# Probability Lecture Notes

Dec 18

## Brownian Motion

**Theorem 1** (Wiener, 1923)**.** *(Check Last Lecture) $t \mapsto B(t)$ is continuous almost surely (a.s.), i.e.,*

$$P(\{\omega \in \Omega : t \mapsto B(t, \omega) \text{ is continuous}\}) = 1$$

*Proof Sketch.* (Continuation) We constructed $B(d)$ for $d \in \mathcal{D}$, where $\mathcal{D}$ is the set of dyadic numbers in $[0, 1]$.

$$\mathcal{D} = \bigcup_{n \geq 0} D_n, \quad D_n = \left\{ \frac{k}{2^n} : 0 \leq k \leq 2^n \right\}$$

**Construction:** Let $Z_d$ be i.i.d. $\sim \mathcal{N}(0, 1)$.

On $D_0 = \{0, 1\}$, assume $B(0) = 0$. Let $B(1) = Z_1$.

For $d \in D_n \setminus D_{n-1}$, we define $B(d)$ assuming $B(d)$ is defined for $D_{n-1}$. Let $d^-, d^+ \in D_{n-1}$ be the neighbors of $d$ in $D_{n-1}$. We define:

$$B(d) = \frac{B(d^+) + B(d^-)}{2} + \frac{Z_d}{2^{\frac{n+1}{2}}}$$

The set $\{B(d) : d \in \mathcal{D}\}$ satisfies the independent, normally distributed increments conditions on $\mathcal{D}$. Condition 1 of BM holds for times in $\mathcal{D}$.

We extend $B(d)$ to $\mathbb{R}$ by linear interpolation. Since $\mathcal{D}$ is countable, we define a function $B(t)$ for $t \in [0, 1]$ by actually interpolating the points $\{B(d) : d \in D_n\}$ and passing to the limit as $n \to \infty$.

Define $F_n : [0, 1] \to \mathbb{R}$. For $n = 0$:

$$F_0(t) = \begin{cases} 0 & t = 0 \\ Z_1 & t = 1 \\ \text{linear in between} & \end{cases}$$

For all $n \geq 0$:

$$F_n(t) = \begin{cases} 0 & t \in D_{n-1} \\ \frac{Z_d}{2^{(n+1)/2}} & t \in D_n \setminus D_{n-1} \\ \text{linear in between} & \end{cases}$$

Note that for $m > n$ and $d \in D_n$, $F_m(d) = 0$ by definition.

**Claim 1.** *For all $d \in \mathcal{D}$:*

$$B_n(d) = \sum_{i=0}^{n} F_i(d)$$

Indeed, using induction. $n = 0$ satisfies the condition. Let $d \in D_n \setminus D_{n-1}$.

$$\sum_{i=0}^{n} F_i(d) = \sum_{i=0}^{n-1} F_i(d) + F_n(d)$$

By hypothesis, the sum up to $n - 1$ is linear on the interval $[d^-, d^+]$, so at the midpoint $d$, it is the average:

$$= \frac{B(d^-) + B(d^+)}{2} + \frac{Z_d}{2^{\frac{n+1}{2}}}$$

$$= B(d) \quad \text{(by definition)}$$

By induction, the hypothesis holds. Therefore,

$$B(d) = \sum_{n=0}^{\infty} F_n(d), \quad \forall d \in \mathcal{D}$$

This suggests the following definition:

$$B(t) = \sum_{n \geq 0} F_n(t), \quad \forall t \in [0,1]$$

We verify if the series converges uniformly. Since $F_n$ are continuous, if the convergence is uniform, $B(t)$ will be continuous.

**Claim 2.** *The series converges almost surely.*

**Observation:** If $X \sim \mathcal{N}(0,1)$, then $\forall \lambda > 0$:

$$P(|X| \geq \lambda) = 2\int_{\lambda}^{\infty} e^{-\frac{u^2}{2}} \frac{du}{\sqrt{2\pi}} \leq 2\int_{\lambda}^{\infty} \frac{u}{\lambda} e^{-\frac{u^2}{2}} \frac{du}{\sqrt{2\pi}} = \frac{2}{\lambda\sqrt{2\pi}} \left[-e^{-\frac{u^2}{2}}\right]_{\lambda}^{\infty} = \frac{2}{\lambda\sqrt{2\pi}} e^{-\frac{\lambda^2}{2}} \leq e^{-\frac{\lambda^2}{2}}$$

(for large $\lambda$, constant adjusted). Since $Z_d \sim \mathcal{N}(0,1)$:

$$P(|Z_d| \geq c\sqrt{n}) \leq e^{-\frac{c^2 n}{2}}$$

We consider the event where $\max_{d \in D_n} |Z_d| \geq c\sqrt{n}$.

$$\sum_{n=0}^{+\infty} P(\exists d \in D_n \mid |Z_d| \geq c\sqrt{n}) \leq \sum_{n=0}^{+\infty} \sum_{d \in D_n} P(|Z_d| \geq c\sqrt{n})$$

$$\leq \sum_{n=0}^{+\infty} 2^n e^{-\frac{c^2}{2} n} = \sum_{n=0}^{+\infty} e^{n(\ln 2 - \frac{c^2}{2})}$$

For convergence, we require $\ln 2 < \frac{c^2}{2} \Rightarrow c > \sqrt{2 \ln 2}$. By the Borel-Cantelli lemma, for almost every $\omega$, there exists an $N(\omega) \in \mathbb{N}$ such that if $n \geq N(\omega)$, then the event $\{\exists d \in D_n : |Z_d| \geq c\sqrt{n}\}$ does not hold.

Therefore, for almost every $\omega$, for $n \geq N(\omega)$:

$$\sup_{t \in [0,1]} |F_n(t)| \leq c\sqrt{n} \frac{1}{2^{\frac{n+1}{2}}}$$

This implies:

$$\sum_{n \in \mathbb{N}} c\sqrt{n} \frac{1}{2^{\frac{n+1}{2}}} < +\infty$$

Since $F_n$'s are continuous on $[0,1]$, by the Weierstrass $M$-test, $\sum_{n=0}^{+\infty} F_n(t)$ converges uniformly to a continuous function $B(t)$.

**Exercise 1.** *Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of Gaussian random vectors, and assume that $X_n \to X$ almost surely. If $\mathbb{E}[X_n] \to b$ and $cov(X_n) \to \Sigma$ (where $X_n$'s are assumed to be d-dimensional), then $X$ is a Gaussian random vector with mean $b$ and covariance matrix $\Sigma$. i.e., $\mathbb{E}[X] = b$ and $cov(X) = \Sigma$. [Hint: For simplicity, assume $b = 0$, and means $\mathbb{E}[X_n] = 0$].*

It remains to verify the increments property for $B(t)$, $t \in [0,1]$. Set $0 < t_1 < t_2 < \cdots < t_{n-1} < t_n \in [0,1]$. Since the set $\mathcal{D}$ is dense in $[0,1]$, there exists a sequence $t_{i,k} \nearrow t_i$ where $t_{i,k} \in \mathcal{D}$. Consider the increment vectors $X_k = (B(t_{i+1,k}) - B(t_{i,k}) \mid 1 \leq i < n)$. $X = (B(t_{i+1}) - B(t_i) \mid 1 \leq i < n)$. We assert that $X_k \to X$ almost surely because $B$ is continuous almost surely. Also $\mathbb{E}[X_k] = 0$, $\text{cov}(X_k) = \text{diag}(t_{i+1,k} - t_{i,k}) \to \text{diag}(t_{i+1} - t_i)$. Remark that we meet the conditions of exercise (!). It follows that $X$ is a Gaussian random vector of mean 0 and $\text{cov}(X) = \text{diag}(t_{i+1} - t_i)$. Since the covariance matrix is diagonal, the entries of $X$ are independent. Therefore, condition 1 of BM holds.

We have constructed BM on $[0,1]$ as a random function $B : \Omega \to (C([0,1]), ||\cdot||_\infty)$. □

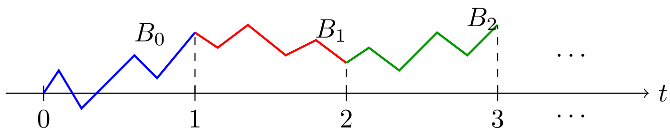


Figure 1: Glueing of independent Brownian Motions

## Glueing Brownian Motions

Let $B_0, B_1, \dots, B_n, \dots$ be a sequence of i.i.d. Brownian motions (i.e., sequence of measurable functions) on $[0,1]$. We glue them continuously into a function on $[0,+\infty)$.

We pose:

$$B(t) = B_{\lfloor t \rfloor}(t - \lfloor t \rfloor) + \sum_{i=0}^{\lfloor t \rfloor - 1} B_i(1)$$

One can check that this is a glueing. $\lfloor t \rfloor$ is the integer part of $t$.

## Properties of Brownian Motion

1. **Hlder Continuity:** Brownian motion $t \mapsto B(t)$ is actually locally Hlder continuous with Hlder exponent $\alpha < 1/2$ (meaning this applies when restricted to a compact interval like $[0,1]$). More precisely, given $\alpha < 1/2$, we have $|B(t) - B(s)| \le C_{\alpha,\omega}|t-s|^\alpha$ almost surely, $\forall t, s \in [0,1]$.
2. **Not 1/2-Hlder:** BM is not 1/2-Hlder continuous.
3. **Nowhere Differentiable:** BM is nowhere differentiable almost surely (like the Weierstrass function). [Theorem of Paley, Wiener, Zygmund 1933].

## Random Walks

A random walk is a process $(S_n)_{n\ge 0}$ where $S_0 = 0$ and the increments $S_n - S_{n-1}$ are independent.

$$S_n = X_1 + \cdots + X_n$$

$S_0 = 0$, $X_n = S_n - S_{n-1}$ are i.i.d. random variables. Assume $\mathbb{E}[X_n] = 0$, $\mathrm{var}(X_n) = \sigma^2 > 0$, $\forall n \in \mathbb{N}$ (we can always assume $\sigma^2 = 1$). A random walk is in some sense a discrete BM in that the increments are independent with mean 0 and variance 1 (but not necessarily normally distributed).

## Definitions

**Definition 1** (Hlder Continuity)**.** *Let $M$ be a metric space. $f : M \to \mathbb{R}$ is called $\alpha$-Hlder continuous if $\forall x, y \in M$:*

$$|f(x) - f(y)| \le C d(x,y)^\alpha \quad (\alpha \in [0,1])$$

*for some $C \in \mathbb{R}_+^*$.*

- **Stopping time filtration:** $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. Replacing discrete time $(n)$ by continuous time, we get a Filtration $\mathfrak{F} = \{\mathcal{F}_t\}_{t\ge 0}$, a family of sub-algebras s.t. $s < t \implies \mathcal{F}_s \subseteq \mathcal{F}_t$.
- **Stopped process:** $\{X(t)\}_{t\ge 0}$ s.t. $X(t)$ is $\mathcal{F}_t$ measurable.
- **Natural filtration** $\mathcal{F}^0 = \{\mathcal{F}^0(t)\}_{t\ge 0}$ of a process $X = \{X_t\}_{t\ge 0}$: $\mathcal{F}^0(t) = \sigma(X(s) \mid s \le t)$. Clearly, $X$ is adapted to $\mathcal{F}^0$.
- **Right continuous filtration** of a random process $X = \{X(t)\}_{t\ge 0}$: $\mathfrak{F} = \{\mathcal{F}(t)\}_{t\ge 0}$, $\mathcal{F}(t) = \bigcap_{s>t} \mathcal{F}^0(s) \supseteq \mathcal{F}^0(t)$. $\mathfrak{F}$ has the property $\bigcap_{s>t} \mathcal{F}(s) = \mathcal{F}(t)$.
- **Adapted stopping time:** $T : \Omega \to [0,+\infty]$ relative to $\mathcal{F} = \{\mathcal{F}(t)\}_{t\ge 0}$: $\{T \le t\} \in \mathcal{F}(t)$.

**Properties of Adapted Process:** 1. $X$ is $\mathcal{F}$-adapted. 2. $\mathbb{E}[|X(t)|] < \infty$, $\forall t$. 3. $\forall s < t, \mathbb{E}[X(t) \mid \mathcal{F}(s)] = X(s)$.

## Theorems

**Exercise 2.** *Brownian motion is a martingale process relative to its natural filtration.*

*Hint:* Independence of increments gives independence from the past.
$\mathbb{E}[B(t) - B(s) \mid \mathcal{F}(s)] = \mathbb{E}[B(t) - B(s)] = 0 \implies \mathbb{E}[B(t) \mid \mathcal{F}(s)] = B(s)$.

**Definition 2.** *We say that a random variable $X$ with $\mathbb{E}[X] = 0$, $var(X) = \sigma^2 < +\infty$ can be embedded into BM if there exists a stopping time $T : \Omega \to [0, +\infty]$ (adapted to the natural filtration $\sigma(B(t))$) s.t. $\mathbb{E}[T] < +\infty$ and $B(T) =^d X$. (Note: $X(\omega)$ is $B(T(\omega))$. When $T$ is fixed, $B(T)$ is normal, but when the time is random, the distribution can be anything reasonable).*

**Theorem 2** (Skorokhod's Embedding Theorem)**.** *Any random variable $X$ with $\mathbb{E}[X] = 0$ and $\mathbb{E}[X^2] < +\infty$ can be embedded into a standard BM.*

**Theorem 3.** *Let $S_n = X_1 + \cdots + X_n$ be a random walk with $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X_1^2] = 1$. Then there exists a sequence of stopping times $T_n$ with respect to $\mathcal{F}^+$ s.t. $\{B(T_n) \mid n \geq 0\}$ has the distribution of $\{S_n \mid n \geq 0\}$.*

# Functional CLT

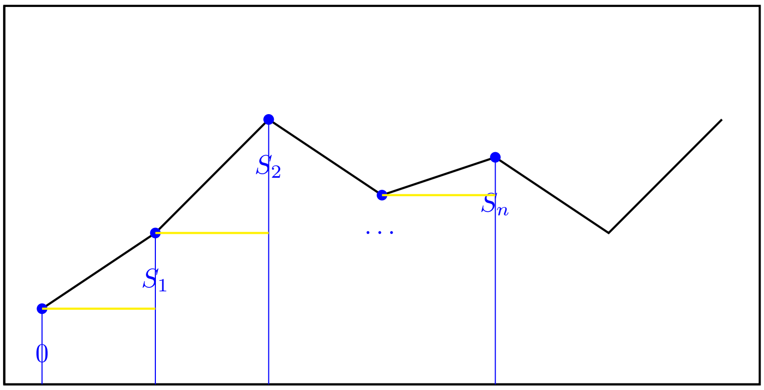


Figure 2: Random Walk Interpolation

The limit of the interpolation between points of the random walk is (in distribution) the BM.

**Setup for Functional CLTs:** Let $S_n = X_1 + \cdots + X_n$ be a random walk where $X_k$ are i.i.d. with $\mathbb{E}[X_k] = 0$ and $\mathbb{E}[X_k^2] = 1$. Let $S(t)$ be a continuous random function on $[0, +\infty)$ that interpolates $\{S_n \mid n \geq 0\}$, that is:

$$S(t) = S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)(S_{\lfloor t \rfloor + 1} - S_{\lfloor t \rfloor})$$

Define a sequence of random functions $S_n^* : [0, 1] \to \mathbb{R}$:

$$S_n^*(t) = \frac{S(nt)}{\sqrt{n}}$$

Observe for $t = 1$, the CLT implies $S_n^*(1) = \frac{S_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$.

**Theorem 4** (Donsker's Invariance Principle / Functional CLT)**.**

$$S_n^*(t) \xrightarrow{d} B(t), \quad \forall t \in [0, 1]$$

**Observations:** If $X_n \xrightarrow{d} X$ and $g$ is continuous, then $g(X_n) \xrightarrow{d} g(X)$. What we mean by this is convergence in distribution in the space $C([0, 1])$, where $B \in C([0, 1])$ and $S_n^*$ are random functions with values in $C([0, 1])$.

$$S_n^* \xrightarrow{d} B$$

as random variables in $C([0, 1])$. i.e., not just convergence for every fixed $t \in [0, 1]$, but as a sequence of functions in $C([0, 1])$. In particular, $g(S_n^*) \xrightarrow{d} g(B)$ for any $g : C([0, 1]) \to \mathbb{R}$ continuous.