What is a Martingale in Mathematics and Why Is It Important for Finance?

written by Mathieu Delehaye on Functor Network
original link: https://functor.network/user/2888/entry/987

In probability theory, an *outcome*, *sample*, or *path* is a specific sequence of values a process can take over time, which can be observed afterward. For example, the daily closing prices observed for Nvidia stock from 1st January to 31st December 2024 form an outcome.

If we travel back in time to 1st January 2024 and try to predict the stock price one year later (i.e., on 31st December), we might consider all the possible paths the Nvidia stock price could follow over that one-year period. Even though the number of such paths is infinite, we can associate a "probability" to each of them (or more precisely: a probability *density*). By multiplying the stock value on 31st December, which would be observed on each path, by the probability of that path, we obtain a forward-looking average for the stock price called the *expectation*:

$E[S(T)] = \sum_{\omega \in \Omega} P(\omega) \cdot S(T, \omega) = \int_{\omega \in \Omega} S(T, \omega) \, f_{S(T)}(\omega) \, d\omega$

where:

- T is the date 31st December 2024,
- $S(T, \omega)$ is the stock price at that time on a given path ω,
- Ω is the set of all possible paths.

Since the number of paths is infinite, we replace the summation $\sum_{\Omega}$ with an integral $\int_{\Omega} \ldots \, d\omega$, and the discrete probability measure $P(\omega)$ is replaced by a probability density function $f_{S(T)}(\omega)$ for the values of $S(T)$.

All paths start from the same value on 1st January 2024, or time 0, which we denote as $S(0)$.

The stock price process is said to be a *martingale* if the following holds:

$E[S(t) \mid S(0) = x] = E[S(t)] = S(0) = x, \quad \forall t \in [0, T]$

Here, the vertical bar $\mid$ in the first expectation indicates that a specific *condition* has already been observed — namely, the initial stock price. Given this condition, the expected future price at any time $t \in [0, T]$, averaged over all paths, equals the initial price x. Therefore, a martingale is a process that shows no tendency to either increase or decrease over time.

The name *martingale* comes from a betting strategy popular in 18th-century France: at each turn, a player bets an amount b and tosses a coin. If she wins, she earns b and the game ends. If she loses, she plays another turn, doubling her bet each time. When she eventually wins, she recovers all her previous losses and secures a net gain of b.

For example, if she loses four times before winning, her final profit will be:

$16 - 8 - 4 - 2 - 1 = 1 = b$

With a fair coin, the probability of *not* winning after n tosses is 0.5^n, which tends toward 0 as $n \to \infty$. Therefore, winning b seems like a sure thing. However, this is only theoretical, as the player will often go bankrupt before recovering her losses. With an initial capital of C and a base bet b, the maximum number of losses she can afford before exhausting her capital is:

$N = \log_2 \left(\frac{C}{b} + 1\right)$

Moreover, casinos typically impose a maximum betting limit (to prevent infinite doubling), and the probability of winning is usually less than 0.5.

If the player's game lasts N turns, there are two possible scenarios:

- She loses all her bets, which corresponds to a total loss of $\sum_{i=0}^{N-1} 2^i(-b) = -b(2^N - 1)$, with probability q^N, where q is the probability of losing a single toss,
- Or she wins once before reaching N tosses, which happens with probability $1 - q^N$.

The expected value of her gains and losses, given a maximum of N tosses, is:

$E[V(N)] = (1 - q^N) \cdot b + q^N \cdot \left(-(2^N - 1)b\right) = b\left(1 - q^N 2^N\right)$

This expectation equals 0 when $q = 0.5$:

$E[V(N)] = b\left(1 - 0.5^N \cdot 2^N\right) = b(1 - 1) = E[V(0)] = 0$

This emphasizes that this is a zero-profit game, regardless of the maximum number of tosses N — except in the purely theoretical case where $N \to \infty$.

Contrary to our earlier example — which was, unfortunately, misleading — stock prices are *not* martingales. They tend to *increase* over time (except during disruptive events like market crashes), leading to a relationship such as:

$E[S(t) \mid S(0) = x] = E[S(t)] \geq x, \quad \forall t \in [0, T]$

This is a special case of a *submartingale*, which more generally refers to a process that is bounded from below but not necessarily from above, often resulting in values that grow over time.

Indeed, stock prices are generally modeled using a *geometric Brownian motion*, given by the expression:

$S(t) = S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W(t)\right)$

where:

- $S(0)$ is the initial stock price on 1st January 2024,
- μ is the drift coefficient, which reflects the average rate of growth of the stock price over time,
- σ is the standard deviation of the stock price, also known as volatility,
- $W(t)$ is a *standard* Brownian motion — that is, a random value observed at time t, which follows a normal distribution with mean 0 and variance t (denoted $W(t) \sim \mathcal{N}(0, t)$).

Therefore, the random value $Y(t) = \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W(t)$ follows a normal distribution with mean $\left(\mu - \frac{\sigma^2}{2}\right) t$ and variance $\sigma^2 t$, that is:

$Y(t) \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right) t, \, \sigma^2 t\right)$

If we compute the expectation of the stock price, we obtain:

$E[S(t) \mid S(0) = x] = E[S(t)] = E[S(0) \cdot \exp(Y(t))] = x \cdot \exp(\mu t)$

We see that $E[S(t)] = x$ when the drift $\mu = 0$, and $E[S(t)] \geq x$ when $\mu \geq 0$. In practice, since μ is generally positive, the stock price process tends to increase over time, which makes it a submartingale.