[Commissioned by: *The International Encyclopedia of Social and Behavioral Sciences*]

Martingales

René L. Schilling

TU Dresden, Dresden, Germany

Abstract

A martingale is the mathematical description of a fair game: The expected net gain or loss from further play, independent of the history, is 0. The rigorous mathematical definition involves conditional expectation. Following earlier work by Paul Lévy and Jean Ville, Joseph Leo Doob developed in the 1940s and 1950s basic martingale theory and many of its applications. Martingales allowed one to study, for the first time, the behaviour of sums and sequences of random variables which are not independent. Martingale theory is one of the cornerstones of modern mathematical probability theory with wide-ranging applications in stochastic analysis and mathematical finance.

Keywords: Banach-space geometry; Berry–Esseen theorem; central limit theorem; conditional probability; fair game; gambling strategy; game theory; harmonic analysis; likelihood ratio; martingale; mathematical finance; optimal stopping; stochastic analysis; stochastic (Itô) integral; stochastic process.

Online Resources: *Oxford English Dictionary* (OED) `www.oed.com`
Journal Eléctronique d'Histoire des Probabilités et de la Statistique `www.jehps.net`, in particular `www.jehps.net/juin2009/Mansuy.pdf`

The notion of *martingale* is intimately connected with that of a *fair game* where the player and the bank have equal chances of winning. In the 18th and 19th centuries, a martingale meant a gambling strategy; in its simplest form a player doubles his stake if he lost in the previous round, and goes away otherwise. Under the assumption that the player can bet indefinitely high stakes, the losses are given by a geometric progression and, after possibly arbitrarily high losses, there is a guaranteed net win of the *initial* stake. This is equivalent to the so-called *St. Petersburg paradox* (Gorroochurn, 2012). Usually this strategy fails due to limited wealth of the player, time constraints, and limits on the stakes on the side of the casino: The mean loss (prior to winning) is infinite.

In mathematical terms, the fair game underlying a simple martingale is a sum of independent, identically distributed Bernoulli random variables $Y_1, Y_2, \dots$ such that $\mathbb{P}(Y_n = -1) = \mathbb{P}(Y_n = 1) = \frac{1}{2}$ (e.g. representing the nth toss of a fair coin); $Y_n = 1$ returns twice your stake, while $Y_n = -1$ means you lose all. Denoting by $e_1 = 1$ and $e_n = e_n(Y_1, \dots, Y_{n-1})$, $n = 2, 3, \dots$, the stake for the nth game, the total net winnings are $X_0 = 0$ and $X_n = e_1 Y_1 + \dots + e_n Y_n$. One easily sees that the conditional expectation (Schilling, 2005; Schilling and Partzsch, 2014; Williams, 1991) satisfies

$$\mathbb{E}(X_{n+1} \mid Y_1, \dots, Y_n) = X_n. \tag{1}$$

Since the events $\{Y_i = \pm 1\}$ have strictly positive probability, this abstract conditional expectation still has its classical meaning as

$$\mathbb{E}(X_{n+1} \mid Y_1 = y_1, \dots, Y_n = y_n) = x_n,$$

where the y_i, $i = 1, \dots, n$, are the outcomes of the rounds $1, \dots, n$, and $x_n = e_1 y_1 + e_2(y_1) y_2 + \dots + e_n(y_1, \dots, y_{n-1}) y_n$ is the net win after the nth round. Iterating (1) yields

$$\mathbb{E}(X_{n+k} \mid Y_1, \dots, Y_n) = X_n \quad k, n = 1, 2, \dots \tag{2}$$

The simple martingale strategy corresponds to the stakes $e_n(-1, -1, \dots, -1) = 2^{n-1}$ and 0 in all other cases. In particular, the expected gains are

$$\mathbb{E}X_n = \mathbb{E}X_{n-1} = \dots = \mathbb{E}X_0 = 0.$$

Conditional expectation can be interpreted as the best mean-square predictor for the future win X_{n+k} conditional on the past outcomes $Y_1, \dots, Y_n$. Since the stakes for the next round depend only on the past results, i.e. e_{n+1} is a function of $Y_1, \dots, Y_n$, the stakes are previsible. The martingale property guarantees that the best prediction is always the current (time n) level of winnings. In this sense, a martingale describes a fair game.

It is not possible to change the character of a game by a previsible gambling strategy: A fair game will stay fair no matter which gambling strategy is being used. This follows immediately from the fact that in the discussion above we have made no further assumptions on the gambling strategy, i.e. the stakes $e_1, e_2, \ldots$, than previsibility. The same assertion is true for favourable [unfavourable] games, that is submartingales [supermartingales].

Definition

The *general definition of a martingale* is as follows: Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, an index set I equipped with a partial order $s \leq t$ and an increasing family of σ-algebras (a *filtration*) $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ if $s \leq t$, then the real-valued random variables $(X_t)_{t\in I}$ are a *martingale* with respect to $(\mathcal{F}_t)_{t\in I}$, if each X_t is $\mathcal{F}_t$ measurable, $\mathbb{E}|X_t| < \infty$ for all $t \in I$, and

$$\mathbb{E}(X_t \mid \mathcal{F}_s) = X_s \quad \text{for all} \quad s, t \in I,\ s \leq t; \qquad (3)$$

if we replace '=' by '≥' [or '≤'], we get a sub- [or super-] martingale.

The expectation $m(t) = \mathbb{E}X_t$ of a martingale is constant, while those of a sub-/supermartingale are increasing/decreasing functions of t.

A *downward*, *backwards* or *reversed martingale* is a martingale where the family of σ-algebras $(\mathcal{F}_t)_{t\in I}$ decreases as t increases. This means that $(X_{-t}, \mathcal{F}_{-t})_{t\in -I}$ is a martingale (with the natural order of $-I$) and all martingale results hold for backwards martingales. Therefore, there is no need to distinguish between martingales and backwards martingales. If I has a smallest element, i_0, then $-I$ has a largest element and the backwards martingale is right-closed, i.e. $X_{-t} = \mathbb{E}(X_{i_0} \mid \mathcal{F}_{-t})$ for all $t \in -I$. Backwards martingales were first studied in (Doob, 1953).

The notion of a (sub-/super-) martingale essentially depends on the given stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. If $(X_t)_{t\in I}$ is a martingale for the filtration $(\mathcal{F}_t)_{t\in I}$, it is also a martingale for the minimal σ-algebras $\mathcal{F}_t^X = \sigma(X_s : s \leq t,\ s \in I)$ which is, by definition, the smallest σ-algebra such that all X_s, $s \leq t$, are measurable. A further reduction is not possible.

On the other hand, enlarging the filtration usually results in the loss of the martingale property: The σ-algebra $\mathcal{F}_t$ has a natural interpretation as 'the information until time t' and enlarging $\mathcal{F}_t$ means increasing the available information (e.g. because of insider knowledge or cheating) which immediately leads to a non-fair situation.

In (1), (2) one uses the σ-algebra $\sigma(Y_1, \ldots, Y_n)$ which is generated by $Y_1, \ldots, Y_n$. Changing to the smaller σ-algebras $\sigma(X_1, \ldots, X_n)$, one would still get a martingale, since both $\sigma(Y_1, \ldots, Y_n)$ and $\sigma(X_1, \ldots, X_n)$ contain the essential information on the game. Indeed, X_n is a function of the $Y_1, \ldots, Y_n$, while

$$e_{n+1}(Y_1, \ldots, Y_n)Y_{n+1} = X_{n+1} - X_n;$$

this allows one to represent $Y_1, Y_2, \ldots$ – on the the sets where $e_1, e_2, \ldots$ are strictly positive – recursively in terms of $X_1, X_2, \ldots$.

For simplicity, we will assume that $I \subset [0, \infty)$. Integer indices are usually denoted by m or n, real indices by s, t.

Examples

In the examples $\mathcal{F}_t^X = \sigma(X_s : s \leq t,\ s \in I)$ denotes the minimal filtration of the process $(X_t)_{t\in I}$.

a) (Sum of independent random variables) $X_n = Y_1 + \ldots + Y_n$, $X_0 = 0$, where $(Y_n)_{n\in\mathbb{N}}$ are real-valued, independent, mean-zero random variables, is a martingale with filtration $\mathcal{F}_n^X = \mathcal{F}_n^Y$. If $\mathbb{E}Y_n \geq 0$ [≤ 0] for all n, then $(X_n)_{n\in\mathbb{N}_0}$ is a submartingale [supermartingale].

   If $\mathbb{E}(X_n^2) < \infty$ for all $n \in \mathbb{N}_0$, then $(X_n^2)_{n\in\mathbb{N}_0}$ is a submartingale and $(X_n^2 - \mathbb{E}(X_n^2))_{n\in\mathbb{N}_0}$ is a martingale with filtration $(\mathcal{F}_n^X)_{n\in\mathbb{N}_0}$.

b) (de Moivre's martingale) Let $(Y_n)_{n\in\mathbb{N}}$ be independent identically distributed random variables such that $\mathbb{P}(Y_n = 1) = p$ and $\mathbb{P}(Y_n = -1) = q = 1 - p$. Then $(X_n)_{n\in\mathbb{N}_0}$ with $X_0 := 1$ and $X_n := (q/p)^{Y_1+\ldots+Y_n}$ is a martingale for the filtration $(\mathcal{F}_n^Y)_{n\in\mathbb{N}}$.

c) (Wald martingale) Let $(Z_n)_{n\in\mathbb{N}}$ be non-negative independent random variables with finite means $\mu_n = \mathbb{E}Z_n > 0$. Then $(X_n)_{n\in\mathbb{N}_0}$ with $X_0 := 1$ and $X_n := \prod_{i=1}^n \mu_i^{-1} Z_i$ is a martingale for the filtration $(\mathcal{F}_n^Z)_{n\in\mathbb{N}}$.

d) (Pólya's urn) An urn contains X_0 blue and Y_0 red balls. Draw randomly a ball from the urn. If a blue (resp. red) ball is drawn, replace the ball and add c more blue (resp. red) balls to the urn. Denote by X_n and Y_n the number of blue and red balls in the urn after the nth round. Then $(V_n)_{n\in\mathbb{N}_0}$ with $V_n = X_n/(X_n + Y_n)$ is a martingale for the filtration $\mathcal{F}_n := \sigma(X_i, Y_i : i \leq n)$. (Pólya invented this scheme in order to model the spread of contagious diseases.)

e) Let $(X_t)_{t\geq 0}$ be a stochastic process with independent increments such that $m(t) = \mathbb{E}X_t$ exists. It is a martingale (with filtration $(\mathcal{F}_t^X)_{t\geq 0}$) if, and only if, $m(t)$ is constant; it is a submartingale if, and only if, $m(t)$ is increasing.

The process $(X_t - m(t))_{t\geq 0}$ is a martingale. In particular, one-dimensional Brownian motion $(W_t)_{t\geq 0}$ and the compensated Poisson process $(N_t - \lambda t)_{t\geq 0}$ with intensity $\lambda > 0$ are martingales.

f) (Exponential martingale) Let $(X_t)_{t\geq 0}$ be a real-valued Lévy process (a process with independent and stationary increments which is continuous in probability). Then $\big(\exp(i\xi X_t)/\mathbb{E}e^{i\xi X_t}\big)_{t\geq 0}$, $\xi \in \mathbb{R}$, is a martingale for the filtration $(\mathcal{F}_t^X)_{t\geq 0}$.

g) Let $(X_n)_{n\in\mathbb{N}_0}$ be a Markov chain with countable state space E and transition matrix $P = (p_{ik})_{i,k\in E}$. Then, for any $u : E \to \mathbb{R}$, $\big(u(X_n) - \sum_{m=0}^{n-1}(\mathrm{id} - P)u(X_m)\big)_{n\in\mathbb{N}_0}$ is a martingale with filtration $(\mathcal{F}_n^X)_{n\in\mathbb{N}_0}$.

If u is invariant for P, i.e. $Pu \equiv u$, then $(u(X_n))_{n\in\mathbb{N}_0}$ is again a martingale.

h) Let $(X_t)_{t\geq 0}$ be a Markov process with transition semigroup $(P_t)_{t\geq 0}$ and generator A. Then $\big(u(X_t) - \int_0^t Au(X_s)\,ds\big)_{t\geq 0}$ is for any $u : \mathbb{R} \to \mathbb{R}$ from the domain of A a martingale with filtration $(\mathcal{F}_t^X)_{t\geq 0}$.

If u is A-harmonic, i.e. $Au \equiv 0$, then $(u(X_t))_{t\geq 0}$ is again a martingale.

i) (Lévy martingale, closed martingale) Let $(\mathcal{F}_t)_{t\geq 0}$ be an increasing family of σ-algebras in $\mathcal{F}$ and X be a real-valued random variable with $\mathbb{E}|X| < \infty$. Then $(X_t)_{t\geq 0}$ where $X_t := \mathbb{E}(X \mid \mathcal{F}_t)$ is a uniformly integrable martingale. Since $X = X_\infty$ is the right endpoint, this martingale is often called a (*right-*)*closed* martingale.

j) Let $\mathbb{Q}$ be a further probability measure on $(\Omega, \mathcal{F})$, let $(\mathcal{F}_t)_{t\geq 0}$ be an increasing family of σ-algebras in $\mathcal{F}$ and denote by $\mathbb{Q}_t$ and $\mathbb{P}_t$ the restrictions of the measures $\mathbb{Q}$ and $\mathbb{P}$ to $\mathcal{F}_t$. The *likelihood ratio* $L_t := d\mathbb{Q}_t/d\mathbb{P}_t$, $t \geq 0$, is a nonnegative martingale.

Properties

(Schilling, 2005; Schilling and Partzsch, 2014; Revuz and Yor, 1999; Rogers and Williams, 1987/94; Williams, 1991; Chow and Teicher, 1997) (Sub-)martingales have good convergence properties: If $\sup_{t\in I}\mathbb{E}|X_t| < \infty$ and $t \to t_0 \in I$, then $X_t \to X_{t_0}$ almost surely; if $(X_t)_{t\in I}$ is a uniformly integrable (sub-)martingale, this convergence holds also in L^1-sense. In particular, all uniformly integrable martingales are of the form of Example i) with $X = \lim_{t\to\infty} X_t$, and vice versa. A sufficient condition for uniform integrability is that $\sup_{t\in I}\mathbb{E}(|X_t|^p) < \infty$ for some $p > 1$. The convergence properties make martingale techniques powerful tools to get càdlàg (i.e. right-continuous with finite left limits) modifications of continuous-time stochastic processes. In the sequel, we will only consider càdlàg modifications.

If $(X_t)_{t\in I}$ is a martingale and $\phi : \mathbb{R} \to \mathbb{R}$ a convex function with $\mathbb{E}|\phi(X_t)| < \infty$ for all $t \in I$, then $(\phi(X_t))_{t\in I}$ is a submartingale. For (positive sub-)martingales we have *Doob's maximal inequalities*

$$\mathbb{E}\left(\sup_{t\in I}|X_t|^p\right) \leq c_p \sup_{t\in I}\mathbb{E}(|X_t|^p)$$

with $c_p = (p/(p-1))^p$ and $1 < p < \infty$.

Martingale convergence theorems were the first convergence results for sums and sequences of random variables which are *not* independent. Even nowadays, martingale techniques are among the most successful methods to deal with dependent random variables. For example the Lindeberg–Feller central limit theorem admits a martingale version: Let $\mathcal{F}_0 := \{\emptyset, \Omega\}$, $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration and $(X_n)_{n\in\mathbb{N}_0}$ be a sequence of L^2-martingale differences (i.e. $\mathbb{E}(X_n^2) < \infty$ and $\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) = 0$ for all $n \in \mathbb{N}_0$), set $\sigma_n^2 = \mathbb{E}(X_n^2)$ and $s_n^2 = \sigma_1^2 + \ldots + \sigma_n^2$. If $(X_n)_{n\in\mathbb{N}_0}$ satisfies the Lindeberg condition and

$$\sum_{i=1}^n \mathbb{E}\left|\mathbb{E}(X_i^2 \mid \mathcal{F}_{i-1}) - \sigma_i^2\right| = o(s_n),$$

then $s_n^{-1}\sum_{i=1}^n X_i$ converges in distribution to a standard normal random variable. An analogous martingale version exists for the *Berry–Esseen theorem.*

A *stopping time* (also: Markov time or optional time) is a random time $\sigma : \Omega \to I \cup \{\infty\}$ that is adapted to a given filtration $(\mathcal{F}_t)_{t\in I}$ in the sense that $\{\omega : \sigma(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \in I$; this means that it is enough to know X_s for all $s \leq t$, $s, t \in I$, in order to decide whether $\sigma \leq t$ or not. If $(X_t)_{t\in I}$ is a uniformly integrable martingale and $\sigma \leq \tau$ are stopping times with values in I, then (3) still holds for random times $\sigma \leq \tau < \infty$ (*optional stopping*); in particular, the stopped process $X_t^\tau := X_{\min(t,\tau)}$ is a martingale both for $(\mathcal{F}_t)_{t\in I}$ and $(\mathcal{F}_{\min(t,\tau)})_{t\in I}$ (*optional sampling*).

The fundamental relationship between martingales and submartingales is the *Doob–Meyer decomposition*: If $(X_t)_{t\geq 0}$ is a càdlàg submartingale of *class D* (i.e. the family $\{X_\tau :$

[1]In a continuous-time setting *previsible* means that the process $(A_t)_{t\geq 0}$ is measurable with respect to the σ-algebra generated by all left-continuous adapted processes. In particular, any left-continuous process $(Y_t)_{t\geq 0}$ such that Y_t is $\mathcal{F}_t$ measurable is previsible.

τ finite stopping time$\}$ is uniformly integrable), then there exists a unique (up to indistinguishability), increasing, previsible[1] process $(A_t)_{t\geq 0}$, $A_0 = 0$, and a uniformly integrable martingale $(M_t)_{t\geq 0}$, $M_0 = 0$, such that $X_t = X_0 + A_t + M_t$. Without the assumption of *class D*, we still have a unique decomposition into an increasing previsible process A_t and a *local* martingale M_t. Local martingales are processes for which there is a sequence of stopping times τ_n converging almost surely to infinity as $n \to \infty$ and such that $(M_t^{\tau_n})_{t\geq 0}$ are, for each $n \in \mathbb{N}$, martingales.

Examples g) and h) point to the connection of martingales with potential theory. Recall that the Laplacian $\frac{1}{2}\Delta$ is the generator of Brownian motion $(W_t)_{t\geq 0}$. Thus, a (classical) harmonic function $u \in C^2(\mathbb{R}^n)$ satisfies $\Delta u \equiv 0$ and $(u(W_t))_{t\geq 0}$ is a martingale. Similarly a sub-harmonic function u produces a submartingale. The Doob–Meyer decomposition is the stochastic analogue of the Riesz decomposition (any superharmonic function can be written as the sum of a harmonic function and a potential) from potential theory.

Martingale theory and the Doob–Meyer decomposition were instrumental for the development of a general theory of stochastic (Itô) integrals. Let $(M_t)_{t\geq 0}$ be an L^2-martingale, i.e. $\mathbb{E}(|M_t|^2) < \infty$ for all $t \geq 0$. By the Doob–Meyer decomposition there is a unique increasing previsible process $\langle M\rangle_t$ such that $M_t^2 - \langle M\rangle_t$ is a martingale; $\langle M\rangle_t$ is called the (*previsible*) *compensator* or *angle bracket* of $(M_t)_{t\geq 0}$. If $t \mapsto M_t$ is almost surely continuous, then there is a simple formula for the angle bracket: Denoting by Π a partition $0 = t_0 < t_1 < \ldots < t_n = t$ of $[0, t]$ with mesh $|\Pi| = \max_{1\leq i\leq n}(t_i - t_{i-1})$, then

$$\langle M\rangle_t = \lim_{|\Pi|\to 0} \sum_{i=1}^{n} (M_{t_i} - M_{t_{i-1}})^2$$

(the limit is in the sense of local uniform convergence in probability). For a martingale with jumps $\Delta M_s = M_s - M_{s-}$, $M_{s-} = \lim_{r<s,\, r\to s} M_r$, and continuous part M_t^c one has

$$\sum_{s\leq t} (\Delta M_s)^2 + \langle M^c\rangle_t = \lim_{|\Pi|\to 0} \sum_{i=1}^{n} (M_{t_i} - M_{t_{i-1}})^2.$$

For instance, one-dimensional Brownian motion $(W_t)_{t\geq 0}$ satisfies $\langle W\rangle_t = t$, and for a Poisson process $(N_t)_{t\geq 0}$ of intensity λ and the associated martingale $M_t = N_t - \lambda t$ one has $\langle M\rangle_t = \lambda t$. In fact, *Lévy's martingale characterization of Brownian motion* says that a one-dimensional Brownian motion is the only (local) martingale in continuous time and with continuous paths such that $\langle W\rangle_t = t$ and $W_0 = 0$. A result by Doeblin, Dambis and Dubins–Schwarz tells us that every L^2-martingale $(M_t)_{t\geq 0}$ with continuous paths can be embedded into a (suitable) Brownian motion by $M_t = M_0 + W_{\langle M\rangle_t}$.

In the guise of Itô's isometry (see below) quadratic variations can be used to define the stochastic integral driven by a martingale $(M_t)_{t\geq 0}$ with filtration $(\mathcal{F}_t)_{t\geq 0}$. A *simple* or *elementary process* $(Y_t)_{t\geq 0}$ is a bounded càdlàg process with finitely many values

$$Y_t(\omega) = \sum_{i=0}^{n-1} \eta_i(\omega)\mathbb{1}_{[t_i,t_{i+1})}(t), \quad t \geq 0,\ \omega \in \Omega,$$

where $t_0 = 0 < t_1 < \ldots < t_n$ and η_i is a $\mathcal{F}_{t_i}$ measurable random variable. Then

$$\begin{aligned} Y \bullet M_t &= \int_0^t Y_{s-}\, dM_s \\ &:= \sum_{i=0}^{n-1} \eta_i \left(M_{\min(t_{i+1},t)} - M_{\min(t_i,t)}\right) \end{aligned}$$

defines the Itô integral and $(Y \bullet M_t)_{t\geq 0}$ is an L^2-martingale; its compensator $(\langle Y \bullet M\rangle_t)_{t\geq 0}$ is given by $\int_0^t |Y_{s-}|^2\, d\langle M\rangle_s$. The compensator property yields the *Itô isometry*

$$\mathbb{E}\left\{\left(\int_0^t Y_{s-}\, dM_s\right)^2\right\} = \mathbb{E}\int_0^t |Y_{s-}|^2\, d\langle M\rangle_s$$

which allows one to extend the stochastic integral by continuity to all càdlàg integrands such that Y_t is $\mathcal{F}_t$ measurable, $t \geq 0$, and $\mathbb{E}\int_0^T |Y_{t-}|^2\, d\langle M\rangle_t < \infty$.

The development of stochastic analysis gave rise to further generalizations of martingales: Local martingales and semi-martingales (Revuz and Yor, 1999; Rogers and Williams, 1987/94). Martingales also turned out to be a powerful tool in characterizing and constructing Markov processes. The key role in these *martingale problem* approaches (Stroock and Varadhan, 1997; Ethier and Kurtz, 1986; Böttcher, Schilling and Wang, 2013) is the observation that the martingales appearing in Examples g) and h) characterize the underlying Markov chains and processes.

Applications

Martingales have found many important applications also outside of probability theory.

One of the most prominent applications is in mathematical finance in the theory of *arbitrage* and *risk-neutral pricing* of financial assets (Bingham and Kiesel, 1998). In a market, an arbitrage opportunity is a trading strategy that has a positive probability of winning and zero probability of losing money. It is a money-making machine, the ultimate unfair game. On the other hand, a martingale represents a fair game and will not allow for arbitrage. Since one can always invest in riskless interest-bearing bonds, it is clear that one has to discount prices by a numéraire, e.g. to divide asset prices by the price of the riskless bond. A *martingale measure* is a probability measure $\mathbb{Q}$ such that the discounted asset prices form a martingale. In a real market, no player (with their own beliefs, expressed by the player's own probability measure $\mathbb{P}$, sometimes called the real-world probability measure) expects the asset prices to be martingales since no one gains by betting on a martingale. On the other hand, in many situations there is an equivalent (with respect to $\mathbb{P}$) martingale measure $\mathbb{Q}$ such that the discounted asset prices become martingales under the conditional expectations induced by $\mathbb{Q}$. By *the fundamental theorem of asset pricing* the existence of a martingale measure guarantees that there is no arbitrage. Moreover, the market is complete (i.e. each derivative can be 'hedged' or 'replicated' by a trading strategy) if the martingale measure is unique. In a discrete-time setting the existence of an equivalent martingale measure is even equivalent to an arbitrage-free complete market. In the classical Black–Scholes model where the asset prices are modelled by a geometric Brownian motion, it is always possible to find a martingale measure $\mathbb{Q}$ with the help of Girsanov's theorem. In order to calculate the fair price of a derivative (e.g. a share option) one performs the following steps: First one determines the pay-off of the derivative, then one constructs an equivalent martingale measure $\mathbb{Q}$, finally the discounted price process of the derivative is the expectation (under $\mathbb{Q}$) of its payoff; since everything is, by construction, a martingale, no arbitrage is possible.

Within mathematics, martingales turn out to be useful in the study of almost everywhere convergence. Typical applications include the *geometry of Banach spaces* and *Littlewood–Paley theory*. An unconditional basis in a Banach space B is a basis $(e_n)_{n\in\mathbb{N}}$ such that every rearrangement of the series $\sum_n a_n e_n$ converges to the same limit $b \in B$. The spaces $L^p((0,1),dx)$ have the Haar functions as unconditional basis if, and only if, $1 < p < \infty$. This can be checked by martingale arguments involving the Burkholder–Davis–Gundy inequalities and so-called martingale difference sequences: $(e_n)_{n\in\mathbb{N}}$ is a martingale difference sequence if $\mathbb{E}(e_{n+1} \mid \sigma(e_1,\ldots,e_n)) = 0$. In fact, the Haar functions are an unconditional basis for every rearrangement-invariant function space which is isomorphic to a Banach space with an unconditional basis. This is the basis for many structure and embedding results for Banach spaces.

In the theory of vector measures, the L^1-martingale convergence theorem (for Banach space valued martingales) is equivalent to the Radon-Nikodým property of a Banach space B, i.e. every function $f : [0,1] \to B$ of bounded variation is almost everywhere differentiable or, equivalently, every separable subspace of B has a separable dual.

If a function space has an orthonormal basis consisting of martingale differences, one can often use the pointwise martingale convergence theorem and maximal inequalities to obtain pointwise almost everywhere convergence results (Schilling, 2005, Chapter 24). In harmonic analysis (Stein, 1993, pp. 184–192) there are deep connections with (maximal) inequalities for the Littlewood–Paley functions and the Hilbert transform leading to proofs of Carleson's famous result that the Fourier series of a square-integrable function converges almost everywhere. In the theory of BMO (bounded mean oscillation) spaces, martingale maximal inequalities are most powerful tools. For example, one can prove the John–Nirenberg estimates for BMO spaces and Fefferman's theorem: BMO is the dual of the Hardy space H^1.

Historical notes

Jean Ville (Ville, 1936, 1939) introduced the term 'martingale' into probability. In his 1936 definition he considered only martingales of the form $X_n = e_1Y_1 + \ldots + e_nY_n$ (as in the second paragraph of this article) which he generalized in the monograph 1939 both in the time-discrete and time-continuous case. Ville used martingales to discuss *the concept of a collective, upon which many mathematicians found the theory of probability.* [...] *This leads* [...] *to a new definition of collective* (Doob, 1939). Ville's approach to von Mises' theory of collectives interprets the complete irregularity in a random sequence as the impossibility of a successful gambling system of 'martingale' type (in the classical sense of the first paragraph). In fact, (von Mises, 1932, pp. 157–158) already contains elements of this line of thought, and martingales are explicitly mentioned – however, as gambling systems only. The first general definition of martingales using conditional probabilities and the first ever martingale convergence theorem (*Lévy's* 0–1 *law*) can be found in (Lévy, 1937, pp. 122–123 (definition), p. 129, Theorème 41). Ville's 1939 definition follows (Lévy, 1937) and it is very close to the first 'modern' definition in (Doob, 1940) who calls (3) *condition* $\mathcal{E}$ (seemingly using Lévy's slightly obscure diction). It was mainly Doob who pioneered the theory and applications of martingales and his groundbreaking monograph (Doob, 1953) contains most of the basic results on martingales. Subsequently, major contributions to the theory, in particular in the connection with stochastic calculus, are mainly due to the French and Japanese schools of probability, see (Meyer, 1966), (Dellacherie and Meyer, 1980), (Kunita and Watanabe, 1967) and (Ikeda and Watanabe, 1981).

Etymology

According to the *Oxford English Dictionary* (OED) `www.oed.com`[2] the etymological origin of the word martingale is unclear. Most likely it derives from *martengala*, feminine of martegal/martengal, an inhabitant of Martigues (a small town in south-eastern France, département Bouches-du-Rhône). The word martingale has been used in the following senses:

1. *Horse riding.* A strap or harness preventing the horse from rearing or throwing its head back.
2. *Nautical.* Any part of a rigging strengthening the jib boom and bowsprit of a ship.
3. A dog collar made of two loops. When the animal tries to pull, the loop around the dog's neck contracts without the choking effect of a slip collar.
4. *Clothing.* (a) A loose half-belt or strap on the back of a garment. (b) A pair of trousers.
5. *Games.* A gambling system as described in the introduction to this article.
6. *Mathematics.* A stochastic process as described in this article.

The earliest use is in French as in 1 and 4(b). The OED states that '[t]he application may arise from a belief that the inhabitants of Martigues, a remote town, were eccentric and naive; hence also the application to an apparently foolish system of gambling. Sense 2, however, is probably attributable to the former importance of Martigues as a port and ship-building centre.' An in-depth discussion on the origins of the word martingale can be found in Mansuy (2009).

Acknowledgements

I am grateful for the help from the following friends and colleagues: Prof. Nick Bingham (London), Dr. Björn Böttcher (Dresden), Dr. Katharina Fischer (Dresden), Prof. Charles Goldie (Brighton), Prof. Niels Jacob (Swansea), Ms. Franziska Kühn (Dresden), Dr. Jan Rudl (Dresden) and Mr. Michael Schwarzenberger (Dresden). Their input helped to improve the presentation of this article.

References

Bingham NH and Kiesel R (1998) *Risk-Neutral Valuation.* London: Springer.

Böttcher B, Schilling RL and Wang J (2013) *Lévy Matters III – Lévy-Type Processes: Construction, Approximation and Path Properties.* Lecture Notes in Mathematics vol. 2099, Cham: Springer.

Chow YS and Teicher H (1997) *Probability Theory. Independence, Interchangeability, Martingales* (3rd ed). New York (NY): Springer.

Dellacherie C and Meyer P-A (1980) *Probabilités et Potentiel. Chapitres V à VIII: Théorie des Martingales.* Actualités Scientifiques et Industrielles 1385, Paris: Hermann. [English translation: (1982) *Probabilities and Potentials B: Theory of Martingales.* Amsterdam: Elsevier.]

Doob JL (1939) Book review of (Ville, 1939). *Bull. Am. Math. Soc.* 45: 824.

Doob JL (1940) Regularity properties of certain families of chance variables. *Trans. Am. Math. Soc.* 47: 455–486.

Doob JL (1953) *Stochastic Processes.* New York (NY): John Wiley & Sons.

Ethier SN and Kurtz TG (1986) *Markov Processes. Characterization and Convergence.* New York (NY): John Wiley & Sons.

Gorroochurn P (2012) *Classic Problems of Probability.* Hoboken (NJ): John Wiley & Sons.

Ikeda N and Watanabe S (1981) *Stochastic Differential Equations and Diffusion Processes.* Amsterdam: North-Holland.

Kunita H and Watanabe S (1967) On square integrable martingales. *Nagoya Math. J.* 30: 209–245.

Lévy P (1937) *Théorie de l'Addition des Variables Aléatoires.* Monographies des Probabilités vol. 1, Paris: Gauthier–Villars.

Mansuy R (2009) The origins of the word "martingale". *Journal Eléctronique d'Histoire des Probabilités et de la Statistique* 5 (1): 1–10.

Meyer P-A (1966) *Probabilités et Potentiel.* Actualités Scientifiques et Industrielles 1318, Paris: Hermann. [English translation: (1966) *Probability and Potentials.* Waltham (MA): Blaisdell.]

Revuz D and Yor M (1999) *Continuous Martingales and Brownian Motion* (3rd ed). Berlin: Springer.

Rogers LCG and Williams D (1987/94) *Diffusions, Markov Processes and Martingales 1, 2.* Chichester: John Wiley & Sons.

Schilling RL (2005) *Measures, Integrals and Martingales.* Cambridge: Cambridge University Press.

[2] Accessed online 20/July/2014.

Schilling RL and Partzsch L (2014) *Brownian Motion* (2nd ed). Berlin: de Gruyter.

Stein EM (1993) *Harmonic Analysis. Real-Variable Methods, Orthogonality, and Oscillatory Integrals.* Princeton (NJ): Princeton University Press.

Stroock DW and Varadhan SRS (1997) *Multidimensional Diffusion Processes.* Berlin: Springer.

Ville J (1936) Sur la notion de collectif. *Comptes Rendus Hebd. Acad. Sci. (Paris)* 203: 26–27.

Ville J (1939) *Étude Critique de la Notion de Collectif.* Monographies des Probabilités vol. 3, Paris: Gauthier–Villars.

von Mises R (1932) Théorie des probabilités. Fondements et applications. *Ann. Inst. H. Poincaré* 3: 137–190.

Williams D (1991) *Probability with Martingales.* Cambridge: Cambridge University Press.

René L. Schilling
Institut für Mathematische Stochastik
Fachrichtung Mathematik
TU Dresden
01062 Dresden, Germany
`rene.schilling@tu-dresden.de`