Research Report on Stochastic Processes: Martingales, Renewal Processes, and Brownian Motion

**What is a Stochastic Process?**

A random process is defined as a family of random variables in a probability space which are indexed by time.

**In other words**: A stochastic process is just a way of describing how something random evolves over time. At each point in time, you can think of the process as producing a random variable, and together these random variables trace out a path that represents one possible realization of the process.

**Martingales:**

**Definition:**

Martingale is a stochastic process that formalizes the idea of a “fair game.” Given a sequence of random variables adapted to an information structure (also known as filtration) the martingale condition is:

In other words, the best prediction of the next outcome, given everything we know so far, is simple the present value.

**Fundamentals:**

To fully understand Martingales, we’ll need to cover some prerequisite concepts.

1. Expectation   
     
   The expectation of a random variable is the theoretical “average” outcome if the random process were repeated infinitely many times. So, for a continuous random variable with density , or for a discrete random variable…

Supposed you flip a fair coin: heads = 1, tails = 0. The expected value of that fair coin is:

This means if you flipped this coin many times, on average half of the outcomes would be heads.

1. Conditional Expectation  
     
   Conditional expectation is the best estimate of given knowledge of . It refines the expectation by restricting the “universe” of possibilities.  
     
   Now imagine you’ve already flipped one coin. If it came up heads, then given that knowledge, the expected value is 1; if tails, its 0. Formally, just says: once you know the past, your estimate of the future reflects that information.

This reliance on past information is the backbone of martingale definitions: tomorrow’s best estimate is today’s value.

1. Filtration

A filtration is a growing collection of -algebras[[1]](#footnote-1), where each represents all information available up to time n.   
  
Continuing with the coin toss example, after two flips, contains the outcomes of both tosses, while contains only the first. As we continue flipping the coin n times, the information pool of the filtration only grows…

1. Integrability

For martingales, we assume . Thise ensures our conditional expectations are well-defined and avoids “exploding” processes.

Bringing this all together, we can express the martingale property:

This expresses the fair game principle alluded to before: given all past information (), your expected next value is exactly the current value. In the coin flipping example, if you bet $1 on each toss, your expected future wealth given the past is always equal to your current wealth. If this seems initially unintuitive, just remember that the expected value of the fair is **not** equivalent to a guaranteed outcome. It just means that’s over many possible futures and iterations of the fair coin game, your future wealth is expected to be your current wealth.

**Martingale Deep Dive**

Martingales go beyond the simple fair game definition we explored above. They provide a framework to bifurcate processes into predictable and unpredictable components. And through Doobs Decomposition we can outline **any** integrable, adapted process by decomposing it below:

Where is a martingale and is a predictable, finite-variation process (or what is called the “drift”). In essence, the martingale functions as the “noise core” from any stochastic process. Once you’ve stripped away all of the systematic, forecastable structure in the model, what you have left is the unpredictable randomness, our martingale, which is only a measure of the information we have up until that point. To clarify the future unpredictability of the martingale condition, consider this difference sequence:

This difference sequence effectively states that given everything we know (the filtration variable), we cannot predict a change between the current and next state. [[2]](#footnote-2)

Where we start to see changes in the n+1 case is when we reintroduce the drift predictor. Upward and downward drifts are referred to as Submartingales and Supermartingales, respectively.

Submartingales: Models processed with downward drift.

Supermartingale: Models processed with downward drift.

**Renewal Processes:**

**Definition:**

A renewal process models system that reset after random intervals. A classic example is machine replacement: when a part fails, it is replaced, and the system renews.

**Fundamentals:**

Just as with Martingales, understanding Renewal Processes requires some prerequisite knowledge.

1. Poisson Processes as a Starting Point

A Poisson process models random arrivals of events over time, with properties:

1. Inter-arrival times are exponential with mean (see Exponential Distribution below)
2. Memoryless property:
3. Future events are independent of the past

Note that the memoryless property implies that future arrivals are independent of the past. Formally since , the probability of waiting at least an additional t unit of time is unaffected by how much time s has already passed without an event. In other words, no matter when the last event occurred, only the next waiting period matters.

1. Exponential Distribution

Now, when we say inter-arrival times are exponential, we mean that they follow an exponential distribution: a process in which events occur continuously and independently at a constant average rate.

The probability density function (pdf) of an exponential distribution is

Here (or the “rate parameter”) is the parameter of the distribution, supported on the interval of .

The mean or expected value of an exponentially distributed random variable with rate parameter is given by

Imagine buses arriving at your bus stop: if buses arrive with exponentially distributed times between them, the process is memoryless. If you’ve already waited 5 minutes, your expected additional wait is the same.

1. Renewal Generalization

Renewal processes actually relax the above exponential assumption: inter-arrival times T1, T2, … are independently and identically distributed (i.i.d.) but from an arbitrary distribution F with mean .

1. This breaks the memorylessness: past events now influence how long until the next event.
2. The process N(t) counts how many renewals have occurred by time t.

Going back to the bus example, now suppose the buses have schedule but variable arrival times – sometimes 10 minutes, sometimes 15, sometimes 5. The inter-arrival times are i.i.d. but not exponential.[[3]](#footnote-3)

1. Renewal Function

captures the expected number of renewals up to time t. The Elementary Renewal Theorem (ERT) states:

So, on average, the long-run renewal rate stabilizes at the reciprocal of the mean inter-arrival time.

**Renewal Process Deep Dive**

As alluded to above, renewal processes extend the Poisson process into settings where inter-arrival times are not exponential, usually for modeling realistic lifetime or failure/replacement systems.

To reiterate, the Elementary Renewal Theorem states that the expected number of renewals, , divided by time , approaches the reciprocal of the mean interarrival time, , as time goes to infinity. [[4]](#footnote-4)

But what if our first inter-arrival time has a different distribution G from the overarching distribution F of later inter-arrival times? In this case, our renewal process is called “delayed”. As you might expect, the delayed start permanently affects the early-time behavior before long-run averages kick in. Formally,

In other words, this states that the total expected number of renewals by time t is the sum of the change that the first renewal happens before t and the expected number of renewals generated after that first renewal, averaged over all possible first-renewal times.

There are many applications to this specific variation of renewal processes, some of which are explored below.

At times, reward systems can be attached to renewal processes, giving us the Renewal Reward theorem.

Where the cumulative reward by time t states…

The Renewal Reward Theorem tells you, on average, you get reward per cycle and cycles occur at rate .

Therefore, your reward per unit time approaches as approaches infinity.

Another important thing to note is that a “*reward*” does not necessarily imply a positive gain. In this context, a “reward” can represent any measurable quantity associated with a renewal cycle. For instance, a “reward” could capture costs incurred when a machine part is replaced, or downtime lost during service.

Now you might have caught on that although the long run stabilization of the average number of renewals per unit time settles at , there will still be fluctuations around that average. And the variance of these fluctuations actually happens to be Gaussian in the long run, with a variance that grows proportionally to t.

The renewal process isn’t deterministic or stuck to it’s long run average; it wiggles due to randomness for each arrival time. The above statement just outlines that the normalization factor scales with t and that the gaussian limit tells us the deviations are symmetric and bell-shaped.

**Brownian Motion:**

**Definition:**

Brownian Motion (also called the Wiener process) is a continuous-time stochastic process that models random continuous movement. It arises as a scaling limit of simple symmetric random walks. Formally, a process is a Brownian motion if:

1. Has independent and stationary increments
2. For each
3. Paths are continuous almost surely[[5]](#footnote-5).

For reason that will be explained later, Brownian motion can actually be thought of as continuous-time martingales.

**Fundamentals:**

1. Random Walk Foundation

Imagine a particle on a number line that takes one step every second.

* With probability ½, it moves right by distance
* With probability ½, it moves left by distance

Formally, this can be defined as:

Here Sn is the position after n steps, and each Xi is an independent, identically distributed step.

But now we want to describe this walk not just after n steps, but as a function of continuous time. Supposed each step occurs after a time increment . Then after the position is:

We’ve now scaled by linking the number of steps n to elapsed time t.

As your continuous, time scaled random walk marches forward, you must take into consideration the walk’s variance growth. To understand variance growth, understand that it is technically possible (although unlikely assuming equal probability to move in any given direction) for your random walk to “drift” or randomly sway in one specific direction. Imagine our particle having a long series of only right or only left movements; this behavior should be captured in variance growth.

The variance of S(t) is defined as:

Where represents the n number of time increments (infinite in the fully continuous case). And as you can see, assuming t is very large, it could capture a wide range of viable locations for our particle, away from .

To get a non-degenerate continuous limit[[6]](#footnote-6), we must choose and so that variance grows proportionally to time t. The critical scaling is:

Here, is a constant (our volatility parameter). If shrinks too slowly or quickly relative to , the variance would blow up or collapse to zero. The square-root scaling balances the two so that the limiting variance is finite and linear in time.

Now let’s assume that and approach 0, the rescaled process now converges in distribution to a continuous-time process with properties:

* 1. Increments are independent and normally distributed:
  2. Paths are surely continuous

This limiting process takes us from jagged steps in our time interval to seemingly smooth, continuous movement, defining Brownian motion. Technically, our path is smooth, yes, but also infinitely rough – Brownian paths are continuous everywhere but differentiable nowhere.

1. Stationary and Independent Increments

Stationarity: The distribution of depends only on the time gap s, not the absolute location t. This follows because the random walk increments are identically distributed at each step.

Independence: Non-overlapping increments of the random walk are independent; in the limit, Brownian increments inherit this property.

Together, these two give the memoryless in distribution feature of Brownian motion: knowing what happened up to time t doesn’t affect the distribution of what happens I the next interval – only the variance grows with elapse time.

1. Gaussian (a.k.a. Normal) Distribution

Brownian increments are normally distributed:

This links Brownian motion with Gaussian theory: the entire path distribution is determined by means, variances, and covariances.

Another way to see the Gaussian nature of Brownian motion is through its covariance. Seeing as covariance measure how two random variables move together, we’re therefore asking how is the value at time s related to the value at time t?

If we look at for some , you can naturally split it into:

As mentioned above, a defining property of Brownian motion is that increments over disjoint intervals are independent.

* So and are independent.
* And the increment has mean 0.

Now, the covariance function

can be expanded using the split above for B(t):

This simplifies:

The entire second term vanishes as the B(s) and B(t) are independent of each other, and the expected value of B(s) reduces to zero, leaving:

Therefore:

If we instead assumed that , the same logic would return , meaning the general covariance formula is:

The covariance between two Brownian motion values depends only on how *shared history* they have. The smaller of the two times, , is the overlap of the intervals [0,s] and [0,t].

**Brownian Motion Deep Dive**

As mentioned earlier, Brownian paths are continuous, and *almost surely* nowhere differentiable; zooming in would reveal their “infinite roughness.” Because of this infinite roughness, ordinary calculus cannot handle Brownian paths, leading to the development of Itô calculus[[7]](#footnote-7).

Brownian paths also exhibit Markov and Martingale properties, hence them being referred to as continuous time Martingales. Nevertheless, Brownian motion in its distribution:

In words, the Markov property states that the probability that the Brownian motion is in some set A at future time , given everything we know up to time t, depends on only its present position B(t), not on the full history of how it got there.

The Martingale property asserts, just as it did above, that the future expected value of a future time given its filtration is equal to the current value of the Brownian path:

Another fascinating property of Brownian motion is that it’s “self-similarity.” No matter how far you in or out you zoom in time and space, the process looks statistically the same (not to be confused with geometrically the same). Below is the mathematical expression that describes this property:

means you look at the Brownian path but speed it up (for c > 1) or slow it down (for c < 1) in the time axis. However, means you rescale the space axis (the vertical displacements) by . The equal sign in the middle indicates that these two statements are “equal in distribution”, they have the same law even though they are not literally the same sample paths.

Take a 1-minute sample of Brownian motion and zoom in on it, then compare it with a 10-minute sample that’s been rescaled appropriately. You’ll notice that they’re indistinguishable in distribution.

1. -algebras are just the formal way to mathematically describe all possible, distinguishable events at this point in time. [↑](#footnote-ref-1)
2. Also consider the Optional Stopping Theorem which is yet another way to digest the martingale condition. The stopping theorem states that, under certain conditions, the expected value of a martingale at a stopping time is equal to its initial expected value: . Imagine modeling the wealth of a gambler participating in a fair game: On average, nothing can be gained by stopping play based on past information. [↑](#footnote-ref-2)
3. This is actually the basis of the Inspection or Waiting-Time paradox. Let’s say the bus arrival times are uniformly distributed between 5 and 15 minutes. If you arrive at the bus stop at a random time, you’re more likely to arrive during a longer bus interval than a shorter one. If this is still confusing, imagine a timeline that records a full day’s bus arrival schedule. Because we can assume that we will have roughly the same number of early and late inter-arrival times during the day, the late arrival windows take up more real estate on the timeline. This means if you were to randomly choose when on that timeline to show up to the bus stop, you’ll be more likely to choose a arrival window that takes up more space on the timeline – a late (> 10 min) bus arrival. [↑](#footnote-ref-3)
4. The Strong Law of Large Numbers (SLLN) also applies here, a more rigorous version of ERT. It’s a semantic distinction where ERT only gives us convergence in expectation, SLLN shows that almost every sample path of the renewal process will exhibit stabilization given enough time. [↑](#footnote-ref-4)
5. The specification of “almost surely” (a.s.) is a probability property that maintains that something is true with probability 1 but might fail on a very specific set of outcomes of probability 0. In the context of Brownian motion, paths are continuous almost surely, any simulated sample path will be continuous in time with probability one. However, in an abstract, theoretical sense, there *do exist* exceptional paths that are not continuous, yet so rare that they are deemed unobservable. [↑](#footnote-ref-5)
6. A non-degenerate continuous limit essentially describes the situation where a sequence of random variables, after appropriate scaling and centering, converges to a probability distribution that is spread out over a range of values and has a continuous probability density function, rather than being concentrated on a single point. [↑](#footnote-ref-6)
7. Itô calculus (which is a stochastic generalization of the Riemann-Stieltjes integral) is designed to handle integration and differentiation with respect to Brownian motion. The framework defines the stochastic integral as a limit of sums evaluated at the left endpoints, making it adapted to the filtration. The fundamental result of Itô’s Lemma states that if follows a stochastic differential equation driven by Brownian motion, then ) evolves with an extra correction term involving the second derivative of f, reflecting the quadratic variation of Brownian motion. This lemma acts as the backbone of modern stochastic differential equations and financial mathematics. [↑](#footnote-ref-7)