4 The Brownian motion

4.1 Finite-dimensional distributions, Gaussian processes

A continuous-time random process with time parameter $t \in \mathbb{R}_+$ is a family of random variables $(X(t),\ t \geq 0)$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The random function $t \mapsto X(t)$ is called a path of the process. For any selection of distinct times $0 \leq t_1 < \cdots < t_k$, $(X(t_1), \dots, X(t_k))$ is a random vector characterised by some joint probability distribution

$$\mu_{t_1,\dots,t_k}(A) = \mathbb{P}((X(t_1), \dots, X(t_k)) \in A), \quad A \in \mathcal{B}(\mathbb{R}^k). \tag{1}$$

(which may or may not have density). Each $\mu_{t_1,\dots,t_k}$ is a probability measure on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, and these measures are *consistent*:

$$\mu_{t_1,\dots,t_k}(A \times \mathbb{R}) = \mu_{t_1,\dots,t_{k-1}}(A), \quad A \in \mathcal{B}(\mathbb{R}^{k-1}).$$

In practice, a starting point for the construction of random process is a family of such consistent probability measures $\mu_{t_1,\dots,t_k}$. Let $\mathbb{R}^{[0,\infty)}$ be the space of functions $x : \mathbb{R}_+ \to \mathbb{R}$, with Borel σ-algebra $\mathcal{B}(\mathbb{R}^{[0,\infty)})$ generated by the cylider sets $\{x : x(t) \in A\}$ where t is fixed and $A \in \mathcal{B}(\mathbb{R})$. By Kolmogorov's extension theorem (see Lecture 1) there exists a unique probability measure $\mathbb{P}$ on the measurable space $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R}^{[0,\infty)}))$ such that (1) holds.

It is more challenging to verify if there exists a *version* of the process whose paths have certain properties like continuity. Two random processes $(X(t),\ t \geq 0)$ and $(Y(t),\ t \geq 0)$ are versions of one another if $\mathbb{P}(X(t) = Y(t)) = 1$ for all t.

Definition 4.1. A random vector $(\xi_1, \dots, \xi_k)$ has *a multivariate normal distribution* (we also call such vector Gaussian) if each linear combination $\sum_{i=1}^k a_i \xi_i$ has a one-dimensional normal distribution.

This definition has advantage over specifying the joint density of $(\xi_1, \dots, \xi_k)$ in that the existence of the density is not required. For instance, (ξ, ξ) is Gaussian if ξ is. Also, the following properties are immediate:

(i) vector $(\xi_1, \dots, \xi_k)M$ is Gaussian for any $k \times n$ matrix M,

(ii) distribution $(\xi_1, \dots, \xi_k)$ is uniquely determined by the mean vector $(\mathbb{E}\,\xi_1, \dots, \mathbb{E}\,\xi_k)$ and the $k \times k$ covariance matrix with entries $\mathrm{Cov}(\xi_i, \xi_j)$,

(iii) if $\mathrm{Cov}(\xi_i, \xi_j) = 0$ for all $i \neq j$, then $\xi_1, \dots, \xi_k$ are mutually independent.

Definition 4.2. A random process $(X(t),\ t \geq 0)$ is called Gaussian, if the vector $(X(t_1), \dots, X(t_k))$ is Gaussian for any selection of distinct times $t_1, \dots .t_k$.

The probability law of a Gaussian process is uniquely determined by the mean function $m(t) = \mathbb{E}\,X(t)$ and the covariance function $c(s,t) = \mathrm{Cov}(X(s), X(t))$; this is because of property (ii) of the multivariate normal distribution of each vector $(X(t_1), \dots, X(t_k))$.

4.2 Processes with stationary independent increments

A stochastic process $(X(t), t \geq 0)$ has stationary increments if the distribution of $X(t) - X(s)$ depends only on $t - s$ for $0 \leq s \leq t$. It has independent increments if $X_{t_{j+1}} - X_{t_j}, 1 \leq j \leq k$, are independent random variables for $0 \leq t_1 < t_2 < \cdots < t_k$ and any $k \geq 1$.

A process with stationary independent increments is called a Lévy process.

Example Poisson process $(N(t), t \geq 0)$ with parameter $\lambda > 0$ is a Lévy process, with $N(t) - N(s)$ having the Poisson$(\lambda(t-s))$ distribution.

A more general type of process is

Example Compound Poisson process $X(t) = \sum_{j=1}^{N(t)} Y_j$ where $(N(t), t \geq 0)$ is a Poisson process, independent of $Y_1, Y_2, \cdots$ which are i.i.d.

These processes have right-continuous paths with jumps at 'arrivals times' of the Poisson process.

Proposition 4.3. *For random process $(X(t),\ t \geq 0)$ the following are equivalent:*

(i) $(X(t),\ t \geq 0)$ *is a Lévy process with* $X(t) \sim \mathcal{N}(0,t)$ *for* $t \geq 0$.

(ii) $(X(t),\ t \geq 0)$ *is Gaussian with* $\mathbb{E}\, X(t) = 0$ *and* $\mathrm{Cov}(X(t), X(s)) = s \wedge t$.

A Gaussian process in Proposition 4.3 is a 'Brownian motion in a wide sense', that is without concern about the continuity of paths.

4.3 Scaled random walk

The general random walk is a sequence of sums of i.i.d. (independent, identically distributed) random variables. Let $\xi_1, \xi_2, \ldots$ be i.i.d. rv's with $\mathbb{E}\, \xi_i = 0, \mathrm{Var}\ \xi_i = 1$. By the Central Limit Theorem (CLT), the random walk $S_n = \xi_1 + \cdots + \xi_n,\ n \in \mathbb{N}$, satisfies

$$\frac{S_n}{\sqrt{n}} \stackrel{d}{\to} B(1), \quad \text{where } B(1) \sim \mathcal{N}(0,1),$$

and the symbol $\sim$ in this context stays for 'the random variable has distribution...'. The convergence in distribution stated in the CLT means that

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \leq x\right) = \Phi(x) := \int_{-\infty}^{x} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy, \quad x \in \mathbb{R}.$$

We can interpolate the random walk to a continuous-time (piecewise-constant) random process

$$X_n(t) := \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}, \quad t \geq 0$$

(setting $X_n(0) = 0$), where $\lfloor z \rfloor$ is the integer part of z. The CLT above concerns the limit distribution of the sequence of random variables $X_n(1)$, as $n \to \infty$, which is the value of the random process at a sole time $t = 1$. This has a natural generalisation. Fix times $0 = t_0 < t_1 < t_2 < \cdots < t_k$, and note that the *increments*

$$X_n(t_i) - X_n(t_{i-1}) = \sum_{j=\lfloor nt_{i-1} \rfloor + 1}^{\lfloor nt_i \rfloor} \xi_j\,, \quad i = 1, \ldots, k,$$

are independent, because the sums for $i = 1, \ldots, k$ are over non-intersecting blocks of independent random variables. We now claim that there is a joint convergence

$$(X_n(t_1), \ldots, X_n(t_k)) \stackrel{d}{\to} (B(t_1), \ldots, B(t_k)) \tag{2}$$

to some Gaussian vector. Changing variables the latter is equivalent to the joint convergence

$$X_n(t_i) - X_n(t_{i-1}) \stackrel{d}{\to} B(t_i) - B(t_{i-1}), \;\; i = 1, \ldots, k,$$

which is indeed true by the CLT, for independent random variables $B(t_i) - B(t_{i-1}) \sim \mathcal{N}(0, t_i - t_{i-1})$.

It is natural to ask if this multivariate convergence (2) can be embedded into a more general theory, which would imply, in particular, that the maximum $\max_{t\in[0,1]} X_n(t)$ converges in distribution to maximum of some random process, appearing as a limit form of scaled random walks.

4.4 Definition of the BM

Definition 4.4. The Brownian motion (BM) *is a continuous-time random process* $(B(t),\ t \geq 0)$*, such that*

(i) $B(0) = 0$ a.s.,

(ii) *the paths* $t \mapsto B(t)$ *are continuous functions, almost surely,*

(iii) *the increments are independent and stationary,*

(iv) $B(t) \sim \mathcal{N}(0, t)$*,* $t > 0$.

For times $0 = t_0 < t_1, \cdots < t_k$, we have that $B(t_i) - B(t_{i-1}) \sim \mathcal{N}(0, t_i - t_{i-1})$ are independent rv's. Passing from the increments to the vector of values $(B(t_1), \ldots, B(t_k))$, we see that the vector is multivariate normal, therefore the BM is a Gaussian process. The mean function is $\mathbb{E}\, B(t) = 0$. The covariance function is $\mathrm{Cov}(B(s), B(t)) = s \wedge t$ (where $\wedge$ =minimum). The proof follows from (iii) and (iv): for $s < t$

$$\mathbb{E}[B(s)B(t)] = \mathbb{E}[B(s)(B(s) + (B(t) - B(s))] = \mathbb{E}[B(s)^2] + \mathbb{E}[B(s)(B(t) - B(s))] = \\ \mathrm{Var} B(s) + \mathbb{E}[B(s)]\, \mathbb{E}[B(t) - B(s)] = s + 0 = s.$$

Conditions (iii), (iv) can be replaced by the equivalent condition (Proposition 4.3)

(v) *the process is Gaussian with mean zero and* $\mathrm{Cov}(B(s), B(t)) = s \wedge t$.

To derive from (v) the independence of increments, compute for $0 \leq u < v \leq s < t$

$$\mathrm{Cov}(B(v) - B(u), B(t) - B(s)) = v \wedge t - v \wedge s - u \wedge t + u \wedge s = v - v - u + u = 0,$$

and recall that, for normal variables, independence holds precisely when the variables are uncorrelated. A similar calculation with account of (v) implies that $\mathrm{Var}(B(t) - B(s)) = t - s$.

There are some useful transformations, which map the BM to another BM.

Proposition 4.5. *The following are Brownian motions:*

(a) *time shift:* $W_1(t) = B(t + c) - B(t),\ \ c > 0$,

(b) *Brownian scaling:* $W_2(t) = B(ct)/\sqrt{c},\ \ c > 0$,

(c) *time reversal on* $[0, 1]$*:* $W_3(t) = B(1 - t) - B(1),\ \ t \in [0, 1]$,

(d) $W_4(t) = tB(1/t)$ *with the convention* $W_4(t) = 0$.

Checking the BM axioms in (a), (b), (c) is easy. For (d) one can use the Gaussian property (v), but there is also need to verify the a.s. continuity at $t = 0$; as a partial check one can verify that $\mathrm{Var}\ W_4(t) \to 0$ whence $W_4 \stackrel{\mathbb{P}}{\to} 0$ as $t \to 0$.

4.5 Existence of the BM

There are various ways to construct the BM by manipulating a countable reserve of independent random variables defined on some probability space. We take first a less constructive approach, just showing that the BM, as a process satisfying (i)-(iv), does exist. With the continuity condition (ii) omitted, the existence of the Gaussian process with required distributional properties (cf Proposition

4.3) follows from Kolmogorov's extension theorem, since the finite-dimensional distributions are consistent. However, the existence of a version with continuous paths needs effort.

Let us first use Kolmogorov's theorem to set up a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to define a process $(B(t), t \in \mathbb{Q} \cap [0,1])$ with time parameter running over the rational numbers within $[0,1]$. If it turns that the paths are continuous on $\mathbb{Q}_1 := \mathbb{Q} \cap [0,1]$, then the paths can be uniquely interpolated to continuous functions on the whole $[0,1]$. Consider

$$\Delta_n := \sup_{s,t \in \mathbb{Q}_1, |s-t| \le 1/n} |B(t) - B(s)|.$$

Proving $\Delta_n \overset{\text{a.s.}}{\to} 0$ would be enough to show the continuity.

It is convenient to consider a better tractable quantity

$$Y_{k,n} = \sup_{s,t \in \mathbb{Q} \cap \left[\frac{k-1}{n}, \frac{k}{n}\right]} |B(t) - B(s)|, \quad k = 1, \dots, n,$$

which gives upper bound

$$\Delta_n \le 3 \max_{1 \le k \le n} Y_{k,n}.$$

As a check, for instance, if $s < k/n \le t$,

$$|B(t) - B(s)| \le \left| B(t) - B\left(\frac{k-1}{n}\right) \right| + \left| B\left(\frac{k}{n}\right) - B\left(\frac{k-1}{n}\right) \right| + \left| B(s) - B\left(\frac{k}{n}\right) \right|,$$

which explains the factor 3 above.

Clearly,

$$\mathbb{P}(\max_{1 \le k \le n} Y_{k,n} \ge \varepsilon) \le \sum_{k=1}^{n} \mathbb{P}(Y_{k,n} \ge \varepsilon) = n\, \mathbb{P}(Y_{1,n} \ge \varepsilon),$$

the latter by the stationarity of increments.

The process $(B(t), t \in \mathbb{Q}_1)$ is a martingale, because the increments have mean zero. To comply with the setting of Lecture 3, where we only discussed discrete-time martingales with time-range $\mathbb{N}$, we may just use the fact that the sequence $Z_1 = B(t_1), \dots, Z_n = B(t_n)$ has the martingale property for arbitrary $t_1 < \dots < t_n$ in $\mathbb{Q}_1$. The function $z \mapsto z^4$ is convex, hence $(B(t)^4, t \in \mathbb{Q}_1)$ is a submartingale. Applying the maximal inequality (Proposition 1.13 of Lecture 3), we have estimate

$$\mathbb{P}(Y_{1,n} \ge \varepsilon) = \mathbb{P}(\max_{t \in \mathbb{Q}_1 \cap [0,\, 1/n]} |B(t)) \ge \varepsilon) \le \varepsilon^{-4}\, \mathbb{E}\, B(1/n)^4.$$

By the scaling property, $\mathbb{E}\, B^4(t) = t^2\, \mathbb{E}\, B(1) = 3t^2$, as a consequence of the familiar moments formula $\mathbb{E}[B(1)]^{2k} = (2k-1)(2k-3) \cdot \ldots \cdot 3 \cdot 1$ for $\mathcal{N}(0,1)$. It follows that

$$\mathbb{P}(Y_{1,n} \ge \varepsilon) \le \frac{3}{n^2 \varepsilon^4} \quad \Rightarrow \quad \mathbb{P}(\max_{1 \le k \le n} Y_{k,n}) < \frac{3}{n \varepsilon^4} \to 0, \quad \text{as } n \to \infty.$$

This implies $\Delta_n \overset{\mathbb{P}}{\to} 0$, but since Δ_n decreases with n, we have a stronger $\Delta_n \overset{\text{a.s.}}{\to} 0$, as wanted[(1)].

Replacing power 4 in our argument by other powers, we can conclude that

$$\sup_{t,s \in [0,1]} |B(t) - B(s)| < C|t-s|^{\alpha},$$

with some random variable C and $0 < \alpha < 1/2$. This property of the Brownian path is known as *the local Hölder continuity* with exponent α.

[(1)]We use that $\xi_1 \le \xi_2 \le \cdots$, and $\xi_n \overset{\mathbb{P}}{\to} \xi$ imply $\xi_n \overset{\text{a.s.}}{\to} \xi$.

If function f is continuously differentiable on $[0,1]$, then f is locally Hölder-continuous with exponent $\alpha = 1$, where for C we can choose $\max_{t\in[0,1]} |f'(t)|$. The functions you studied in Analysis are all continuously differentiable, perhaps excluding some set of isolated singular points. Compared with that, the Brownian path is a very odd function, because it is *nowhere differentiable with probability one*.

To see that there is no derivative at $t = 0$, note that the process $(B(t)/t,\ t \geq 0)$ has the same distribution as $(tB(1/t)/t,\ t \geq 0) = (B(1/t),\ t \geq 0)$ (see Proposition 1.4, part (d)). But $B(1/t)$ has no limit value, as $t \to \infty$, with probability one (see Exercises), thus $B(t)/t$ does not converge for $t \to 0$, which along with $B(0) = 0$ means that the derivative at 0 does not exist. The same argument together with the shift property (a) shows that the BM is not differentiable in any other point $t > 0$, and with some more effort nowhere differentiability follows.

If a differentiable function has $f' \neq 0$ then the set of solutions to $f(t) = 0$ is discrete. To highlight the complexity of Brownian path, we mention here that the zero set $\mathcal{Z} = \{t : B(t) = 0\}$ of BM is a closed set of Lebesgue measure 0, with no isolated points (similarly to the Cantor set). It is possible to define on a measure (so called local time) supported by $\mathcal{Z}$ that characterises the time spent by a path near 0.

4.6 A random series construction

We sketch a construction of the BM based on a random orthogonal series of functions. The space of measurable, square-integrable functions $L^2([0,1], \mathcal{B}([0,1], \lambda)$ (λ - the Lebesgue measure) endowed with the scalar product

$$\langle f, g\rangle = \int_0^1 f(s)g(s)ds$$

is a Hilbert space, in many respects analogous to finite-dimensional Euclidean spaces $\mathbb{R}^n$. A sequence of functions $\psi_k,\ k \in \mathbb{Z}_+$, is an orthonormal basis of the Hilbert space if $\langle \psi_i, \psi_i\rangle = 1$, $\langle \psi_i, \psi_j\rangle = 0$ for $i \neq j$, and the completeness property holds, that is $\langle f, \psi_j\rangle = 0$ for all $j \in \mathbb{Z}_+$ implies $f = 0$ a.s. For instance, the system of trigonometric functions $\{1, \sqrt{2}\sin(2\pi js), \sqrt{2}\cos(2\pi js); j = 1, 2, \cdots\}$ is an orthonormal basis.

Then, for $\xi_0, \xi_1, \ldots$ i.i.d. $\mathcal{N}(0,1)$-distributed r.v.'s, and $\psi_0, \psi_1, \ldots$ orthonormal basis, the series

$$B(t) = \sum_{j=0}^{\infty} \xi_j \int_0^t \psi_j(s)ds, \ \ t \in [0,1]$$

converges uniformly (almost surely) and defines a BM on $[0,1]$. To show the convergence one needs to verify that the series of variances converges.

4.7 The arcsine laws

A random variable ξ has the arcsine distribution if its cumulative distribution function is

$$\mathbb{P}(\xi \leq x) = \frac{2}{\pi} \arcsin \sqrt{t}, \ \ 0 \leq t \leq 1.$$

Equivalently, the density is

$$f_\xi(x) = \frac{1}{\pi\sqrt{t(1-t)}}, \ \ \ 0 < t < 1.$$

This belongs to the family of beta distributions.

The arcsine distribution apppears as

(i) the distribution of the last zero Z of the BM on $[0, 1]$

$$Z = \max\{t \in [0, 1] : B(t) = 0\}.$$

(ii) the distribution of the time spent on the positive side

$$T := \int_0^1 1(B(t) > 0)\mathrm{d}t.$$

4.8 BM as Markov process and a martingale

Define $\mathcal{F}_t$ to be the σ-algebra generated by $(B(s),\ s \leq t)$. The family $(\mathcal{F}_t,\ t \geq 0)$ is the natural filtration of the BM.

For $g : \mathbb{R} \to \mathbb{R}$ any measurable function and $s < t$ we have the identity

$$\mathbb{E}[g(B(t))|\mathcal{F}_s] = \mathbb{E}[g(B(t))|B(s)],$$

which is one way to express that the BM is a *Markov process* taking values in the 'continuous state-space' $\mathbb{R}$.

The distribution of future values of the BM after time s depends on $(B(u),\ u \leq s)$ only through $B(s)$. That is to say, conditionally on $B(s) = x$ (a present, time-s state), the future (after-s) values of the process are independent of the BM values $B(u_1), \dots, B(u_k)$ for arbitrary choice of past times $u_i < s$. Note that given $B(s) = x$, the conditional distribution of $B(t)$ is $\mathcal{N}(x, t - s)$, as is seen from the decomposition

$$B(t) = B(s) + (B(t) - B(s))$$

into independent terms $B(s)$ and $B(t) - B(s)$, therefore

$$\mathbb{E}[g(B(t))|B(s) = x] = \int_{-\infty}^{\infty} g(y)p(t - s, x, y)dy.$$

Here, the function of three variables

$$p(\tau, x, y) := \frac{1}{\sqrt{2\pi\tau}} \exp\left(\frac{-(y - x)^2}{2\tau}\right)$$

is the *transition density* of the BM. Viewed as a function of y for fixed $\tau > 0$ and fixed $x \in \mathbb{R}$, this is the conditional p.d.f. (probability density function) of $B(s + \tau)$ given $B(s) = x$.

Axioms (iii), (iv) in the definition of the BM can be equivalently replaced by the condition that

(vi) $(B(t),\ t \geq 0)$ *is a Markov process with transition density function* $p(\tau, x, y)$ *for moving from state* x *to state* y *over time interval of length* $\tau > 0$.

In our notation, τ is a dummy variable for the time increment $t - s$. That the transition density depends on s, t (with $s < t$) only through $t - s$ is the feature called time-homogeneity of the Markov process. The BM is also state-homogeneous, in the sense that $p(\tau, x, y)$ depends on x, y only through $|x - y|$.

Choosing for g the identity function we arrive at the identity

$$\mathbb{E}[B(t)|\mathcal{F}_s] = \mathbb{E}[B(t)|B(s)] = B(s), \quad 0 \leq s < t,$$

which means that $(B(t),\ t \geq 0)$ is a continuous-time martingale, considered along with its natural filtration.

4.9 Finite dimensional distributions

Let $0 = t_0 < t_1 < \cdots < t_k$. The joint p.d.f. of the vector of increments $(B(t_1) - B(t_0), \ldots, B(t_k) - B(t_{k-1}))$ is clear from the axioms (i), (iii), (iv):

$$(y_1, \ldots, y_k) \mapsto \prod_{j=1}^{k} \frac{1}{\sqrt{2\pi(t_j - t_{j-1})}} \exp\left(\frac{-y_k^2}{2(t_k - t_{k-1})}\right),$$

which is just a product of one-dimensional normal p.d.f.'s. Changing variables to $x_1 = y_1, x_2 = y_1 + y_2, \ldots, x_k = y_1 + \cdots + y_k$ (and so $y_j = x_j - x_{j-1}$) and observing that the Jacobian of the linear transformation is 1, we arrive at the joint p.d.f. of $(B(t_1), \ldots, B(t_k))$

$$(x_1, \ldots, x_k) \mapsto \prod_{j=1}^{k} p(t_j - t_{j-1}, x_{j-1}, x_j),$$

written in terms of the transition density function. This formula for p.d.f.'s of the finite-dimensional distributions is yet another way to express axioms (iii), (iv).

4.10 The quadratic variation

For function $f : [a, b] \to \mathbb{R}$ the *variation of order* $\beta > 0$ *on the interval* $[a, b]$ is defined as the supremum

$$V_\beta(f; a, b) := \sup \sum_{i=1}^{k} |f(t_i) - f(t_{i-1})|^\beta,$$

over all partitions $a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b$ of the interval $[a, b]$. We call $V_1(f; a, b)$ the variation of f, and $V_2(f; a, b)$ the quadratic variation of f on $[a, b]$.

If f is continuously differentiable, its variation is

$$V_1(f; a, b) = \int_a^b |f'(t)| dt,$$

but $V_\beta(f; a, b) = 0$ for $\beta = 2$ or any other $\beta > 1$.

For the Brownian motion the situation is radically different. To save notation, let us focus on $[0, 1]$, and consider the uniform partition by points $t_j = j/n,\ j = 0, 1, \ldots, n$. We have as $n \to \infty$

$$\mathbb{E} \sum_{i=1}^{n} |B(i/n) - B((i-1)/n)| = n\, \mathbb{E}\, |B(1/n)| = \frac{n}{\sqrt{n}}\, \mathbb{E}\, |B(1)| = \sqrt{n}\sqrt{\frac{\pi}{2}} \to \infty.$$

This can be pursued to show that $V_1(B; a, b) = \infty$ a.s., the variation of the BM is infinite.

Let us assess the quadratic variation of the BM. We have

$$\mathbb{E} \sum_{i=1}^{n} |B(i/n) - B((i-1)/n)|^2 = n\, \mathbb{E}\, B(1/n)^2 = 1,$$

and

$$\text{Var} \sum_{i=1}^{n} |B(i/n) - B((i-1)/n)|^2 = n \text{Var}[B(1/n)^2] = \frac{n}{n^2} \text{Var} B(1)^2 \to 0.$$

Thus, as $n \to \infty$,

$$\sum_{i=1}^{n} |B(i/n) - B((i-1)/n)|^2 \xrightarrow{\mathbb{P}} 1,$$

and with some effort one shows that the limit holds in the sense 'almost surely'.

Using the scaling property, we see that the quadratic variation of the BM on $[0, t]$ is equal to t. In Probability Theory the quadratic variation is also denoted by angular brackets, thus

$$\langle B\rangle(t) = t.$$

Warning. One should not confuse the variance and the quadratic variation. Formula $\mathrm{Var}B(t) = t$ features a single rv $B(t)$. But $\langle B\rangle(t) = t$ is the quadratic variation accumulated over the time interval $[0, t]$, which is the property of a path $(B(s),\ s \leq t)$.

Written symbolically, the quadratic variation property of the BM is the differential rule

$$(dB(t))^2 = dt.$$

This underlies the formulas for stochastic integrals like

$$\int_0^t B(s)dB(s) = \tfrac{1}{2}B(t)^2 - \tfrac{1}{2}t,$$

where the second term appears due to the nontrivial quadratic variation. To compare with classic Calculus, for differentiable f with $f(0) = 0$

$$\int_0^t f(s)df(s) = \int_0^t f(s)f'(s)ds = \tfrac{1}{2}f(t)^2.$$

Theorem 4.6. (Lévy's characterisation) *Let $(M(t),\ t \geq 0)$ be a martingale (relative to the natural filtration of the process) satisfying*

(i) $M(0) = 0$,

(ii) *the paths are continuous,*

(iii) $\langle M\rangle(t) = t,\ t \geq 0$.

Then $(M(t),\ t \geq 0)$ is a Brownian motion.

Thus the the normal distribution and independence of increments can be concluded from the martingale property and formula (iii) for the quadratic variation.

Literature

1. T. Liggett, Continuous time Markov processes, AMS 2010.

2. S. Shreve, Stochastic calculus for finance, vol. II, Springer 2004.

Exercises

1. Write in terms of the density functions, what it means for distributions of $B(1/2)$ and $(B(1/2), B(3/2))$ to be consistent.

2. Show that the process $(X^k(t) - t,\ t \geq 0)$ is a martingale for (a) $k = 1$ and X the Poisson process, (b) $k = 2$ and X the BM.

3. Give a detailed proof of Proposition 1.3.

4. For constants $\mu, \sigma > 0$ the process $B_{\mu,\sigma}(t) := \mu t + \sigma B(t)$ is 'a Brownian motion with drift μ and diffusion/volatility σ'. Show that $B_{\mu,\sigma}$ is Gaussian, find its mean and covariance functions. Also find the quadratic variation of the process on $[0, t]$.

5. The process $B^\circ(t) := B(t) - tB(1), \quad t \in [0,1]$, is known as the Brownian bridge. Find the covariance function of B°. Is this process Gaussian? Markov? Martingale? Explain your answers.

6. Prove that $B(t)$ has no limit as $t \to \infty$ almost surely. Hint: it is enough to show that $B(n) - B(n-1)$ is not a Cauchy sequence, $n \in \mathbb{N}$.

7. Show that $\limsup_{t\to\infty} B(t)/\sqrt{t} = \infty$. [Hint: use Kolmogorov's 0-1 law.