 0$ and for $n \geqslant 2$, $E(d_n \mid Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}) = 0$ for all values $z_1, \ldots, z_{n-1}$ of $Z_1, \ldots, Z_{n-1}$. This implies that for $n \geqslant 2$, $E(d_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) = 0$ and hence $E(f_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) = x_{n-1}$ for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. This shows that $f = (f_1, f_2, \ldots)$ is a martingale.

In Example 1, if for $n \geqslant 2$, $EZ_n \geqslant 0$ (respectively $\leqslant 0$), then $E(f_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \geqslant x_{n-1}$ (respectively $\leqslant x_{n-1}$) for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. In Example 3, if the games are favourable (respectively unfavourable) to the gambler, that is $Ed_1 \geqslant 0$ (respectively $\leqslant 0$) and for $n \geqslant 2$, $E(d_n \mid Z_1 = z_1, \ldots, Z_{n-1} = z_{n-1}) \geqslant 0$ (respectively $\leqslant 0$) for all values $z_1, \ldots, z_{n-1}$ of $Z_1, \ldots, Z_{n-1}$, then it can be shown that for $n \geqslant 2$, $E(d_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \geqslant 0$ (respectively $\leqslant 0$) for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. Consequently, $Ef_1 \geqslant f_0$ (respectively $\leqslant f_0$) and for $n \geqslant 2$, $E(f_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \geqslant x_{n-1}$ (respectively $\leqslant x_{n-1}$) for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. This leads us to the notion of submartingales and supermartingales which we defined as follows.

A sequence of discrete random variables $f = (f_1, f_2, \ldots)$ is a submartingale (respectively supermartingale) if for $n \geqslant 1$, $E \mid f_n \mid < \infty$ and for $n \geqslant 2$,

$$(11) \qquad E(f_n \mid f_1 = x_1, \ldots, f_{n-1} = x_{n-1}) \geqslant x_{n-1} \; (\text{respectively} \leqslant x_{n-1})$$

for all values $x_1, \ldots, x_{n-1}$ of $f_1, \ldots, f_{n-1}$. It follows from this definition that if $f = (f_1, f_2, \ldots)$ is a submartingale then $-f = (-f_1, -f_2, \ldots)$ is a supermartingale and vice versa. Also, a martingale is both a submartingale and a supermartingale. The reader should verify for himself that for a submartingale (respectively supermartingale), $Ef_1 \leqslant Ef_2 \leqslant \ldots$ (respectively $Ef_1 \geqslant Ef_2 \geqslant \ldots$).

Practically all games in a casino are unfavourable to the gambler. For example, in the game of roulette, there are 37 numbers, namely 0, 1, 2, . . . , 36. (This is the European roulette; the American roulette has one more number, namely the double zero 00.) For every \$1 which a gambler bets on a number, he will get a payoff of \$35. So his expected gain is $35(1/37) - 1(36/37) = -1/37$ which is negative. Hence the supermartingale is a more realistic model for the fortunes of a casino gambler.

3. Fundamental Theorems

Two fundamental theorems of martingales are the optional sampling theorem and the martingale convergence theorem. Both theorems are due to Doob. Before we can discuss the optional sampling theorem we need the notion of a stopping time.

Let $f = (f_1, f_2, \ldots)$ be a submartingale or a supermartingale. A random variable τ is a stopping time if it takes values $1, 2, \ldots, \infty$ such that for $n \geqslant 1$, the event $\{\tau = n\}$ is determined by $f_1, \ldots, f_n$ and $\{\tau = \infty\}$ determined by $f_1, f_2, \ldots$. For example, τ is a stopping time if

$$\tau = \begin{cases} \inf \{n \geqslant 1 : |f_n| > c\} \\ \infty \quad \text{if } |f_1| \leqslant c, |f_2| \leqslant c, \ldots \end{cases}$$

where c is some positive number. Here τ is the first time the absolute value of f_n exceeds c.

In Examples 2 and 3, assume that the gambler is playing in a casino which has unlimited capital. Suppose the gambler decides to quit when his fourtune falls below c. Then the stopping time τ defined by

$$\tau = \begin{cases} \inf \{n \geqslant 1 : f_n < c\} \\ \infty \quad \text{if } f_1 \geqslant c_1, f_2 \geqslant c_1, \ldots, \end{cases}$$

is the time when the gambler quits.

We now state a special case of the optional sampling theorem.

Theorem 1

Let $f = (f_1, f_2, \ldots)$ be a bounded martingale, that is, a martingale for which there exists a constant M such that for $n \geqslant 1$, $|f_n| \leqslant M$. Then for any stopping time τ such that $P(\tau < \infty) = 1$ (or equivalently $P(\tau = \infty) = 0$), we have

$$(12) \qquad Ef_1 = Ef_\tau.$$

If f is a bounded submartingale (respectively bounded supermartingale), then the equality in (12) is replaced by $\leqslant$ (respectively $\geqslant$).

In terms of gambling, equation (12) simply means that the games remain fair up to and including the time the gambler decides to quit. We will see later how equation (12) can be used to calculate the probability of "ruin" of a gambler. The next theorem is the martingale convergence theorem.

Theorem 2

Let $f = (f_1, f_2, \ldots)$ be an L_1-bounded martingale, that is,a martingale such that $\sup_{1 \leqslant n < \infty} E \mid f_n \mid < \infty$. Then

(13) $P(\ \{\omega \in \Omega : \lim_{n \to \infty} f_n\ (\omega) \text{ exists and is finite}\}\ ) = 1$, where Ω is the sample space on which all $f_1, f_2, \ldots$ are defined.

We may interpret statement (13) as follows. Since each $\omega \in \Omega$ is associated with a sequence $(f_1(\omega), f_2(\omega), \ldots)$, we may regard the sample space Ω as consisting of all such sequences $(f_1(\omega), f_2(\omega), \ldots)$. Each such sequence is called a sample path. If $f_1, f_2, \ldots$ represent the fortunes of a gambler, then the sample space may be thought of as consisting of all possible sequences of fortunes of the gambler. The statement (13) simply says that the set of sample paths which converge has probability 1. For brevity, we abbreivate statement (13) to $P(\lim_{n \to \infty} f_n$ exists and is finite) = 1 or simply, P(f converges) = 1.

If $f = (f_1, f_2, \ldots)$ is a non-negative supermartingale (this includes the martingale case), then $-f = (-f_1, -f_2, \ldots)$ is a submartingale and $E \mid -f_n \mid = Ef_n$. Since $Ef_1 \geqslant Ef_2 \geqslant \ldots$, it follows that $\sup_{1 \leqslant n < \infty} E \mid -f_n \mid = \sup_{1 \leqslant n < \infty} Ef_n \leqslant Ef_1 < \infty$. By the martingale convergence theorem (Theorem 2), P(−f converges) = 1 which implies that P(f converges) = 1. Hence a non-negative supermartingale always converges with probability 1.

In 1922, H. Steinhaus posed the following problem: Given a sequence of real numbers $a_1, a_2, \ldots$, what is the probability that $\sum_{n=1}^{\infty} \pm a_n$ converges where the signs of the terms are chosen independently and each with probability ½? In the same year, H. Rademacher proved that if $\sum_{n=1}^{\infty} a_n{}^2 < \infty$ then $\sum_{n=1}^{\infty} \pm a_n$ converges with probability 1. In probabilistic notation, the series should be written as $\sum_{n=1}^{\infty} a_n Z_n$ where $Z_1, Z_2, \ldots$ are independent random variables such that for $n \geqslant 1$, $P(Z_n = 1) = P(Z_n = -1) = \frac{1}{2}$.

A classical limit theorem in probability theory states that if $X_1, X_2, \ldots$ is a sequence of independent random variables with means 0 and variances $\sigma_1^2, \sigma_2^2, \ldots$, then $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$ implies that $P(\sum_{n=1}^{\infty} X_n \text{ converges}) = 1$. Since $a_1 Z_1, a_2 Z_2, \ldots$ are independent random variables with $E(a_n Z_n) = 0$ and $Var(a_n Z_n) = a_n^2$, Rademacher's result is a special case of the classical limit theorem.

According to the argument in Example 1, the sequence $f = (f_1, f_2, \ldots)$ defined by $f_n = X_1 + \ldots + X_n$ is a martingale. (This is true in general but in this article we may assume that $X_1, X_2, \ldots$ are discrete random variables.) By independence, for $n \geqslant 1$, $\sum_{k=1}^{n} \sigma_k^2 = Ef_n^2$ and so $\sum_{n=1}^{\infty} \sigma_n^2 = \sup_{1 \leqslant n < \infty} Ef_n^2$. Since for $n \geqslant 1$, $E\,|f_n| \leqslant (Ef_n^2)^{1/2}$ (by an application of the Cauchy-Schwarz inquality), it follows that $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$ implies $\sup_{1 \leqslant n < \infty} E\,|f_n| < \infty$. Hence the martingale convergence theorem is a generalization of the classical limit theorem and hence of Rademacher's result.

4. Application to Gambling

Suppose a gambler has an initial fortune f_0 and plays in a casino which for all practical purposes may be assumed to have an unlimited capital. Assume that his bet in each game does not exceed his current fortune. Then by the argument in Section 2, the sequence of fortunes $f = (f_1, f_2, \ldots)$ is a non-negative supermartingale and, by the martingale convergence theorem, f must converge with probability 1. Suppose further that there must be a change of his fortune in each game. Then due to the existence of a smallest denomination in any given currency and the imposition by the casino of a minimum bet for each game, there exists a positive number δ such that for $n \geqslant 0$, $|f_{n+1} - f_n| \geqslant \delta$. Since f converges with probability 1, we must have

$$P(\,|f_{n+1} - f_n| \geqslant \delta \text{ infinitely often}) = 0.$$

This is because every sample path which converges cannot fluctuate by no less than any preassigned positive number indefinitely. Therefore the set of sample paths for which $|f_{n+1} - f_n| < \delta$ for all sufficiently large n has probability 1. But $|f_{n+1} - f_n| < \delta$ implies $f_{n+1} = f_n$. It follows that the set of sample paths which become constant after some time has probability 1. This implies that whatever strategy the gambler uses, he will quit after some time with probability 1.

If the gambler persists in playing as long as his fortune does not fall below the amount of minimum bet c imposed by the casino, then the only time he will quit is when his fortune falls below c. Since he will eventually quit with probability 1, it follows that his fortune will fall below c with probability 1. Hence we have a theorem that a persistent gambler will go broke with probability 1 regardless of the strategy he adopts.

We now turn to Example 2 of Section 2. Let us call the gambler Peter and assume that he is playing against another gambler called Paul. So, in each game, Peter loses his bet to Paul with probability ½ and wins an equal amount from him with probability ½. Let the initial fortune of Peter be a dollars (that is $f_0 = a$) and that of Paul be b dollars, where a and b are positive integers. Assume further that Peter's bet in each game does not exceed both his and Paul's current fortunes and is at least \$1 unless one of them has been "ruined" (that is, one of them has lost all his fortune). Then the sequence of fortunes $f = (f_1, f_2, \ldots)$ of Peter is given by (6), while Peter's bet $b_n(Z_1, \ldots, Z_{n-1})$ in the nth game, $n \geqslant 1$, is now governed by

(14) $0 \leqslant b_n(Z_1, \ldots Z_{n-1}) \leqslant \min(f_{n-1}, a+b-f_{n-1})$

instead of (5), where $f_0 = a$. Furthermore, $0 \leqslant f_n \leqslant a+b$ for $n \geqslant 1$, and $|f_{n+1} - f_n| \geqslant 1$ until either $f_n = 0$ or $a + b$, after which time f_n will remain constant. As in Example 2, Peter's fortunes form a non-negative martingale and therefore must converge with probability 1. Since f cannot converge as long as $|f_{n+1} - f_n| \geqslant 1$, it follows that f_n must eventually become 0 or a + b with probability 1, that is, with probability 1 either Peter will be ruined or Paul will be ruined. This implies

(15) P(Peter will be ruined) + P(Paul will be ruined) = 1.

Now let τ be the first time $f_n = 0$ or $a+b$, that is,

$$\tau = \begin{cases} \inf\{n \geqslant 1 : f_n = 0 \text{ or } a+b\} \\ \infty \quad \text{if } 1 \leqslant f_n \leqslant a+b-1 \text{ for all } n \geqslant 1. \end{cases}$$

Then we have $P(\tau < \infty) = 1$. Since $0 \leqslant f_n \leqslant a+b$ for $n \geqslant 1$, it follows from Thereom 1 that

(16) $Ef_1 = Ef_\tau$.

Now f_τ equals 0 or a + b. Therefore

(17) $Ef_\tau = 0 \cdot P(\text{Peter will be ruined}) + (a+b)P(\text{Paul will be ruined})$
$= (a+b)P(\text{Paul will be ruined})$.

By (6), $f_1 = a + b_1 Z_1$ and so

(18) $Ef_1 = a + b_1 \, EZ_1 = a$.

Combining (16), (17) and (18), we obtain

(19) P(Paul is ruined) = a/(a + b).

From this and (15), we also obtain

(20) P(Peter is ruined) = b/(a + b).

Note that if $b_n = 1$ (unless one of the gamblers is ruined), then the problem of finding the probabilities of ruin reduces to the classical gambler's ruin problem. It is also the problem of finding the probability of absorption at 0 in a symmetric simple random walk with two absorbing boundaries (that is, the problem of finding the probability that the particle, starting from a, will reach 0 before a + b). In this special case, the classical method of finding the probabilities of ruin consists of solving a difference equation.

In conclusion, martingale theory makes it possible to tell the fate of a persistent gambler under very general conditions. It provides a solution to a more general gambler's ruin problem in which the bets may vary from game to game. It also gives an insight into the reason why the probabilities of ruin must be given by (19) and (20) — the games remain fair up to and including the time of a gambler's ruin. These applications of martingale theory shed some light on the nature of martingales, but provide only a glimpse of the elegance and power of martingale theory.

FURTHER READING

J. L. DOOB, What is a martingale?, American Mathematical Monthly 76 (1971), 451-463.